Supporting Information

Markov et al. 10.1073/pnas.1218972110

SI Methods

Source and Target Locations. Source and target locations first were identified on section outlines of the individual macaque brains and then transferred to the M132 reference atlas using sulcal landmarks (1). Source and target locations were the areal geometric centers, except for areas V1, V2, and V4. The vertex defining the geometric center of each area was calculated in Caret (2) using the inflated (smoothed) atlas surface (Fig. 1).

Measurement of Interareal Distance. The 3D coordinates of the vertices in the atlas midthickness surface configuration were used to compute interareal distances. Because V1, V2, and V4 are much larger in surface area, distances to them as targets were measured with respect to the injection sites, which in all three cases were located in the representation of the fovea. We used a 3D reconstruction of the M132 atlas to measure the distances between geometric centers. Map3D (Explora Nova) software was used to define a trajectory between centers of gravity in a 3D reconstruction of the cortex. Pathways were limited to the white matter and forced to take the shortest route, thereby approximating the trajectory of axons linking the geometric centers.

Similarity Index

Our database provides systematic information about all 91 areas in terms of the outgoing connections projecting to the 29 target areas. We compared the similarity of the output pattern of pairs of areas by evaluating the number of target areas to which both project or to which neither projects (i.e., similarity implies both projections exist or are absent; dissimilarity implies one is absent and the other is present). We define a normalized out-link similarity measure, S^{out}_{xy} as follows: For any pair of areas (x,y), let n^{out}_{xy} denote the number of injected areas into which both x and y have a projecting out-link or neither x nor y has a projecting out-link. Because $n^{out}_{xy} \le 29$, we compute the ratio $n^{out}_{xy}/29$ for every area pair (x,y). Clearly, this number will depend on the out-degrees of x and y, denoted by k^{out}_x and k^{out}_y $(0 \le k^{out}_{x(y)} \le 91)$. We define the out-link similarity as $S^{out}_{xy} = n^{out}_{xy}/29 - p^{out}_{xy}$, where p^{out}_{xy} is the expected value of the ratio $(n^{out}_{xy}/29)$ if the outgoing connections of x and y were distributed uniformly at random across the 29 injected areas. Thus, $p^{out}_{xy} = (k^{out}_x/29)(k^{out}_y/29) + (1-k^{out}_x/29)(1-k^{out}_y/29)$, where the first term is the probability that both x and y project to a given target and the second term is the probability that neither of them projects to a given target. A similar measure may be devised for the in-link similarity S^{in}_{xy}. In this case, given two target areas, x and y, we count the number of common-source areas and areas that have inputs into neither of the two targets x and y. If we denote their sum by n^{in}_{xy}, we can write $S^{in}_{xy} = n^{in}_{xy}/91 - p^{in}_{xy}$. Here, we divide by 91 because there is a total of 91 possible source areas and $0 \le k^{in}_{x(y)} \le 91$. Here, $p^{in}_{xy} = (k^{in}_x/91)(k^{in}_y/91) + (1-k^{in}_x/91)(1-k^{in}_y/91)$. The out-(in-)link similarity is positive or negative according to whether there is correlation or anticorrelation between the targets of areas x and y. Because simple pairwise comparisons are noisy, we calculate a regional similarity.

Weighted Similarity

To evaluate the weighted similarity between two target areas A and B, we have computed a similarity measure based on the dot product. Let the FLNe (fraction of labeled neurons in the source area with respect to the total number of labeled neurons extrinsic to the target area) weights of source areas be recorded as ordered vectors $A = (a_1, a_2, \ldots, a_{91})$ and $B = (b_1, b_2, \ldots, b_{91})$ such that each entry represents the FLN of a given source area to the target area. The dot product between A and B is $A \bullet B = \sum a_i b_i$. We measure the cosine of the angle, θ, $cos\theta = \frac{A \bullet B}{\|A\|\|B\|}$, where $\|A\| = (\sum a_i^2)^{\frac{1}{2}}$, the lengh of A, between pairs of areas, and report the average, except in Fig. S2.

In Fig. S2, we consider all pairs of areas according to whether they are within the same region (Fig. S2*B*) or from two different regions (Fig. S2*A*). We then compute the $cos\theta$ values of their inputs. In Fig. S2 *A* and *B*, we show the cosine similarity distribution for FLN. In Fig. S2*C*, we overlay the two datasets as a density plot.

Fig. S2*D* is a plot of the average $cos\theta$ correlation matrix with FLN data in which the pairs of targets are split according to regions. Note the very high $cos\theta$ within regions (Fig. S2*D*). In this correlation matrix, the occipital region is the one with the highest self-similarity and all the other regions have greater self-similarity than similarity with other regions.

Taking out NFPs (Tables S1–S3)

Two areas (subiculum and piriform) have only NFP-type connections to the rest of the network; therefore, after removing the NFPs, there will not be a completely dominating set (those two areas become isolated). Hence, we did two types of dominating set analyses. In the first, we still considered all 91 areas; thus, the 100% domination value has zero representation, but the table itself still carries information. The second type of analysis was on the connected component only, i.e., with the 89 areas (the isolates were removed). In the second case, there is full domination (within the graph of 89 areas) but the size of the minimum dominating set (MDS) is much larger without the NFPs, MDS = 5, showing the key role the NFP plays in the domination aspect of the network.

Nodes with Strong Domination (Table S4)

We have analyzed the connectivity of the dominating sets for all three-node target groups (dominating triples) and all four-node target groups (dominating quadruplets) with strong domination. Recall, the number of dominating n-tuplets at a given domination percentage $q\%$ is found in Table S1; for example, for triplets, there are 3,654 combinations. There are 1.88%—that is, q = 0.0188 fraction of those, i.e., $3654 \times 0.0188 = 69$—that dominate $q = 100\%$ of the network (full domination). Thus for dominating triplets, we find 28 of the 69 fully dominating triplets (40%) have all possible directed links between them (six edges in total); 22 (32%) have only one edge missing; 17 (25%) have two links missing (i.e., four edges of a possible six, giving a 66% density); and only 2 (3%) have just three links (50% density). The average density within the fully dominating 69 triplets is 85%. Table S4 shows the amount of connectivity (expressed as the average density within the triplet) for triplets that dominate a $q\%$ of all the areas. We can see a tendency to have larger densities for sets that dominate a larger percentage of all areas (the larger q is, the larger the density). Although we also may see large densities for small q values, there the number of triplets is very small and the statistics are not necessarily meaningful. Thus, the smallest groups of areas with large domination also have tight connections among them, meaning information exchanges are possible within the nodes of the group. Table S4 shows these densities for quadruplets. The average density of the fully dominating quadruplets is 73.6%, and the tendency is very similar to that of the triplets. These observations further support the statement that groups of areas with high domination play an important role in global integration.

Common-Source Signatures

We have used the number of common sources as an indicator of the specificity of the projections pattern in Figs. 3*F* and 4*A*. We start from the assumption that a source area found to project to all the targets in a given region is an event unlikely to occur in a random network but indicates a very specific localization of interactions. To test this, we have evaluated the number of areas not including the new-found projections (NFPs) that project to all targets in a given lobe (white bars in Fig. 3*F*). Each area is considered to have a self-loop; therefore, if it projects to all the other targets of its own region, it is considered a common-source area. Simply by adding connections to the network, the NFPs are expected to increase the number of common sources; this is the case when we quantify the number of common sources of the network that includes the known + NFPs (blue bars in Fig. 3*F*). Accordingly, we have tested the effect of adding to each area, instead of the actual NFPs, a random selection of the same number of projections as there are NFPs but at positions chosen at random from all areas that do not constitute a previously known source (gray bars). Repeating this procedure 20,000 times leads to a much lower increase, on average, in the number of common sources than observed when the actual NFPs are added to the known connections (with the exception of the parietal cortex, in which the numbers of common sources between the actual and randomized network were not significantly different). Finally, to test the effects of specificity of NFPs, we randomized the known and the NFP labels within each weight bin of Fig. 3*F* and recomputed the number of known common areas (the total number would be unchanged). We repeated this 20,000 times (orange bars in Fig. 3*F*).

In Fig. 4*A*, we again calculate the common sources but in the frame of within- and between-region projections. In Fig. 4*A*, we compute the number of common sources from within the target lobe (all projections are taken into account). We then randomly reattribute for each of the targets the same number of source areas but taken from the entire list of areas within the region. This amounts to randomizing the matrix with degree preservation. In the randomized within-region matrix, we reevaluate the number of common sources and repeat the operation 20,000 times. The average and 90% confidence intervals are given as gray bars in Fig. 4*A*. Fig. 4*A*, *Right*, represents the same type of analysis but is restricted to the projections originating from outside the target region.

1. Markov NT, et al. (2013) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex*, 10.1093/cercor/bhs1270.

2. Van Essen DC, Ugurbil K (2012) The future of the human connectome. *Neuroimage* 62(2):1299–1310.

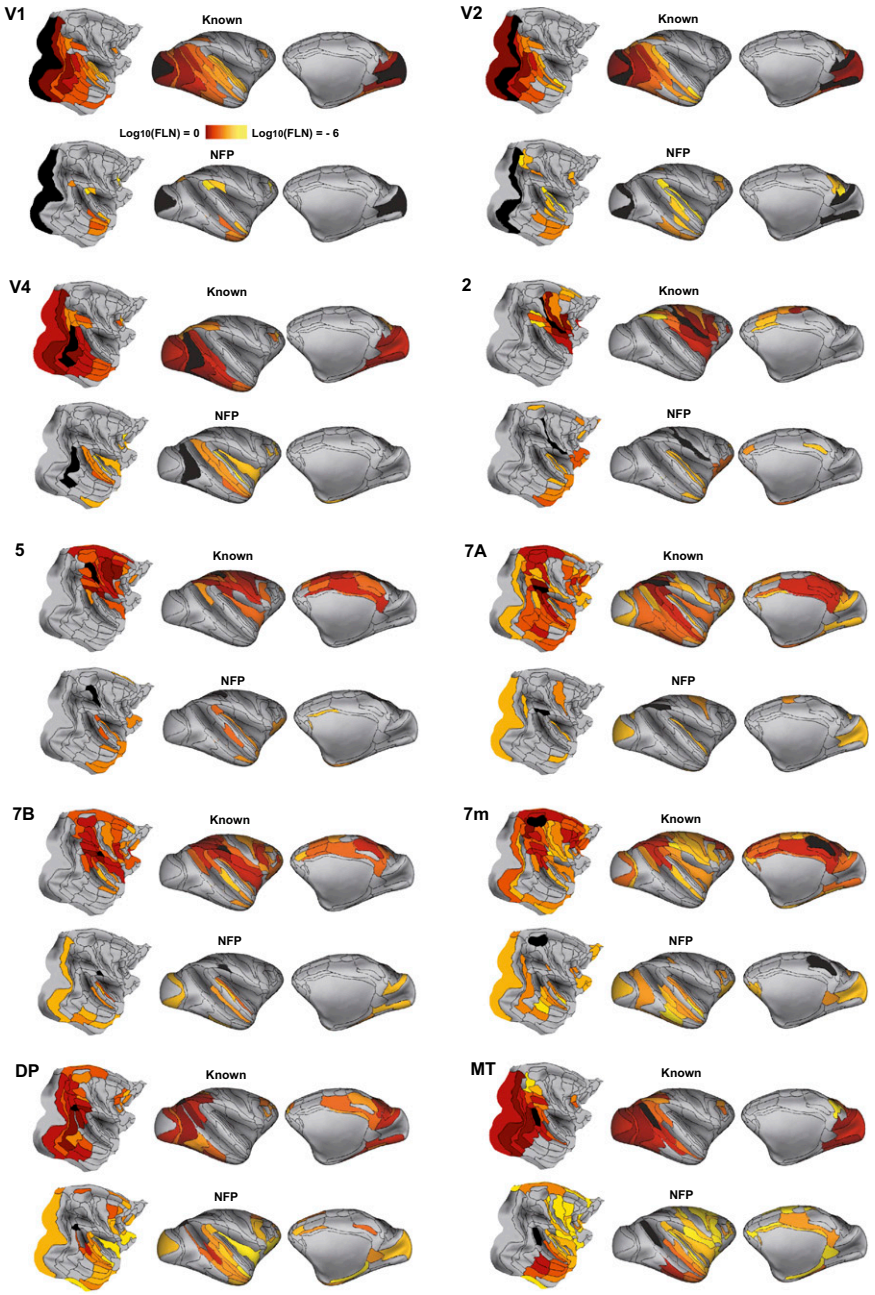

Fig. S1. (Continued)

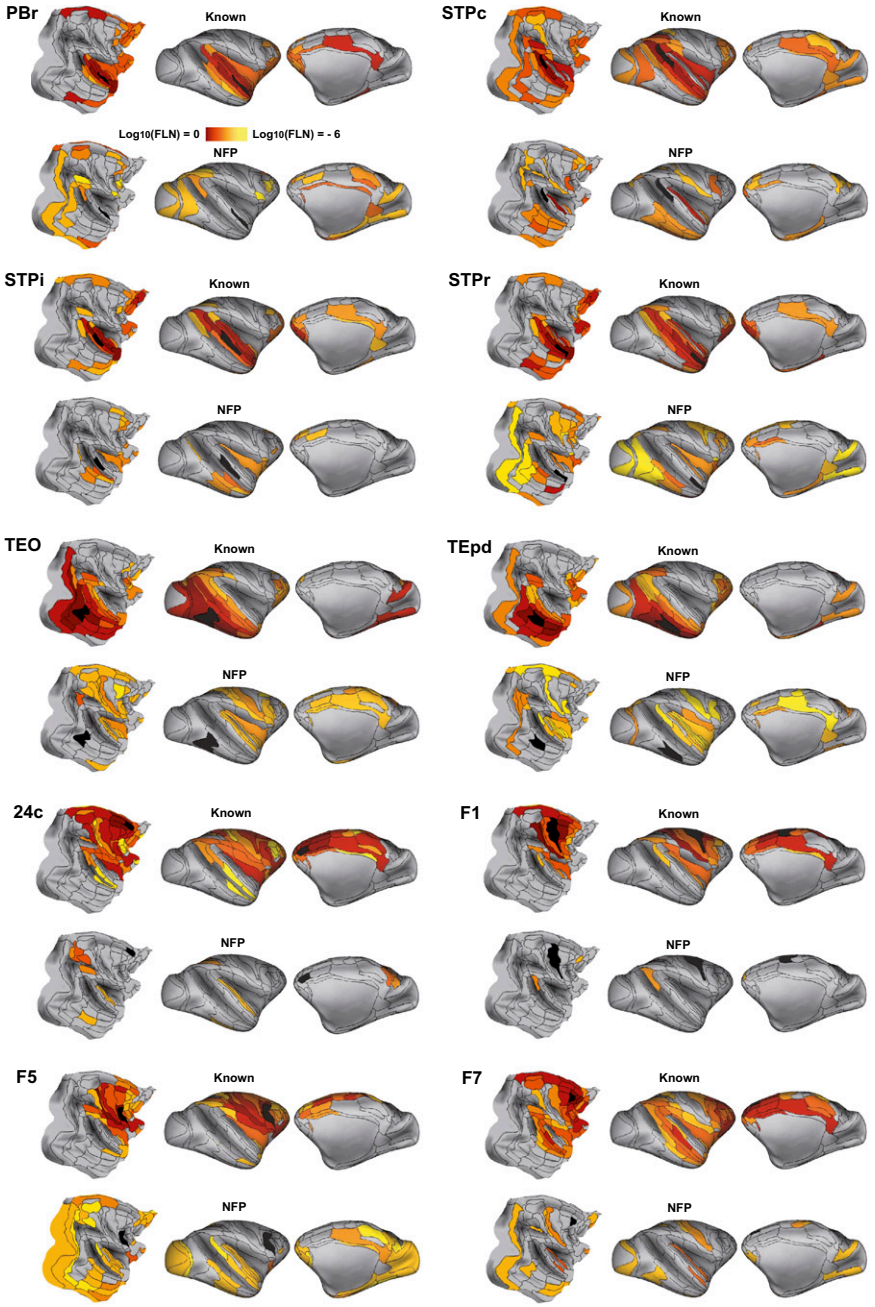

Fig. S1. (Continued)

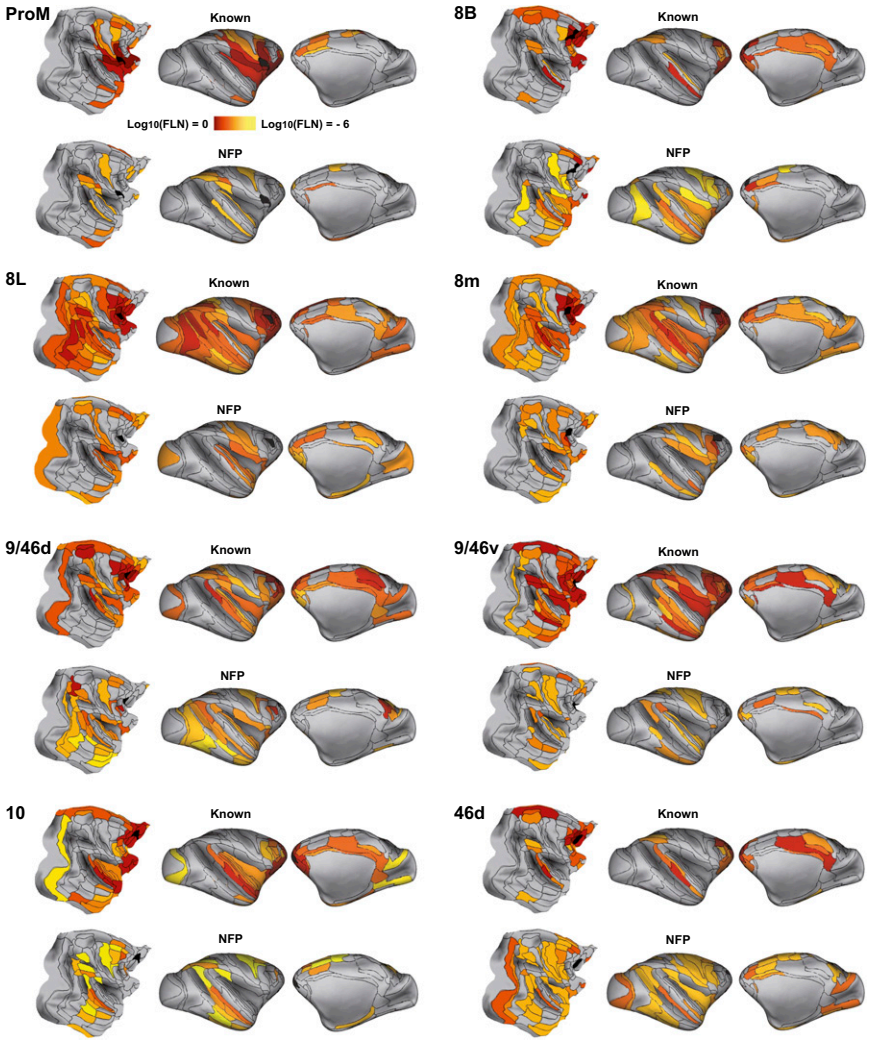

Fig. S1. Surface connectivity maps of 28 cortical injections. Flat map (*Left*) plus medial (*Center*) and lateral inflated maps (*Right*). Connection strengths are encoded as weak, moderate, or strong (reddest hues strongest).

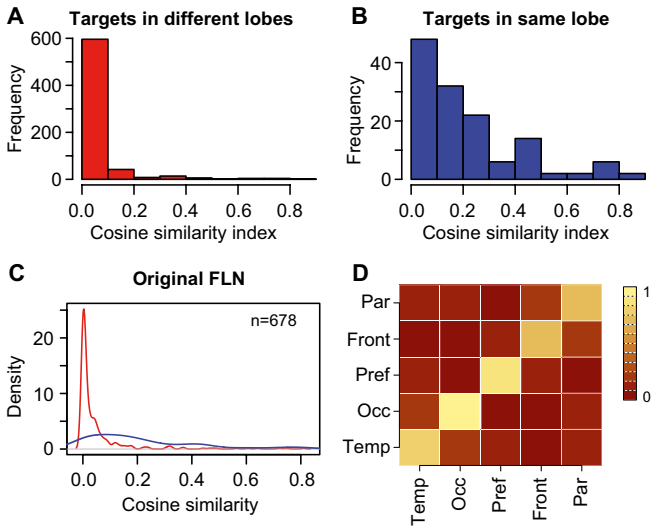

Fig. S2. Cosine similarity of targets within vs. between lobes calculated over the $\mathbf{G}_{29 \times 91}$ matrix. (*A*) *Cosθ* between pairs of targets found in different lobes. (*B*) *Cosθ* between pairs of targets found in the same lobe. (*C*) Density plot overlay of the data from *A* and *B*. FLNe values are not transformed, and the color code from *A* and *B* is preserved. (*D*) Correlation matrix of average cosine angle between pairs of target areas within a lobe, using FLNe directly. Front, frontal region; Occ, occipital region; Par, parietal region; Pref, prefrontal region; Temp, temporal region.

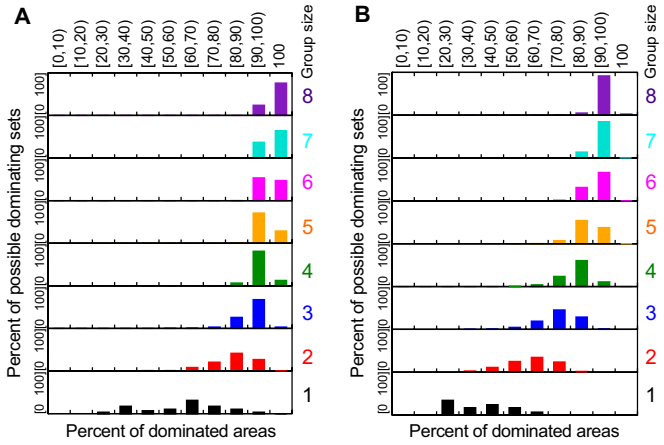

Fig. S3. The NFP significantly shapes the input statistics to area groups. Shown is a graphical representation of Tables S1 and S3 ((*A*, all links and *B*, no NFPs; respectively). Given a group size, it shows what percent of all possible groups of target combinations (*y* axis) of that size will receive inputs from a given percent (indicated by the intervals on the *x* axis) of all areas. The last column corresponds to full (100%) domination, which is severely suppressed without NFP links. In generating *B*, two areas were removed because they became isolated after removing the NFPs. For full comparisons, see Tables S1–S3.

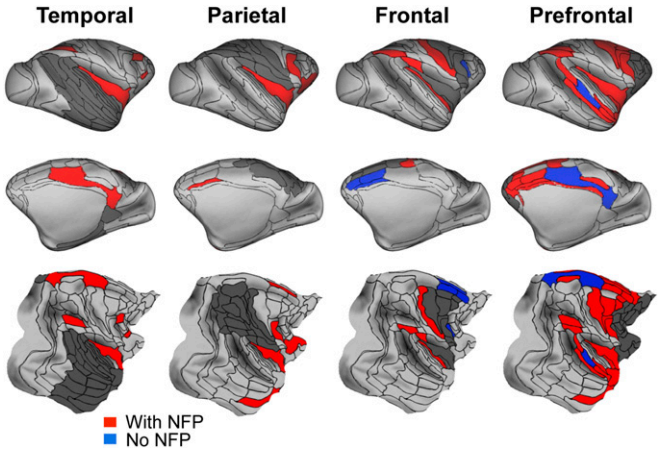

Fig. S4. Surface maps showing spatial relationships of the common projection territories. Flat maps and lateral and medial inflated maps. Regions are shaded in dark gray; areas projecting to all the injected areas of a region via known connections are in blue; areas projecting to all injected areas of a lobe via known connections plus NFPs are in red. Occipital cortex: See main text (Fig. 3*C*) for the description of the common input to this region. Temporal region: The injected areas of the temporal cortex do not share long-distance input via known connections. Inclusion of NFPs reveals common input from six areas dispersed in the occipital (area prostriata), parietal (area LIP), frontal (area insula), prefrontal (areas 45A and 8m), and limbic (area 23) regions. Inclusion of the NFPs reveals common input from five areas dispersed in the temporal (area perirhinal), frontal (areas F5 and insula), prefrontal (area 12), and limbic (area 24b) regions. Parietal region: Similarly, the injected areas of the parietal cortex share no long-distance input from known connections. Inclusion of the NFPs reveals common input from five areas dispersed in the temporal (area perirhinal), frontal (areas F5 and insula), prefrontal (area 12), and limbic (area 24b) regions. Frontal region: The injected areas in the frontal cortex share input via known connections from four areas in the prefrontal (area 44) and limbic (areas 24b, 24c, and 24d) regions. Inclusion of NFPs reveals an additional four common areas in the temporal (area MB) and parietal (areas 3, 7A, and 7op) regions. Prefrontal region: The prefrontal cortex has common input via known projections from two areas in the temporal (area STPi) and limbic (area 23) regions. Inclusion of the NFPs reveals an additional 23 common source areas concentrated in the temporal (temporal pole, perirhinal, PGa, PBr, STPc, and STPr), parietal (areas 7A and LIP), frontal (areas F1, F2, F4, F5, F6, F7, parainsula, and insula), and limbic (areas 24a, 24b, 24c, 24d, 29/30, 31, and 32) regions (see Table S5 for the complete nomenclature of the cortical areas).

Table S1. Percentage of dominating sets of all connections

	Targets*							
	1	2	3	4	5	6	7	8
No. of sets	29	406	3,654	23,751	118,755	475,020	1,560,780	4,292,145
Dominated size (%)								
0–10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10–20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20–30	3.45	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30–40	17.24	0.0	0.0	0.0	0.0	0.0	0.0	0.0
40–50	6.90	1.23	0.03	0.0	0.0	0.0	0.0	0.0
50–60	10.34	1.23	0.11	0.0	0.0	0.0	0.0	0.0
60–70	31.03	8.37	0.68	0.02	0.0	0.0	0.0	0.0
70–80	20.69	19.95	3.50	0.28	0.02	0.0	0.0	0.0
80–90	6.90	42.36	25.12	6.33	1.05	0.12	0.01	0.0
90–100	3.45	26.60	68.66	85.02	78.40	63.67	47.88	34.04
100	0.0	0.24	1.88	8.33	20.52	36.20	52.10	65.96

*"Targets" is the dominating set *D* size (number of injected sets).

Table S2. Percentage of dominating sets without NFPs and considering all 91 areas

	Targets*							
	1	2	3	4	5	6	7	8
No. of sets	29	406	3,654	23,751	118,755	475,020	1,560,780	4,292,145
Dominated size (%)								
0–10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10–20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20–30	34.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30–40	17.24	3.44	0.10	0.0	0.0	0.0	0.0	0.0
40–50	24.13	13.30	0.79	0.05	0.0	0.0	0.0	0.0
50–60	17.24	25.12	5.77	0.48	0.03	0.0	0.0	0.0
60–70	6.89	37.43	21.72	5.28	0.82	0.1	0.0	0.0
70–80	0.0	20.19	47.40	27.91	9.49	2.50	0.55	0.10
80–90	0.0	0.49	23.72	59.35	61.12	40.65	20.77	8.83
90–100	0.0	0.0	0.46	6.92	28.54	56.75	78.66	91.07
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

*"Targets" is the dominating set *D* size (number of injected sets).

Table S3. Percentage of dominating sets without NFPs and considering only 89 areas (without subiculum and piriform)

	Targets*							
	1	2	3	4	5	6	7	8
No. of sets	29	406	3,654	23,751	118,755	475,020	1,560,780	4,292,145
Dominated size (%)								
0–10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10–20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20–30	34.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30–40	17.24	2.96	0.08	0.0	0.0	0.0	0.0	0.0
40–50	24.13	11.33	0.65	0.03	0.0	0.0	0.0	0.0
50–60	17.24	25.12	4.76	0.38	0.02	0.0	0.0	0.0
60–70	6.89	34.73	19.29	4.14	0.55	0.06	0.0	0.0
70–80	0.0	23.89	45.46	23.73	7.38	1.76	0.35	0.06
80–90	0.0	1.97	28.92	60.67	54.24	31.46	14.02	5.23
90–100	0.0	0.0	0.82	11.03	37.77	66.51	84.69	91.87
100	0.0	0.0	0.0	0.0	0.02	0.02	0.95	2.84

*"Targets" is the dominating set *D* size (number of injected sets).

Table S4. Connectivity of dominating triples and quadruples for various domination sizes (percentages)

Dominated size *q* (%)	No. of triplets	Average edge density (%)	No. of quadruplets	Average edge density (%)
0–10	0	—	0	—
10–20	0	—	0	—
20–30	0	—	0	—
30–40	0	—	0	—
40–50	1	100.00	0	—
50–60	4	75.00	0	—
60–70	25	73.33	6	77.78
70–80	128	57.33	68	73.04
80–90	918	59.59	1,487	61.21
90–100	2,509	68.15	20,221	65.59
100	69	85.02	1,969	73.63

Table S5. Allocation of areas to regions

Abbreviation	Area name	Region
TEO*	Temporal area TE, occipital part	Occipital
V1*	Visual area 1, primary visual cortex	Occipital
V2*	Visual area 2	Occipital
V3	Visual area 3	Occipital
V3A	Visual area 3A	Occipital
V4*	Visual area 4	Occipital
CORE	Core auditory area	Temporal
ENTORHINAL	Entorhinal	Temporal
FST	Fundus of the superior temporal sulcus	Temporal
IPa	Intraparietal sulcus associated area in the superior temporal sulcus	Temporal
LB	Belt region of the auditory cortex, lateral part	Temporal
MB	Belt region of the auditory cortex, medial part	Temporal
MST	Medial superior temporal area	Temporal
MT*	Middle temporal area	Temporal
PBc	Parabelt region of the auditory cortex, caudal part	Temporal
PBr*	Parabelt region of the auditory cortex, rostral part	Temporal
PERIRHINAL	Perirhinal	Temporal
PGa	PG associated area of the superior temporal sulcus	Temporal
PIRIFORM	Piriform	Temporal
Pro.St.	Prostriata	Temporal
STPc*	Superior temporal polysensory area, caudal subdivision	Temporal
STPi*	Superior temporal polysensory area, intermediary subdivision	Temporal
STPr*	Superior temporal polysensory area, rostral subdivision	Temporal
SUBICULUM	Subiculum	Temporal
TEa/ma	Superior temporal sulcus ventral bank area, anterior part	Temporal
TEa/mp	Superior temporal sulcus ventral bank area, posterior part	Temporal
TEad	Temporal area TE, anterodorsal part	Temporal
TEav	Temporal area TE, anteroventral part	Temporal
TEMPORAL_POLE	Temporal pole	Temporal
TEOm	Temporal area TE, occipitomedial part	Temporal
TEpd*	Temporal area TE, posteriodorsal part	Temporal
TEpv	Temporal area TE, posterioventral part	Temporal
TH/TF	Areas TH and TF of the parahippocampal cortex	Temporal
TPt	Temporoparietal area	Temporal
V4t	Visual area 4, transitional part	Temporal
1	Area 1	Parietal
2*	Area 2	Parietal
3	Area 3	Parietal
5*	Area 5	Parietal
7A*	Area 7A (caudal inferior parietal lobule area)	Parietal
7B*	Area 7B (rostral inferior parietal lobule area)	Parietal
7m*	Area 7m (medial parietal area)	Parietal
7op	Area 7op (parietal operculum)	Parietal
AIP	Anterior intraparietal area	Parietal
DP*	Dorsal prelunate area	Parietal
LIP	Lateral intraparietal area	Parietal
MIP	Medial intraparietal area	Parietal
PIP	Posterior intaparietal area	Parietal
V6	Visual area 6	Parietal
V6A	Visual area 6A	Parietal
VIP	Ventral intraparietal area	Parietal
F1*	Agranular frontal area 1, primary motor cortex	Frontal
F2*	Agranular frontal area 2	Frontal
F3	Agranular frontal area 3	Frontal
F4	Agranular frontal area 4	Frontal
F5*	Agranular frontal area 5	Frontal
F6	Agranular frontal area 6	Frontal
F7*	Agranular frontal area 7	Frontal
Gu	Gustatory cortex	Frontal
INSULA	Insula	Frontal
ProM*	Area ProM (promotor)	Frontal

Table S5. Cont.

Abbreviation	Area name	Region
SII	Secondary somatosensory area	Frontal
10*	Area 10	Prefrontal
11	Area 11	Prefrontal
12	Area 12	Prefrontal
13	Area 13	Prefrontal
14	Area 14	Prefrontal
44	Area 44	Prefrontal
45A	Area 45A	Prefrontal
45B	Area 45B	Prefrontal
46d*	Area 46, dorsal part	Prefrontal
46v	Area 46, ventral part	Prefrontal
8B*	Area 8B	Prefrontal
8L*	Area 8, lateral part	Prefrontal
8m*	Area 8, medial part	Prefrontal
8r	Area 8, rostral part	Prefrontal
9	Area 9	Prefrontal
9/46d*	Area 9/46, dorsal part	Prefrontal
9/46v*	Area 9/46, ventral part	Prefrontal
OPAI	Orbital periallocortex	Prefrontal
OPRO	Orbital proisocortex	Prefrontal
23	Area 23	Limbic
24a	Area 24a	Limbic
24b	Area 24b	Limbic
24c*	Area 24c	Limbic
24d	Area 24d	Limbic
25	Area 25	Limbic
29/30	Area 29/30	Limbic
31	Area 31	Limbic
32	Area 32	Limbic

* Areas injected.