Supporting Information

Crossley et al. 10.1073/pnas.1220826110

1. SI Results

As can be seen in Fig. S5*B*, the rich-club organization of the connectivity network peaks earlier than in the coactivation network. Although we present the results of the top 21 strongest nodes of this network to make it comparable to the coactivation analysis, here we report the results from the extended group of regions displaying a rich-club behavior for completeness.

To make it consistent with our method of choosing the rich-club nodes in the main coactivation analyses (*SI Materials and Methods*, *Network Metrics*), we selected the point of the normalized rich-club curve where there was a significant departure from 1 (point of no difference to random networks). As can be seen in Fig. S5*B* (blue vertical line), we selected the sharp increase at the top 28.8 percentile, including 184 regions. Fig. S6 shows that, although this rich club included more regions, they were still mostly located in central and occipital modules, with some medial frontal regions from the default-mode module. Similarly, they had higher participation coefficients than non–rich-club regions ($P < 10^{-4}$, permutation test), and this configuration was costly to the system, with connection distances between rich-club nodes, and between peripheral and rich-club nodes, being longer than connection distances between peripheral nodes (both $P < 10^{-4}$, permutation tests). Although all these characteristics were very similar to the more selective rich club shown in the primary analysis, edges within the extended rich-club group of regions were less central than peripheral or feeder edges (both $P < 10^{-4}$, permutation tests). In other words, although edges within the core of this extended rich-club group of regions were very central for the network (as shown in the main analysis), the more extensive definition of this group led to the inclusion of edges that were less used by the rest of the brain to pass information within the network.

2. SI Discussion

It is important to consider more closely the methodological differences between graph theoretical analysis of modules, as used here, and independent-component analysis (ICA) as used by Smith et al. (1, 2) in a prior analysis of functional coactivation also based on the BrainMap database.

The modularity algorithm that we applied partitions a graph into a set of modules such that there is a high density of intramodular edges (between nodes in the same module) and a low density of intermodular edges (between nodes in different modules). Note that each brain region (or node of the network) can be assigned to only one module, but all nodes will be assigned to a module. Many complex systems, including brain networks (3, 4), demonstrate hierarchical modularity such that it is possible to decompose the nodes within each module into a community of submodules. There are many different algorithms available for modularity analysis; we used a well-established option, the Newman algorithm (5), to find the partition of the network that maximizes its modularity, defined quantitatively as follows:

$$Q = \frac{1}{4m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) s_i s_j, \qquad \text{[S1]}$$

where A_{ij} is the edge between nodes i and j, with degrees denoted k_i and k_j, m denotes the total number of edges of the network, and s_i s_j is 1 if the nodes belong to the same group or −1 if not.

ICA finds statistically independent patterns of coactivations, so that a region's activity is described as a combination of all of the different activity patterns. Visualization of individual components is possible after arbitrarily thresholding brain maps of independent component scores at each voxel so that only voxels or regions that load strongly on a particular component are represented as part of the network of regions that shares the same profile of (co)activation. Thus, it is possible that the same brain regions might belong to more than one independent component; note, for example, the anatomical overlap between the components denoted $BM1_{20}$, $BM2_{20}$, and $BM3_{20}$ shown in Fig. S8. It is also possible as a result of ICA that a region might not load strongly on any single component, and might therefore not appear in any network map, i.e., the coactivation profile for some regions may be represented similarly across components and thus not exceed the arbitrary threshold for visualization as part of any component map.

These preliminary considerations indicate that it is difficult to rigorously compare the similarity of network partition by a modular decomposition to the results obtained by ICA. Nonetheless, by inspection (Fig. S8), we can see that the four modules identified by graph analysis correspond closely to one or more of the components identified by ICA. A major difference is also evident by inspection: whereas the ICA implemented by Smith et al. (1) found 20 independent components, our modularity analysis found four larger modules. In the analysis reported by Smith et al., the maximum number of components was set a priori to be 20 (with a secondary analysis of 70 components); post hoc analysis then retained 10 components that had high reliability between coactivation and connectivity matrices, and discarded 10 components that were judged to be artifactual or not interpretable. The modular partitioning algorithm we applied to similar [resting-state functional MRI (fMRI)] or nearly identical (BrainMap meta-analytics) data, identified four major modules. In contrast to ICA, the Newman algorithm (5) does not find a prespecified number of modules; instead, it maximizes a global modularity function. The number of modules is thus more data driven than the number of independent components; but as noted above, prior studies have shown that fMRI and other information processing networks have a hierarchical modular structure, each one of a few large modules (such as identified by the Newman algorithm) being nearly decomposable into a larger number of submodules, sometimes with submodules that are themselves decomposable into subsubmodules, and so on (3, 4). Thus, both hierarchical modularity analysis and higher dimensional ICA have the potential to define variable numbers of modules (or components) in brain systems and "the true number" of modules or components is not yet known. What remains striking about the comparative results presented in Fig. S8 is the extent to which the larger number of anatomically smaller components identified by ICA can be superimposed to approximate closely the smaller number of anatomically larger modules identified by the Newman algorithm.

Similarly, both approaches can be tuned to study the relationship between modules or components and more-or-less specific aspects of cognition. BrainMap database describes the cognitive characteristics of the paradigms of the studies included in a hierarchical way, starting with generic behavioral domains such as "action" or "cognition," followed by subdomains such as action–execution speech, or cognition–working memory. Researchers can then choose the level of specificity of the cognitive classification they want to explore, balancing the gain of information with the reliability of the results, because dividing the studies in more classes or subdomains leads to some cognitive subdomains being represented by very few primary studies. The level of

resolution used in the modular or ICA analysis should also be taken into account: partitions focusing on the higher levels of the hierarchy (bigger modules or fewer components) probably will not gain much from relating them to highly specific aspects of cognition, because many subdomains (possibly those within the same behavioral domain) will map to one of the big modules or components. Smith et al. (1) described the relationship of their independent components to 66 subdomains, of which they highlighted 20 that corresponded most strongly to their independent components. We instead mapped our four topological modules to the five major behavioral domains defined by BrainMap. Essentially, we preferred reliability of a relatively coarse-grained cognitive analysis over the potentially greater specificity of a more fine-grained analysis. This allowed us to test important basic predictions about the functional specialization of topological modules and visualize the interactions between modules and the rich club of the network.

3. SI Materials and Methods

3.1. Further Details on Method Used to Build the Coactivation Network. Studies were extracted from the BrainMap database, which at the time of the search (October 5, 2011) included 2,155 neuroimaging papers reporting data on 40,569 subjects. After excluding studies including patients, subjects under 18 or over 65 y of age, pharmacological interventions, or less than five subjects, we were left with 1,641 studies reporting 6,884 contrasts on 21,713 subjects. Peak activations were extracted and modeled as 1-cm^3 spheres (Fig. S10*A* in white), which "activated" a region from the template if 20% of its area was in it. The need to choose a threshold was based on activations falling in boundaries between regions, and at 20% the number of modeled links in the bipartite graph was at least the number of activations extracted from the database (see Table S1 for results without modeling the activation as sphere). The two-mode network is then transposed to a one-mode network using Jaccard indices, defined as the intersection of contrasts activating both regions divided by the union of contrasts activating any. In Fig. S10*A*, the edge between blue to green region has a Jaccard index of 2/3, because two contrasts coactivate both regions (intersection of contrasts is 2), but an extra contrast activates the blue region only (union of contrasts is 3). This matrix was then thresholded statistically as described below. One limitation of this approach is that, similarly to previous studies (1, 6), we modeled peak activations without considering the effect of sample size in the primary studies (apart from excluding small studies with less than five subjects), or the between- and within-study variances (7). This made the analysis of more than 6,000 contrasts more expedient, and allowed for an easier binary solution when building a bipartite graph. Note that we did not consider the direction of the contrast used (whether it could be classified as an activation or deactivation), because we were interested in collaborative behavior, irrespective of the direction. See also Movie S2 for an illustrative brief summary of this approach.

For the deactivations analysis, a subgroup of studies that used rest or fixation as baseline and reported both directions of contrasts were included. Deactivations and activations from each pair of contrasts (task > rest, rest > task) were modeled separately, and a directed line drawn from all of the activations to all deactivations reported in a pair of contrasts (Fig. S10*B*). The direction of the edge is arbitrary, it only intends to highlight that the relationship is not reversible. For purposes of the analyses, we first obtained the network characteristics of the nodes in the activation/deactivation dyads from the coactivation network. We then looked at whether certain network characteristics from the group of all dyads, built from all of the 110 pairs of contrasts included, were significantly different from the null model (see below).

3.2. Statistical Correction for Nonsignificant Connections. Although the weight of the edge was defined by the Jaccard index as detailed above, the existence of an edge was based on a statistical approach. To identify those significant connections in the coactivation network, we followed Toro et al. (6) and examined whether activations in two regions X and Y were best modeled as independent phenomena (null hypothesis H_0, regions activating independently from each other) or dependent (alternative hypothesis H_1, regions being functionally connected). From the data, we can compute the maximum-likelihood estimate for an activation in a region under the null hypothesis:

$$p = m/N, \qquad \textbf{[S2]}$$

where m is the number of contrasts that activated region X and N is the total number of contrasts. We can then say that the likelihood of the null hypothesis (activation in X being independent from activation in Y) is as follows:

$$L(H_0) = B(k;n,p)B(m-k;N-n,p), \qquad \textbf{[S3]}$$

where B refers to the binomial distribution, n is the number of contrasts that activated region Y, and k is the number of contrasts that activated both regions X and Y. Likewise, the values for p_1 and p_0 under the alternative hypothesis that there is a dependence between activations in both regions is as follows:

$$p_1 = k/n, \qquad \textbf{[S4]}$$

$$p_0 = (m-k)/(N-n), \qquad \textbf{[S5]}$$

and their likelihood

$$L(H_1) = B(k,n,p_1)B(m-k;N-n,p_0). \qquad \textbf{[S6]}$$

P values were estimated by calculating

$$-2\log\left(\frac{L(H_0)}{L(H_1)}\right), \qquad \textbf{[S7]}$$

which follows χ^2 distribution. Connections significant at $P < 0.01$ corrected for false-discovery rate were maintained.

3.3. Statistical Comparisons of Network Properties. To see whether certain topological properties of the coactivation network were nontrivial, we compared their medians to a null distribution. Unless otherwise stated, this null distribution was built from measuring the topological property in 1,000 random graphs, which had the same degree and weight distribution as the coactivation network (8). All P values from permutation tests reported refer to the probability of the value in question to be part of this null distribution.

Three other null models should be mentioned. First, comparison of the modularity organization in the coactivation and connectivity network did not compare it to modularity assignment of random networks. Instead, we randomly permuted the module assignment of the nodes of the coactivation network 10,000 times. Likewise, null distributions for comparisons between rich-club nodes and peripheral nodes, or rich-club connections, peripheral connections, and feeder connections, were based on 10,000 iterations of random permutation of assignment of these labels to the edges of the network. For the purposes of the topological distribution of the activation/deactivation pairs, the null model was built by keeping the activations but randomly assigning deactivations to other nodes, iterated 10,000 times.

3.4. Alternative Null Models for Comparative Analyses of Brain Networks. A recent study has shown that networks constructed from a matrix of pairwise correlations between fMRI time series tend to overestimate the clustering of the network, and can thus potentially bias estimation of its small-world properties (9). This is particularly a problem when the network is compared with a random graph or a null model created by random link permutation (topological randomization) of the observed graph. Our coactivation metric was not correlation but the Jaccard index, a different measure of similarity or association between nodes. We were not aware of any prior studies to suggest that an association matrix comprising multiple measures of pairwise association in terms of the Jaccard index was as susceptible as a correlation matrix to this potential source of bias in graph analysis. Nonetheless, to address this possible issue, we repeated our analysis of the topological properties of the brain coactivation network by comparing it to an ensemble of random graphs generated by permuting the regional activations associated with each contrast (data randomization), to generate a set of random association matrices, each of which was then thresholded to construct a random graph (Fig. S11). This differed from our original approach whereby a set of random graphs was generated by permuting the edges of the network constructed from the observed association matrix. However, we found that the results of topological analysis of the coactivation network were not substantively affected by the choice of null model, suggesting that the results reported in the main text (based on random edge permutation) are not attributable to any topological bias inherent in the use of the Jaccard index as a measure of association between nodes (Fig. S1*B*).

For completeness, we also show results from the connectivity network with a null model using data randomization (Fig. S1*D*). In this case, we calculated small-world parameters for each subject after randomizing the fMRI time series by taking its Fourier transform, randomly permuting the phases, and then reconstructing the time series, as discussed in Zalesky et al. (9).

3.5. Network Metrics.

Degree: number of significant connections that a node has.

Weighted degree (strength): nodal characteristic that describes the sum of all of the connectivity indices connecting it to the network. For the coactivation network, this is highly correlated to degree ($r = 0.97$).

Density: percentage of existing edges in the network compared with all possible edges.

Shortest path: the shortest topological distance (geodesic) between two nodes, where distances for the meta-analytical network are the sum of the inverse of the Jaccard indices of edges traversed. The shortest path of a network is the average of all of the shortest paths between all pairs of nodes (10).

Clustering: the proportion of existing links between the neighbors of a node from the possible ones, where neighbors of a node are the nodes directly connected to them (10).

Global efficiency: the average of the inverse of the shortest paths between all nodes. Unlike shortest path, it can be used in weighted or disconnected networks (11).

Local efficiency: the inverse of the shortest paths between neighbors of a node, averaged across all nodes for the network. Unlike clustering, it can be used in weighted networks (11).

Small world (σ): this is a ratio between γ and λ. γ is the ratio between the clustering of the network analyzed and the clustering of comparable random networks (same number of nodes and edges, same degree and weight distribution), and λ is the ratio between the shortest path of the network and comparable random networks (12). For disconnected networks, we used a conceptually similar idea and used the ratio between global efficiency of the network and random networks, divided by the inverse of the ratio between local efficiency of the network and random.

Modularity: the best partition of the network into subgroups that are highly connected between themselves. This is obtained by maximizing a parameter Q, described in Eq. **S1** (5).

Participation coefficient: measures how well a node in a given module is connected to other modules, defined as follows:

$$P_i = 1 - \sum \left(\frac{k_{is}}{k_i}\right)^2, \qquad \textbf{[S8]}$$

where k_{is} are the sum of the connections from node i to module s, and k_i its degree (13).

Rich-club configuration: for weighted networks, this is the proportion of the strongest edges of the network that connect high-degree nodes. Formally, this is defined for a cutoff richness factor r as follows:

$$\phi^w(r) = \frac{W_{>r}}{\sum_{l=1}^{E>r} w_l^{\text{rank}}}, \qquad \textbf{[S9]}$$

where $W_{>r}$ is the sum of the weights of the connections between nodes with weighted degree higher than r, $E_{>r}$ is the number of these edges, and the denominator describes the sum of the top $E_{>r}$ strongest edges of the network (14). This value can be normalized by the same parameter from randomized networks. Although the rich club thus defined is a curve, we explored the characteristics of the top hub nodes defined by a steep and significant departure from null (e.g., steep rise in the curve and 95% confidence interval of normalized rich-club coefficient above 1). Degree can be used as the richness factor to rank the nodes, rather than weighted degree (strength). Weighted degree and degree were correlated in this network, so it is unsurprising that results are similar if using this definition.

Diversity of contrasts coactivating a pair of regions, V_{edge}:

$$V_{\text{edge}} = 1 - \left(\sum_d \left(\frac{k_d}{k_t}\right)^2\right), \qquad \textbf{[S10]}$$

where k_d is the normalized frequency of coactivation by domain d, and k_t is the sum of the normalized coactivation frequencies across domains. Values of V_{edge} near zero mean that the edge is only coactivated by one domain of cognitive contrasts, whereas values near 1 describe edges that are coactivated by contrasts in several domains.

3.6. Effect of Head Motion in Resting-State fMRI Data. We investigated the effect of head motion in the fMRI blood oxygenation level-dependent time series following Power et al. (15). As Fig. S12*A* shows, there was little impact of movement on data quality, with only nine subjects requiring removal of some corrupted frames (frames with a framewise displacement higher than 0.5 mm; correction included removing one back and two forward of the marked frame), in all cases being less than 25% of the data. For those subjects with significant movement, removing (scrubbing) the affected frames did not uncover a differential effect of movement on long- and short-range connections in our data (Fig. S12*B*). In this context, we did not consider it necessary to work with censored data in the main analysis presented.

3.7. Reliability Analysis. We performed several reliability analyses, looking into the effect of the parcellation scheme we used, the possible effect of overrepresentation of certain types of tasks in the database, and the assumption of the size and overlap of the activation sphere. For the parcellation problem, we reanalyzed the data using two alternative functionally based templates, one with regions of similar size (594 regions) and another with bigger regions (190 regions), excluding the cerebellum (16). The possible publication bias was tackled by including a random subgroup of all contrasts included, balanced across the different behavioral domains of BrainMap database. Finally, the assumption of the activation sphere was reviewed by defining a region as activated only when the coordinates of the peak activation voxels fell in it. Table S1 summarizes all these analyses. Weighted networks for alternative parcellation scheme and modeling the peak activation only are analyzed at the same density as main findings. Controlling for publication bias led to a decrease in the number of studies included, and therefore we used a less stringent statistical threshold of $P < 0.05$, false-discovery rate corrected, when building the weighted network (density of network, 2.5%). As shown in Table S1, overall the main results are qualitatively the same as those reported in the main analysis. Note that the modularity comparison between different templates was performed at the voxel level, so differences might be partly due to the lack of overlap between regions.

1. Smith SM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* 106(31):13040–13045.
2. Laird AR, et al. (2011) Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci* 23(12):4022–4037.
3. Meunier D, Lambiotte R, Bullmore ET (2010) Modular and hierarchically modular organization of brain networks. *Front Neurosci* 4:200.
4. Bassett DS, et al. (2010) Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Comput Biol* 6(4):e1000748.
5. Newman ME (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103(23):8577–8582.
6. Toro R, Fox PT, Paus T (2008) Functional coactivation map of the human brain. *Cereb Cortex* 18(11):2553–2559.
7. Eickhoff SB, et al. (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp* 30(9):2907–2926.
8. Rubinov M, Sporns O (2011) Weight-conserving characterization of complex functional brain networks. *Neuroimage* 56(4):2068–2079.
9. Zalesky A, Fornito A, Bullmore E (2012) On the use of correlation as a measure of network connectivity. *Neuroimage* 60(4):2096–2106.
10. Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. *Nature* 393(6684):440–442.
11. Latora V, Marchiori M (2001) Efficient behavior of small-world networks. *Phys Rev Lett* 87(19):198701.
12. Humphries MD, Gurney K (2008) Network "small-world-ness": A quantitative method for determining canonical network equivalence. *PLoS ONE* 3(4):e0002051.
13. Guimerà R, Nunes Amaral LA (2005) Functional cartography of complex metabolic networks. *Nature* 433(7028):895–900.
14. Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: The weighted rich-club effect. *Phys Rev Lett* 101(16):168702.
15. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59(3):2142–2154.
16. Craddock RC, James GA, Holtzheimer PE, 3rd, Hu XP, Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum Brain Mapp* 33(8):1914–1928.

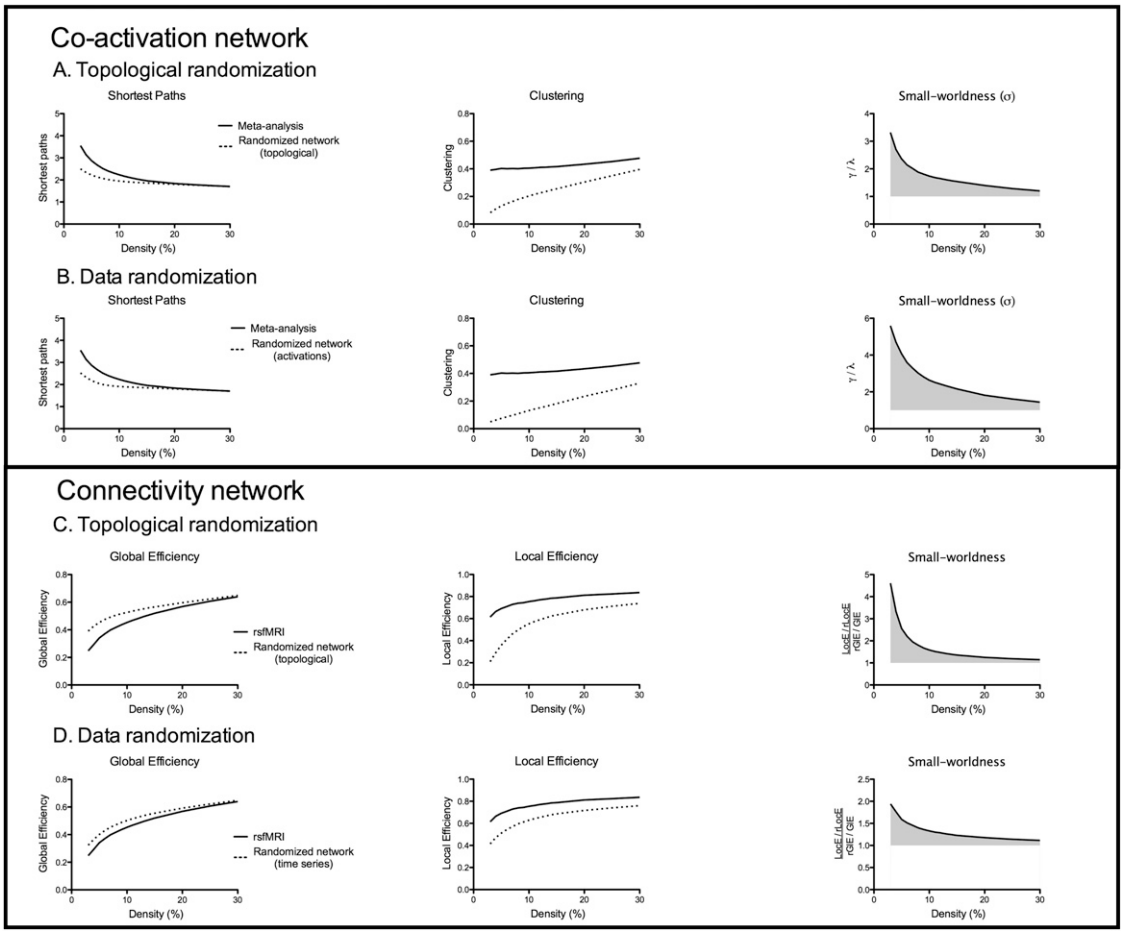


Fig. S1. Small-world behavior of coactivation and connectivity networks. *A* and *B* show path length, clustering, and small-world organization of coactivation network, thresholded at different connection densities. At sparse densities, the network has a path length similar to random but higher clustering, having a small-world configuration (area in gray). Null model in *A* is based on random permutation of links of the network (topological randomization), whereas in *B* the null graphs are constructed by random permutation of activations (data) reported for each contrast. Both analyses show similar results, demonstrating that, unlike networks based on correlation matrices, networks based on the Jaccard index as a measure of association do not overestimate the clustering and small worldness of the network. (*C* and *D*) Global (GlE) and local (LocE) efficiency in the connectivity network (resting-state fMRI), and the relationship between their ratios and random networks (LocE/rLocE/rGlE/GlE). Metrics are conceptually similar to path length, clustering, and small worldness, but can be used on disconnected networks such as the functional connectivity network (disconnected at densities below 12%). In *C*, the null model is based on topological randomization, whereas in *D* the null graphs are constructed by randomly permuting the time series in the Fourier domain, as described by Zalesky et al. (1), estimating the correlations between all pairs of permuted time series, and then thresholding the correlation matrices. Note that the local efficiency (a metric similar to clustering) of the null graphs built using topological randomization (*C*) is lower than when using data randomization (*D*), replicating the bias described by Zalesky in correlation networks. Nevertheless, both analyses show that the network has a small-world organization (ratio of LocE/rLocE and rGlE/GlE is higher than 1).

1. Zalesky A, Fornito A, Bullmore E (2012) On the use of correlation as a measure of network connectivity. *Neuroimage* 60(4):2096–2106.

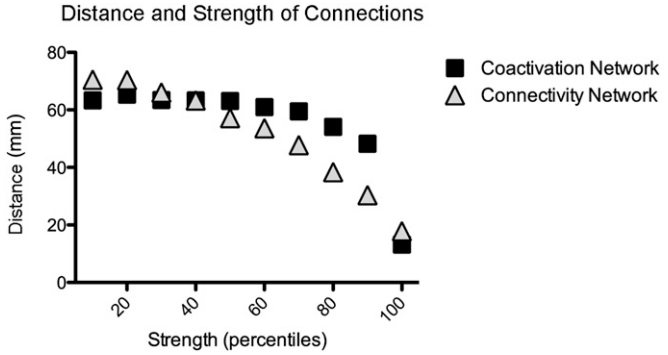


Fig. S2. Connection distance and strength in coactivation and connectivity networks. Median lengths of edges in the two networks divided into 10 percentile bins according to their strength (weight). Note that the excess of long connections in the coactivation network seen in Fig. 1*D* is mainly driven by strong long-range connections.

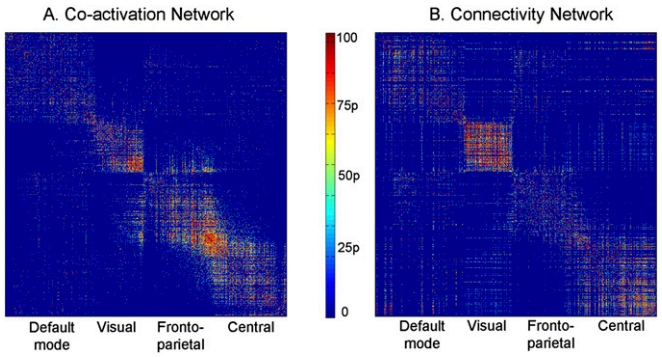


Fig. S3. Association matrices for functional coactivation (*Left*) and functional connectivity (*Right*) datasets. Heat maps show the ranked strengths (percentile scores) of the pairwise associations between regional nodes. Regions were ordered according to their membership of one of the four modules of the co-activation network, highlighting the existence of more connections within modules (see the block diagonal clustering of hot colors in *A*), with a similar modular configuration for the connectivity network (similar clustering formed in *B*). Some differences are also present, particularly in the level of connectivity within a module, with the visual module having greater intramodular density of connections in the fMRI connectivity network, whereas the frontoparietal module had more dense intramodular connectivity in the coactivation network.

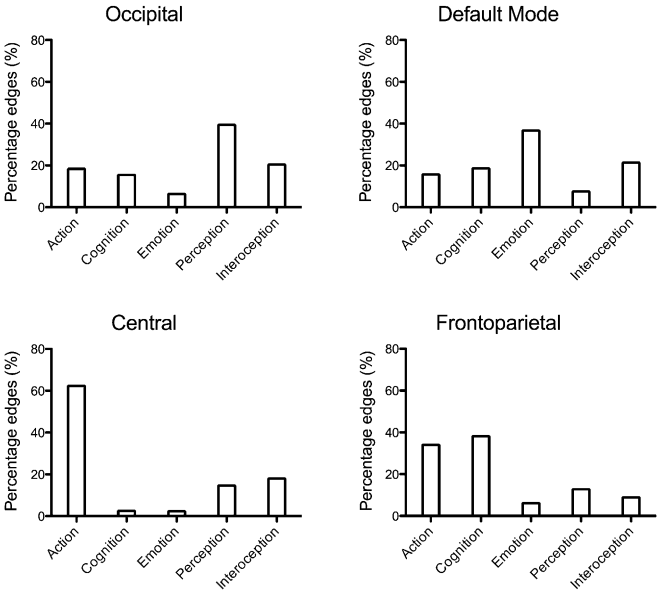


Fig. S4. Behavioral domains of edges in modules. Proportion of intramodular edges classified according to the five behavioral domains studied.

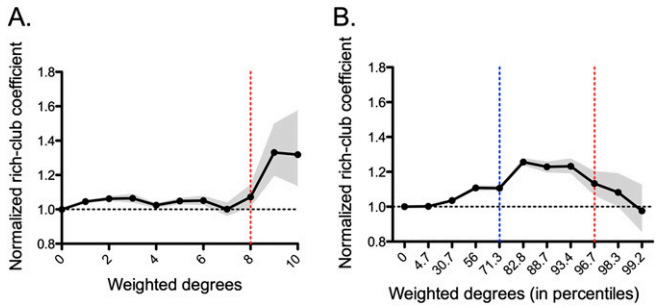


Fig. S5. Rich-club organization of coactivation and connectivity networks. Normalized rich-club coefficients (normalized by rich-club coefficients from random graphs) with 95% confidence interval for coactivation network (*A*) and connectivity network (*B*). Rich-club nodes were defined from the coactivation rich-club curve (*A*): at weighted degrees above 8, there is a clear and statistically significant increase in the number of strong edges shared by these nodes (marked with red vertical line). Note that weighted degrees are described in percentiles for the connectivity network (*B*) to make them comparable to the coactivation network. Although the connectivity network peaks earlier, for the purposes of comparing the two networks, a subgroup of the strongest nodes was included (red vertical line; see *SI Results* and Fig. S6 for characteristics of rich club defined from blue vertical line).

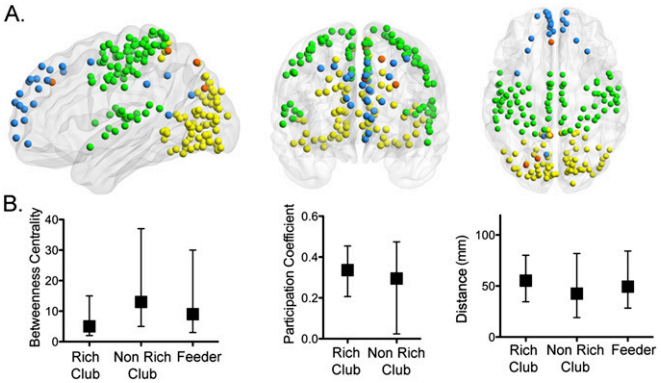


Fig. S6. Characteristics of extended rich club of the connectivity network. (*A*) Anatomical location of the nodes belonging to the extended rich club (blue line in Fig. S5*B*). Color of nodes corresponds to their module as in Fig. 1. (*B*) Betweenness centrality of edges, participation coefficient of nodes, and connection distance of edges defined according to their relationship to the extended rich club. Median and interquartile range are shown.

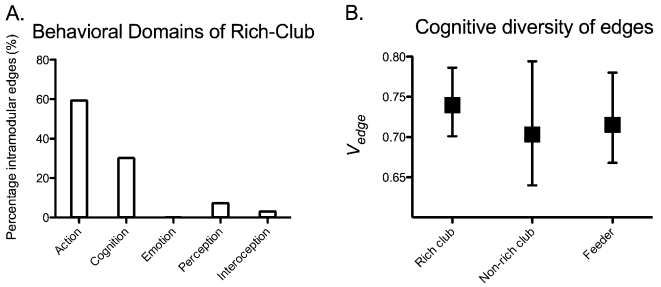


Fig. S7. Behavioral domains of rich-club edges, and their diversity. (*A*) Percentage of edges classified according to the five behavioral domains. Note that each edge is labeled according to the behavioral domain most frequently driving coactivation of the pair of regional nodes that it connects. (*B*) Diversity of rich-club edges compared with peripheral and feeder edges. Median and interquartile range are shown.

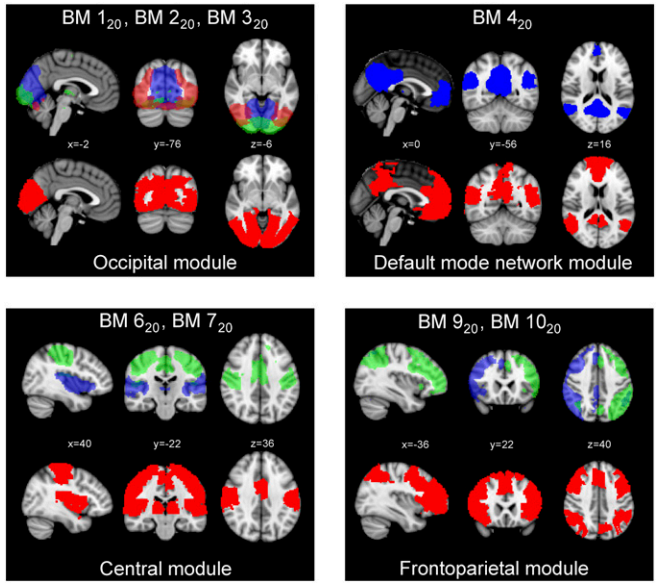


Fig. S8. Comparison between graph theoretical modules and independent components defined by Smith et al. (1). In each panel, the upper row shows independent components labeled according to the notation in Smith et al. and uniquely colored to show how the superposition of one or more components closely approximates the anatomical distribution of one of the four modules from the coactivation network represented in the lower row (colored red and labeled in accordance with our notation). Note that each of the four modules can be closely approximated by the superposition of one or more of the independent components. Only two components identified by Smith et al. are not easily reconciled with our results: a cerebellar component, denoted BM 5_{20}, does not correspond to any of our modules because we chose not to include cerebellar regions in our construction of the coactivation network; and a frontal component, denoted BM 8_{20}, overlaps both the default-mode and frontoparietal modules defined by our analysis.

1. Smith SM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* 106(31):13040–13045.

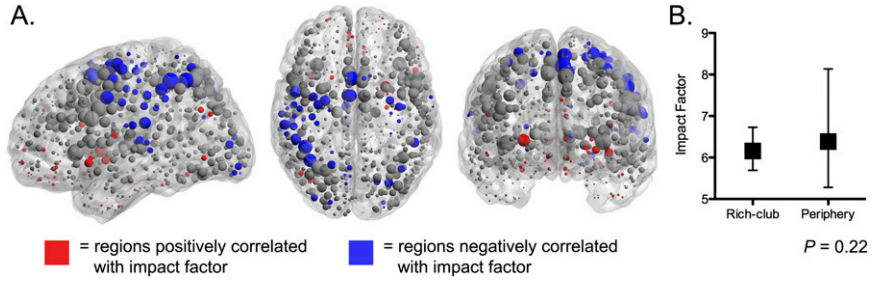


Fig. S9. Relationship between journal impact factor and degree in coactivation network. As described in the main text, a possible confounding factor could be the uneven interest of the community in certain tasks or regions. Although in Table S1 we show the effect of balancing the contrasts across behavioral domains, we here explored the relationship of the degrees of the coactivation network and the impact factor of the studies reporting the activations, with the impact factor used as a proxy measure for the interest of the community in this area (1). (*A*) Nodes plotted in anatomical space, with their size being proportional to their degree in the coactivation network. Most of the high-degree hubs, which constitute the rich club, are located in frontal, central and parietal regions. Following Behrens et al. (1), the nodes have been colored according to the sign of significant correlation between activation and journal impact factor ($P < 0.05$, uncorrected). Red nodes are positively correlated, and blue nodes are negatively correlated, with impact factor. It can be seen that most of the nodes positively correlated with impact factor are not high-degree hubs of the coactivation network. (*B*) Mean impact factor and 95% confidence intervals of studies reporting activations in the rich-club nodes and in the peripheral nodes; there is no significant difference in impact factor (*t* test, $P = 0.22$). In short, we do not consider that the hubs of the coactivation network are biased by the impact factor of the journals in which the primary data have been reported.

1. Behrens TE, Fox P, Laird A, Smith SM (2013) What is the most interesting part of the brain? *Trends Cogn Sci* 17(1):2–4.

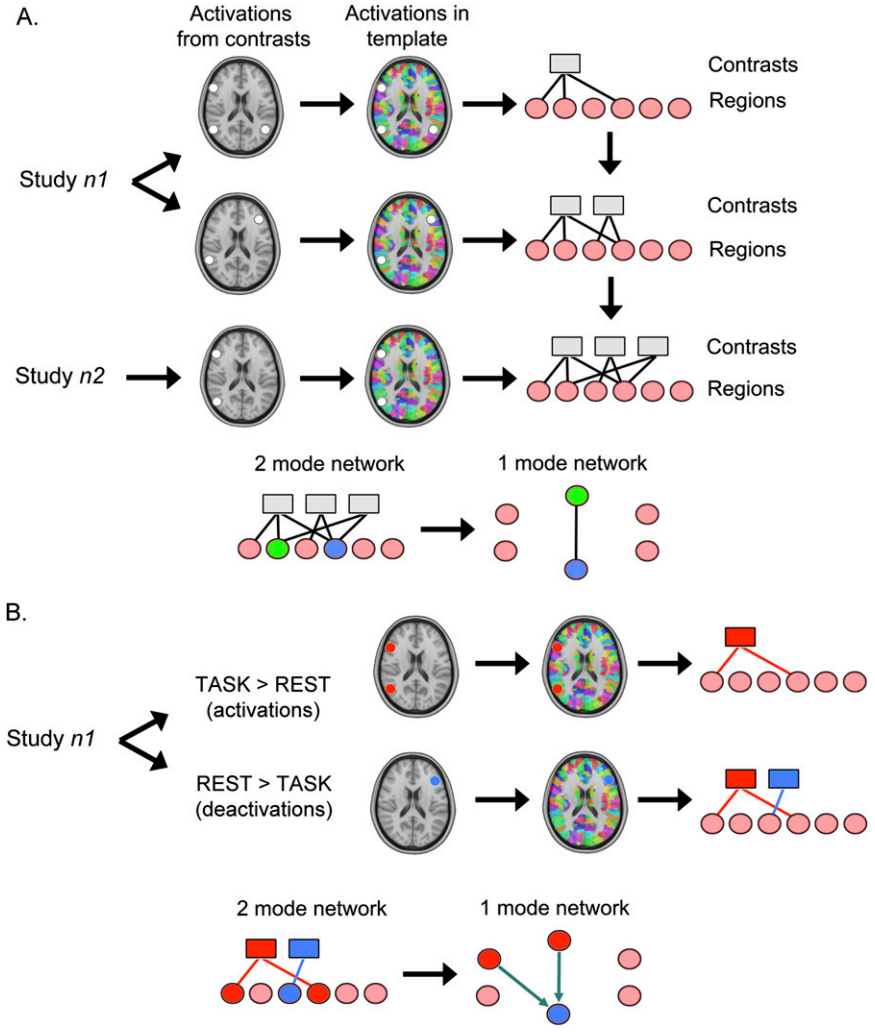


Fig. S10. Building the coactivation network. Diagram showing in stages the construction of the coactivation network (*A*) and the analysis of deactivations (*B*).

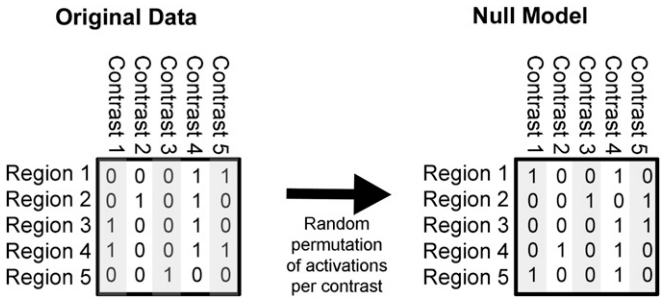


Fig. S11. Null model based on data randomization for coactivation networks. Columns represent different contrasts and rows regions, with 1s denoting activations. For each contrast included, activations were randomly permuted. Null model has therefore the same distribution of activations per contrasts. On this null model, we calculated the Jaccard indices looking at coactivation patterns between regions.

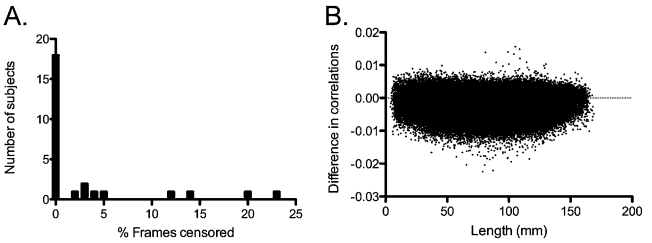
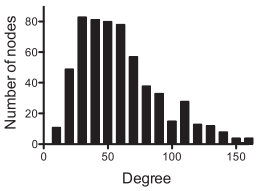
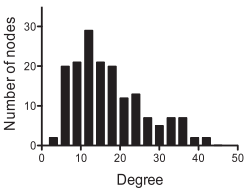
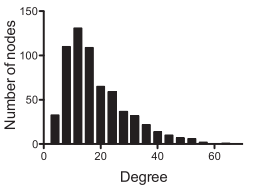
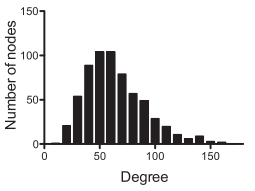
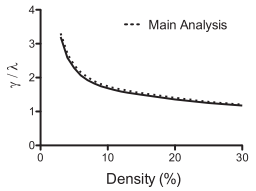
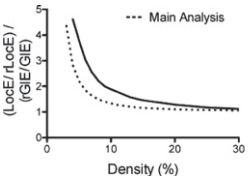
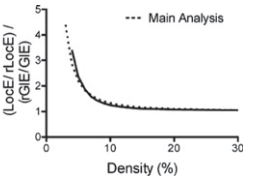
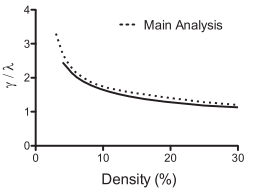
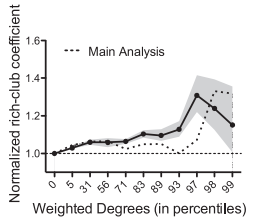
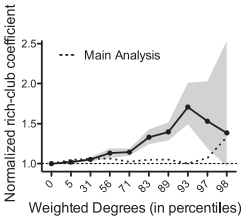
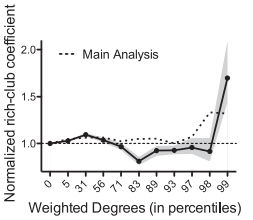
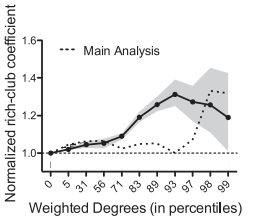
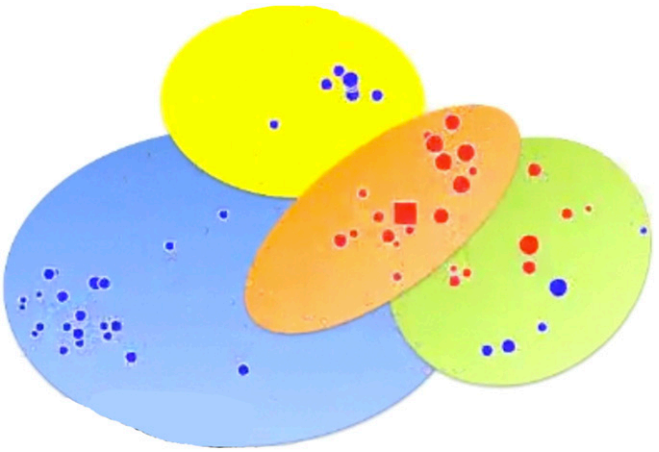


Fig. S12. Motion effect on fMRI data. (*A*) Number of subjects and percentage of noisy frames requiring scrubbing. (*B*) Differences in all pairwise correlation indices before and after scrubbing for the whole group of subjects, and its relationship to distance between regions. Note that there is no clear relationship to distance, suggesting no systematic bias in the group matrix introduced by head movement.

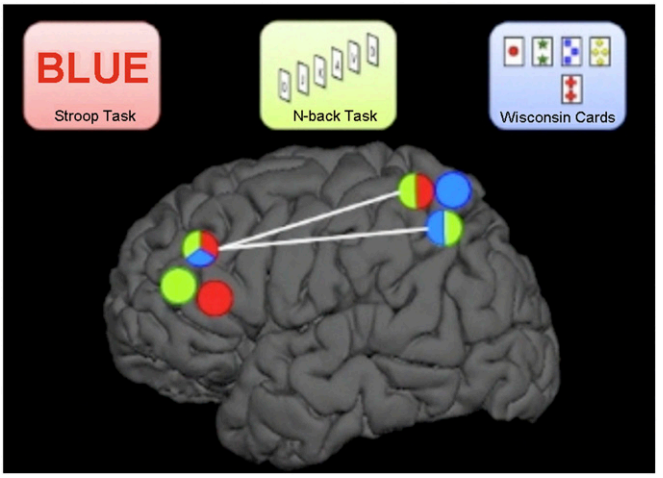
Table S1. Robustness analysis

Network analysis	Functional template 594 regions	Functional template 170 regions	Publication bias corrected	Peak activation modeled
Degree distribution				
Small worldness (σ)				
Modularity (Q) and its correspondence with main analysis (Rand index)	Q = 0.44, 4 modules. Rand index = 0.85	Q = 0.49, 5 modules. Rand index = 0.82	Q = 0.54, 7 modules. Rand index = 0.79	Q = 0.37, 4 modules. Rand index = 0.87
Rich-club configuration				



Movie S1. Activations (in red) and deactivations (in blue) from 110 contrasts mapped dynamically onto the coactivation network. Colors represent approximate location of modules as in Fig. 3*A*.

Movie S1



Movie S2. Brief schematic summary of network construction and analysis.

Movie S2