Network dimensionality underlies flexible representation of cognitive information

Takuya Ito[1,2,*], Michael W. Cole[1]

[1]Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102
[2]Behavioral and Neural Sciences PhD Program, Rutgers University, Newark, NJ 07102

Contact:
Takuya Ito
Center for Molecular and Behavioral Neuroscience, Rutgers University
197 University Avenue, Newark, NJ 07102
taku.ito1@gmail.edu

Abstract

The functional networks in human cortex that most flexibly represent cognitive information are hubs with widespread connectivity throughout the brain. Going beyond simple hub measures, we hypothesized that the dimensionality of each network's global connectivity pattern (its global dimensionality) underlies its ability to produce highly diverse task activation patterns (its representational flexibility). Supporting our hypothesis, we report that the global dimensionality estimated during resting state correlates with the representational flexibility estimated across a variety of cognitive tasks. Demonstrating the robustness of this relationship, each network's global connectivity pattern could be used to predict its representational flexibility. Additionally, we found that the frontoparietal cognitive control network had the highest dimensionality and flexibility, and that individuals with higher network dimensionality had higher representational flexibility. Together, these findings suggest that a network's global dimensionality contributes to its ability to represent diverse cognitive information, implicating dimensionality as a network mechanism underlying flexible cognitive representation.

Introduction

The human brain's network organization is thought to contribute to its ability to process information, but the mechanisms linking network organization to information processing remain unclear. Recent studies have provided links between the brain's intrinsic network architecture and representations of task-related information (in the form of task activation patterns)[1,2], yet the large-scale network properties that underlie the human brain's ability to flexibly perform a wide range of tasks remains unknown. Studies at the single and multi-cell level have begun to elucidate the neurophysiological mechanisms underlying such cognitive flexibility. For example, neurons with mixed selectivity (i.e., complex tuning) have been shown to flexibly represent a range of stimuli and task rules[3,4]. However, these studies were often limited to specific brain regions (e.g., dorsolateral prefrontal cortex), rather than identifying the contribution of large-scale network organization. Computational studies have provided abstract models for how various tasks might be executed[5,6], yet such abstract models leave many questions with regard to biological mechanisms. Thus, it remains unclear how the human brain's large-scale network organization might contribute to the flexible implementation of cognitive tasks.

Several studies have provided clues that hub connectivity – a large-scale network property in which regions have extensive connectivity throughout the brain – supports high cognitive flexibility[7–11]. For instance, regions within the frontoparietal cognitive control network (FPN) are hubs[7,9,12] that systematically shift their global functional connectivity patterns across a variety of tasks[8]. This combination of hub connectivity that is flexible across tasks led these regions to be termed "flexible hubs". Critically, however, it has remained unclear whether hub flexibility is only a region-level or is also a network-level property. Casting doubt on the region-level flexible hub hypothesis, there is evidence that no cortical regions are "super" hubs in the sense of individual regions having strong connectivity to all or even most other regions[13]. This suggests that regions with widespread connectivity would have to pool their connections to achieve strong hub status. Further, it is unclear why various flexible hubs would be integrated within the FPN if they have redundant connectivity patterns. We therefore hypothesized that strong flexible hub properties emerge at the network level, with each FPN region contributing limited connectivity and flexibility that is integrated within FPN to collectively produce strong flexible hub properties. More generally, we hypothesized that network-level dimensionality – the tendency for individual-region connectivity patterns to be differentiated – would contribute to network-level representational flexibility (the tendency for a network's activation patterns to be diverse across tasks).

To test our hypothesis, we developed a network-level graph theoretical property – global dimensionality. Global dimensionality characterizes how pattern-separated the global (i.e., out-of-network) connections of a network are (Fig. 1a). Recent evidence has suggested robust statistical relationships between resting-state network organization and task-evoked activations[2,14], with activity flow – the movement of task-evoked activations between brain regions – over resting-state connections providing a potential mechanistic explanation[1]. We sought to build on these findings to investigate whether the organizational properties of large-scale intrinsic brain networks play a role in the production of flexible neural representations. We hypothesized that a hub network with high intrinsic global dimensionality would have a computational advantage in processing

task information flexibly, in part by reducing interference between task-relevant cognitive representations. Providing concrete evidence that links a network's global dimensionality with flexible task representation would suggest a role for intrinsic network organization in providing the space of possible computations (cognitive, or otherwise) performed by the human brain. Given recent evidence suggesting that the FPN acts as a flexible hub network for adaptive task control[8,10,15,16], we hypothesized that the dimensionality of the FPN's global connectivity patterns estimated during resting-state underlies its ability to flexibly represent a diverse range of tasks.

We tested this hypothesis using functional magnetic resonance imaging (fMRI) data collected as part of the Human Connectome Project (HCP). Evidence linking a network's global dimensionality estimated during resting-state fMRI and representational flexibility estimated during task-state fMRI would suggest that such a network can integrate distributed sets of task-relevant information in an organized fashion, reducing pattern overlap/interference and producing highly decodable representations underlying task performance (Fig. 1).

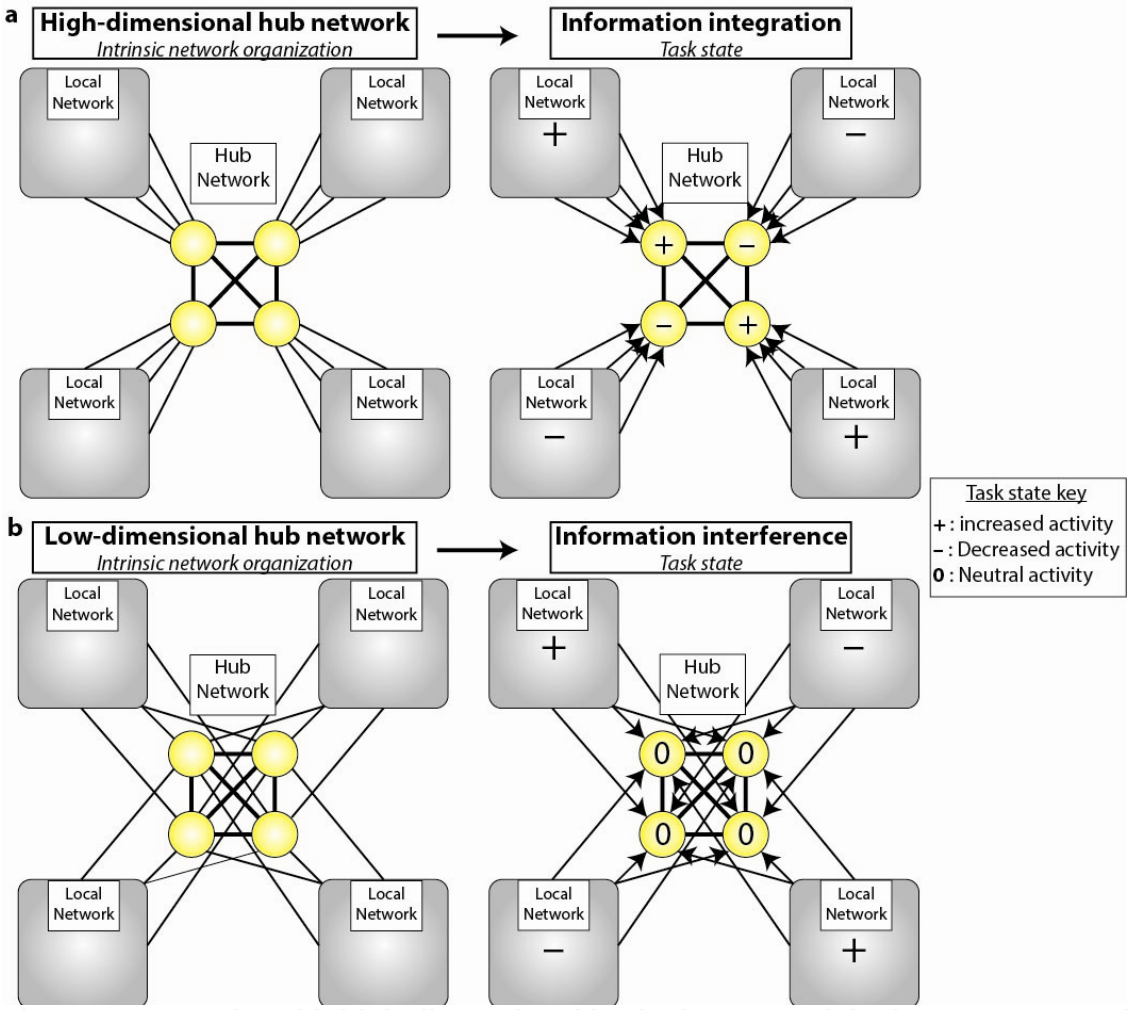

Figure 1. Networks with high-dimensional intrinsic connectivity have a computational advantage over low-dimensional networks. a) A schematic example of a hub network with high network dimensionality, due to pattern-separated global connections. Regions within the

hub network are yellow. The high-dimensional connectivity organization allows for information integration from diverse brain systems in a pattern-separated manner. Activity from local networks mapped to hub network regions produces decodable patterns of activation. **b)** A schematic example of a hub network with low network dimensionality, due to the lack of pattern-separated global connections. Every region in the hub network has the same global connectivity pattern, leading to low network dimensionality. Activity from local networks mapped to hub network regions produces a net activity of 0 in each region of the hub network. This is due to the lack of connectivity separation (low-dimensional connections), leading to an interference of information-bearing signals. Note that the regions in the hub network in panels **a** and **b** have the same weighted degree centrality.

Methods

Data collection

Data were collected as part of the Washington University-Minnesota Consortium of the Human Connectome Project (HCP; Van Essen et al., 2013). The data from the "100 Unrelated Subjects" (*n*=100) of the greater "500 Subjects" HCP release was used for empirical analyses. Specific details and procedures of subject recruitment and data collection can be found in [45]. 100 human participants (54 female) were recruited from Washington University in St. Louis and the surrounding area. The mean age of the human participants was 29 years of age (range=24 – 36 years of age). Whole-brain multiband echo-planar imaging acquisitions were collected on a 32-channel head coil on a modified 3T Siemens Skyra with TR=720 ms, TE=33.1 ms, flip angle=52º, Bandwidth=2,290 Hz/Px, in-plane FOV=208x180 mm, 72 slices, 2.0 mm isotropic voxels, with a multiband acceleration factor of 8. Data for each subject were collected over the span of two days. On the first day, anatomical scans were collected (including T1-weighted and T2-weighted images acquired at 0.7 mm isotropic voxels) followed by two resting-state fMRI scans (each lasting 14.4 minutes), and ending with a task fMRI component. The second day consisted with first collecting a diffusion imaging scan, followed by a second set of two resting-state fMRI scans (each lasting 14.4 minutes), and again ending with a task fMRI session. Each of the seven tasks was collected over two consecutive fMRI runs. Further details on the resting-state fMRI portion can be found in [46], and additional details on the task fMRI components can be found in [47].

Task paradigms

The data set was collected as part of the HCP project, which included both resting-state and seven task fMRI scans[45]. The seven collected task scans consisted of an emotion cognition task, a gambling reward task, a language task, a motor task, a relational reasoning task, a social cognition task, and a working memory task. Briefly, the emotion cognition task required making valence judgements on negative (fearful and angry) and neutral faces. The gambling reward task consisted of a card guessing game, where subjects were asked to guess the number on the card to win or lose money. The language processing task consisted of interleaving a language condition, which involved answering questions related to a story presented aurally, and a math condition, which involved basic arithmetic questions presented aurally. The motor task involved asking subjects to either tap their left/right fingers, squeeze their left/right toes, or move their tongue. The reasoning task involved asking subjects to determine whether two sets of objects differ from each other in the same dimension (e.g., shape or texture). The social cognition task was a theory of mind task, where objects (squares, circles, triangles) interacted with each other in a video clip, and subjects were subsequently asked whether the objects interacted in a social manner. Lastly, the working memory task was a variant of the N-back task. A complete description of these task paradigms and scans can be found in [47].

fMRI Preprocessing

Minimally preprocessed data for both resting-state and task fMRI were obtained from the publicly available HCP data. We performed additional preprocessing steps for

resting-state fMRI, which included removing the first five frames of each run and performing nuisance regression on the minimally preprocessed data. Nuisance regression included removing the mean of each run, linear detrending, and regressing out 12 motion parameters (six motion parameter estimates and their derivatives), the mean white matter time series and its derivative, the mean ventricle time series and its derivative, and the mean global signal time series and its derivative.

Task data for task activation analyses were additionally preprocessed using a standard general linear model (GLM) for fMRI analysis. The first five frames of each run were removed prior to fitting the GLM. Nuisance regressors included 12 motion parameters, regressors for the mean ventricles, white matter, and global signals and their derivatives. In addition, for each task paradigm, we estimated the task-evoked activations of each task condition by fitting the task timing for each condition convolved with the SPM canonical hemodynamic response function. Two regressors were fit for the emotion cognition task, where coefficients were fit to either the face condition or shape condition. For the gambling reward task, one regressor was fit to trials with the punishment condition, and the other regressor was fit to trials with the reward condition. For the language task, one regressor was fit for the story condition, and the other regressor was fit to the math condition. For the motor task, six regressors were fit to each of the following conditions: (1) cue; (2) right hand trials; (3) left hand trials; (4) right foot trials; (5) left foot trials; (6) tongue trials. For the relational reasoning task, one regressor was fit to trials when the sets of objects were matched, and the other regressor was fit to trials when the objects were not matched. For the social cognition task, one regressor was fit if the objects were interacting socially (theory of mind), and the other regressor was fit to trials where objects were moving randomly. Lastly, for the working memory task, 8 regressors were fit to the following conditions: (1) 2-back body trials; (2) 2-back face trials; (3) 2-back tool trials; (4) 2-back place trials; (5) 0-back body trials; (6) 0-back face trials; (7) 0-back tool trials; (8) 0-back place trials. Given that all tasks were block designs, we fit one regressor for each task condition mentioned above.

FC estimation

Resting-state FC was estimated using standard Pearson correlations on preprocessed resting-state fMRI (Fig. 2b). Whole-brain, region-to-region resting-state FC was estimated by computing the pairwise Pearson correlation between the mean time series of every pair of regions in the Glasser et al. (2016) atlas. Network dimensionality and NPS were carried out on the unthresholded, whole-brain FC matrix. Participation coefficient was computed on three different weighted thresholds: all positive FC weights, top 10% FC threshold, top 2% FC threshold.

It has been previously shown that resting-state FC estimated with multiple linear regression better predicts task-evoked activity flow over standard Pearson correlations[1]. Thus, when predicting activity flow over resting-state FC), we estimated the resting-state connectivity-based mapping using multiple regression FC. Using ordinary least squares regression, we calculated whole-brain, region-to-region FC estimates by obtaining the regression coefficients from the equation

$$\vec{x}_i = \beta_0 + \sum_{j \neq i}^{N} \beta_{ji} \vec{x}_j + \epsilon \qquad (1)$$

for all regions x_i. We define $\overrightarrow{x_i}$ as the time series in region x_i, β_0 as the y-intercept of the regression model, β_{ji} as the FC coefficient for the jth regressor/region (which we use as the element in the jth row and the ith column in the FC adjacency matrix), and ϵ as the residual error of the regression model. N is the total number of regressors included in the model, which corresponds to the number of all brain regions. This provided an estimate of the contribution of each source region in explaining unique variance in the target region's time series. This approach of estimating FC is also described in[1,9].

Estimating basic network properties

To first test the integrity of the network partition on the HCP data set, we estimated the averaged within-network FC for each subject (Supplementary Figure 1). To ensure that only strong FC values were contributing to our estimate of within-network connectivity, we applied a 2% FC threshold, a previously used threshold for graph analyses[11]. Only 10% of subjects had a non-zero within-network FC for the ORA, and only 1% of subjects had a non-zero within-network FC for the VMM. In other words, for the majority of subjects, these networks had no functional connections that survived a 2% FC threshold.

To establish whether a network had the basic property of being a hub (i.e., high inter-network connectivity), we used several graph-theoretic techniques. We first used participation coefficient (Supplementary Figure 4), which measures the degree of inter-network connectivity at a given region/node. Given the difficulty in estimating participation coefficient with an unthresholded FC matrix, we used three different FC thresholds largely consistent with previous studies[11,38,40]: weighted positive-only FC values, 10% FC threshold, and 2% FC threshold. Participation coefficient estimated for each region was then averaged across regions within a network (for each subject separately) to obtain network level statistics for participation coefficient. Participation coefficient was implemented using the python version of Brain Connectivity Toolbox[20] (bctpy version 0.5.0).

We next estimated whether each network had a statistically significant functional connection (estimated using Pearson correlation during resting state) to every other network (Supplementary Figure 5). For all subjects, we performed the Fisher's *z*-transformation on all FC values, and performed a cross-subject, one-sided t-test for every functional connection. We then corrected for multiple comparisons using FWE permutation testing using 1000 permutations[48]. Statistical significance was assessed using an FWE-corrected $p<0.05$. For each network, we counted whether or not that network contained a statistically significant connection to every other network.

Network dimensionality measure

We adapted a previously-developed measure used to study the dimensionality of activations across space[18,49] and applied it in a graph theoretical context. Specifically, we applied it to the out-of-network connectivity patterns of functional networks estimated using resting-state fMRI. The network dimensionality measure estimates the dimensionality of the out-of-network global connectivity space for each functional network. We first obtain the correlation matrix of the Fisher's *z*-transformed out-of-network connectivity space

$$A_C = corr\left(z\left(W_{i\in C, j\notin C}\right)\right) \qquad (2)$$

where $W_{i\in C, j\notin C}$ is the m x n connectivity matrix (i.e., a subset of a the whole-brain, region-to-region adjacency matrix), where m refers to all regions within network C, and n refers to all regions not in network C. z refers to the Fisher's *z*- transform, and $corr$ performs pairwise Pearson correlations between all rows of $W_{i\in C, j\notin C}$, resulting in a A_C, which is the m x m correlation matrix from which we obtain eigenvalues. We then calculate

$$\mathbf{dim_C} = \frac{(\sum_{i=1}^{m} \lambda_i)^2}{\sum_{i=1}^{m} \lambda_i^2} \qquad (3)$$

where $\mathbf{dim_C}$ corresponds to the statistical dimensionality of network C, and λ_i corresponds to the eigenvalues of the matrix A_C[18,49].

Network pattern separation measure

We developed a new graph-theoretic measure – network pattern separation (NPS) – that characterizes the dissimilarity of global connectivity patterns between brain regions belonging to the same network (i.e., pattern-separated connectivity of a network). Using a recently defined set of functional network assignments of the Glasser et al. (2016) parcels[19], we measured the NPS for each functional network. Mathematically, we defined the NPS of a network *C* as

$$NPS_c = \sum_{i\in C} \left(\sum_{j\in C, j\neq i} 1 - scorr(\vec{\boldsymbol{w}}_{i,k\notin C}, \vec{\boldsymbol{w}}_{j,k\notin C}) \right) / (N_C^2 - N_C) \qquad (4)$$

where *scorr* refers to the Spearman's rank correlation, $\vec{\boldsymbol{w}}_{i,k\notin C}$ refers to the connectivity vector for brain region *i* to all other brain regions *k* not in network *C* (i.e., the out-of-network connectivity vector), and N_C refers to the number of regions in network *C*. NPS was computed for each subject separately using the subject's whole-brain Fisher's *z*-transformed FC matrix estimated with Pearson correlation. No threshold was applied to the matrix prior to computing NPS for each network. We compared the NPS values between pairs of functional networks by performing cross-subject t-tests for every pair of networks. We corrected for multiple comparisons using a False Discovery Rate-corrected (FDR) *p*-value of $p<0.05$[50].

Decoding task information in functional networks using multivariate pattern analysis

We performed multivariate pattern analysis[51] to decode task condition information for each of the seven HCP tasks. Whole-brain task condition activations were obtained via task GLM estimates as described above in the *fMRI preprocessing* subsection. We then segmented the whole-brain activation pattern for each subject into separate activation patterns for each functional network.

To estimate how much task information each functional network contained in its activation pattern, we performed a cross-validated *n*-way classification for each task separately, where *n* refers to the number of experimental conditions within each task (Supplementary Figure 2; Supplementary Table 1). We employed a leave-one-subject-out cross-validation scheme using random splits of the training set, which has been shown to produce more stable and robust decoding accuracies[23]. For each held-out subject, we used 100 random splits of the training data, each time randomly sampling with replacement 49 subjects to train on (approximately half of the training data), and

classifying a held-out subject's data. Thus, for each held-out subject, we generated 100 x *n* classification accuracies, from which we calculated a subject's average decoding accuracy. This approach had the advantage of allowing us to perform a random effects cross-subject t-test against chance (given the multiple decoding accuracies from each random split) rather than a fixed effects binomial test to calculate statistical significance.

Our decoder was trained using logistic regression. For tasks which had $n > 2$ conditions, we employed a multiclass classification approach with a one versus rest strategy for each class label. Logistic regression was implemented using the scikit-learn package (version 0.18) in Python (version 2.7.9). We then performed a cross-subject t-test to test whether the decoder could classify each condition within a task using a functional network's activation pattern significantly greater than chance. Since we ran classifications on all functional networks, we corrected for multiple comparisons using FDR. Statistical significance was assessed using an FDR-corrected $p<0.05$.

Estimating the representational flexibility of each functional network

The above analysis illustrated that every functional network could decode task condition information significantly above chance. However, to better quantify the degree of decodability for each task, we measured the multivariate pattern distance between the activation patterns for each task condition using Mahalanobis distance[22]. We used Mahalanobis distance as opposed to decoding statistics (e.g., accuracy) given the more intuitive interpretation of distance between activation patterns to infer highly distinct (and therefore decodable) task representations.

We used the same cross-validation scheme as the above section for this analysis. To estimate the pattern distinctness of each condition for a subject using the distribution of activation patterns from all other subjects, for each task condition $\mathbf{C} = \{c_1, \ldots, c_n\}$, we calculate the pattern distinctness $\mathbf{PD}_{c_x}$ of condition c_x as

$$\mathbf{PD}_{c_x} = \mathbf{D_M}\left(\mathbf{v}_{c_x}, \bar{\mathbf{v}}_{c_x}\right), - \mathbf{D_M}\left(\mathbf{v}_{c_x}, \bar{\mathbf{v}}_{\mathbf{C}\setminus\{c_x\}}\right) \quad (5)$$

where $\mathbf{D_M}(\mathbf{x}, \mathbf{y})$ is the Mahalanobis distance of observation $\mathbf{x}$ from the set of observations $\mathbf{y}$, $\mathbf{v}_{c_x}$ corresponds to the activation pattern during condition c_x for the held-out subject, $\bar{\mathbf{v}}_{c_x}$ corresponds to the set of activation patterns during condition c_x for all subjects in the training sample determined by the random split, and $\bar{\mathbf{v}}_{C\setminus\{c_x\}}$ corresponds to the set of activation patterns in the training sample for all conditions $\mathbf{C}$ excluding c_x. In other words, we measured the difference between matched conditions and mismatched conditions, for a held-out subject and a set of training subjects determined by the random split. For each subject, we then averaged the pattern distinctness of each condition across all random splits. This provided us with a single measure of how distinct the network's task activation patterns were across task conditions for each subject.

We performed this procedure for each task separately. To adjust for differences in distances across tasks (due to the possibility that certain tasks contain more distinct task conditions relative to others), we *z*-normalized the pattern distinctness (i.e., $\mathbf{PD}$) across networks. This allowed us to compare the pattern distinctness of each network across tasks, while preserving the relative $\mathbf{PD}$ of each network during a given task. We then computed the representational flexibility of each network by averaging the normalized $\mathbf{PD}$ across tasks (Fig. 3a). The representational flexibility for each network score was calculated within subject.

Mahalanobis distance was calculated using SciPy version 1.0.0 (the "cdist" function) with Python version 2.7.9.

Mapping whole-brain representations to functional networks via information transfer mapping

We recently developed a new procedure to characterize the role of resting-state FC in transferring task information[9]. Based on the concept of activity flow – the movement of activity between areas of the brain – via channels described by resting-state FC[1], we constructed a connectivity-based mapping that predicts the activation pattern of a target network using activity from the rest of the brain. Mathematically, we define this mapping between a target network and regions outside that network as

$$\overline{\boldsymbol{B}}_k = \boldsymbol{A}_k \bullet W_{RSFC} \qquad (6)$$

where $\overline{\boldsymbol{B}}_k$ is a *1 x n* vector corresponding to the predicted activation pattern for a target network (with *n* regions) for some task condition k, $\boldsymbol{A}_k$ is a *1 x m* vector corresponding to the activation pattern for the rest of the brain (with *m* regions), W_{RSFC} corresponds to the *m x n* matrix representing the region-to-region resting-state FC (estimated using multiple linear regression) between all regions outside the target network and regions inside the target network. Lastly, the operator • refers to the dot product. This formulation allowed us to project activation patterns to a target network using activity from regions outside that network (i.e., a spatial transformation represented as matrix multiplication).

We tested whether the connectivity-based mapping could predict the transfer of information from regions outside the target network to the target network (Fig. 4b). This required a two-step process: (1) generating predicted activation patterns for each experimental condition in the target network by estimating the activity flow to the target network from the rest of the brain; (2) training a decoder on the activity flow-predicted activation patterns of that network, and then subsequently classifying the actual (non-activity flow-predicted) activation patterns of that network using a held-out subject's data. Note, the training set did not include any data from the to-be-predicted subject's data set, and also were exclusively generated from the activity flow-predicted activations of the target network using the connectivity-based mapping in equation 6. This approach ensured that the analyses were not circular and the predictions were two-fold: (1) predicting a held-out target network's activity; (2) predicting a held-out subject's data. We used the same cross-validation scheme as in the previous section. This involved a leave-one-subject out cross-validation with random splits on the training set using logistic regression. Success of this analysis would suggest that the connectivity-based mapping from out-of-network regions to a target network could accurately predict the target network's actual activation patterns for conditions within a task. This would demonstrate the role of a network's global connectivity organization in transferring information between out-of-network regions and a target network.

To assess the statistical significance of the activity flow-predicted activation patterns, we performed a one-sided t-test to assess whether decoding accuracies were greater than chance (where chance is $1/n$, and n corresponds to the number of task conditions). Statistical significance was assessed with an FDR-corrected $p<0.05$ (Supplementary Figure 3; Supplementary Table 2).

As in the previous section, we used the scikit-learn package (version 0.18) in Python (version 2.7.9) to implement these analyses. Visualizations were mapped onto the parcellated surface using HCP's Connectome Workbench version 1.2.3[52].

Predicting representational flexibility using activity flow estimates

We wanted to demonstrate a direct relationship between the intrinsic global connectivity organization of functional networks with representational flexibility across a variety of tasks. Thus, we used the activity flow predictions of a target network across all tasks to predict the representational flexibility. In this way, the predicted representational flexibility was exclusively dependent on the combination of the intrinsic global connectivity organization of the target network and out-of-network task activations.

To predict the representational flexibility of a network using activity flow estimates from out-of-network regions, we first predicted a target network's activation pattern for each condition within a task as described above. Then, instead of training a decoder for classification, we estimated the activity flow-predicted representational flexibility using the same cross-validated Mahalanobis distance approach as when we calculated the actual representational flexibility of each network. This was done by calculating the Mahalanobis distance between a held-out subject's actual sample and the set of all other activity flow-predicted samples. In other words, we modified equation 5 and substituted the set of vectors $\bar{\mathbf{v}}_{c_x}$ and $\bar{\mathbf{v}}_{C\setminus\{c_x\}}$ with the set of activity flow-predictions of the target network (Fig. 5a).

To quantify the correspondence between the actual and activity flow-predicted representational flexibility across networks, we performed a cross-network rank correlation between the actual and predicted representational flexibility scores for each subject (Fig. 5b). To test for statistical significance, we performed a Fisher *z*-transformation on the rho values for each subject and performed a cross-subject t-test against 0.

Correlating intrinsic network properties to representational flexibility

To see if variability in intrinsic network dimensionality could explain variability in network level representational flexibility, we performed several correlation analyses relating the two measures. We first evaluated whether cross-network variance in network dimensionality related to cross-network variance in representational flexibility. For each subject, we obtained statistics for every network for both network dimensionality and NPS. In addition, to compare these two measures with a more traditional graph-theoretic measure of inter-modular connectivity, we obtained network statistics for participation coefficient[20]. However, since participation coefficient is typically used after thresholding the whole-brain FC matrix, we measured the weighted participation coefficient using three different thresholds: positive threshold (all positive FC values), 10%, and 2% (Supplementary Figure 4)[20]. Then, for each subject, we correlated the cross-network representational flexibility with the cross-network network dimensionality, NPS, and participation coefficient at each of the FC matrix thresholds (Supplementary Tables 4 and 5). To test if FC dimensionality was significantly greater than the other measures, we computed a cross-subject t-test assessing if network dimensionality was greater than each of the other measures. We corrected for multiple

comparisons using FDR-correction, and assessed significance using an FDR-corrected $p<0.05$.

We next tested if cross-subject variability in network dimensionality could explain cross-subject variability in a network's representational flexibility. Thus, for each functional network, we performed a cross-subject rank correlation between a network's representational flexibility, and each of the graph-theoretic measures mentioned above. However, to ensure that the correlations were not confounded by mean differences of any of the graph-theoretic measures (e.g., mean network dimensionality across networks), we *z*-normalized the cross-network scores for network dimensionality, NPS, and participation coefficient within subject. For each graph-theoretic measure, we obtained a rank correlation and corresponding *p*-value for each functional network. We corrected for multiple comparisons using family-wise error correction (FWE) using permutation testing (with 1000 permutations; Nichols and Holmes, 2002). Statistical significance was assessed using a FWE-corrected $p<0.05$.

Data and code availability

All data is made publicly available through the HCP[45]. All code related to analyses conducted in this manuscript will be made publicly available upon publication. In the interm, code can be made available by request.

Code to compute participation coefficient was implemented by bctpy (version 0.5.0; https://github.com/aestrivex/bctpy)[20].

Code to control for FWE rates using permutation tests can be found here: https://github.com/ColeLab/MultipleComparisonsPermutationTesting

Results

Estimating the dimensionality of a network's global connectivity patterns

We first sought to estimate the specific network properties that we hypothesized might contribute to flexible cognitive processing. We hypothesized that high-dimensional hub networks (i.e., networks with high inter-network connectivity containing pattern-separated global connections) would demonstrate high involvement during a wide range of tasks. We reasoned that the combination of high inter-network connectivity and pattern-separated global connections would lead to both increased integrative network function while limiting information interference (Fig. 1a).

We used two complementary graph-theoretic measures to target the theoretical construct of a network's global dimensionality. First, we used *network dimensionality*, which was adapted from a previously-developed measure to study the dimensionality of spatial activation patterns in cerebellum[18]. Network dimensionality measures the dimensionality of a network's out-of-network global connections. However, given the possibility that the network dimensionality statistic could be biased by the size of each network, we also devised a novel graph-theoretic measure – *network pattern separation* (NPS) – that accounts for network size. Briefly, NPS measures the dissimilarity of out-of-network FC patterns between pairs of regions belonging to the same functional network, and then averages across dissimilarities within a network. Each of these measures targeted the theoretical concept of global dimensionality in complementary ways. NPS measures the dissimilarity of global connections between every pair of regions, and can be biased by smaller, ill-defined networks. In contrast, network dimensionality looks at the dimensionality of the collective global connections of a network, and can potentially be biased by the size of the network.

We computed the network dimensionality and NPS for every functional network (Fig. 2d,e). Though network dimensionality and NPS target a distinct theoretical construct relative to region-level measures such as participation coefficient, we ran a control analysis to demonstrate the uniqueness of these measures. We computed the participation coefficient for each network using weighted participation coefficient for each subject at three FC thresholds: all positive weights, 10% FC threshold, and 2% FC threshold (Supplementary Figure 4). To test the relationship between global dimensionality measures and participation coefficient across networks, we computed the cross-network rank correlation of network dimensionality and participation coefficient, as well as NPS and participation coefficient for each subject separately. We found no significant positive correlation between participation coefficient and either network dimensionality or NPS (all average rho<0.04; all t_{99}<1.70; all p>0.05), suggesting that the measures targeting global dimensionality provide distinct information relative to participation coefficient.

Though we were interested in the broad relationship between global dimensionality and flexible activity-based representations, we also focused on differences between the FPN and other networks given our a priori hypothesis of the FPN as a flexible hub network. When comparing the FPN and other networks for each of the two graph-theoretic measures, we found that the FPN had the highest network dimensionality (pairwise t-test for FPN versus other networks, averaged t_{99}=20.12; FDR-corrected p<0.0001) and second highest NPS (pairwise t-test for FPN versus other

networks, averaged t_{99}=14.08; FDR-corrected p<0.0001, except for FPN versus ORA FDR-corrected p>0.05). The orbital affective (ORA) network had the highest NPS, but is a poorly defined network, as evidenced by extremely weak within-network FC (Supplementary Figure 1). In addition, it has previously been shown to be a poorly defined network, potentially due to low signal-to-noise ratio[19]. (In Spronk et al. 2017, the authors showed that the ORA had a network assignment confidence score that was two standard deviations below the mean.) Thus, we found that FPN had consistently high global dimensionality in the form of pattern-separated global connections, which we hypothesized to be a characteristic network property of an integrative, flexible hub network.

In addition to high global dimensionality, we wanted to ensure that FPN had the basic hub property of high inter-network connectivity. Thus, we computed the participation coefficient for all networks[11,20]. Using a weighted 2% FC threshold, we found that FPN had a significantly higher participation coefficient relative to the whole-brain average (t_{99}=15.37; FDR-corrected p<0.0001), indicating that the FPN is indeed a hub network. To next assess whether FPN's connectivity was truly global, we calculated whether FPN had at least one statistically significant functional connection to every other network. (Note, we define functional connection as a statistically significant correlation across all subjects.) Indeed, we found that FPN had at least one statistically significant functional connection to every other network estimated at the group level (significant connections, averaged r=0.13; t_{99}=13.47; FWE-corrected p<0.05). Further, when calculating this statistic for all other networks, we found that almost every network (excluding VIS1, VMM, and PMM) had at least one significant functional connection to every other network (Supplementary Figure 5). This indicates that most networks are hub networks, in the simplistic sense that they have a functional connection to every other network. These findings suggest that simple hub measures alone cannot explain the dimensionality of a network's global connectivity patterns; instead, the global dimensionality of a network collectively emerges as a function of the differences of node-specific global connectivity patterns, a property not captured by existing network statistics.

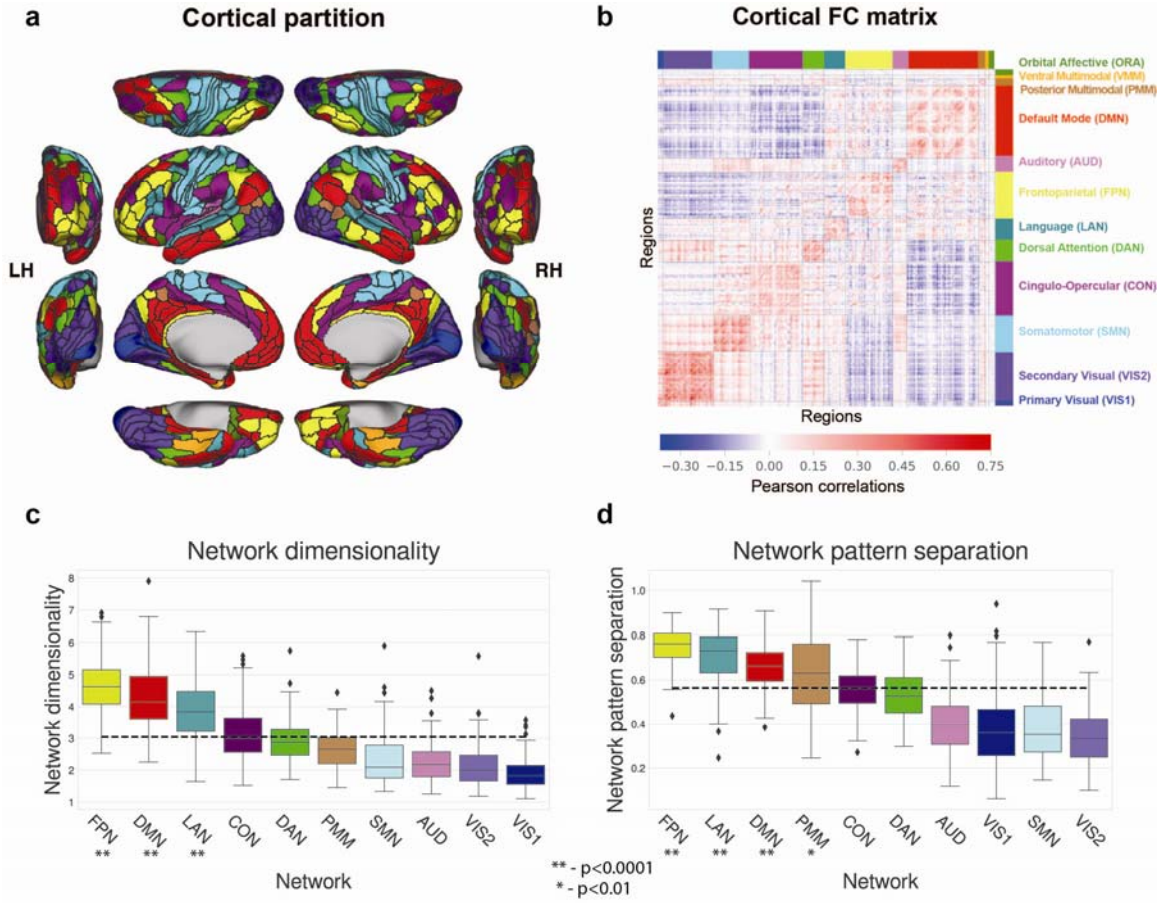

Figure 2. Measuring the intrinsic network dimensionality of human functional brain networks estimated with resting-state fMRI. a) We used a cortical parcellation atlas published in[21] with 360 parcels and a network partition of 12 functional networks estimated during resting-state fMRI[19]. **b)** A whole-brain, functional connectivity matrix estimated with standard Pearson correlations for every pair of brain regions. Regions are sorted by functional networks, and are ordered according to the colored labels along the rows and columns. **c)** We computed the network dimensionality for every functional network in the network partition. We found that the FPN had the highest network dimensionality relative to all other functional networks. **d)** We computed the network pattern separation (NPS) for every functional network in the network partition. We found that the FPN had the highest NPS of all networks containing strong within-network connectivity (Supplementary Figure 1). Note that the VMM and ORA networks were not included since they had especially low within-network connectivity, raising questions about their status as coherent networks. Asterisks denote that group-level t-statistics were significantly greater than the cross-network mean. Boxplots represent the cross-subject distribution and are organized as follows: lower and upper bounds of the box indicate the quartiles of the distribution; whiskers extend to show the rest of the distribution with outliers determined by the inter-quartile range; the line indicates the median. The dashed black lines in panels **c** and **d** indicate the cross-network mean.

Estimating the representational flexibility of functional networks using multivariate pattern analysis

We next sought to characterize a network's ability to flexibly represent task information (i.e., representational flexibility). To estimate a network's representational

flexibility, we rely on the notion that patterns of task-related activity can represent task information[22]. We performed multivariate pattern analysis to decode task conditions within each task using network-level activation patterns. We used a leave-one-subject out cross-validation scheme with random splits on the training set, allowing us to generate an averaged decoding accuracy for each subject across the random splits[23]. We then performed a cross-subject t-test against chance to assess whether we could decode task conditions significantly above chance for each task. We found that across all seven HCP tasks, data from every network could be used to decode task information significantly above chance (Supplementary Figure 2; FDR-corrected $p<0.05$ for each task). This was unsurprising, since we had many subjects ($n=100$) and trained each decoding model using distributed regions across large-scale networks. This suggested task-relevant information was widely distributed across many brain regions and functional networks, which is consistent with previous findings[9,24,25].

Since all networks could decode task information with respect to statistical significance, we instead quantified the pattern distinctness of the activation patterns associated with each task condition. Using the same cross-validation scheme, we measured the average representational distance of each task condition (relative to the other task conditions within each task) using Mahalanobis distance[26]. This provided a measure for how distinct each network's task representations were, which allows for greater decodability. We then took the averaged Z-scored pattern distinctness across all tasks to obtain the measure of representational flexibility (Fig. 3a). Consistent with our hypothesis that FPN is a flexible hub network, we found that the FPN had the highest representational flexibility across all networks (averaged t-statistic for FPN versus each network $t_{99}=11.74$; all FDR-corrected $p<0.0001$; Supplementary Table 3). These findings suggest FPN can flexibly represent task information, providing highly decodable task representations across a wide variety of tasks.

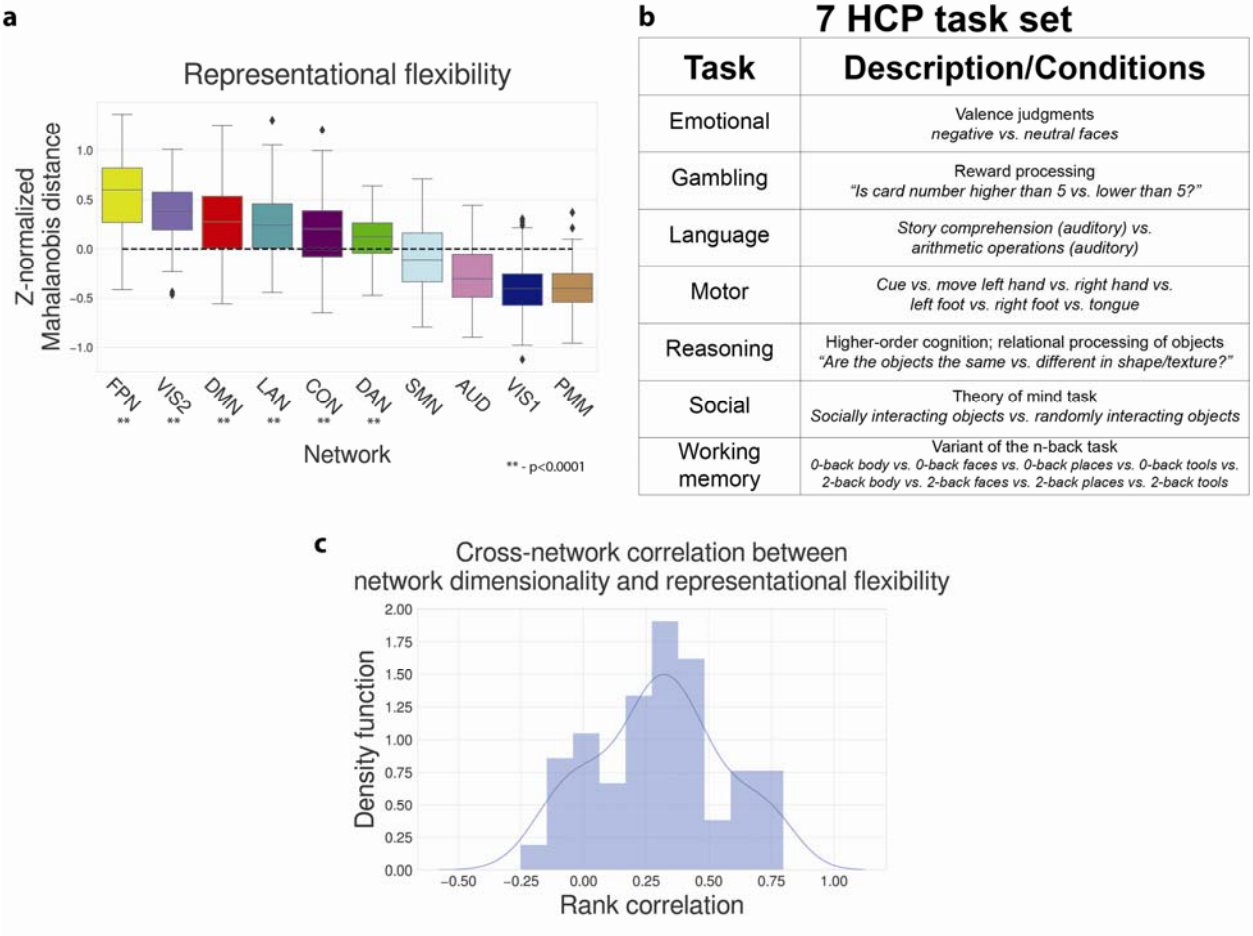

Figure 3. Network level representational flexibility and its relation to network dimensionality. Boxplots represent the cross-subject distribution and are organized as follows: lower and upper bounds of the box indicate the quartiles of the distribution; whiskers extend to show the rest of the distribution with outliers determined by the inter-quartile range; the line indicates the median. Asterisks denote that group-level t-statistics were significantly greater than the cross-network mean. **a)** Representational flexibility of each functional network across the seven HCP tasks, sorted by mean from highest to lowest. We calculated the representational flexibility for each network by averaging the Z-scored Mahalanobis distance between each task condition within a task, across all seven tasks. This estimates the flexibility (i.e., pattern separation) of network level activation patterns between task conditions. The dashed black line indicates the cross-network mean. **b)** The 7 tasks used from the HCP data set. All tasks contained two experimental conditions, excluding the motor task (six conditions) and the working memory task (8 conditions). **c)** Cross-network correlation of representational flexibility with network dimensionality. For each subject, we performed a cross-network rank correlation between representational flexibility and network dimensionality. The histogram represents the distribution of correlations across subjects estimated with a Gaussian kernel density function. Network dimensionality significantly correlated with representational flexibility (average rho=0.33; t_{99} vs. 0=10.77; p<0.0001).

Relating global dimensionality to representational flexibility

We hypothesized that networks with high-dimensional global connectivity patterns would produce flexible representations that are highly decodable. The

preceding results identified these two properties of functional networks using independent data: resting-state data was used to identify the global dimensionality of networks, and task data was used to estimate the representational flexibility of networks. We next sought to determine whether these two independent measures are related to one another.

We first performed a simple cross-network rank correlation between network dimensionality and representational flexibility, and NPS and representational flexibility. As a comparison, we also correlated representational flexibility and participation coefficient. We computed the cross-network rank correlation of every subject's representational flexibility with each graph-theoretic measure separately (Fig. 3c). We found that network dimensionality significantly explained cross-network variance in representational flexibility (cross-subject mean rho=0.33; t-test versus 0, t_{99}=10.77; p<0.0001; Supplementary Table 4). We further demonstrate that network dimensionality significantly explains more cross-network variance of representational flexibility than all other measures (Supplementary Figure 5), including participation coefficient (averaged t_{99} across all FC thresholds=9.87; FDR-corrected p<0.05). This suggests that the dimensionality of a network's global connectivity patterns can explain a network's ability to flexibly represent task information more than a previously method used to infer integrative network function (i.e., participation coefficient).

While the above analysis describes a simple correlative relationship between task-based representational flexibility and the intrinsic network properties estimated from resting-state fMRI, the analysis does not implicate a network mechanism relating the two properties. Thus, we next wanted to test whether the organization of a network's intrinsic global connectivity patterns could – via a mechanistic model of how connectivity influences task activations[1,9] – predict the representational flexibility of a network. Explicit prediction of a network's representational flexibility using the network's global connectivity organization would more rigorously test the hypothesis that its global connectivity organization is critical to its ability to flexibly integrate a wide variety of task-relevant information.

Recent work has demonstrated that the intrinsic FC architecture estimated during resting-state fMRI accurately describes the routes of activity flow – the movement of task-evoked activations between regions – during tasks[1] (Fig. 4a). We recently validated a new procedure – information transfer mapping – to infer the transfer of task information between two brain areas by mapping task representations between those regions[9]. Briefly, the procedure involves two steps: (1) mapping estimated activity flow from a source area to a target area using a resting-state connectivity-based mapping, and (2) information decoding of the actual activation pattern by a decoder trained on the activity flow-predicted activation patterns. We sought to build on these findings to demonstrate that the organization of a network's intrinsic global connections can explain a network's ability to integrate diverse sets of task-evoked information for flexible task representation.

To map activity to a target network using brain regions outside of that network, we first estimated a connectivity-based mapping by obtaining the resting-state FC patterns between regions in the target network and regions outside the network. We then predicted the task activation pattern in the target network by transforming activations from out-of-network regions into the spatial dimensions of the target network

(Fig. 4b). Briefly, this involved calculating the weighted sum of all out-of-network regions' activations weighted by the to-be-predicted region's connections. To see how well these connectivity-based mappings preserved task information in the target network, we trained a decoder using the activity flow-predicted activation patterns, and tested that decoder with the network's actual activation pattern for a held-out subject. By training the decoder using predicted activation patterns and testing on the actual activation patterns, this approach required that the activity flow-predicted activation patterns retained representations that were in the same representational geometry as the original activation pattern. Success with this procedure would suggest that the network's intrinsic global connectivity organization was responsible for its ability to integrate widespread information from the rest of the brain.

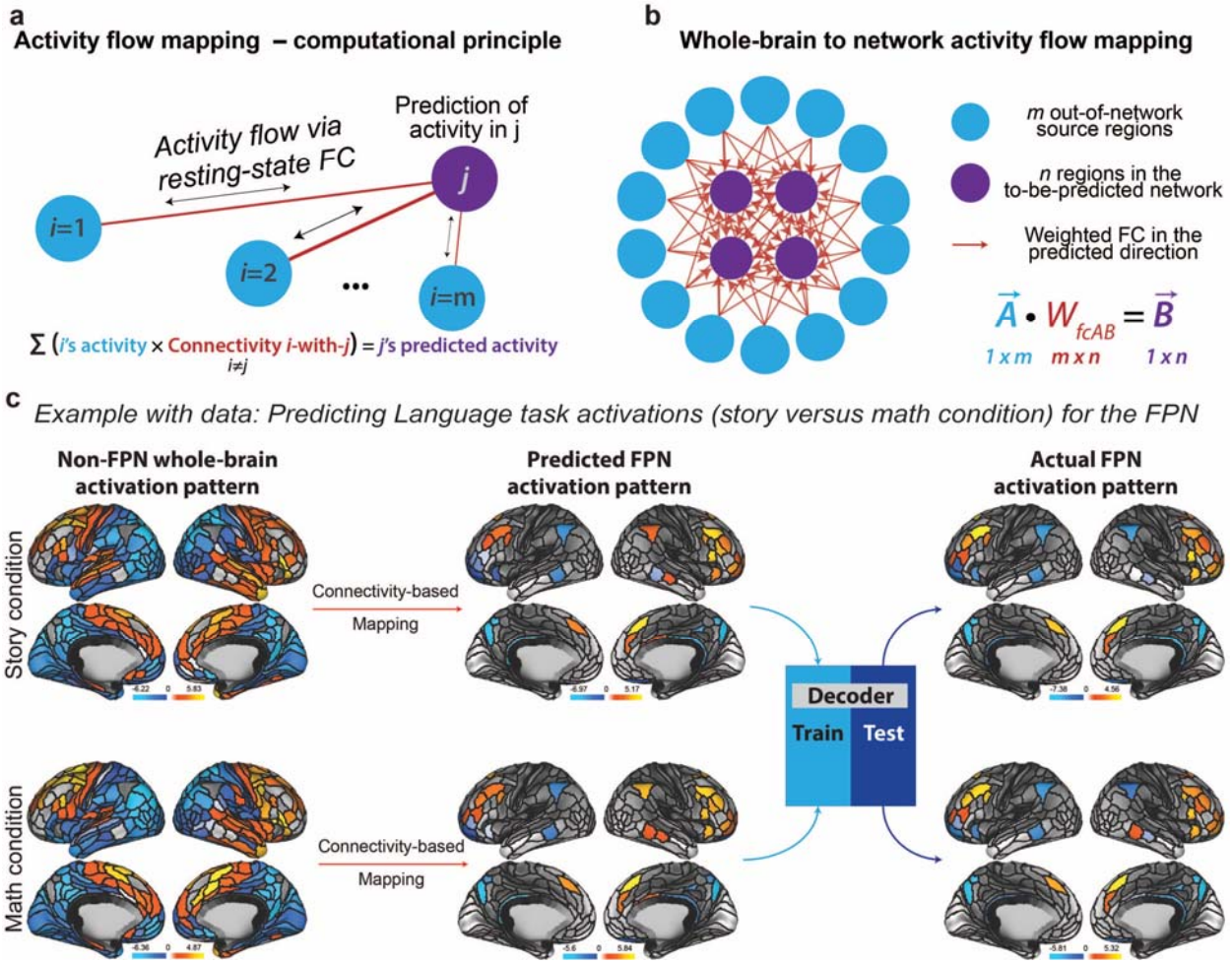

Figure 4. Network level task activation patterns can be predicted by estimating activity flow (the movement of task activations) over each network's intrinsic global connectivity pattern. a) The basic computational principle of activity flow mapping[1]. Activity flow to a held-out region can be estimated by computing the linear weighted sum of a set of regions' activity weighted by the resting-state FC to the to-be-predicted region. Figure adapted with permission from [1]. **b)** Whole-brain to target network activity flow mapping. We predicted a network's task-evoked activation pattern by estimating the activity flow from each region outside the to-be-predicted network to each region in the network. **c)** As an example, we illustrate the prediction of the FPN's activation pattern for task conditions within the Language task. The activity of each

FPN region was predicted by computing the linear weighted sum of the activity of all non-FPN regions weighted by the resting-state FC value to each FPN region. This procedure was performed to produce a predicted activation pattern for the FPN for each task condition. We subsequently trained a decoder using the predicted FPN activation patterns, and then classified the actual FPN activation patterns using a held-out subject's data. This procedure was repeated for each network and every HCP task.

We performed the information transfer mapping procedure using activations from out-of-network regions to a target network for every functional network (see Fig. 4c for an example). We then computed a network's representational flexibility based on the predicted activation pattern for that network (Fig. 5a). To see how well the activity flow-predicted representational flexibility scores recapitulated the actual representational flexibility scores for each network, we performed a cross-network rank correlation between the actual and predicted representational flexibility scores for each subject (Fig. 5b). We found that the activity flow-predicted representational flexibility accurately recapitulated its representational flexibility across networks (mean rho=0.71; t_{99}=22.39; p<0.0001). These findings suggest that the inter-network variability of representational flexibility can be explained, in part, by the intrinsic global connectivity organization of these networks. More broadly, this implicates a network mechanism for flexible representation, suggesting that the dimensionality of a network's intrinsic global connections takes part in determining the flexibility of task representation in large-scale networks.

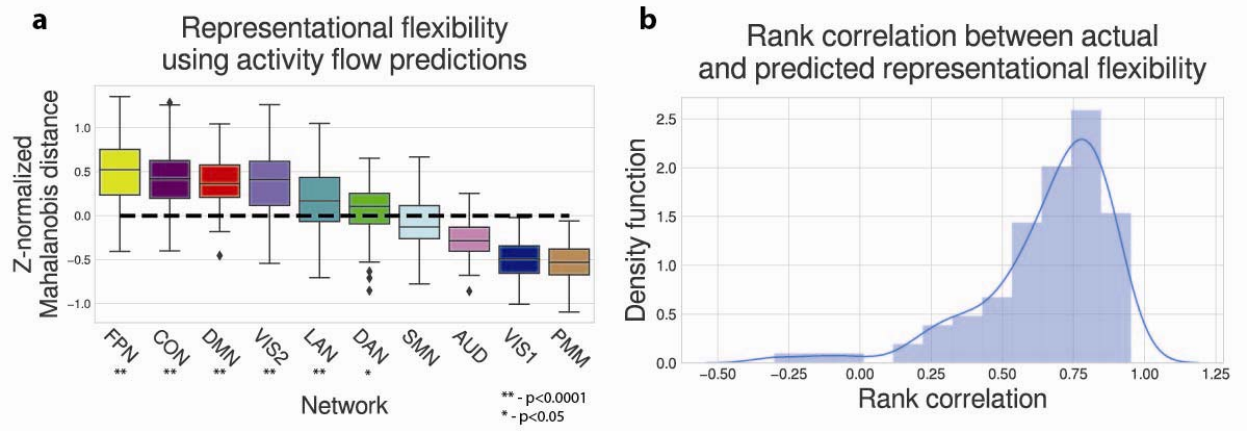

Figure 5. Predicting a network's representational flexibility by estimating activity flow over its intrinsic global network organization. a) The predicted representational flexibility for each functional network using the estimated activity flow over its global network connections estimated during resting-state. Using the activity flow approach, we generated predicted activation patterns for each network. (Note that the actual activation pattern of each functional network was excluded in generating the predicted activation pattern.) Consistent with our hypothesis, FPN had the highest predicted representational flexibility across all networks (pairwise t-test for FPN versus every network, averaged t_{99}=11.74; all FDR-corrected p<0.05). Asterisks denote that group-level t-statistics were significantly greater than the cross-network mean. Boxplots represent the cross-subject distribution and are organized as follows: lower and upper bounds of the box indicate the quartiles of the distribution; whiskers extend to show the rest of the distribution with outliers determined by the inter-quartile range; the line indicates the median. The dashed black line indicates the cross-mean network. **b)** The subject-wise

699 distribution of cross-network correlations between the actual and predicted representational
700 flexibility, estimated with a Gaussian kernel density function. For each subject, we performed a
701 rank correlation across networks between the actual and predicted representational flexibility.
702 The average rank correlation across subjects was rho=0.71 (t-test versus 0, t_{99}=22.39;
703 p<0.0001).
704
705 *Global dimensionality of FPN correlates with its representational flexibility across*
706 *individuals*
707 The preceding results demonstrated that across networks, global dimensionality
708 correlates with representational flexibility. We next sought to better establish this
709 relationship between global dimensionality and representational flexibility by additionally
710 testing for individual difference correlations between them. This would demonstrate that
711 individuals having especially high global dimensionality tend to have especially high
712 representational flexibility.
713 For each network, we performed a cross-subject rank correlation of
714 representational flexibility with each of the measures targeting global dimensionality, as
715 well as participation coefficient (Fig. 6a,c,e). For participation coefficient, we used a 2%
716 thresholded weighted FC matrix, based on previous success using this threshold with
717 participation[11,13]. Similar results were found at a 10% threshold and without any
718 threshold. We found that representational flexibility did not significantly correlate across
719 individuals with participation coefficient for any network (cross-network average
720 rho=0.01; all p>0.05; Fig. 6e). However, we found that the FPN's representational
721 flexibility significantly correlated across individuals with both its network dimensionality
722 (rho=0.30; p=0.003; FWE-corrected p=0.014) and NPS (rho=0.34; p=0.0005; FWE-
723 corrected p=0.001), though this relationship did not hold with other networks (all FWE-
724 corrected p>0.05; Fig. 6b,d). In other words, the network dimensionality and NPS of the
725 FPN, which both target the theoretical concept of global dimensionality (i.e., pattern-
726 separated global connections), relate to the inter-individual variability of the FPN's
727 representational flexibility. This suggests that our notion of global dimensionality
728 accurately provides an explanatory relationship to the network-level representational
729 flexibility of the FPN.
730

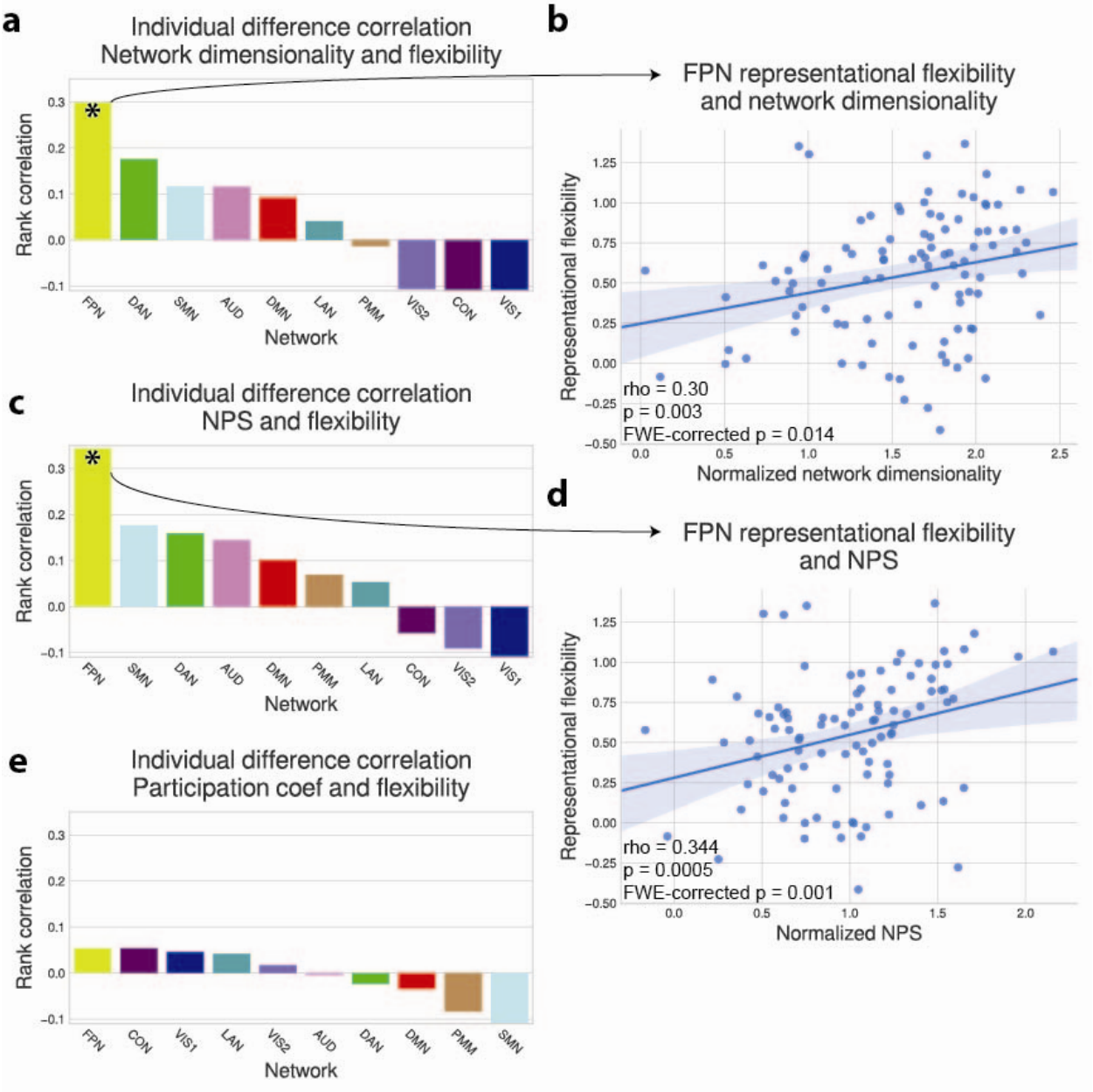

Figure 6. Individual difference correlation between network-level representational flexibility and corresponding intrinsic network properties. We performed rank correlations between a network's representational flexibility with an intrinsic network property (i.e., network dimensionality, NPS, and participation coefficient) across individuals. **a)** Individual difference rank correlation between representational flexibility and network dimensionality demonstrated a significant correlation with the FPN. **b**) Scatter plot and best-fit line between FPN's representational flexibility and network dimensionality (rho=0.30; FWE-corrected *p*=0.014). Data points represent individual subjects. The translucent blue band around the best-fit line represents the 95% confidence interval. **c)** Individual difference correlation between representational flexibility and NPS across individuals. **d)** Scatter plot and best-fit line between FPN's representational flexibility and NPS (rho=0.34; FWE-corrected *p*=0.001). **e)** Individual difference correlation between representational flexibility and participation coefficient. No networks had a significant cross-subject correlation.

Discussion

Flexible representation of cognitive information likely requires the integration of diverse signals with minimal interference. Though recent studies have characterized the neurophysiological mechanisms underlying flexible cognitive control at the single and multi-cell level[4,25,28], it has been unclear what mechanisms might allow for flexible representation at higher levels of organization. In this study, we identified a theoretical property of large-scale networks likely involved in the ability to integrate diverse sets of information with minimal signal interference: high-dimensional, pattern-separated global connectivity (i.e., high global dimensionality). Related measures of dimensionality have been previously used to study the complexity of the brain's activation spaces[18,29], and have also been hypothesized to be related to conscious integration of information[30]. Additionally, studies in the hippocampus have demonstrated the importance of pattern-separated representations for episodic recall[31]. However, a direct relationship between the human brain's large-scale network organization and flexible decoding of task-evoked activity has remained elusive. The present results provide a concrete link that suggests a mechanism of flexible representation of task information via high-dimensional global connectivity.

A recent study provided computational evidence demonstrating that the local connectivity densities of neuronal ensembles are closely related to their representational capacity in cerebellum[18]. Here we demonstrate that analogous principles also apply at the large-scale network level. However, rather than focusing on synaptic connectivity densities and cellular mechanisms such as synaptic plasticity, we used large-scale network analyses using spontaneous fluctuations to target intrinsic global network properties. Additionally, to study the representational flexibility of these large-scale networks, we used the decoding of multivariate task representations, which have been linked to the successful performance of cognitive tasks[4,16,32]. We reasoned that networks that had highly decodable activation patterns across a variety of tasks most flexibly represented task information. By directly related intrinsic network organization with activation-based representational flexibility, our findings implicate a network mechanism that contributes to the emergence of flexible hub networks via intrinsic network organization.

Recent findings have implicated the FPN as a flexible hub network for adaptive task control[8], providing evidence that regions within this network are functionally flexible[15]. Further, the intrinsic properties of the FPN have been shown to correlate across individuals with cognitive ability[33–35]. Consistent with the flexible hub theory of the FPN, we found that the FPN contained highly flexible representations across tasks. However, our results provide a link between the static intrinsic network organization of the FPN and its ability to flexibly represent cognitive information. This finding suggests that the flexible nature of the FPN is driven by a static network property, global dimensionality, which is estimated during a separate cognitive state (resting state).

Previous work has shown that the large-scale network architecture estimated at rest is largely preserved during task states[36,37]. Given this strong correspondence of intrinsic and task-evoked network architectures, the contributions of static resting-state network properties to flexible cognitive representations (in the form of flexible activation patterns) has remained unclear. Recent evidence has suggested that the intrinsic network connections estimated from spontaneous activity likely reflect the channels by

which task-evoked activity propagates between brain regions[1,2,9], providing evidence that estimated intrinsic functional connections reflect the capacity for inter-region communication. Building on these findings, the present results provide evidence that a static property of intrinsic functional networks – global dimensionality – contributes to a network's ability to flexibly represent cognitive task information.

The finding that the global dimensionality of networks contributes to their ability to flexibly represent cognitive information has several broader implications. First, it suggests that a network's global dimensionality estimated during resting state reflects the representational capacity of that network during task states. Second, it provides a specific property of network organization that can be leveraged to design future network models and architectures that can maximize representational ability. Lastly, it improves upon the previously described notion that rich club networks (or diverse club networks) underlie integrative network function[38–40]. In contrast to previous studies focusing on rich and diverse club networks, which typically characterized networks by averaging region-level connectivity properties such as weighted degree centrality[12,41,39] or participation coefficient[38,40], we sought to further characterize specific topological features emergent at the network level that might contribute to flexible representations. Global dimensionality takes into account the collective global connections of a network and the degree to which they target distinct sets of regions. Thus, global dimensionality refines the concept of an integrative hub network by taking into account the collective dimensionality of all global connections belonging to a network.

Though most studies in cognitive neuroscience are limited to a single experimental paradigm, we leveraged the HCP's multi-task dataset to investigate the brain-behavior relationship underlying flexible cognitive representation. Despite this advantage, our measure of representational flexibility was still constrained by the seven cognitive tasks included in the HCP dataset. As a particularly prominent example of a limitation of this dataset, all but the Language task used only visual stimuli. Thus, while neuroimaging studies with human participants becomes more difficult as the number of tasks increases (largely due to the experimental duration), recent advances in computational modeling has made it tractable to study the computational properties of models able to perform large number of tasks[42]. It will thus be important for future work to find converging evidence from both empirical and computational studies to study the neural and computational basis of flexible task representation.

Another limitation of this study is that the information transfer mapping procedure used to link intrinsic FC organization and task activation patterns assumes a linear relationship between sets of regions. While this provides a simple approach to approximate the flow of activity between brain regions with minimal assumptions, neural processing is typically thought to rely on nonlinear information transformation through a sequence of processing pipelines, such as in the ventral visual stream[43]. Further, transformation of information via recurrent network connections is also thought to be crucial for many cognitive tasks[42,44], as well as for pattern completion in hippocampal networks[17]. Thus, future work elucidating the contribution of nonlinear neural transformations through either feedforward or recurrent network architectures will be important to understand how information is transformed between brain systems.

In summary, we used graph-theoretic analysis of resting-state networks and information decoding across a wide range of tasks to show the co-occurrence of a

network's global dimensionality and its ability to flexibly represent task information. We then demonstrated that information from the whole brain can be mapped to specific networks by inferring the transfer of information over a network's global connectivity organization. These results demonstrate the close relationship between global dimensionality and representational flexibility at the large-scale network level, implicating a network mechanism underlying flexible representation for adaptive task control. We expect these findings to prompt further research into the relationship between network properties and their ability to produce cognitive representations, providing a deeper insight into the mechanisms underlying flexible cognitive control.

Author Contributions

T.I. and M.W.C. designed the study and the methodological tools. T.I. preprocessed and analyzed the data. T.I. and M.W.C. wrote the manuscript.

Conflict of Interest: None

Acknowledgements: We thank Travis E. Baker, Vincent B. McGinty, Joan I. Morrell, and Laszlo Zaborszky for feedback on earlier drafts of the manuscript. We also thank Miguel Vivar Lazo for helpful discussions pertaining to the integrated pattern diversity measure. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: D. Van Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors acknowledge support by the US National Institutes of Health under awards R01 AG055556 and R01 MH109520. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

References

1. Cole, M. W., Ito, T., Bassett, D. S. & Schultz, D. H. Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* (2016). doi:10.1038/nn.4406
2. Tavor, I. *et al.* Task-free MRI predicts individual differences in brain activity during task performance. *Science (80-.).* **352,** 1773–1776 (2016).
3. Fusi, S., Miller, E. K. & Rigotti, M. Why neurons mix: high dimensionality for higher cognition. *Curr. Opin. Neurobiol.* **37,** 66–74 (2016).
4. Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497,** 1–6 (2013).
5. Cohen, J. D., Dunbar, K. & McClelland, J. L. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* **97,** 332–61 (1990).
6. Rumelhart, D. E., Hinton, G. E., McClelland, J. L. & others. A general framework for parallel distributed processing. *Parallel Distrib. Process. Explor. Microstruct. Cogn.* **1,** 45–76 (1986).
7. Cole, M. W. & Schneider, W. The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage* **37,** 343–360 (2007).
8. Cole, M. W. *et al.* Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* **16,** 1348–1355 (2013).
9. Ito, T. *et al.* Cognitive task information is transferred between brain regions via resting-state network topology. *Nat. Commun.* (2017). doi:10.1038/s41467-017-01000-w
10. Dosenbach, N. U. F. *et al.* Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 11073–11078 (2007).
11. Power, J. D. *et al.* Functional Network Organization of the Human Brain. *Neuron* **72,** 665–678 (2011).
12. Cole, M. W., Pathak, S. & Schneider, W. Identifying the brain's most globally connected regions. *Neuroimage* **49,** 3132–3148 (2010).
13. Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. & Petersen, S. E. Evidence for hubs in human functional brain networks. *Neuron* **79,** 798–813 (2013).
14. Smith, S. M. *et al.* Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 13040–13045 (2009).
15. Yeo, B. T. T. *et al.* Functional specialization and flexibility in human association cortex. *Cereb. Cortex* **25,** 3654–3672 (2015).
16. Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J. & Wagner, A. D. Frontoparietal representations of task context support the flexible control of goal-directed cognition. *J. Neurosci.* **34,** 10743–55 (2014).
17. Yassa, M. A. & Stark, C. E. L. Pattern separation in the hippocampus. *Trends Neurosci.* **34,** 515–525 (2011).
18. Litwin-Kumar, A., Harris, K. D., Axel, R., Sompolinsky, H. & Abbott, L. F. Optimal Degrees of Synaptic Connectivity. *Neuron* **0,** 1153–1164.e7 (2017).
19. Spronk, M. *et al.* Mapping the human brain's cortical-subcortical functional network organization. *bioRxiv* (2017).

20. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52,** 1059–1069 (2010).
21. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* 1–11 (2016). doi:10.1038/nature18933
22. Diedrichsen, J. & Kriegeskorte, N. Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLOS Comput. Biol.* **13,** e1005508 (2017).
23. Varoquaux, G. *et al.* Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *Neuroimage* **145,** 166–179 (2017).
24. Gonzalez-Castillo, J. *et al.* Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 5487–92 (2012).
25. Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science (80-.).* **348,** 1352–55 (2015).
26. Kriegeskorte, N. & Kievit, R. A. Representational geometry: Integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17,** 401–412 (2013).
27. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1,** 1—47 (1991).
28. Stokes, M. G. *et al.* Dynamic coding for cognitive control in prefrontal cortex. *Neuron* **78,** 364–375 (2013).
29. Tononi, G., Edelman, G. M. & Sporns, O. Complexity and coherency: integrating information in the brain. (1998).
30. Tononi, G., Sporns, O. & Edelman, G. M. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Neurobiology* **91,** 5033–5037 (1994).
31. Leal, S. L. & Yassa, M. A. Integrating new findings and examining clinical applications of pattern separation. *Nat. Neurosci.* (2018). doi:10.1038/s41593-017-0065-1
32. Cole, M. W., Ito, T. & Braver, T. S. The Behavioral Relevance of Task Information in Human Prefrontal Cortex. *Cereb. Cortex* bhv072- (2015). doi:10.1093/cercor/bhv072
33. Hilger, K., Ekman, M., Fiebach, C. J. & Basten, U. Intelligence is associated with the modular structure of intrinsic brain networks. *Sci. Rep.* 1–12 (2017). doi:10.1038/s41598-017-15795-7
34. Cole, M. W., Yarkoni, T., Repovs, G., Anticevic, A. & Braver, T. S. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* **32,** 8988–99 (2012).
35. Woolgar, A. *et al.* Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proc. Natl. Acad. Sci.* **107,** 14899 LP-14902 (2010).
36. Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83,** 238–251 (2014).
37. Krienen, F. M., Yeo, B. T. T. & Buckner, R. L. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos. Trans. R. Soc. B Biol. Sci.* **369,** 20130526–20130526 (2014).

38. Bertolero, M. A., Yeo, B. T. T. & D'Esposito, M. The diverse club. *Nat. Commun.* **8,** 1277 (2017).
39. van den Heuvel, M. P. & Sporns, O. Rich-Club Organization of the Human Connectome. *J. Neurosci.* **31,** 15775–15786 (2011).
40. Bertolero, M. A., Yeo, B. T. T. & D'Esposito, M. The modular and integrative functional architecture of the human brain. *Proc. Natl. Acad. Sci.* 201510619 (2015). doi:10.1073/pnas.1510619112
41. Zuo, X.-N. *et al.* Network Centrality in the Human Functional Connectome. *Cereb. Cortex* **22,** 1862–1875 (2012).
42. Yang, G. R., Song, H. F., Newsome, W. T. & Wang, X.-J. Clustering and compositionality of task representations in a neural network trained to perform many cognitive tasks. *bioRxiv* (2017).
43. Yamins, D. L. K. *et al.* Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci.* **111,** 8619–8624 (2014).
44. Song, H. F., Yang, G. R. & Wang, X.-J. Training Excitatory-Inhibitory Recurrent Neural Networks for Cognitive Tasks: A Simple and Flexible Framework. *PLOS Comput. Biol.* **12,** e1004792 (2016).
45. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80,** 62–79 (2013).
46. Smith, S. M. *et al.* Resting-state fMRI in the Human Connectome Project. *Neuroimage* **80,** 144–168 (2013).
47. Barch, D. M. *et al.* Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* **80,** 169–89 (2013).
48. Nichols, T. E. & Holmes, A. P. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum. Brain Mapp.* **15,** 1–25 (2002).
49. Abbott, L. F., Rajan, K. & Sompolinsky, H. Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks. *Dyn. Brain An Explor. Neuronal Var. Its Funct. Significance* 1–16 (2011). doi:10.1093/acprof:oso/9780195393798.003.0004
50. Genovese, C. & Wasserman, L. Operating characteristics and extensions of the false discovery rate procedure. *J. R. Stat. Soc. Ser. B (Statistical Methodol.* **64,** 499–517 (2002).
51. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10,** 424–430 (2006).
52. Glasser, M. F. *et al.* The Human Connectome Project's neuroimaging approach. *Nat. Neurosci.* **19,** 1175–87 (2016).