

The most relevant human brain regions for functional connectivity: Evidence for a dynamical workspace of binding nodes from whole- brain computational modelling

Gustavo Deco^{1,2}, Tim J. Van Hartevelt^{3,4}, Henrique M. Fernandes^{3,4}, Angus Stevner^{3,4}, Morten L. Kringelbach^{3,4}

¹ Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat 138, Barcelona, 08018, Spain

² Institució Catalana de la Recerca i Estudis Avançats (ICREA), Universitat Pompeu Fabra, Passeig Lluís Companys 23, Barcelona, 08010, Spain

³ Department of Psychiatry, University of Oxford, Oxford, UK

⁴ Center for Music in the Brain (MIB), Dept. of Clinical Medicine, Aarhus University, DK

Running title: Binding nodes for functional connectivity

Corresponding author:

Professor Morten L. Kringelbach, Department of Psychiatry, University of Oxford, Oxford, UK.
morten.kringelbach@psych.ox.ac.uk, tel: 0044-1865613118

Conflict of interest: the authors declare to have no conflict of interest.

Abstract

In order to promote survival through flexible cognition and goal-directed behaviour, the brain has to optimize segregation and integration of information into coherent, distributed dynamical states. Certain organizational features of the brain have been proposed to be essential to facilitate cognitive flexibility, especially hub regions in the so-called *rich club* with shows dense interconnectivity. These structural hubs have been suggested to be vital for integration and segregation of information. Yet, this has not been evaluated in terms of resulting functional temporal dynamics. A complementary measure covering the temporal aspects of functional connectivity could thus bring new insights into a more complete picture of the integrative nature of brain networks. Here, we use causal whole-brain computational modelling to determine the functional dynamical significance of the rich club and compare this to a new measure of the most functionally relevant brain regions for binding information over time (“dynamical workspace of binding nodes”). We found that removal of the iteratively generated workspace of binding nodes impacts significantly more on measures of integration and encoding of information capability than the removal of the rich club regions. While the rich club procedure produced almost half of the binding nodes, the remaining nodes have low degree yet still play a significant role in the workspace essential for binding information over time and as such goes beyond a description of the structural backbone.

Introduction

Every conscious human experience of perceptions, memories and emotions relies on the flexible integration and segregation of information (Baars, 1989; Dehaene et al., 1998; Tononi et al., 1998). Importantly, this information integration is above and beyond the information available from the sum of its parts (Balduzzi and Tononi, 2008; Griffith and Koch, 2012), and as such has been linked to consciousness (but can also proceed without awareness (Mudrik et al., 2014)).

The integration of information is likely to take place in a functionally coherent, yet distributed network of brain regions, where computations are highly segregated (Power et al., 2011; Sporns, 2013). However, these computations must also be integrated *globally* and *over time* (Hansen et al., 2015). Large-scale projects such as the Human Connectome Project (Van Essen et al., 2012) and Human Brain Project (Gerstner et al., 2012) have begun to map the brain structurally and functionally, providing the experimental tools to aid a deeper understanding of how the brain integrates and segregates relevant information.

These projects use primarily non-invasive neuroimaging methods to map the human connectome (Sporns et al., 2005) as the complete map of the brain's neural elements and their structural interactions that allow complex integration and segregation of relevant information (Sporns, 2013). Structural anatomical connectivity can be precisely extracted by diffusion weighted/tensor imaging (DWI/DTI) measuring the white-matter fiber tracts constrained by the diffusion of water molecules (Basser and Pierpaoli, 1996; Beaulieu, 2002), from which the connectivity can be reconstructed by deterministic or probabilistic tract-tracing methods (Hagmann et al., 2010; Johansen-Berg and Rushworth, 2009). Similarly, functional information can be mapped *in vivo* in humans on the scale of millimetres by magnetic resonance imaging (MRI) at a temporal resolution of seconds or by magnetoencephalography (MEG) allowing a much finer temporal resolution of milliseconds. Typically, functional neuroimaging studies have measured task-related activity, but in the past decade studies have started to measure spontaneous resting-state activity over several minutes (Snyder and Raichle, 2012). These resting state studies have found highly reproducible and organized patterns of brain activity (Damoiseaux et al., 2006; Greicius et al., 2003), which overlap with task-related activity patterns (Fox and Raichle, 2007; Smith et al., 2009).

Mapping the human connectome has identified some of the crucial structural topological features of human brain architecture relevant for the functional integration and segregation of information. Graph-theoretical measures have shown that the brain is structured as a small-world network (Watts and Strogatz, 1998) around a large number of spatially distributed network communities with

clustered connectivity in which the local computations are likely to be highly segregated (Power et al., 2011; Sporns, 2013) (although see (Markov et al., 2013)). Furthermore, network hubs linking those network communities could potentially serve to ensure efficient communication and information integration (Bullmore and Sporns, 2009). Notably, some hubs show rich, dense interconnectivity and form a central core of structural connectivity (static domain) or ‘rich club’ that has been suggested to play an important role for global brain integration (Van Boven and Loewenstein, 2003; van den Heuvel and Sporns, 2011). This raises the question whether the brain regions from the rich club are sufficient for optimal information flow over time. Misic and colleagues have shown that the hippocampus is a critical convergence zone for information flow, despite its modest degree profile (Mišić et al., 2014). This suggests that a region that is structurally not densely connected could still be crucial for information flow. In this paper we suggest a new degree-naive measure to describe the functional relevance of brain regions in dynamically integrating information over time.

In order to gain more insight into functional connectivity and network measures, significant progress has been made using whole-brain computational models. Such models have shown to be able to reflect and reproduce much of the dynamics and complexity of the real brain. The models typically use various mesoscopic top-down approximations of brain complexity with dynamical networks of local brain area attractor networks (Cabral et al., 2014a), where the more advanced models use a dynamic mean-field model derived from a proper reduction of a detailed spiking neuron model (Deco et al., 2013b). Furthermore, the dynamics of whole-brain models rely on reducing the complexity of brain networks by using a given macro scale brain parcellation, which historically has been carried out based on careful studies of the properties of the underlying brain tissue (Zilles and Amunts, 2010), which has been supplemented with modern neuroimaging parcellations that typically range from tens to several hundreds of regions (Craddock et al., 2013). A current popular choice for whole brain parcellations is the automated anatomical labeling (AAL) parcellation with 116 regions in the cerebellum, cortical and subcortical regions (Tzourio-Mazoyer et al., 2002). Whole-brain computational models have been able to provide a mechanistic explanation of the origin of resting-state networks, as e.g. shown for resting-state networks derived from resting-state MRI data (Deco and Jirsa, 2012; Honey et al., 2007) and for resting-state networks derived from MEG data (Cabral et al., 2014b). Such models have been successfully used to show that both spontaneous and task-related neural activity are strongly dependent on the properties of the underlying structural connectivity and the dynamical working point, where the working point refers to the oscillatory and bifurcation properties of any given node (Deco and Corbetta, 2011).

The optimal functioning of the brain in terms of information processing depends on its ability to balance the amount of spatial segregation and integration of relevant information (Deco and Kringelbach, 2014; Sporns, 2013). Although information capability (entropy) and integration are not directly related to brain function, both measures represent fundamental aspects of brain organization and are indeed related to the exploration of the dynamical repertoire (Deco et al., 2015; Tononi et al., 1994). Moreover, emergent research has revealed that the richness of that exploration is related with task activity (Palva and Palva, 2012). Hence, measures of information capability and integration can be used to characterise the richness of the functional dynamics.

Here we used whole-brain computational modelling to clarify the functional role of the rich club and the functional impact on resting state activity. First, we compared this topological measure to a novel dynamical measure of the temporal binding of information, which we call the *dynamical workspace of binding nodes*. It is proposed that evaluation of temporal binding reveals which regions within the network are more integrative, or binding, across both space (over spatial segregated brain regions) and time (as defined by the grand average of functional activity over time) (as carefully described in Deco et al., 2015). This measure is related to the *binding problem* – that is how distributed information is bound and made available for awareness and consciousness (Crick and Koch, 1990). We compared the top ranked brain regions belonging to the rich club and to the workspace of binding nodes, and used whole-brain computational modelling to measure the impact of removing the members of either clubs on the integration and information capabilities of the human brain. This direct comparison is included to illustrate the conceptual difference of the dynamical workspace of binding nodes as a measure capable of finding functionally relevant nodes that may not exhibit high structural connectivity.

Experimental Procedures

Ethics

This study was approved by the internal research board at CFIN, Aarhus University, Denmark. Ethics approval was granted by the Research Ethics Committee of the Central Denmark Region (De Videnskabetiske Komitéer for Region Midtjylland). Written informed consent was obtained from all participants prior to participation.

Overall analysis pipeline

In order to determine the impact of brain regions on functional dynamics, we analyzed the functional and structural connectivity of neuroimaging data together with whole-brain

computational modelling. We defined a new measure that allowed us to capture the key regions for *binding* functional connectivity over time. These regions together form the *dynamical workspace of binding nodes*, comparable to the rich club in the domain of structural connectivity networks. The following pipeline outlines the necessary and sufficient steps to calculate this workspace of binding nodes, and compare the impact of these regions to the rich club. For all subsequent analysis and calculations we used both the structural (DTI) and functional (rs-fMRI) imaging data from 16 healthy subjects:

1. Structural connectivity. A group-averaged structural connectome (see Figure 1A and *Structural connectivity analysis*) was constructed and subsequently used to estimate the regions forming the rich club (see *Measure of Rich Club*).
2. Dynamical whole-brain computational modelling. The group-averaged structural and functional connectomes were further combined (see Figure 1B and *Functional connectivity analysis*) to optimize the whole-brain computational model (see Figure 1C and *Whole-Brain Dynamic Mean Field Model*). As shown in previous studies (Cabral et al., 2014a), the outcome of this model captures many significant features of the resting-state dynamics of functional connectivity.
3. Rich club membership was determined using a standard method on the structural connectivity data (van den Heuvel and Sporns, 2013a), shown in Figure 3.
4. Membership to the workspace of binding nodes was established by combining whole-brain computational modelling and functional connectivity neuroimaging data. In order to get a measure of binding we first quantified the entropy of the resting state activity (see Figure 1D and *Measure of Resting Entropy*), based on the modeled connectivity, i.e. the resting state dynamics. The workspace of binding nodes was then defined as the regions with largest impact on resting entropy by sequentially removing regions from the modeled functional connectivity network (see *Measure of Workspace of Binding nodes*).
5. The functional impact on brain dynamics of the binding nodes and rich club was determined by removing the top-ranked members and comparing their impact on integration and information capability of the modeled resting-state dynamics with perturbations (see *Whole-Brain Dynamic Mean Field Model*). Hence, novel measures of integration and information capability (see Figure 1E, *Measure of integration*, and Figure 1D, *Measure of Information Capability*) were used to evaluate the impact that removing each set of nodes had on the resting-state and stimulated dynamics. The evaluation of these measures over the stimulated dynamics corresponds to estimating what we call the *perturbational integration* and *perturbational information capability*.

Empirical neuroimaging data collection

All participants were recruited through the online recruitment system at Aarhus University. Data were collected at CFIN, Aarhus University, Denmark, from 16 healthy right-handed participants (11 men and 5 women, mean age: 24.75+/-2.54). Participants with psychiatric or neurological disorders (or a history thereof) were excluded from participation in this study.

Data collection parameters

The MRI data (structural MRI, rs-fMRI and DTI) were collected in one session on a 3T Siemens Skyra scanner at CFIN, Aarhus University, Denmark. The parameters for the structural MRI T1 scan were as follows: voxel size of 1 mm³; reconstructed matrix size 256x256; echo time (TE) of 3.8 ms and repetition time (TR) of 2300 ms. The resting-state fMRI data were collected using whole-brain echo planar images (EPI) with TR = 3030 ms, TE = 27 ms, flip angle = 90°, reconstructed matrix size = 96x96, voxel size 2x2 mm with slice thickness of 2.6 mm and a bandwidth of 1795 Hz/Px. Approximately seven minutes of resting state data were collected per subject.

The DTI data were collected using TR = 9000 ms, TE = 84 ms, flip angle = 90°, reconstructed matrix size of 106x106, voxel size of 1.98x1.98 mm with slice thickness of 2 mm and a bandwidth of 1745 Hz/Px. Furthermore, the data were collected with 62 optimal nonlinear diffusion gradient directions at $b=1500$ s/mm². Approximately one non-diffusion weighted image ($b=0$) per 10 diffusion weighted images was acquired. Additionally, the DTI images were collected with different phase encoding directions. One set was collected using anterior to posterior phase encoding direction and the second acquisition was performed in the opposite direction.

Functional connectivity analysis

The functional connectivity analysis of the resting state fMRI from the 16 participants progressed using the following three-step process:

1) Brain parcellation

We used three versions or sub-parcellations of the automated anatomical labeling (AAL) template to parcellate the entire brain into 76 cortical regions (AAL76), 90 regions (adding 14 subcortical regions, AAL90) and 116 regions (adding another 26 cerebellar regions, AAL116) (Tzourio-Mazoyer et al., 2002). The linear registration tool from the FSL toolbox (www.fmrib.ox.ac.uk/fsl, FMRIB, Oxford) (Jenkinson et al., 2002) was used to coregister the EPI image to the T1-weighted structural image. The T1-weighted image was co-registered to the T1 template of ICBM152 in MNI

space (Collins et al., 1994). The resulting transformations were concatenated and inversed and further applied to warp the AAL template (Tzourio-Mazoyer et al., 2002) from MNI space to the EPI native space, where interpolation using nearest-neighbor method ensured that the discrete labelling values were preserved. Thus the brain parcellations were conducted in each individual's native space.

2) Functional connectivity matrices

Preprocessing of the functional fMRI data was carried out using MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) Version 3.14 (Beckmann and Smith, 2004), part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). We used the default parameters of this imaging pre-processing pipeline on all participants: motion correction using MCFLIRT (Jenkinson et al., 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor and high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50.0s$).

We used tools from FSL to extract and average the time courses from all voxels within each AAL cluster. We then used Matlab (The MathWorks Inc.) to compute the pairwise Pearson correlation between all 76 regions, applying Fisher's transform to the r-values to get the z-values for the final 76x76 FC_fMRI matrix (AAL76, including only cortical regions). For further analysis, we also created a 90x90 correlation matrix (AAL90, including subcortical regions), as well as a 116x116 correlation matrix (AAL116, including all cortical, subcortical and cerebellar regions).

3) Group average

We averaged the correlation matrices for all 16 participants to get three group averaged FC_fMRI matrices: FC76 (for AAL76), FC90 (for AAL90) and FC116 (for AAL116).

Structural connectivity analysis

We generated the structural connectivity (SC) maps for each participant using the DTI data acquired. We processed the two datasets acquired (each with different phase encoding to optimize signal in difficult regions). The construction of these structural connectivity maps or structural brain networks consisted of a three-step process. First, the regions of the whole-brain network were defined using the AAL template as used in the functional MRI data. Secondly, the connections between regions in the whole-brain network (i.e. edges) were estimated using probabilistic tractography. Thirdly, data were averaged across participants.

1) Brain parcellation

Identical to the procedure used for analysing the rs-fMRI data, we used the AAL template to parcellate the entire brain into three sub-parcellations: AAL76, AAL90 and AAL116. We used the FLIRT tool from the FSL toolbox (www.fmrib.ox.ac.uk/fsl, FMRIB, Oxford) to coregister the b0 image in diffusion MRI space to the T1-weighted structural image and then to the T1 template of ICBM152 in MNI space (Collins et al., 1994). The two transformation matrices from these coregistration steps were concatenated and inverted to subsequently be applied to warp the AAL templates (Tzourio-Mazoyer et al., 2002) from MNI space to the diffusion MRI native space.

2) Analysis of interregional connectivity

We used the FSL diffusion toolbox (Fdt) in FSL to carry out the various processing stages of the diffusion MRI data. We used the default parameters of this imaging pre-processing pipeline on all participants. Following this preprocessing, we estimated the local probability distribution of fibre direction at each voxel (Behrens et al., 2003). We used the probtrackx tool in Fdt to provide automatic estimation of crossing fibres within each voxel. This has been shown to significantly improve the tracking sensitivity of non-dominant fibre populations in the human brain (Behrens et al., 2007).

The connectivity probability from a seed voxel i to another voxel j was defined by the proportion of fibres passing through voxel i that reach voxel j using a sampling of 5000 streamlines per voxel (Behrens et al., 2007). This was extended from the voxel level to the region level, i.e. in an AAL parcel consisting of n voxels, $5000 \times n$ fibres were sampled. The connectivity probability P_{ij} from region i to region j is calculated as the number of sampled fibres in region i that connect the two regions divided by $5000 \times n$, where n is the number of voxels in region i .

For each brain region, the connectivity probability to each of the other 115 regions within the AAL was calculated. Due to the dependence of tractography on the seeding location, the probability from i to j is not necessarily equivalent to that from j to i . However, these two probabilities are highly correlated across the brain for all participants (the least Pearson $r=0.70$, $p < 10^{-50}$). As directionality of connections cannot be determined based on diffusion MRI, the unidirectional connectivity probability P_{ij} between regions i and j was defined by averaging these two connectivity probabilities. This unidirectional connectivity was considered as a measure of the structural connectivity between the two areas, with $C_{ij}=C_{ji}$. We implemented the calculation of regional connectivity probability using in-house Perl scripts. For both phase encoding directions, 76x76, 90x90 and 116x116 symmetric weighted networks were constructed based on the AAL76, AAL90

and AAL116 parcellations, and normalised by the number of voxels in each AAL region; thus representing the structural connectivity network organization of the brain.

3) Group averaging

We averaged across all 16 healthy participants to get three group averaged structural connectivity network matrices: SC76, SC90 and SC116.

Whole-Brain Dynamic Mean Field Model

The computation of the dynamics sustained by the underlying empirical structural connectivity was based on the model of Deco et al. (2014b). This model describes the functional dynamics of local regions in a given brain parcellation; each composed by excitatory–inhibitory subnetworks (E–I networks), through the underlying anatomical SC which is gained from diffusion-imaging data from healthy human subjects as described above (SC76, SC90 and SC116). The excitatory synaptic currents are mediated by NMDA receptors and the inhibitory currents are mediated by GABA-A receptors. The dynamics of the model describes the time evolution of the mean synaptic activity of each local node (i.e., brain region) by following synaptic equations:

$$u_i^{(E)} = I_{0,E} + w_{EE}S_i^{(E)} + G \sum_j C_{ij} S_j^{(E)} - w_{EI,i}S_i^{(E)} + I_{ext,i} \quad (1)$$

$$u_i^{(I)} = I_{0,I} + w_{IE}S_i^{(E)} - w_{II}S_i^{(I)} \quad (2)$$

$$r_i^{(E)} = \Phi^{(E)}(u_i^{(E)}) = \frac{a_E u_i^{(E)} - b_E}{1 - \exp(-d_E(a_E u_i^{(E)} - b_E))} \quad (3)$$

$$r_i^{(I)} = \Phi^{(I)}(u_i^{(I)}) = \frac{a_I u_i^{(I)} - b_I}{1 - \exp(-d_I(a_I u_i^{(I)} - b_I))} \quad (4)$$

$$\frac{dS_i^{(E)}}{dt} = -\frac{S_i^{(E)}}{\tau_E} + (1 - S_i^{(E)})y r_i^{(E)} + \beta \eta_i(t) \quad (5)$$

$$\frac{dS_i^{(I)}}{dt} = -\frac{S_i^{(I)}}{\tau_I} + r_i^{(I)} + \beta \eta_i(t) \quad (6)$$

where $S_i^{E,I}$ denotes the average excitatory or inhibitory synaptic fraction of open channels at the local area i ($i \in [1, \dots, N]$). $r_i^{E,I}$ denotes the population firing rate of the excitatory (E) or inhibitory (I) population in the brain area i . The population firing rates are sigmoid functions ($\Phi^{(I)}$ and $\Phi^{(E)}$) of the input synaptic currents to the excitatory or inhibitory population i is given by $u_i^{E,I}$. Synaptic currents are the sum of local currents within the local E–I networks, excitatory currents from the other local nodes, and external inputs I_{ext} . The local currents in node i are the sum of constants inputs to excitatory and inhibitory populations, noted $I_{0,E}$ and $I_{0,I}$, respectively, local excitatory-to-excitatory currents $w_{EE}S_i^{(E)}$, local inhibitory-to-excitatory currents $w_{EI,i}S_i^{(I)}$, local excitatory-to-inhibitory currents $w_{IE}S_i^{(E)}$, and local inhibitory-to-inhibitory currents $w_{II}S_i^{(I)}$. The weights of these local connections are given by: $w_{EE} = 1.4$; $w_{IE} = 0.15$; $w_{II} = 1$; and the feedback inhibition weight, $w_{EI,i}$, is adjusted for each node i so that the firing rate of the local excitatory neural population is clamped around 3Hz, whenever nodes are connected or not — this regulation is known as Feedback Inhibition Control (FIC) and the algorithm to achieve it is described in Deco et al. (2014b). It has been demonstrated that the FIC constrain leads to a better prediction of the resting functional connectivity and a more realistic network evoked activity (Deco et al., 2014b). The excitatory pools are coupled by long-range connections based on the DTI Structural Matrix \mathbf{C} . The structural matrix \mathbf{C} denotes the density of fibres between cortical areas i and j and is scaled by a global scaling factor G (global conductivity parameter scaling equally all excitatory synapses). The global scaling factor is a control parameter that is adjusted to move the system to its optimal working point, defined by the point where the simulated functional connectivity maximally fits the empirical functional connectivity (i.e. the point where the Pearson correlation between the simulated and empirical FC is maximal, (Deco and Jirsa, 2012; Deco et al., 2013a; Deco et al., 2014a)). In our case, for the experimental data described above, the maximal fit (Pearson correlation between the empirical and simulated FC equal 0.56) was obtained at $G=1.1$. I_{ext} denotes the external stimulation for simulating stimuli evoked activity and its values are extracted from a Gaussian distribution with mean 0 and standard deviation of 0.02. Under resting state conditions, the external stimulations are set to zero for all regions. In equations (5) and (6) $\eta_i(t)$ is Gaussian noise and the noise amplitude at each node is $\beta=0.01$. The values of all parameters are taken from Deco et al. (2014b) and are presented in Table 1.

Analytical FC: The Moments Method

In order to obtain an analytical reduction of the whole-brain dynamical mean field model, we used the Moments Method for deriving approximated equations for the statistics of the gating variables and the synaptic activity (Deco et al., 2014b). To estimate the network’s statistics, we assume that

the noise is sufficiently weak so that the state variables fluctuate around their mean value and, by linearizing the equations, we focus on the linear fluctuations. In this way, we express the system of stochastic differential equations (1–6) in terms of the first- and second-order statistics of the distribution of synaptic gating variables: $\mu_i^{(m)}$, the expected mean gating variable of a given local neural population of type m (where $m=E$ or I) of the cortical area i , and $P_{ij}^{(mn)}$, the covariance between gating variables of neural populations of type m and n of local cortical areas i and j , respectively. The statistics are defined as:

$$\mu_i^{(m)}(t) = \langle S_i^{(m)}(t) \rangle \quad (7)$$

$$P_{ij}^{(mn)}(t) = \langle [S_i^{(m)}(t) - \mu_i^{(m)}(t)][S_j^{(n)}(t) - \mu_j^{(n)}(t)] \rangle \quad (8)$$

Where the angular brackets $\langle . \rangle$ denote the average over realizations (i.e. average over simulated trials). In vector form, the system of equations is written as:

$$\frac{d}{dt} \begin{pmatrix} \vec{S}^{(E)} \\ \vec{S}^{(I)} \end{pmatrix} = \begin{pmatrix} f^{(E)}(\vec{S}^{(E)}, \vec{S}^{(I)}) \\ f^{(I)}(\vec{S}^{(E)}, \vec{S}^{(I)}) \end{pmatrix} + \vec{\eta}(t) \quad (9)$$

where $\vec{S} = \{\vec{S}^{(E)}, \vec{S}^{(I)}\} = \{S_1^{(E)}, \dots, S_N^{(E)}, S_1^{(I)}, \dots, S_N^{(I)}\}$, $\vec{\eta}$ is Gaussian noise, $f_k^{(E)}(\vec{S}^{(E)}, \vec{S}^{(I)}) = -\frac{S_k^{(E)}}{\tau_E} + (1 - S_k^{(E)})\gamma\Phi^{(E)}(I_k^{(E)})$, and $f_k^{(I)}(\vec{S}^{(E)}, \vec{S}^{(I)}) = -\frac{S_k^{(I)}}{\tau_I} + \Phi^{(I)}(I_k^{(I)})$ for $k=1, \dots, N$.

In the following we use a linear approximation of the fluctuations. As shown in Deco et al. (2014b), Taylor expanding \vec{S} around $\vec{\mu} = \langle \vec{S} \rangle$, i.e. $S_i^{(m)} = \mu_i^{(m)} + \delta S_i^{(m)}$, up to the first order, we obtain the motion equations for the means of the gating variables and the covariance of the fluctuations around the mean. For the mean values:

$$\frac{d\mu_i^{(E)}}{dt} = \frac{d}{dt} \langle \vec{S}^{(E)} \rangle = -\frac{\mu_i^{(E)}}{\tau_E} + (1 - \mu_i^{(E)})\gamma\Phi^{(E)}(u_i^{(E)}) \quad (10)$$

$$\frac{d\mu_i^{(I)}}{dt} = \frac{d}{dt} \langle \vec{S}^{(I)} \rangle = -\frac{\mu_i^{(I)}}{\tau_I} + \Phi^{(I)}(u_i^{(I)}) \quad (11)$$

Where $u_i^{(m)}$ is the mean input current to the neural population $m=E, I$ of cortical area i , defined as:

$$\vec{u} = \begin{pmatrix} \vec{u}^{(E)} \\ \vec{u}^{(I)} \end{pmatrix} = \mathbf{W}\vec{S} + \vec{I}_0 + \vec{I}_{ext} \quad (12)$$

Where \mathbf{W} is a block matrix defined as:

$$\mathbf{W} = \begin{bmatrix} W_{EE}\mathbf{I}_N + G \cdot \mathbf{C} & \mathbf{D}(\vec{w}_{IE}) \\ W_{EI}\mathbf{I}_N & W_{II}\mathbf{I}_N \end{bmatrix}, \text{ where } \mathbf{C} \text{ is the } N \times N \text{ anatomical matrix, } G \text{ the global coupling parameter, } \mathbf{I}_N \text{ is}$$

the $N \times N$ identity matrix, $\mathbf{D}(\vec{w}_{IE})$ is a $N \times N$ diagonal matrix containing the weights of the feedback inhibition $w_{EI,i}$ as diagonal elements, and \vec{I}_0 and \vec{I}_{ext} are the vectors containing the constant and external inputs, i.e. leaving the possibility

to add external Gaussian perturbation. The perturbation for a single trial is defined by assigning random amplitudes of the external noise for each node. Thus for each trial, when perturbing the system, we use random amplitudes for all nodes which are extracted from a scaled normal distribution.

Let \mathbf{P} being the covariance matrix between gating variables \vec{S} . \mathbf{P} is a block matrix defined as:

$$\mathbf{P} = \begin{bmatrix} P^{(EE)} & P^{(EI)} \\ P^{(IE)} & P^{(II)} \end{bmatrix}. \text{ The motion equation of the covariance matrix:}$$

$$\frac{d\mathbf{P}}{dt} = \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T + \mathbf{Q}_n \quad (13)$$

Where the superscript T is the transpose, \mathbf{Q}_n is the covariance matrix of the noise, given by $\mathbf{Q}_n = \langle \vec{\eta}(t)\vec{\eta}(t)^T \rangle$, and \mathbf{A} is the Jacobian matrix of the system. \mathbf{A} is a block matrix defined as:

$$\mathbf{A} = \begin{bmatrix} A^{(EE)} & A^{(EI)} \\ A^{(IE)} & A^{(II)} \end{bmatrix}, \text{ where } A_{ij}^{(mn)} = \left\{ \frac{\partial f_i^{(m)}(\vec{\mu})}{\partial s_j^{(n)}} \right\}.$$

The synaptic input variables \vec{u} are a linear combination of the gating variables \vec{S} and, thus, the covariance matrix is given analytically by:

$$\mathbf{C}_v = \mathbf{W}\mathbf{P}\mathbf{W}^T \quad (14)$$

Note that all differential equations after using the moments reduction are deterministic and therefore were solved using the Euler method with a time step equal to $dt=0.1$ ms (after testing for convergence with several time steps of $dt=0.02, 0.05, 0.1$ and 0.15).

Measure of Integration

Integration refers to the ability of the brain to share and broadcast information across brain regions. This can be measured in different ways, either by considering the resting state functional connectivity or by measuring the effect of systematic perturbations to the resting state.

In both cases, integration can be measured using the length of the largest connected component (the so-called giant component) in the functional connectivity matrix obtained from a whole-brain computational model. The giant component is a measure of the dispersion of information and communication.

Specifically, for a given absolute threshold θ between 0 and 1 (scanning the whole range), the functional connectivity matrix, FC (using the criteria $|FC_{ij}| < \theta$, i.e. a value of 0 and 1 otherwise) can be binarized and the largest component extracted as a measure of integration. Here, the largest component is the length of the connected component of the undirected graph defined by the binarized matrix considered as an adjacency matrix. This is the largest sub-graph in which 1) any

two vertices are connected to each other by paths, and 2) that connects to no additional vertices in the super-graph. Note that we use the absolute value since this is a measure of functional communication. Finally, in order to get a measure independent of the threshold, this curve can be integrated in the range of the threshold between 0 and 1, which defines the measure of *integration*. We normalize this measure by the maximal number of connected brain areas (i.e. all N areas) for each integration step and by the number of integration steps such that the maximal integration is normalized to one.

This integration measure is computed for the resting state functional connectivity with and without perturbation. For the perturbation case, the integration is computed for each possible external stimulation (i.e. providing random amplitudes of the noise for each node as specified in the Moments method above) given rise to a perturbed resting state functional connectivity. The *perturbational integration* is given as the average of the *integration* over a large amount of instantiations of external stimulations.

Measure of (Perturbational) Information Capability

In order to study the encoding capabilities of the whole-brain network for responding to different stimuli, we analyzed its (perturbational) information capability (Deco et al., 2015). We can obtain a measure of perturbational information capability by perturbing any node of the brain network (as described in the Moments Method) and measure the functional consequences in a whole-brain computational model. After adjusting the dynamical working point of the model according to the empirical measures of resting functional connectivity (Deco et al., 2014a), we perturb the model and the overall statistical dependence among all the nodes can be estimated from the mutual information between nodes for each of the random set of inputs. Thus, the information capability is calculated as the n -dimensional (where n is the number of brain regions) entropy of the set of evoked patterns in terms of firing rate activity (assuming a Gaussian distribution), H_S , and is defined as:

$$H_S = 0.5 \log((2\pi e)^n \det C_v)$$

$$H_S = 0.5(n \log(2\pi e) + \sum_{i=1}^n \log(\lambda_i)) \quad (15)$$

Where λ_i are the eigenvalues of the covariance matrix (C_v) of the evoked activity of the excitatory units following perturbation (see Moments method). To estimate the statistical significance of perturbational information capability, we used 1000 perturbation patterns, repeated 10 times. In order to check convergence, we repeated the same procedure with 10 repetitions of 500 patterns.

Please note that the Gaussian assumption is of course an approximation but in the worst case provides an upper bound for the entropy (Cover and Thomas, 2005).

In order to avoid numerical problems in the estimation of the information capability we introduced obfuscating noise of variance $\sigma_{noise}^2 = 0.001$ so that the information capability is finally given by this formula (Norwich, 2003):

$$I_C = 0.5 \sum_{i=1}^n \log\left(1 + \frac{\lambda_i}{\sigma_{noise}^2}\right) \quad (16)$$

This equation represents the information and is derived from the difference between the entropy of the signal with noise, H_{SN} and the entropy of the noise, H_N :

$$I_C = H_{SN} - H_N$$

$$I_C = 0.5(n \log(2\pi e) + \sum_{i=1}^n \log(\lambda_i + \sigma_{noise}^2)) - 0.5(n \log(2\pi e) + \sum_{i=1}^n \log(\sigma_{noise}^2))$$

$$I_C = 0.5 \sum_{i=1}^n \log\left(1 + \frac{\lambda_i}{\sigma_{noise}^2}\right)$$

Measure of Resting Entropy

Similar to the perturbational information capability we calculate the *resting entropy* by assuming that the resting evoked patterns of activity are characterised by a Gaussian distribution. In this case, we get the following equation (again using the obfuscating noise approach as above (Norwich, 2003)):

$$E = 0.5 \sum_{i=1}^n \log\left(1 + \frac{\lambda_i}{\sigma_{noise}^2}\right) \quad (17)$$

Note however, that λ_i are the eigenvalues of the covariance matrix of the unperturbed resting patterns of activity of the excitatory units, i.e. the normal covariance matrix resulting from the model without perturbation.

Measure of Workspace of Binding nodes

We define binding as the relevance of a single brain region for the resting entropy E . The resting entropy reduction resulting from the elimination of a single region is calculated analytically by using the moments method and equation (17) described above. We test sequentially the single

elimination of each region and select the one that maximally reduces the resting entropy. We use a greedy strategy such that after that first elimination, we search for the second region that maximally reduces the resting entropy in the same way as before, and we continue so sequentially until only a single region remains. The dynamical workspace of binding nodes was defined as the 12 top-ranked brain regions according to the binding measure. Note, that we took 12 members to match the 12 rich club members obtained from the empirical data, and thus produce a fair and number-matched sequence of eliminated nodes, which is the main goal of this study. For the workspace of binding nodes, in contrast with the rich club, we did not check for significant relevance of its members but instead we selected the equivalent number of members of rich club nodes that are top ranked. In this paper, it is beyond our focus to detect the crucial number of members in the dynamical workspace of binding nodes independently. We consider that this is not trivial given the fact that a surrogate technique, such as the one used for the rich club, cannot be applied here given that the measure of binding is dynamical and depends not only on the topology but also on the working point.

Results

The main aim was to study the dynamical relevance of each brain region for the integration and segregation of information in the brain over time. In order to do this we combined functional and structural human neuroimaging data with whole-brain computational modelling, to investigate the role of brain regions in *binding* functional connectivity (see Experimental Procedures). As shown below in detail, we implemented a new measure of temporal binding of information in the brain, which is then used to define a ranking of brain regions participating in this *dynamical workspace of binding nodes*. We compared the workspace of binding nodes to the rich club and found that the top-ranked members of the workspace of binding nodes were considerably different from rich club members, albeit with some overlap. Furthermore, we found that removing the top-ranked rich club members had a significant impact on resting state activity and on measures of perturbational integration and segregation, while removal of the top-ranked members of the workspace of binding nodes had significantly greater functional impact.

The full data analysis processing pipeline is described in details in the Experimental Procedures. First, we determined the structural connectivity (SC) in a group of 16 healthy participants using probabilistic tractography on DTI data, combined with the AAL sub-parcellations of 76, 90 and 116 regions (AAL76, AAL90, AAL116; see Experimental Procedures and Figure 1A). We also used standard tools to determine the functional connectivity (FC) of the spontaneous resting state fluctuations using the same group of participants (Figure 1B).

Second, we used a standard method with graph theoretical tools to determine the rich club members (van den Heuvel and Sporns, 2013a). We probed the structural and functional first order and second order characteristics of the various sub-parcellations, and specifically we extracted the rich club members at the rich club coefficient shown in Figure 2A (van den Heuvel and Sporns, 2013a). A set of brain areas in the network shows a rich club structure if its level of interconnectivity exceeds the level of connectivity that can be expected on basis of chance alone. In particular, the weighted rich club coefficient $\Phi(k)$ is computed as the ratio between 1) the weights of connections present within the subnetwork S of regions with a degree $>k$ and 2) the total sum of weights present within a subset of same size of the top ranking connections in the network. The normalized rich club coefficient $\Phi_{norm}(k)$ (red curve) is computed by dividing $\Phi(k)$ (continuous black curve) by $\Phi_{random}(k)$ (dashed black curve), with $\Phi_{random}(k)$ computed as the average rich club coefficient for each k of a set of 1000 randomized graphs (acquired by randomizing the adjacency matrix A preserving the degree sequence) (Rubinov and Sporns, 2010). Membership of the rich-club is determined if the normalized rich club coefficient $\Phi_{norm}(k)$ is larger than one, for a range of increasing degrees level k . We determined the degree (Figure 2B), mean resting functional connectivity (rFC, Figure 2D) and variance for each brain area respectively (Figure 2F). The qualitative comparisons of those curves show a weak relationship between the structural (degree) and functional (mean rFC and variance), indeed the level of correlations between those curves was moderate to low (correlation rFC and degree: 0.53; variance and degree:0.37). The mean rFC and variance for each brain area were calculated by summing those values (shown in Figure 2C and 2E) over all the pairs connecting that region with the rest of other regions.

Third, we used whole-brain computational modelling (using the Moments reduction of a stochastic mean field model, see Experimental procedures, which confers the advantage of handling the stochasticity in a deterministic and semi-analytical way) combined with empirical functional and structural connectivity to determine the *workspace of binding nodes* for the data (Figures 1C and 3, and Experimental Procedures). The workspace of binding nodes is defined as the 12 top-ranked brain regions according to binding. Here the top 12 was used to allow a comparison of the functional workspace of binding nodes to the 12 hubs previously shown in the structural rich club (see Methods for a discussion of this definition). In Figure 3A we show five cases of sequential elimination according to the lowest degrees (blue curve), to the largest degree (black curve), to the largest nodal variance (dashed red curve), to a random selection (dashed green curve), and to using the workspace of binding nodes strategy (red curve), i.e. searching for the brain region that maximally reduces the resting entropy. The largest degree and largest variances reduces the resting entropy very efficiently (as intuitively expected), much more than the sequential elimination of the

lowest degree regions, but not so efficiently as the optimal elimination of workspace of binding nodes regions (as expected by definition). To further probe this, we provide an inset which shows that eliminating the top-ranked workspace of binding nodes regions indeed (by definition) maximally reduces the resting entropy and therefore defines a ranking of regions that are maximally relevant for the processing of information (as we will show in Figure 6). Many of the regions in the workspace of binding nodes are also high degree “rich club” regions, but not all high degree regions are ranked in the same way and also low degree regions are part of the workspace of binding nodes. This overlap demonstrates that the rich club is of course related to the binding nodes but that the latter definition is better at capturing the functional relevance of the functional integration and information capability of the brain. In order to study the intersection between high degree and workspace of binding nodes regions, we show in Figure 3B the degree of the sequentially eliminated regions for the rich club case (black curve) and for the workspace of binding nodes case (red curve). It is clear that for the sequential elimination of regions according to largest degree, we obtain a monotonically decreasing curve, whereas for the elimination according the maximal reduction of resting entropy not necessarily the highest degrees regions are selected. Nevertheless, a tendency of selecting high degree regions is also shown in the stochastic decreasing of the curve. Figure 3C shows the intersection between those two groups (high degree and workspace of binding nodes). This figure clearly shows that the regions in the dynamical workspace of binding nodes do not have a one to one correspondence with the high degree regions, but that the later are a subset of the first. Figure 3D illustrates, for each brain area, the degree and the percentage of reduced resting entropy by elimination of that region. Figure 3E shows their correlation ($r=0.56$, $p<0.001$). Again, even if there is a correlation (as expected and speculated) between high degree regions and its relevance for the global dynamics, as evaluated by the resting entropy, the relationship is not deterministic (one to one). Not all high degree regions are relevant for the global dynamics. In summary, the results show that ranking regions according to their temporal binding, i.e. impact on the resting entropy, is a more direct and effective way of ranking the “dynamical” relevance of regions. Also note that the structural characterization of the rich club property does not rank each single region, as it is describing a property of a set of regions. Only if a regions has a k-level above a certain threshold, then the Φ_{norm} is calculated to identify rich club members.

Fourth, we constructed four different whole-brain networks (see Experimental Procedures). The first case (SC) corresponds to the use of the full structural connectivity matrix. The second case (SC_NRC, SC minus Non-Rich Club regions) corresponds to the SC matrix but without 12 brain regions with the lowest degree, and therefore not exhibiting the rich club structure as the low degree means they do not reach the k-level threshold necessary to qualify for the rich club. The third case

(SC_RC, SC minus rich club regions) corresponds to the SC matrix but without the top-ranked 12 regions in the rich club with maximal degree (i.e. number of structural connections with other regions). The fourth case (SC_BC, SC minus regions in the workspace of binding nodes), corresponds to the main finding of this paper, namely the use of the SC matrix without the top-ranked 12 brain regions in the dynamical workspace of binding nodes, i.e. showing the largest reduction of the information capability under resting state conditions. In Figure 4 we plot these four networks with the corresponding structural connectivity matrices (top row, Figure 4), the covariances matrices under resting conditions (middle row, Figure 4), and the degree of each brain regions (bottom row, Figure 4). In Figure 5 we show the effects of different node removal strategies, including topological (top-ranked brain regions in degree, betweenness centrality and participation coefficient) and dynamical features (top-ranked brain regions in workspace of binding nodes), in the entropy of resting state. The figure shows that the resting entropy is maximally reduced for the case of the greedy removal of the workspace of binding nodes, when compared to all other strategies tested in this study, including degree and betweenness centrality.

Finally, we used whole-brain computational modelling to study the integration and segregation capabilities of these four structural networks, which allowed us to detect which brain regions are more relevant for the global processing of modelled external information. In particular we were careful to use our previous whole-brain computational model which re-balances the level of excitation/inhibition of each brain region in order to maintain the negligible short-range correlations and introduce through the underlying neuroanatomical coupling the proper large-range functional correlations (i.e. the resting state networks, see Experimental Procedures). This dynamic mean field, derived from a proper reduction of a detailed spiking neuron model, describes the local brain dynamics of each brain area. After adjusting the dynamical working point of the model (see Methods) to the empirical resting FC, we study the integration and information capability of the whole brain model, first by considering its resting state dynamics, and second by considering its evoked activity obtained by stimulating off-line *in silico* the whole brain model by a random set of Gaussian inputs (see Methods). In Figure 6, we show the perturbational integration and information capability for the resting state dynamics (dashed curve) and for the stimulated case, i.e. the corresponding perturbational integration and information capability. To estimate the statistical significances of perturbational integration and information capability, we used 1000 different random stimulations of all the nodes in the model over time (repeated 10 times), and in order to check convergence we repeated the procedure for 10 repetitions of 500 different patterns (see Experimental Procedures). As described in the Moments method, the perturbation for a single trial is defined by assigning random amplitudes of the external noise for each node. Thus for each trial,

when perturbing the system, we use random amplitudes for all nodes which are extracted from a scaled normal distribution. In both cases we got the same highly significant results, with error intervals (bars not visible in the graphic) below 0.002 and corresponding p values below 0.000001. We show the results of using three sub-parcellations of the AAL (AAL76, AAL90 and AAL116). In all cases, the maximal reduction of integration and information capability is observed for the SC_BC case where the members of the workspace of binding nodes were eliminated. Equally, the figure also shows that removing the high degree rich club brain regions (SC_RC) is damaging in terms of functional integration and segregation – although not as damaging as removing the regions from the workspace of binding nodes. In Figure 7 we show 3D renderings of the members of the workspace of binding nodes and rich club. Furthermore, Figure 7 shows a comparison of the workspace of binding nodes and rich club nodes with in the centre of the panel a representation of the overlap of nodes common to both.

Discussion

We used whole-brain computational modelling to address a key question in neuroscience: namely what brain regions are most important for integration of information processing. While some progress has been made through defining topological measures such as the so-called *rich club*, the functional consequences of these brain regions have not been assessed and in particular the importance of time in brain processing has not been fully explored. Here we used whole-brain computational modelling on neuroimaging data from a group of healthy participants to study the functional role of the rich club on information processing in the healthy brain, by measuring the impact on perturbational integration and segregation (information capability) upon removing the rich club. We found that there was a significant influence of the rich club on these measures but we also compared this to our novel measure of binding, which ranks brain regions according to their binding capability. This new binding measure is significantly more computationally expensive compared to the simple rich club algorithm but also performs significantly better. There is some overlap between nodes (almost half) but the remaining binding nodes were non-rich club members, with even some of them having low degree – a third of the binding nodes have degree below average. This could be interpreted in the context in the important findings by Misic and colleagues (Misic et al., 2014) who found that the hippocampus, despite having a relatively low degree, acts as a critical convergence zone and its functional capacity is shaped by its embedding in the large-scale connectome.

Still, compatible with the main notion of the rich club we showed that degree is the topological feature with maximal impact in resting entropy, when compared to other key metrics such as participation coefficient and betweenness centrality (Figure 5). We also found that removal of the top-ranked rich club members significantly impaired the integration and information capability of the brain compared to non-rich club members (compare SC_NRC with SC_RC in Figure 6). This significant reduction following removal of rich club members strengthens recent analytic results (Goni et al., 2014) and results from a psychiatric context (van den Heuvel et al., 2013). Yet, we also found that removal of members of the dynamical workspace of binding nodes further significantly reduced the integration and information capability of the brain compared to removing the rich club members (compare SC_BC with SC_RC, Figure 6). As the dynamical workspace of binding nodes is defined by the measure of resting entropy, this metric is by definition most influenced by removing the top ranking nodes on this measure. However, this clearly shows that the binding club, defined in functional connectivity, is conceptually very different from the rich club, defined in structural connectivity. Furthermore, comparing both the rich club and the workspace of binding nodes on the functional measure of perturbational integration and information capability, shows that the removal of the binding club has a maximal effect on this measure when compared to removal of the structurally defined rich club. While there is some overlap between the members of the workspace of binding nodes and the rich club, there are also clear differences (see Figure 3). This finding that an area that is not well connected structurally can be crucial in information flow is in line with a recent study on information flow (Misic et al., 2014). The results show that approximately 30% of the binding nodes reveal a below average degree and are thus considered to not be well connected.

Evaluating temporal binding reveals which regions within the network are more integrative, or binding, across both space and time. These regions can be said to comprise the brain's *dynamical workspace of binding nodes*. The brain regions included in the workspace of binding nodes are straightforwardly ranked according to their level of temporal binding.

Our novel measure of binding could potentially offer new information on the *temporal binding* of information which is related to the *binding problem* (Engel and Singer, 2001), i.e. how distributed information is bound and made available to consciousness (Crick and Koch, 1990). We hypothesize that regions belonging to the temporal 'dynamical workspace of binding nodes' could be important for mediating the concatenation of different brain states during cognitive sequences, and as such may be important for facilitating awareness.

Yet, future studies will have to rigorously assess any claims for the utility of the dynamical workspace of binding nodes in participating in and mediating between different state changes, including sleep and wakefulness, and for the ability to distinguish between the human brain in health and disease (Deco and Kringelbach, 2014).

It should be noted that in this study the widely used and accepted AAL parcellation was employed. There are however other parcellation schemes available, e.g. the parcellations described by Craddock and colleagues consisting of 200 and 400 cortical, subcortical and cerebellar areas (Craddock et al., 2013). Although these areas described by Craddock are uniform and unbiased, the AAL is based on functional properties of anatomical data.

The analysis has also brought forward a rather interesting finding in terms of lateralisation. The rigorous analysis using the greedy removal of regions in the workspace of binding nodes that have a maximal impact on the brain dynamics, regardless of interconnectivity, has shown that regions that are most important for the dynamics do not in fact have to be symmetrical between hemispheres. Although some brain regions in the workspace of binding nodes are lateralised, there are also brain regions that are included in this workspace of binding nodes that appear in both hemispheres. The question of lateralisation should be further explored in future work.

An interesting finding in this study is that the left hemisphere appears more pronounced in terms of binding nodes. This of course could open a discussion regarding lateralisation and what this could possibly mean. It has been found previously in structural studies that the left hemisphere presents more central or indispensable regions for the whole-brain structural network compared to the right hemisphere (Iturria-Medina et al., 2011). In light of this research it may not be surprising that this study indeed finds a left hemispheric bias in terms of binding nodes.

One potential limitation of this paper is the choice of the whole-brain model. Here we used an established dynamic mean field model, which has been successful in describing empirical resting data (xx Deco). Yet, other future models might perform better in terms of fitting with the empirical data and thus could potentially yield more precise results. It is an open question if the binding nodes obtained with other models would be stable. Still, all models can be approximated through linearisation, and a simple spatial autoregressive model yields similar results (Messe et al., 2014). We thus speculate that the binding nodes will be stable across models but this remains an open empirical question.

Overall, in this article we use whole-brain computational modelling to introduce a promising new measure of binding in the human brain, by which brain regions are ranked according to their level of temporal binding. We show how the top-ranked brain regions in this workspace of binding nodes are different from the brain regions in the structurally defined rich club (albeit with some overlap) (see Figure 7). The overlap found between the workspace of binding nodes and the rich club consists of both cortical and subcortical regions. Specifically, we found the following regions in both the workspace of binding nodes and the rich club: the left putamen, left precuneus, and left mid-occipital cortex and the right putamen, right hippocampus and right superior frontal gyrus. We use whole-brain computational modelling to show that the workspace of binding nodes outperforms the rich club in terms of integration and information capability. Whereas it is obvious that structural and functional connectivity are two different measures, this study reveals that only some of the areas from the structural backbone of the brain play a key role in information integration. Some crucial areas are not captured by this structural measure and as such we suggest the use of the dynamical workspace of binding nodes for functional connectivity analyses. These findings suggest that node degree alone may not be a sufficient player to describe the complexity of integrative processes across the brain. Hence, in order to be considered an important region in regulating information flow across the brain (high functional richness), it is essential to comprise additional topological features, such as measures of centrality (van den Heuvel and Sporns, 2013b) or network embedding or context (Misic et al., 2014), to the original concept of *rich club*. This could potentially lead to a high degree of convergence between the two metrics, and as such significantly expand our knowledge on which intrinsic network mechanisms are driving crucial aspects of the brain's complex dynamics. Furthermore, a comprehensive understanding of the key topological drivers of temporal binding may help to elucidate which local and global network behaviours should be 'induced' to promote optimal dynamical balance, in different neuropsychiatric disorders.

Acknowledgements

GD was supported by the ERC Advanced Grant: DYSTRUCTURE (n. 295129), by the Spanish Research Project SAF2010-16085 and the FP7-ICT BrainScales. MLK was supported by the ERC Consolidator Grant: CAREGIVING (n. 615539), TrygFonden Charitable Foundation and by Center for Music in the Brain, funded by the Danish National Research Foundation (DNRF117).

References

- Baars, B.J., 1989. *A Cognitive Theory of Consciousness*. Cambridge University Press, Cambridge, MA.
- Balduzzi, D., Tononi, G., 2008. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput Biol* 4, e1000091.
- Basser, P.J., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson B* 111, 209-219.
- Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system - A technical review. *NMR Biomed* 15, 435-455.
- Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* 23, 137-152.
- Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34, 144-155.
- Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 50, 1077-1088.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10, 186-198.
- Cabral, J., Kringelbach, M.L., Deco, G., 2014a. Exploring the network dynamics underlying brain activity during rest. *Prog Neurobiol* 114, 102-131.
- Cabral, J., Luckhoo, H., Woolrich, M., Joensson, M., Mohseni, H., Baker, A., Kringelbach, M.L., Deco, G., 2014b. Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. *Neuroimage* 90, 423-435.
- Collins, D., Neelin, P., Peters, T., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography* 18, 192-205.
- Cover, T.M., Thomas, J.A., 2005. *Elements of Information Theory*.
- Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P., 2013. Imaging human connectomes at the macroscale. *Nat Methods* 10, 524-539.
- Crick, F., Koch, C., 1990. Towards a Neurobiological Theory of Consciousness. *Seminars in the Neurosciences* 2, 263-275.

Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the United States of America* 103, 13848-13853.

Deco, G., Corbetta, M., 2011. The dynamical balance of the brain at rest. *Neuroscientist* 17, 107-123.

Deco, G., Jirsa, V.K., 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J Neurosci* 32, 3366-3375.

Deco, G., Jirsa, V.K., McIntosh, A.R., 2013a. Resting brains never rest: computational insights into potential cognitive architectures. *Trends Neurosci* 36, 268-274.

Deco, G., Kringelbach, M.L., 2014. Great Expectations: Using Whole-Brain Computational Connectomics for Understanding Neuropsychiatric Disorders. *Neuron* 84, 892-905.

Deco, G., McIntosh, A.R., Shen, K., Hutchison, R.M., Menon, R.S., Everling, S., Hagmann, P., Jirsa, V.K., 2014a. Identification of optimal structural connectivity using functional connectivity and neural modeling. *J Neurosci* 34, 7910-7916.

Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G.L., Mantini, D., Corbetta, M., 2014b. How local excitation-inhibition ratio impacts the whole brain dynamics. *J Neurosci* 34, 7886-7898.

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., Corbetta, M., 2013b. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J Neurosci* 33, 11239-11252.

Deco, G., Tononi, G., Boly, M., Kringelbach, M.L., 2015. Rethinking segregation and integration: contributions of whole-brain modelling *Nat Rev Neurosci* 16, 430-439.

Dehaene, S., Kerszberg, M., Changeux, J.P., 1998. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14529-14534.

Engel, A.K., Singer, W., 2001. Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci* 5, 16-25.

Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700-711.

Gerstner, W., Sprekeler, H., Deco, G., 2012. Theory and simulation in neuroscience. *Science* 338, 60-65.

Goni, J., van den Heuvel, M.P., Avena-Koenigsberger, A., Velez de Mendizabal, N., Betzel, R.F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.P., Sporns, O., 2014. Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences of the United States of America* 111, 833-838.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl Acad. Sci. USA* 100, 253-258.

Griffith, V., Koch, C., 2012. Quantifying synergistic mutual information. *arXiv*, 1205.4265.

Hagmann, P., Cammoun, L., Gigandet, X., Gerhard, S., Ellen Grant, P., Wedeen, V., Meuli, R., Thiran, J.P., Honey, C.J., Sporns, O., 2010. MR connectomics: Principles and challenges. *J Neurosci Methods* 194, 34-45.

Hansen, E.C., Battaglia, D., Spiegler, A., Deco, G., Jirsa, V.K., 2015. Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105, 525-535.

Honey, C.J., Kotter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc Natl Acad Sci U S A* 104, 10240-10245.

Iturria-Medina, Y., Perez Fernandez, A., Morris, D.M., Canales-Rodriguez, E.J., Haroon, H.A., Garcia Penton, L., Augath, M., Galan Garcia, L., Logothetis, N., Parker, G.J., Melie-Garcia, L., 2011. Brain hemispheric structural efficiency and interconnectivity rightward asymmetry in human and nonhuman primates. *Cereb Cortex* 21, 56-67.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825-841.

Johansen-Berg, H., Rushworth, M.F., 2009. Using diffusion imaging to study human connectional anatomy. *Annu Rev Neurosci* 32, 75-94.

Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy, H., 2013. Cortical high-density counterstream architectures. *Science* 342, 1238406.

Messe, A., Rudrauf, D., Benali, H., Marrelec, G., 2014. Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Comput Biol* 10, e1003530.

Misic, B., Goni, J., Betzel, R.F., Sporns, O., McIntosh, A.R., 2014. A network convergence zone in the hippocampus. *PLoS Comput Biol* 10, e1003982.

Mudrik, L., Faivre, N., Koch, C., 2014. Information integration without awareness. *Trends Cogn Sci*.

Norwich, K.H., 2003. *Information, sensation and perception*. Academic Press, San Diego.

Palva, S., Palva, J.M., 2012. Discovering oscillatory interaction networks with M/EEG: challenges and breakthroughs. *Trends Cogn Sci* 16, 219-230.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional network organization of the human brain. *Neuron* 72, 665-678.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059-1069.

Smith, S.M., 2002. Fast robust automated brain extraction. *Human Brain Mapping* 17, 143-155.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America* 106, 13040-13045.

Snyder, A.Z., Raichle, M.E., 2012. A brief history of the resting state: the Washington University perspective. *Neuroimage* 62, 902-910.

Sporns, O., 2013. Network attributes for segregation and integration in the human brain. *Curr Opin Neurobiol* 23, 162-171.

Sporns, O., Tononi, G., Kotter, R., 2005. The human connectome: A structural description of the human brain. *PLoS Comput Biol* 1, e42.

Tononi, G., Edelman, G.M., Sporns, O., 1998. Complexity and coherency: integrating information in the brain. *Trends Cogn Sci* 2, 474-484.

Tononi, G., Sporns, O., Edelman, G.M., 1994. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci U S A* 91, 5033-5037.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273-289.

Van Boven, L., Loewenstein, G., 2003. Social Projection of Transient Drive States. *Personality and Social Psychology Bulletin* 29, 1159-1168.

van den Heuvel, M.P., Sporns, O., 2011. Rich-club organization of the human connectome. *J Neurosci* 31, 15775-15786.

van den Heuvel, M.P., Sporns, O., 2013a. An anatomical substrate for integration among functional networks in human cortex. *J Neurosci* 33, 14489-14500.

van den Heuvel, M.P., Sporns, O., 2013b. Network hubs in the human brain. *Trends Cogn Sci* 17, 683-696.

van den Heuvel, M.P., Sporns, O., Collin, G., Scheewe, T., Mandl, R.C., Cahn, W., Goni, J., Hulshoff Pol, H.E., Kahn, R.S., 2013. Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry* 70, 783-792.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen,

S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., Consortium, W.U.-M.H., 2012. The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222-2231.

Watts, D., Strogatz, S., 1998. Collective dynamics of 'small-world' networks. *Nature* 393, 440-442.

Zilles, K., Amunts, K., 2010. Centenary of Brodmann's map--conception and fate. *Nat Rev Neurosci* 11, 139-145.

Table 1 - Parameters of dynamic mean-field model.

Excitatory gating variables
$a_E = 310$ (nC ⁻¹)
$b_E = 125$ (Hz)
$d_E = 0.16$ (s)
$\tau_E = \tau_{NMDA} = 100$ (ms)
$I_{0,E} = 0.3820$ (nA)
Inhibitory gating variables
$a_I = 615$ (nC ⁻¹)
$b_I = 177$ (Hz)
$d_I = 0.087$ (s)
$\tau_I = \tau_{GABA} = 10$ (ms)
$I_{0,I} = 0.2674$ (nA)
Fixed local connectivity parameters
$W_{EE} = 1.4$
$W_{II} = 1$
$W_{IE} = 0.15$
kinetic parameter
$\gamma = 0.641/1000$

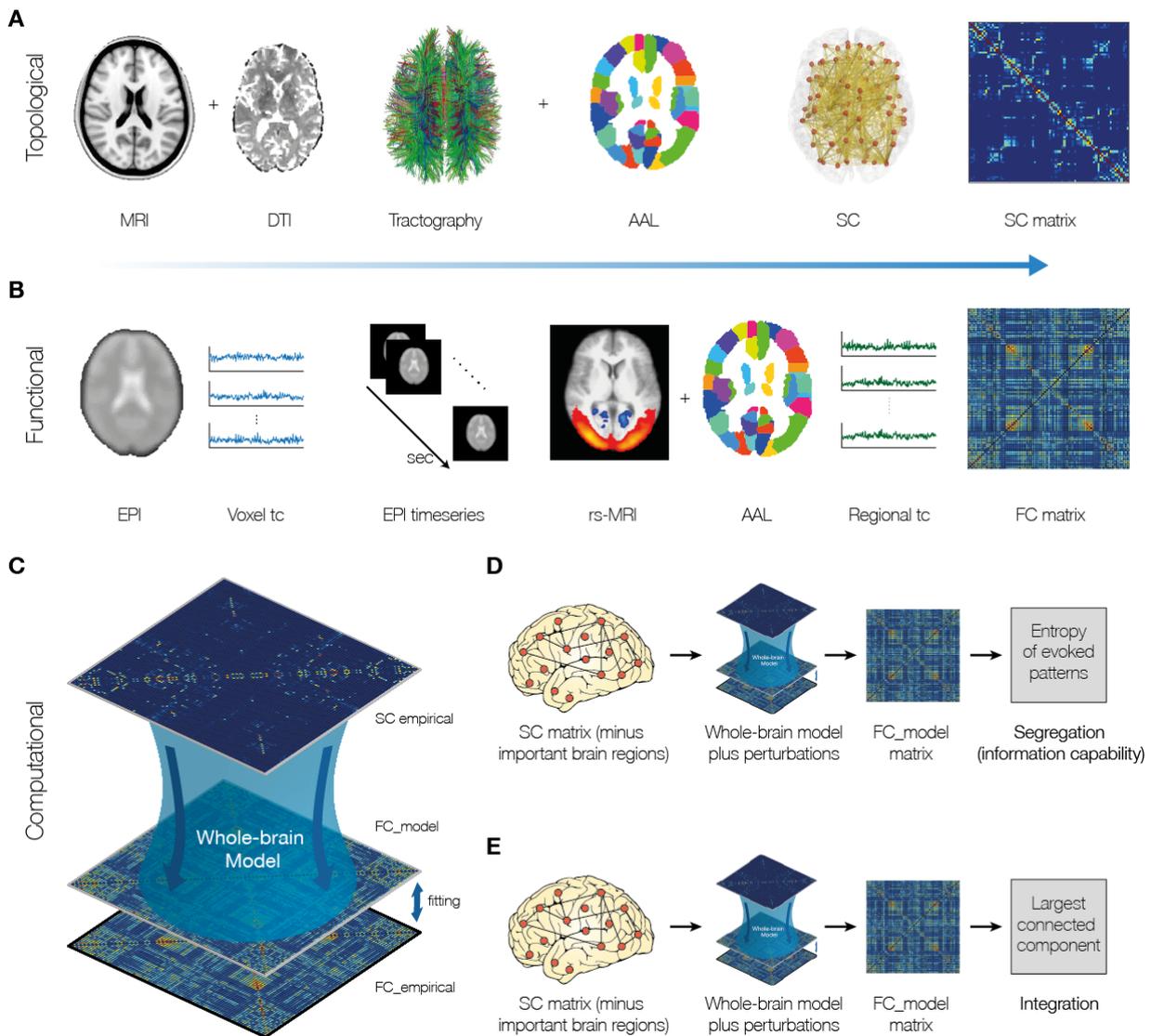


Figure 1. Overview of methods. In order to address the question of which brain regions are most important in integrative processing we analyzed the data from 16 healthy participants using topological, functional and computational methods (see Experimental Procedures for detailed information). **A)** As shown schematically, we generated the topological measure of a structural connectivity (SC) matrix by combining structural MRI, DTI, the AAL parcellation scheme and tractography. **B)** We also generated the functional connectivity (FC) matrix by analysing the BOLD resting state data with the AAL parcellation and correlating the time courses. **C)** We generated a whole-brain computational model using the empirical SC and FC to fit the FC_model output. **D)** Using the model allowed us to estimate the segregation (or *information capability*) of a given SC as the entropy of the set of evoked patterns assuming a Gaussian distribution averaged over a large number of external stimulations. **E)** Similarly, we used the model to estimate the *integration* as the length of the largest connected component in the functional connectivity matrix, averaged over a large number of external stimulations.

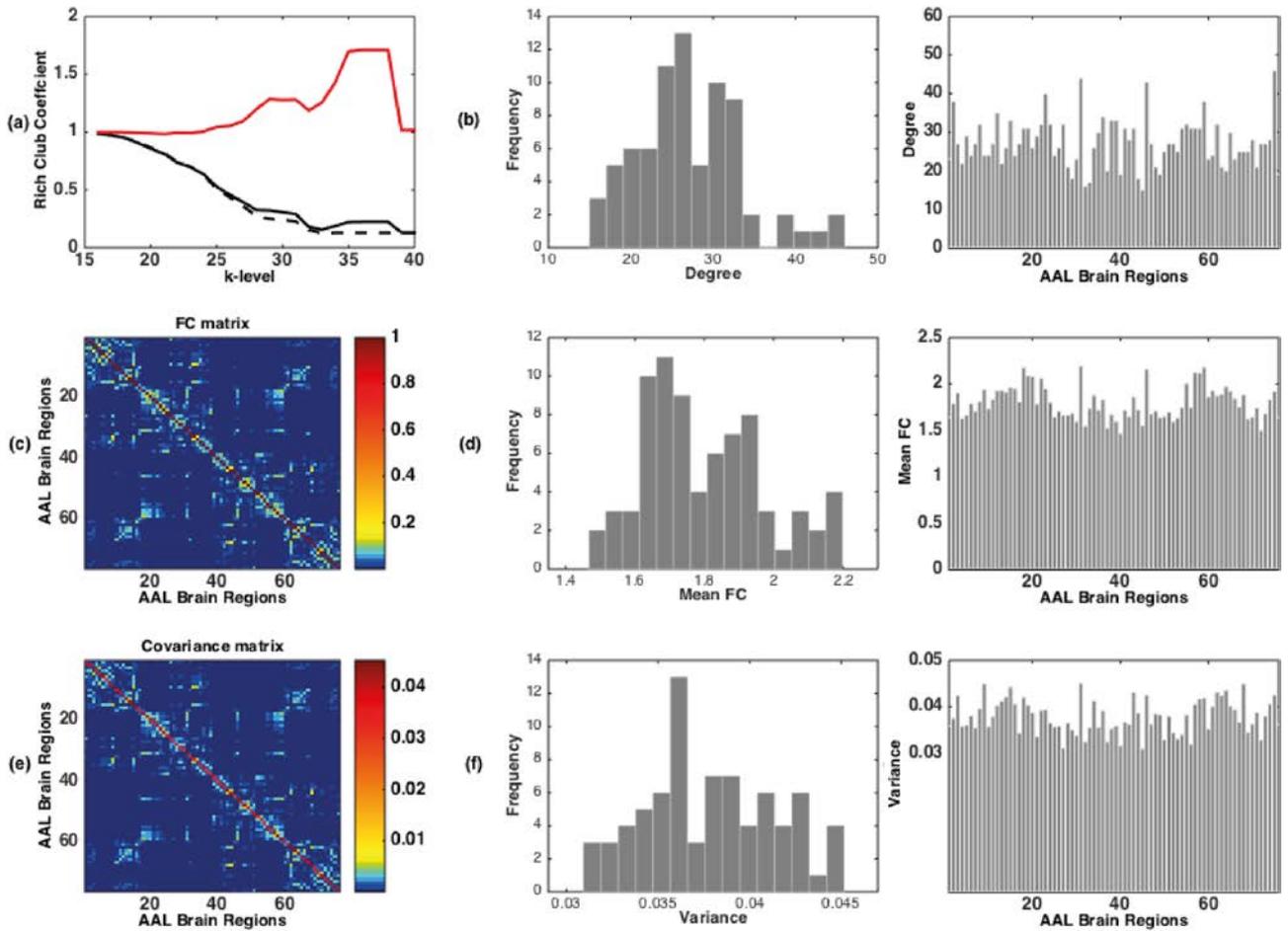


Figure 2. Determining the rich club members. The figure shows the structural and functional first order and second order characteristics of the brain network obtained experimentally in a group of 16 participants and with parcellation of 76 cortical AAL regions (as detailed in the Experimental Procedures). **A)** This structural anatomical network exhibits a rich club organisation, as shown by the normalised rich club coefficient curve (van den Heuvel and Sporns, 2013a). By definition, a set of brain areas in the network shows a rich club structure if its level of interconnectivity exceeds the level of connectivity that can be expected on basis of chance alone. The normalized rich club coefficient $\Phi_{norm}(k)$ (red curve) is computed by dividing $\Phi(k)$ (continuous black curve) by $\Phi_{random}(k)$ (dashed black curve), with $\Phi_{random}(k)$ computed as the average rich club coefficient for each k of a set of 1000 randomized graphs (acquired by randomizing the adjacency matrix A preserving the degree sequence) (Rubinov and Sporns, 2010). **B)** For each brain region the histogram and degree are plotted. **C)** The average resting functional connectivity (FC) matrix was calculated by summing the mean FC resting functional connectivity over all the pairs connecting that region with the rest of other regions. **D)** The mean FC in each brain region is plotted as well as the histogram. **E)** A similar procedure was used to create the average covariance matrix as the sum of the variance between all the pairs connecting that region with the rest of other regions. **F)** The variance in each brain region is plotted as well as the histogram. Note that a qualitative comparison

of the curves shows a moderate to weak relationship between the degree and mean rFC (**G**) and between degree and variance (**H**).

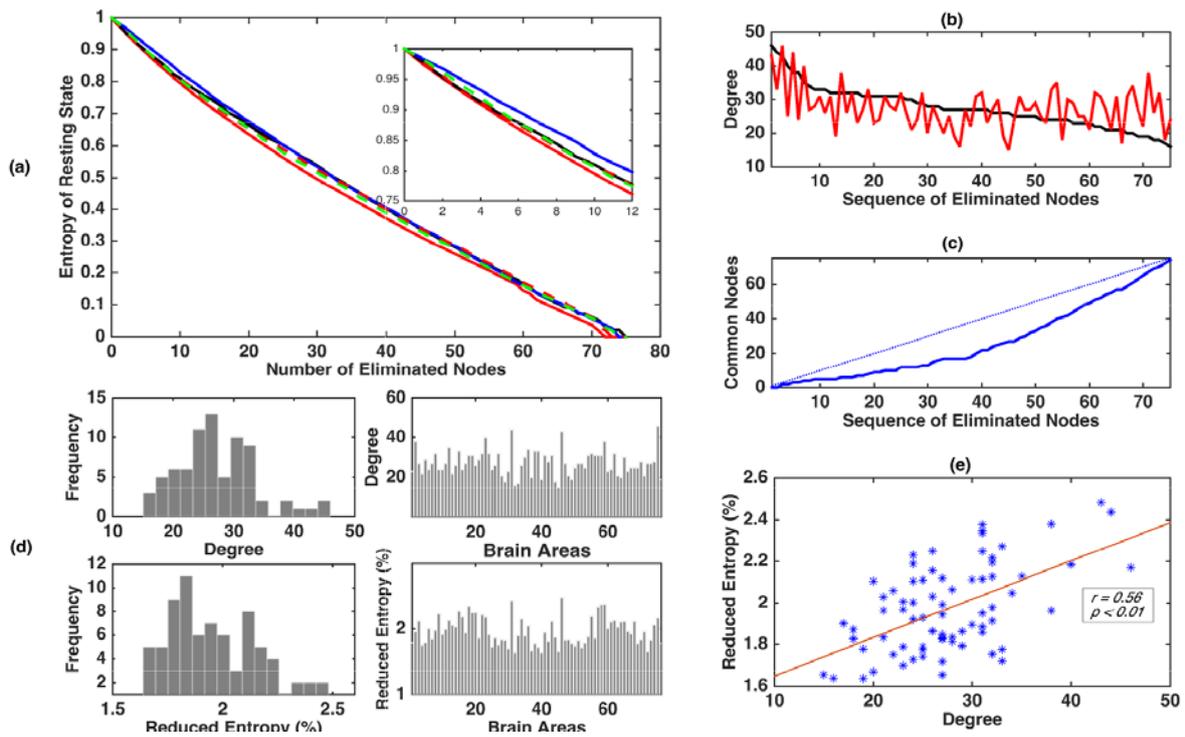


Figure 3. Determining the dynamical workspace of binding nodes and comparing it to the rich club.

A) The reduction of information capability is measured by considering resting state conditions as a function of a sequence of eliminated brain regions for the AAL76 parcellation. Five cases are shown: sequential elimination according to the lowest degrees (blue curve), largest degree (rich club, black curve), largest nodal variance (dashed red curve), random order (dashed green curve) and the workspace of binding nodes strategy (red curve), i.e. searching for the region that maximally reduces the information capability. The largest degree and largest variances reduces the information capability very efficiently (as intuitively expected), and much more than the sequential elimination of the lowest degree regions. Yet, this is not as efficiently as the optimal elimination of regions from the workspace of binding nodes. The inset shows that the workspace of binding nodes maximally reduces the measure of information capability and therefore defines a ranking of regions that are maximally relevant for the functional dynamics (see Figure 4). Some of the regions from the workspace of binding nodes are also high degree “rich club” regions, but not all high degree regions are ranked in the same way and also low degree regions are included in the workspace of binding nodes. **B)** To further study the intersection between the workspace of binding nodes and rich club regions, we show the degree value of the sequentially eliminated regions for the rich club (black curve) and for the workspace of binding nodes (red curve). It is clear that the sequential elimination of regions according to the largest degree results in a monotonically decreasing curve, whereas for the elimination according to the maximal reduction of information capability does not necessarily lead to the selection of the highest degrees regions. Nevertheless, a tendency of selecting high degrees regions is also shown in the stochastic decreasing of that curve. **C)** The

intersection between those two groups (high degree and workspace of binding nodes) is shown and this figure clearly shows that the regions from the dynamical workspace of binding nodes do not have a one-to-one correspondence with the high degree regions, but partially overlapping. **D)** The figure plots for each brain region the degree (top) and the percentage of reduced information capability (bottom) by elimination of that region. The first column shows a histogram representation hereof. **E)** The figure shows the correlation between the degree and information capability (shown in D) as a function of the percentage of reduced entropy ($r=0.56$, $p<0.001$), though it should be noted that there is no deterministic relationship.

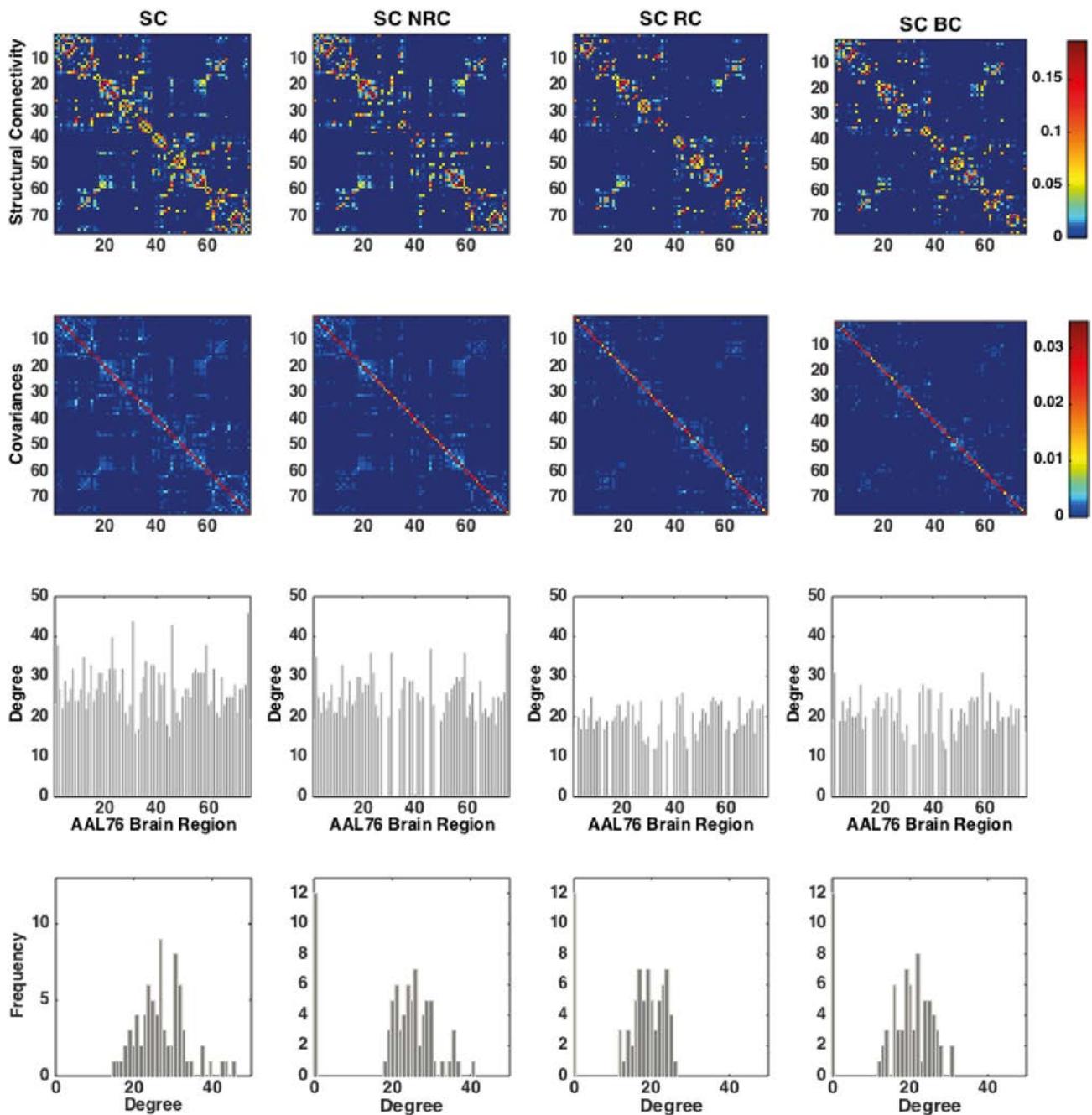


Figure 4. Structural connectivity matrices with and without members of the non-rich club, rich club, and workspace of binding nodes. The top row figure shows four different structural connectivity matrices using the AAL76 parcellation. The middle row shows the covariance matrices and the bottom two rows show the degree and histogram for each of the 76 brain regions, respectively. The x and y axes are corresponding to brain areas and the colour bars represent the strength of structural connectivity and covariance, respectively. The bottom row shows the degree (y-axis) for the brain areas in the AAL76 parcellation (x-axis). First column shows corresponds to the full structural connectivity (SC) matrix from 16 healthy participants. The second column corresponds to the SC matrix minus 12 brain regions which has the lowest ranked degrees, and therefore regions not exhibiting the rich club structure (SC NRC). The third column corresponds to the SC matrix minus 12 regions in the rich club with maximal degree (i.e. number of structural

connections with other regions) (SC RC). The fourth column corresponds to the SC matrix minus the top-ranked 12 brain regions in the workspace of binding nodes showing the largest reduction of the resting entropy under resting state conditions (SC_BC).

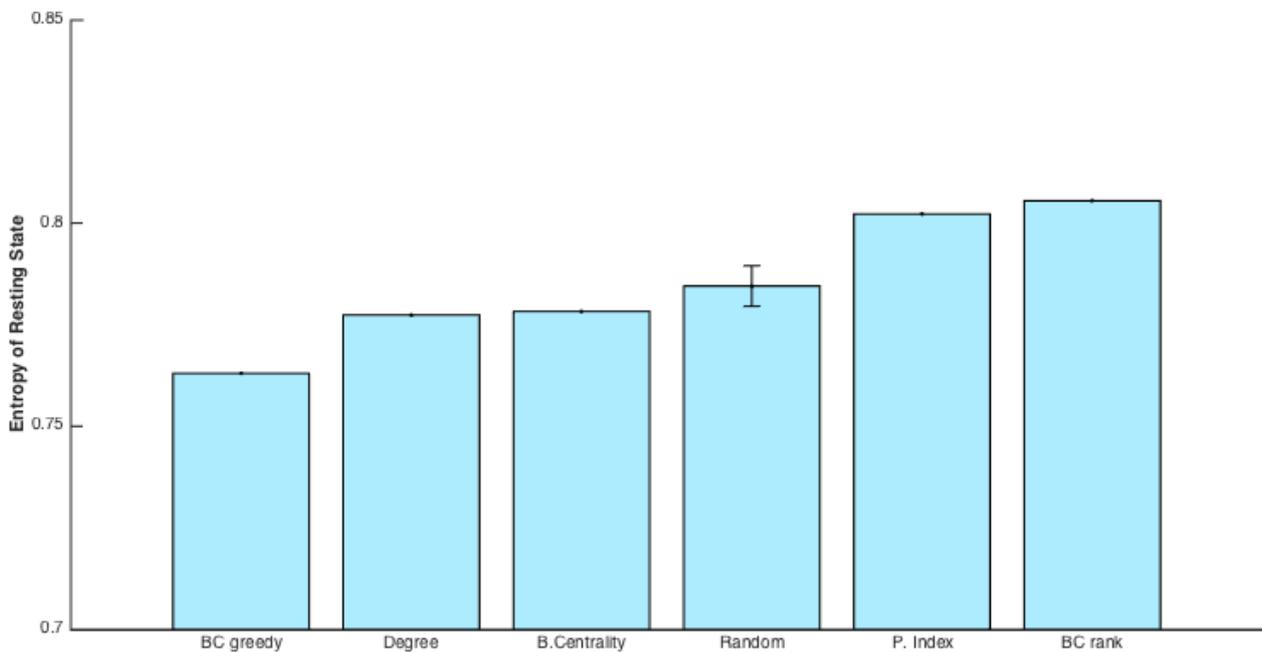


Figure 5. Impact of different node removal strategies on the entropy of resting state. The reduction of entropy is affected by different strategies for node-removal. Here we are using a greedy node-removal strategy of top-ranked nodes of the workspace of binding nodes (BC greedy), of the node degree (Degree), of betweenness centrality (B. Centrality), of random nodes (Random), participation coefficient (P.Index) or just the initial ranked nodes of the workspace of binding nodes (BC rank). In all cases, 12 nodes were removed after which the entropy and corresponding standard deviation were recalculated. The results clearly show that the greedy removal of nodes from the workspace of binding nodes significantly outperforms the other node removal strategies, i.e. entropy is reduced maximally when using the greedy node-removal strategy.

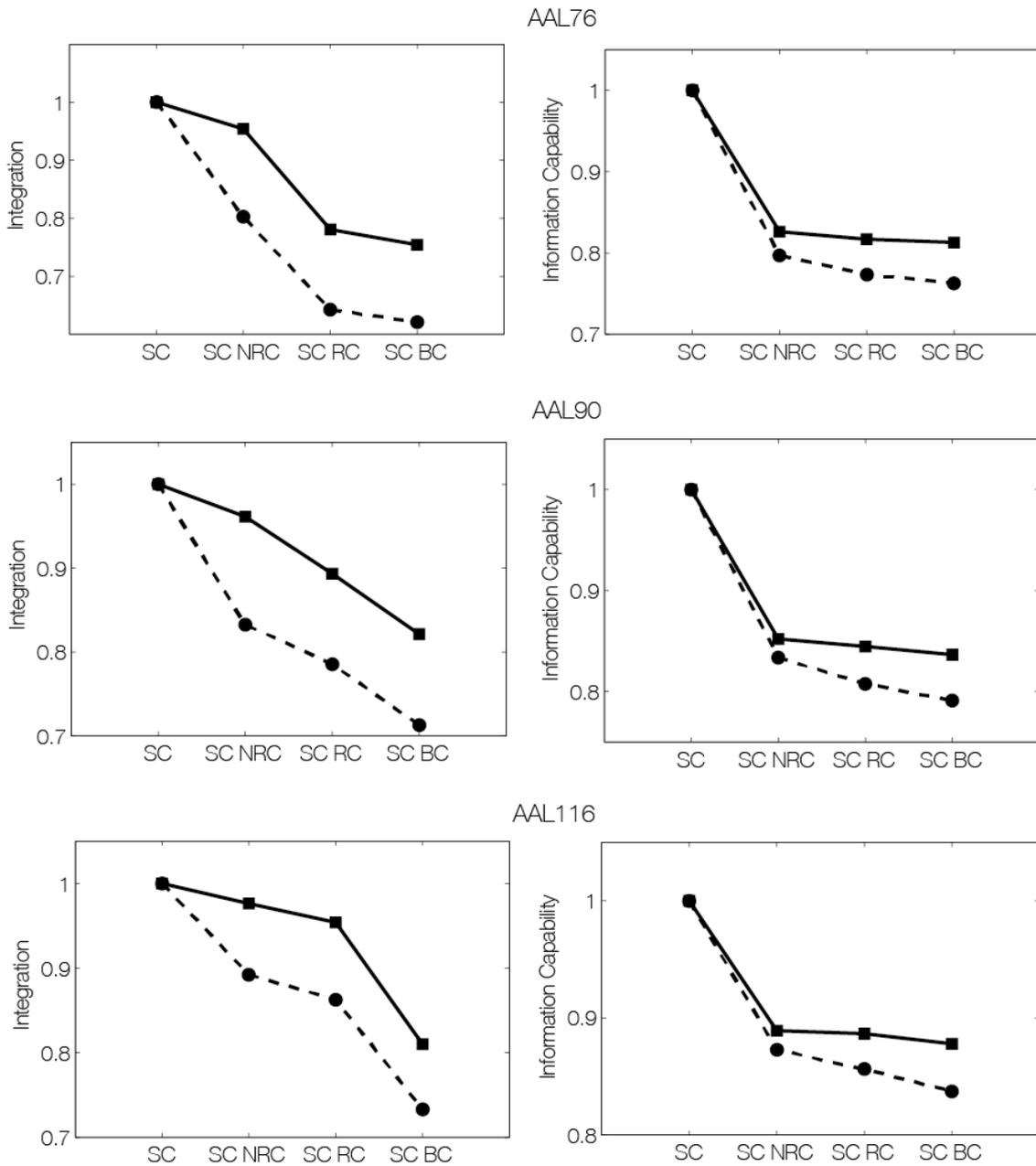


Figure 6. Directly measuring the causal impact of the rich club and the workspace of binding nodes. We show the results of using whole-brain computational modelling to measure the impact on integration and information capability resulting from the different strategies of brain region elimination shown in figure 3. The left column shows the impact on integration and the right column shows the impact on information capability for the resting state dynamics (dashed curve) and for the stimulated case (1000 different random stimulations repeated 10 times), i.e. the corresponding perturbational integration and segregation (see Experimental Procedures). We show this for three different sub-parcellations of the AAL: AAL76 which include only cortical brain regions (top row), AAL90 which includes cortical and subcortical regions (middle row), and AAL116 which includes the cortical, subcortical and cerebellar regions. In all cases, removing the rich club members significantly impacts on both measures. Yet, the maximal reduction of integration and perturbational integration, as well as information capability, is observed for the case

where the members of the workspace of binding nodes were eliminated. (SC = Structural Connectivity; SC NRC = Structural Connectivity – Non-Rich Club members; SC RC = Structural Connectivity – Rich Club members; SC BC = Structural Connectivity – members of the workspace of binding nodes).

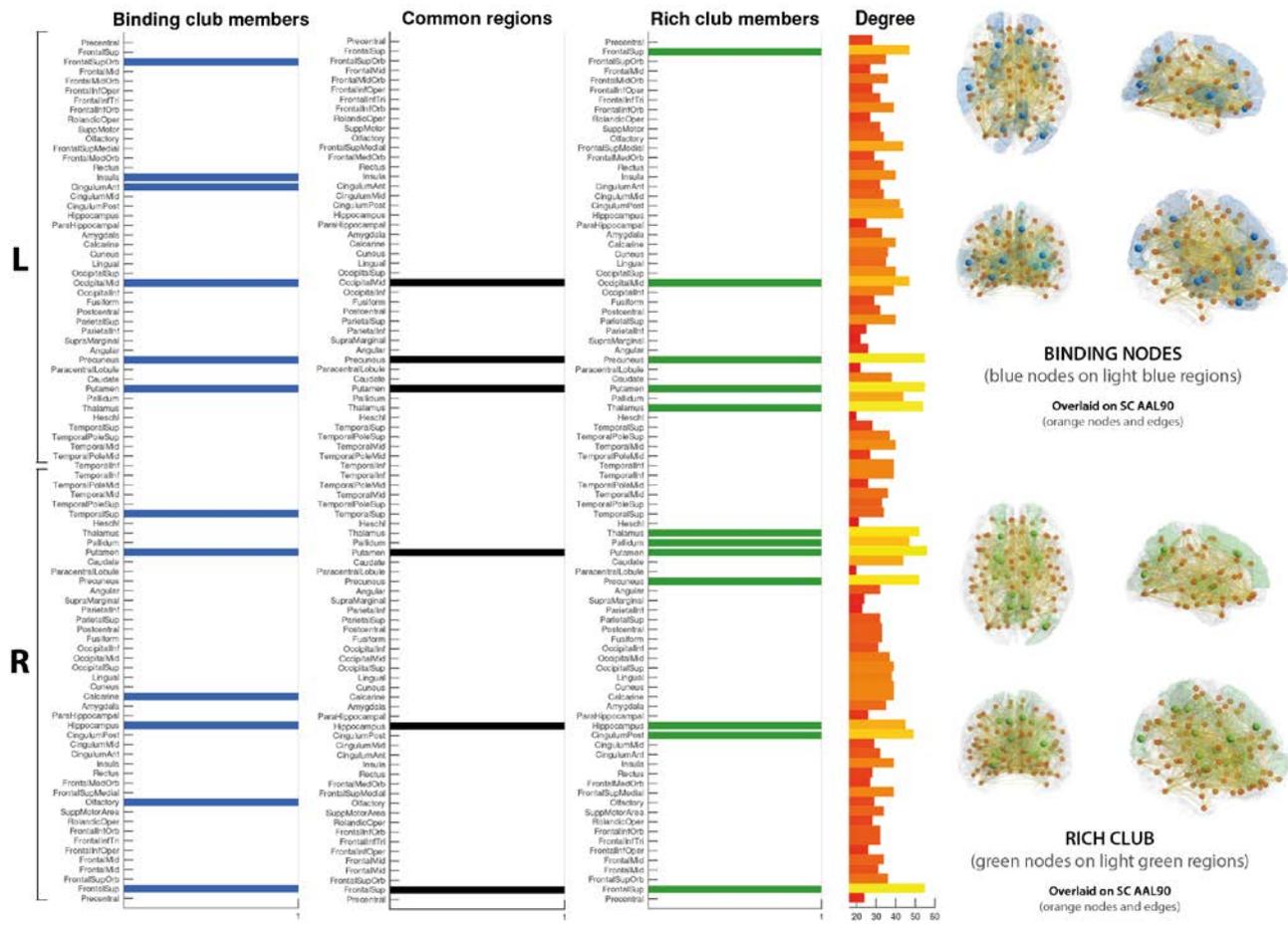


Figure 7. Three-dimensional renderings of the workspace of binding nodes and the rich club. The figure shows 3D renderings of the AAL90 parcellation with the weighted structural connectivity between brain regions. In the top right panel we show four renderings of the 12 top-ranked regions of the workspace of binding nodes (in transparent blue with a blue node in the centre of mass) in top, side, front and perspective views. Similarly, in the bottom right panel we show four renderings of the rich club regions (in green). In the left panel we show the regions that are present in the workspace of binding nodes (blue) and in the rich club (green). The horizontal bar plot shows the node degree colour-coded according to its strength (yellow – high degree; red – low degree) Additionally we highlight the brain areas present in both the rich club and the workspace of binding nodes in the centre of this panel in black.