Supporting Information

van den Heuvel et al. 10.1073/pnas.1203593109

SI Materials and Methods

Subjects. High-resolution structural connectome magnetic resonance imaging (MRI) was carried out in 80 healthy adults after obtaining written informed consent. The original set was divided into two sets of 40 subjects each. One set constitutes the principal data set [set 1: age 28.6 (7.9) y, 21 male], and the other served as a replication dataset [set 2: age 27.0 (6.9) y, 27 male].

MRI Acquisition. MRI data were acquired on a 3-Tesla Philips Achieva Clinical scanner. For each subject, diffusion tensor imaging (acquisition parameters: SENSE-p = 3; two sets of 30 different weighted directions, and 2×5 $b = 0$ images (1); repetition time (TR)/echo time (TE) = 7,035/68 ms, $2 \times 2 \times 2$ mm, 75 slices covering whole brain, b weighting of 1,000 s/mm^2, second set with reversed k-space readout and an anatomical T1 image for anatomical reference [3D fast field echo (FFE) using parallel imaging; TR/TE = 10 ms/4.6 ms; field of view (FOV) = 240 × 240 mm, 200 slices covering whole brain, $0.75 \times 0.75 \times 0.75$ mm] data were acquired.

Data Processing. Processing of the diffusion tensor imaging (DTI) images included the following steps (2, 3): DTI images were corrected for susceptibility and eddy current distortions and a tensor was fitted to the diffusion profile within each voxel using a robust tensor fitting method. The preferred diffusion direction within each voxel was computed as the principal eigenvector of the eigenvalue decomposition of the fitted tensor. Next, the preferred diffusion direction was used for streamline tractography [fiber assignment by continuous tracking (FACT)] to reconstruct white matter pathways. Within each voxel in the brain mask, eight seeds were started, evenly distributed over the volume of the voxel. A streamline was started from each seed following the main diffusion direction from voxel to voxel and was stopped when the fiber track reached a voxel with a fractional anisotropy (FA) value <0.1, when the fiber trajectory left the brain mask, or when the fiber track made a sharp turn of more than 45° (3). T1 images were processed using freesurfer (V5) (3, 4).

Connectome Reconstruction. Structural brain networks of cerebral cortex were constructed by dividing the cortex into 1,170 randomly placed and equally sized parcels (network nodes) and estimating their connectivity strength as the number of connecting white matter streamlines (network edges) (2, 5, 6). Connectome reconstruction included the following steps. First, individual brain networks were modeled on the basis of the set of reconstructed fiber tracts combined with the segmented brain regions (3, 5). For each subject, the brain network was mathematically described as a graph consisting of the set of 1,170 brain parcels (called nodes) and a set of connections describing the number of streamlines between the nodes. For each pair of nodes i and j it was determined whether there were streamlines present in the total collection of reconstructed tracts F that interconnected region i and region j. The number of streamlines between i and j was taken as the connectivity strength (or network density) between nodes i and j in the network and included in the connectivity matrix. In addition to the strength, the physical length of each edge was calculated from the average length of the interconnecting streamlines. Edges comprising fewer than 10 streamlines were considered potentially spurious and were deleted from the connection matrix (5). Next, for both the principal and the replication datasets, a group connectome was constructed by selecting all connections that were present in at least one-third of the group of subjects (3). The group connectome was formed by averaging the "strength" and "length" values over existing (nonzero) edges (3, 5–7).

Network Visualization. To visualize possible differences in the topological ordering of connections for each of the connection classes, nodes were ordered along a circle's perimeter, using an optimization algorithm based on simulated annealing. This algorithm attempts to rearrange the rows and columns (i.e., nodes) of the connection matrix in such a way that the elements of the matrix are squeezed toward the main diagonal by minimizing the "cost" of the system, defined as the total cost of placing the nodes on the perimeter of the ring. At the first iteration, the nodes were randomly placed on a circle. Next, at each iteration, two nodes were randomly replaced and the new configuration was preserved (i) if the cost of the new configuration was a new minimum or (ii) with a certain level of probability p defined as $p = \exp\left(\frac{C - C_{\min}}{T}\right)$ with C the cost of the current configuration, $C_{\min}$ the current cost minimum, and T the "temperature" describing to what extent the system allows for the examination of regimes with a cost higher than the current setting. The initial setting of the temperature was decreased at each iteration following an exponential decay, gradually "cooling off" the system (hence the term "annealing"). In total, 10^6 iterations were run.

"Percentage of cost" ("% of cost") as shown in Fig. 2C expresses the proportion of total network cost taken on by connections in close proximity on the ring (proximal connections), determined as nodes within distance of 5% circumference on the ring (to left and right) and by more long-distance connections (distal connections, >5% of circumference).

Rich Club Detection. The resulting network was analyzed with methods from graph theory (8, 9). The so-called "rich club" phenomenon in networks is said to be present when the high-degree nodes of a network tend to be more densely connected among themselves than expected by chance (10). The rich club was identified by comparing the density of connections among nodes with a minimum degree k to the density found in randomized networks preserving degree sequence. To confirm the presence of a rich club in the brain, first, for the group connectome, the degree k of each node was computed as the number of binary connections. Next, for each value of k, the subset of nodes with a degree larger than k was selected from the network. From this subset, consisting of n nodes and the collection of $E_{>k}$ connections, their total sum of weights $W_{>k}$ was determined, with weights W defined as number of streamlines (network density) of the edges (2). The weighted rich club parameter $\Phi^w(k)$ was then computed as the ratio between $W_{>k}$ and the sum of the weights of the strongest $E_{>k}$ connections of the whole network, given by the top $E_{>k}$ number of connections of the collection of ranked connections W^{ranked} in the network. Formally, $\Phi^w(k)$ is given by (11)

$$\phi^w(k) = \frac{W_{>k}}{\sum_{l=1}^{E_{>k}} w_l^{\text{ranked}}}. \qquad \text{[S1]}$$

$\Phi(k)$ is typically normalized relative to a set of comparable random networks of equal size and degree sequence, giving a normalized rich club coefficient $\Phi^w{}_{\text{norm}}$ (10, 12). Ten thousand random networks were generated, preserving the degree distri-

bution and sequence of the original network (13). For each level of k, for each of these randomized networks, the rich club coefficient $\Phi^{w}{}_{\text{random}}$ was computed. $\Phi^{w}{}_{\text{random}}(k)$ was computed as the average rich club coefficient over the m random networks. The normalized rich club coefficient $\Phi^{w}{}_{\text{norm}}(k)$ was computed as

$$\phi^{w}{}_{\text{norm}}(k) = \frac{\phi^{w}(k)}{\phi^{w}_{\text{random}}(k)}. \qquad \textbf{[S2]}$$

Ten thousand random networks were generated, preserving the degree distribution and sequence of the original network (13). A network is said to have a rich club organization if $\Phi^{w}{}_{\text{norm}}(k) > 1$, for a continuous range of k. To assess statistical significance of the results, permutation testing was used. Rich club measures were statistically assessed with permutation testing (2, 6). The distribution of $\Phi^{w}{}_{\text{random}}(k)$ yielded a null distribution of rich club coefficients obtained from random topologies. Using this null distribution, $\Phi^{w}(k)$ was assigned a P value by computing the percentage of random values that were found to be more extreme than the observed rich club coefficient $\Phi^{w}(k)$. All tests were conducted using Bonferroni-adjusted α-levels of 0.0025 per test (0.05/tests performed).

Path Metrics. Communication cost was based on the path length (topological distance) between any node i and any node j in the network (13). First, all 683,685 unique [i.e., $N^{*}(N - 1)/2$] shortest paths between all $n = 1{,}170$ nodes in the network were traced. Second, the total communication cost of the shortest path between nodes i and j was computed as the sum of the product of the physical lengths and density (i.e., number of streamlines) of the edges that were used while traveling from node i to j in the network. Note that related to the bidirectional character of all edges in the graph the communication cost traveling from node j to i is similar to the communication cost when traveling from node i to node j. Subsequently, for each path, the proportion of the total communication cost spent on rich club, feeder, and local connections was computed. Once aggregated across all paths, the communication cost expresses the total amount of signal traffic carried along each edge of the network, taking into account its density, its length, and the number of short paths it supports.

1. Jones DK (2004) The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study. *Magn Reson Med* 51:807–815.
2. van den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. *J Neurosci* 31:15775–15786.
3. van den Heuvel MP, Mandl RCW, Stam CJ, Kahn RS, Hulshoff Pol HE (2010) Aberrant frontal and temperal network structure in schizophrenia: A graph theoretical analysis. *J Neurosci* 30:5915–5926.
4. Fischl B, et al. (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14:11–22.
5. Hagmann P, et al. (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol* 6:e159.
6. Bassett DS, et al. (2008) Hierarchical organization of human cortical networks in health and schizophrenia. *J Neurosci* 28:9239–9248.
7. Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* 3:e17.
8. Bullmore E, Sporns O (2009) Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10:186–198.
9. Wig GS, Schlaggar BL, Petersen SE (2011) Concepts and principles in the analysis of brain networks. *Ann N Y Acad Sci* 1224:126–146.
10. Colizza V, Flammini A, Serrano MA, Vespignani A (2006) Detecting rich-club ordering in complex networks. *Nat Phys* 2:110–115.
11. Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: The weighted rich-club effect. *Phys Rev Lett* 101:168702.
12. McAuley JJ, da Fontoura Costa L, Caetano TS (2007) Rich-club phenomena across complex network hierachies. *Appl Phys Lett*, 91: 084103.
13. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* 52:1059–1069.
14. Milo R, et al. (2002) Network motifs: Simple building blocks of complex networks. *Science* 298:824–827.

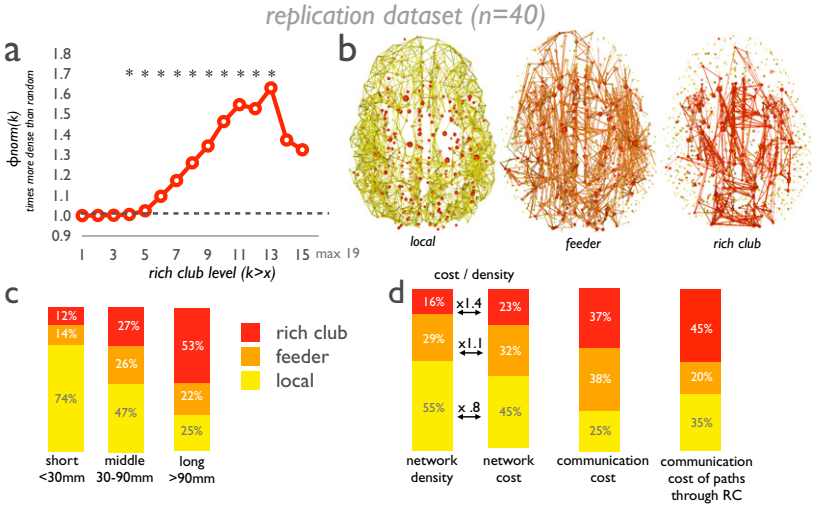

Fig. S1. Results of replication dataset (n = 40). (*A*) Rich club curve of the replication dataset. (B) *Left*, local connections (yellow); *Center*, feeder connections (orange); *Right*, rich club connections (red). Consistent with the results of the principal dataset, the brain's rich club contains parts of the precuneus, superior frontal cortex, anterior cingulate cortex, posterior cingulate cortex, superior parietal cortex, and the insula, all in both cerebral hemispheres. (*C*) Proportions of short-, middle-, and long-range connections that belong to the categories of short (<30 mm), middle (30–90 mm), and long-range connections (>90 mm). (*D*) Proportions of network density, network cost, communication cost, and communication cost of paths through the rich club for the categories of rich club, feeder, and local connections.

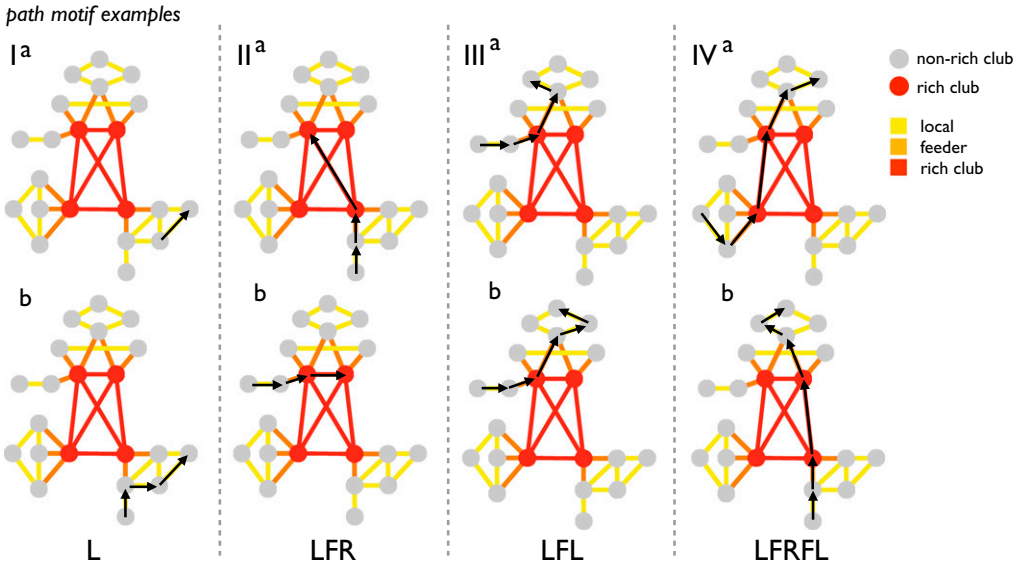

Fig. S2. Each communication path was classified according the specific sequence of connections that were crossed along the path traveled, referred to as "path motifs." Shown are eight examples of paths contributing to the path motifs "L" (column I, *a* and *b*), "LFR" (column II, *a* and *b*), "LFL" (column III, *a* and *b*), and "LFRFL" (column IV, *a* and *b*), respectively.