Supporting Information

Margulies et al. 10.1073/pnas.1608282113

SI Materials and Methods

Connectivity Data. Connectivity data for the human and macaque monkey brains were assembled from openly available sources. Because acquisition procedures and the method for assessing connectivity varied between species, the first step was to transform each species' dataset into an affinity matrix describing the pairwise similarity of connectivity values.

Human. The connectivity matrix for the human brain was based on 1 h of resting-state fMRI data acquired through the HCP (39) and made publicly available for download on ConnectomeDB (71). In brief, for each individual, a functional connectivity matrix was calculated using the correlation coefficient across four minimally preprocessed (40, 72–74), spatially normalized, and concatenated 15-min resting-state fMRI scans. We began with the group-averaged functional connectivity data, which include 820 individuals coregistered using MSMAll. More information about this dataset can be found at www.humanconnectome.org/documentation/S900/820_Group-average_rfMRI_Connectivity_December2015.pdf. The "dense" functional connectome matrix consists of pairwise z-transformed correlation values between all grayordinates (91,282 rows). Cortical data are represented in HCP 32k_LR surface space (75), consisting of 32,492 total nodes per hemisphere (59,412 excluding the medial wall). Subcortical structures are represented in volumetric space and included 31,870 voxels.

We began by transforming the z to r correlation values with a hyperbolic tangent function, which scales them between −1 and 1. For each row in the matrix, the values of the top 10% of connections were retained, whereas all others were zeroed. Remaining connections were almost all positive, except for ∼5,000 (less than 10^{-6}% of all connections) with negative values from 23 voxels. The voxels with negative connections were all located in ventral subcortical regions. These connection were zeroed as well. Because this procedure rendered the connectivity matrix A asymmetric, similarity between all pairs of rows was calculated using cosine distance, resulting in the positive, symmetric affinity matrix L. The full affinity matrix, L, consisted of weights between zero and one, representing similarity of connectivity profiles among over 4 billion node pairs.

Macaque Monkey. Macaque monkey connectivity data were derived from the axonal tract-tracing connectivity database CoCoMac (67, 68) and presented in the F99 template space (76) using cortical area labels based on the Bonin–Bailey parcellation (77, 78). One cortical area was removed, because it displayed no connectivity with other areas, leaving 25 areas remaining. The 25 × 25 connectivity matrix initially consists of sources (rows) and targets (columns), where edges were values between one (weak) and three (strong) connections. Of the possible edges, 56% were present in the connectivity matrix. To create a similarity matrix that accounted for bidirectional connectivity, the connectivity matrix, A, was transposed and concatenated with the initial matrix: $A = (A, A^{\top})$. A was then transformed into a positive, symmetric affinity matrix L by calculating the Euclidean distance between each pair of rows.

Connectivity Embedding. We used diffusion embedding (42), a nonlinear dimensionality reduction technique, to recover a low-dimensional embedding from high-dimensional connectivity data. Connectivity data include both local and long-range connections. Diffusion maps translate these relationships into distances and represent the global connectivity structure as a distribution of cortical points in an embedding space. Cortical points that are strongly connected by either many connections or few very strong connections are close in this space, whereas points without connections are far apart. Linear techniques, such as principal component analysis, are unable to project such data without appropriate kernel manipulations. Among the numerous nonlinear dimensionality reduction algorithms currently in use, we chose diffusion map embedding (42), because the diffusion process limits the distances of influence to the graph neighborhood, thereby ensuring a stable representation of connections, regardless of the graph size (79). Other similar approaches have been previously applied to whole-brain structural connectivity data (80, 81).

Diffusion maps are a one-parameter (α) family of graph Laplacians that integrate local information into a global description. A diffusion map embedding reduces high-dimensional data to a low-dimensional representation that combines geometry with probability distribution of data points. Using an appropriately normalized random walk process, the parameter, α, can control whether the low-dimensional embedding reflects the geometry of the set, regardless of density of points or the long-time dynamics of the samples without uniform sampling. In other words, the parameter α controls the influence of density of sampling points on the underlying manifold (α = 0, maximal influence of sampling density; α = 1, no influence of sampling density). Their relationship to other embedding methods based on graph Laplacians is detailed in ref. 82.

Diffusion maps, which use an α of 0.5, are well-suited for the analysis of brain connectivity data for multiple reasons. First, they retain the global relations between data points in the embedded space. Second, they are more robust to noise in the connectivity matrix, unlike other techniques, such as Isomap (83). Third, by using the appropriate choice of α, they can be made less sensitive to the distribution of data. Fourth, decreasing eigenvalues reflect a natural ordering of the diffusion process; the largest eigenvalues correspond to the slowest processes and thereby, represent the slowest variance in connectivity patterns. Fifth, by relying on local distances, they address the curse of dimensionality, because smaller distances are more meaningful than larger distances as the number of dimensions increases.

Technically, following ref. 42, the species-specific connectivity matrices (represented as $\mathbf{L}$) described previously can be transformed to define a Markov chain with transition matrix $\mathbf{P}$. The eigenvectors of this transition matrix define an embedding that results in a representation of each vertex or region as a point in the embedding space. The mutual Euclidean distances between these representations reflect how closely connected the corresponding vertices are in the graph with regard to the diffusion process defined by the transition matrix. See algorithm 1 for pseudocode. Increasing the diffusion time (t) allows examination of the intrinsic geometric structure of the data at larger and larger scales, allowing a balance between details in the input data and the scale of assessment (79). In this implementation, we used an automated estimation of the diffusion time using a damped regularization process. In this approach, the eigenvalues λ_i are divided by $1-\lambda_i$ to provide robustness against small, noisy eigenvalues. Such diffusion map embedding has been previously applied to functional task and resting MRI data (84–89). The source code for this method is available at https://github.com/satra/mapalign.

Geodesic Distance Along the Cortical Surface. Geodesic distance along the cortical surface was calculated using an algorithm that approximates the exact distance along the shortest path between

two nodes on a triangular surface mesh (69, 70). Seed nodes were selected from seven peak nodes from the principal gradient located within clusters corresponding to independent regions of the DMN (Fig. 2). A distance value was then assigned to each cortical node based on the minimum geodesic distance along the "midthickness" surface to any of the seed nodes. Results were then visualized on the "very_inflated" surface to enable visibility within sulci. Additional results for macaque monkey are available in Fig. S3.

The advantage of using exact geodesic distance rather than the more common and less computationally expensive Dijkstra's algorithm is that the exact calculation returns a distance value that is independent of the mesh density. In addition, to ensure that the shortest paths would only pass through cortex, nodes representing the medial wall were removed from the triangular mesh during the geodesic distance calculation as implemented here: https://github.com/NeuroanatomyAndConnectivity/surfdist.

Ordering of Canonical Intrinsic Connectivity Networks. For each network of the seven-network parcellation (2), we extracted the principal gradient values and sorted them by the mean (Fig. 3*A*). Additional analysis conducted on the 17-network parcellation (2) is presented in Fig. S4.

Metaanalytic Functional Gradients. We used the NeuroSynth metaanalytic database (www.neurosynth.org) (47) to assess topic terms associated with the principal connectivity gradient. Regions of interest masks were created by projecting the cortical surface principal gradient map to the 2-mm volumetric MNI152 standard space (available at neurovault.org/images/24346/). The volumetric map was then binned into five-percentile increments and binarized. Each of these 20 maps ranging from 0–5% to 95–100% were used as inputs to the metaanalysis. For each region of interest map, the output of the analysis was a z statistic associated with the feature term. The terms were then ordered based on the weighted mean for visualization.

NeuroSynth. Feature terms were derived from the 50 set of topic terms (v3) available at https://github.com/NeuroanatomyAndConnectivity/gradient_analysis/blob/master/gradient_data/neurosynth/v3-topics-50-keys.txt. Of the 50, 30 were above the threshold of $z - statistic > 3.1$, and 6 were removed as "noise" terms, because they did not capture any coherent cognitive function, leaving 24 topic terms (presented in Fig. 4).

BrainMap. Functional characterization of the delineated zones was performed by reference to the BrainMap database [www.brainmap.org (90)]. We limited our analysis to functional neuroimaging studies in the healthy human brain (i.e., without interventions and without group comparisons), which reported results as coordinates in stereotaxic standard space. These inclusion criteria yielded 8,400 eligible experiments at the time of analysis. The functional characterization of each zone was then based on the behavioral domain (BD) and paradigm class (PC) metadata of those experiments that featured activation within each distance zone. That is, we first identified all experiments in BrainMap that activate a given zone and then, analyzed the BD and PC metadata of these experiments. BDs code the mental processes isolated by the statistical contrasts (91) and comprise the main categories cognition, action, perception, emotion, and interception as well as their related subcategories. PCs categorize the specific task used (www.brainmap.org/scribe/ has the complete BrainMap taxonomy). Only BDs/PCs for which more than 24 individual experiments were stored in the database were considered further.

The association between each BD/PC and a given distance zone was expressed as the likelihood ratio between the probability of finding this zone activated given a particular BD/PC label relative to the a priori chance (across the entire database) of finding activation in that particular cluster (56, 92). That is, we tested whether the conditional probability of activation given a particular label $[P(Activation|Task)]$ was higher than the baseline probability of activating the region in question per se $[P(Activation)]$. We then only retained significant associations as established using a binomial test ($P < 0.01$).

Algorithm 1: Pseudocode for diffusion map embedding of a connectivity matrix representing similarity

Require: $\mathbf{L}$ is a symmetric, similarity matrix with $L_{ij} \geq 0$, diffusion time $t \geq 0$, and number of components ($N \geq 1$)
  $\mathbf{D} \leftarrow Diagonalize(Degree(\mathbf{L}))$
  $\mathbf{L}^{\alpha} \leftarrow \mathbf{D}^{-\alpha}\mathbf{L}\mathbf{D}^{-\alpha}$ {where $\alpha = 0.5$ recovers long-term dynamics}
  $\mathbf{D}^{\alpha} \leftarrow Diagonalize(Degree(\mathbf{L}^{\alpha}))$
  $\mathbf{P} \leftarrow (\mathbf{D}^{\alpha})^{-1}\mathbf{L}^{\alpha}$
  $\lambda_{0..N}, \mathbf{v}_{0..N} \leftarrow EigenDecomposition(\mathbf{P}, N+1)$ {where $1 \geq \lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_N$}
  For $i = 1$ to N **do**
    If $t = = 0$, then
      $\lambda_i \leftarrow \frac{\lambda_i}{1-\lambda_i}$
    else
      $\lambda_i \leftarrow \lambda_i^t$
    end if
    $\psi_i \leftarrow \frac{\mathbf{v}_i}{\mathbf{v}_0}$
    $\phi_i \leftarrow \lambda_i \psi_i$
  end for
  return $\mathbf{\Phi}_{1...N} = 0$

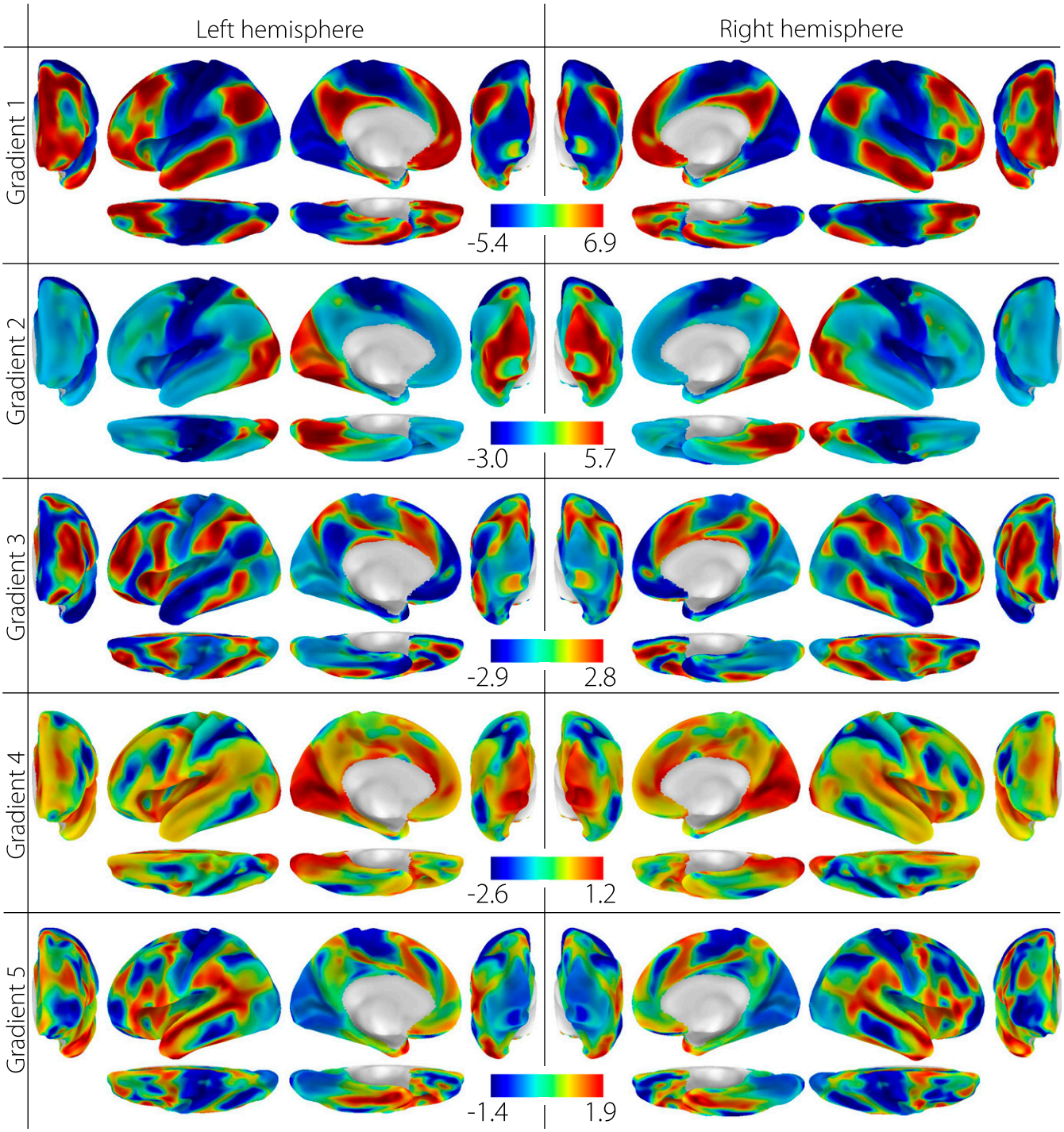

Fig. S1. Human connectivity gradients 1–5. The first five components result from diffusion embedding of the human connectivity matrix. The first five are shown because of the drop in variance explained after the fifth component (Fig. S2).

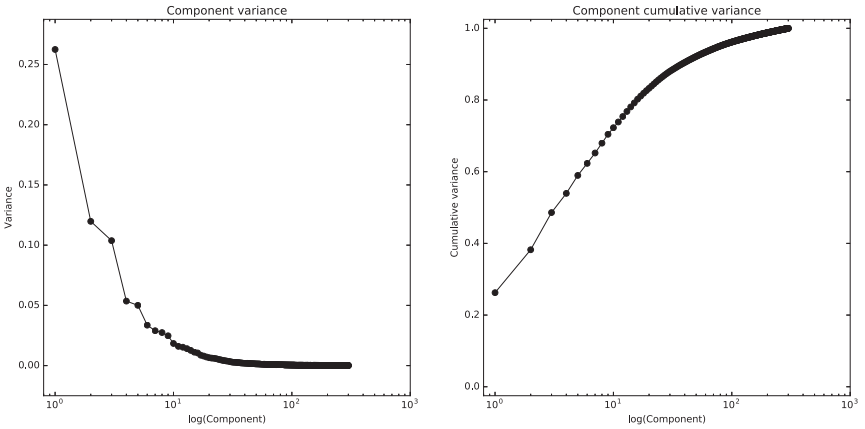

Fig. S2. The λ values of the diffusion embedding components in humans. The λ values of the first 300 components in the human connectivity data (*Left*) and cumulative λ values (*Right*). The gap after the first component indicates that the sorting of the principal gradient is stable.

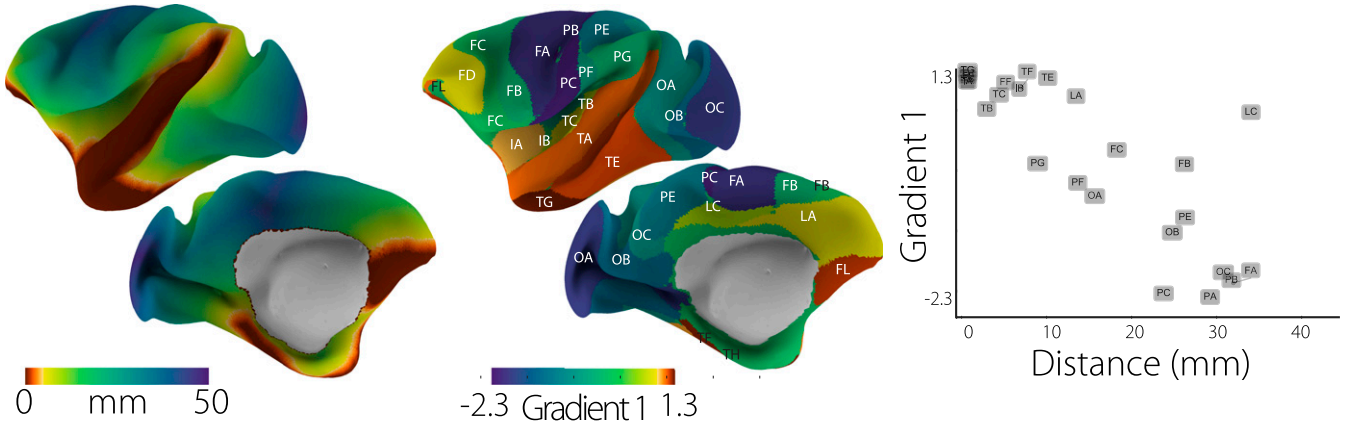

Fig. S3. Geodesic distance along the macaque monkey cortical surface from transmodal peak areas. The minimum geodesic distance along the cortical surface (*Left*) from the top five regions of the principal gradient (*Center*) (Fig. 1*B*). The scatter plot (*Right*) depicts the average geodesic within each region with respect to the principal gradient. Cortical areas are labeled using the Bonin–Bailey parcellation (77, 78).

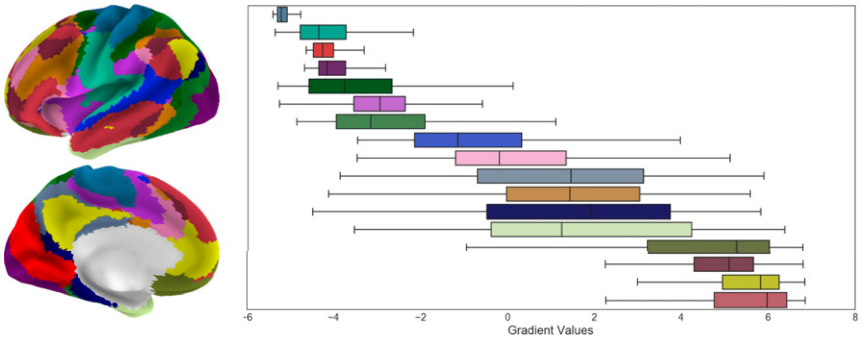

Fig. S4. Distribution of the principal gradient across 17-network parcellation from ref. 2. Each boxplot presents the principal gradient values within the respective network of interest. The network parcellation includes the 17 networks from ref. 2.

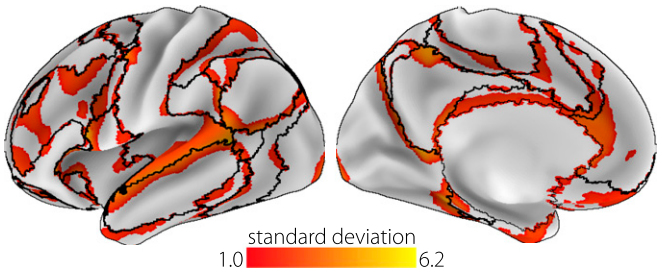

Fig. S5. Principal gradient outliers are proximate to the boundaries between networks. The standard deviation (SD) values of principal gradient values are calculated for each node with respect to other values within the same network [seven networks (2)]. SD values greater than 1.0 are presented. Outliers predominantly neighbor the network boundaries (black lines), indicating that the principal gradient captures gradual transitions between network-level connectivity patterns.

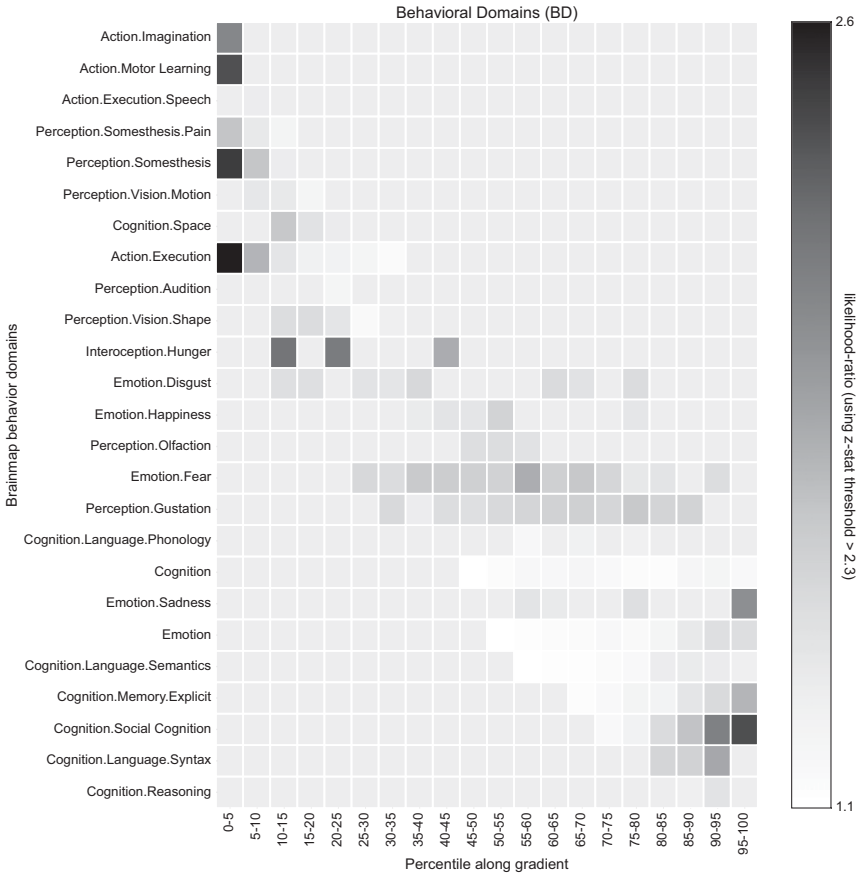

Fig. S6. BrainMap behavioral domains. Metaanalysis of the BrainMap behavioral domains with respect to the principal gradient.

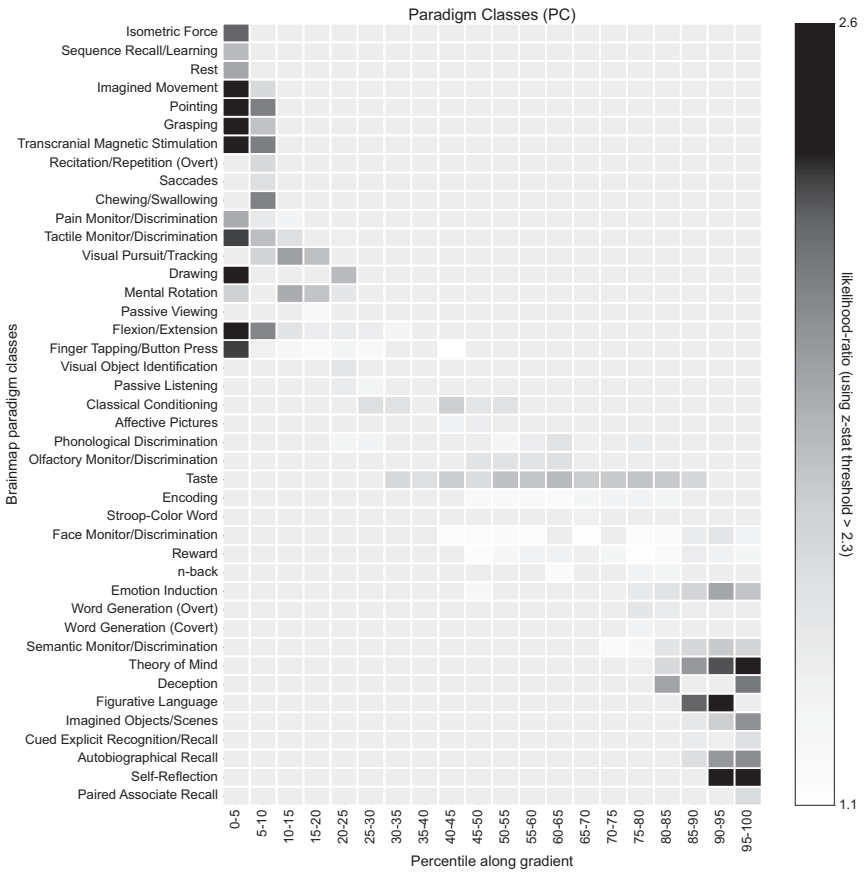

Fig. S7. BrainMap paradigm classes. Metaanalysis of the BrainMap paradigm classes with respect to the principal gradient.