

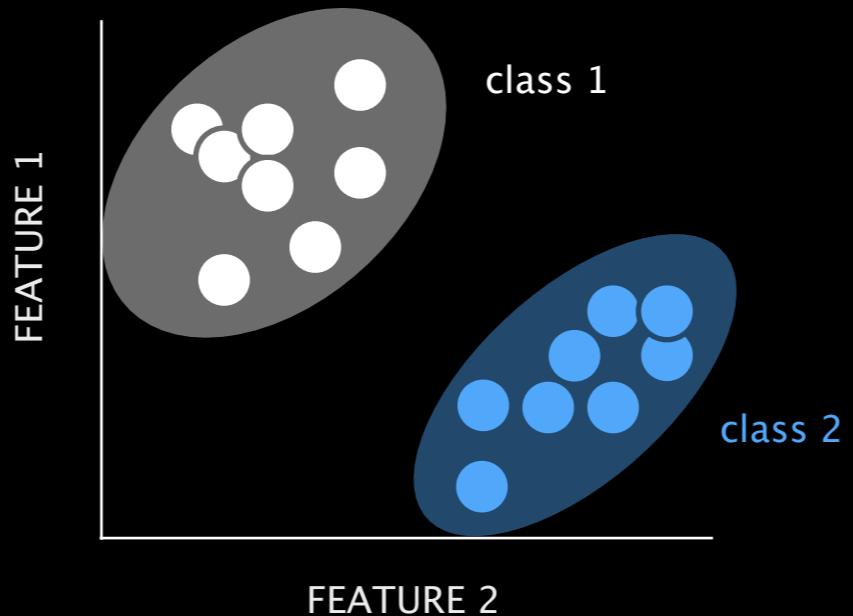
# W05- CLUSTERING TECHNIQUES

Boris Bernhardt, PhD

<http://mica-mni.github.io>



IDEA:  
GROUP DATAPoints INTO SUBCLASSES  
BASED ON SIMILARITY

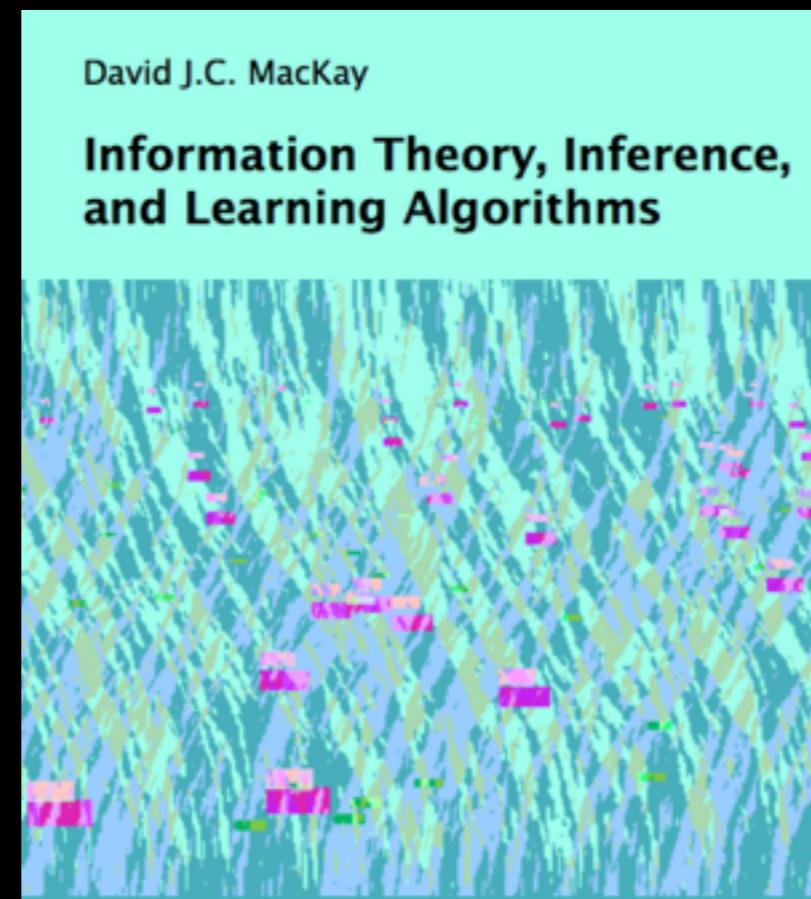


EXAMPLES IN NEUROIMAGING:

IMAGE SEGMENTATION  
PATIENT 'BIOTYPING'  
REGIONAL PARCELLATIONS



## THE K-MEANS



<http://www.inference.org.uk/itprnn/book.pdf>

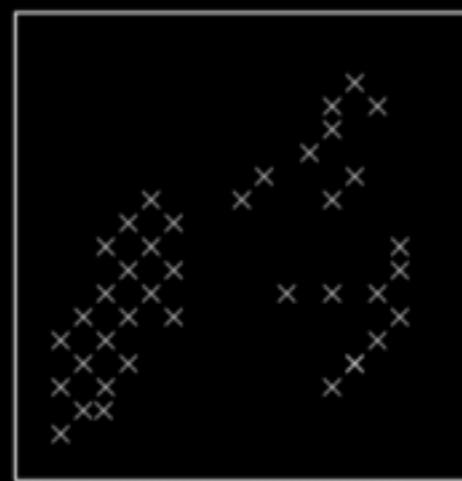


Figure 20.1.  $N = 40$  data points.

HOW CAN WE DESCRIBE THIS DATA MORE EFFICIENTLY?

## THE IDEAL CASE

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \sum_i (x_i - y_i)^2.$$

FIRST STEP: DEFINE A DISTANCE METRIC

**Initialization.** Set  $K$  means  $\{\mathbf{m}^{(k)}\}$  to random values.

**Assignment step.** Each data point  $n$  is assigned to the nearest mean.

We denote our guess for the cluster  $k^{(n)}$  that the point  $\mathbf{x}^{(n)}$  belongs to by  $\hat{k}^{(n)}$ .

$$\hat{k}^{(n)} = \operatorname{argmin}_k \{d(\mathbf{m}^{(k)}, \mathbf{x}^{(n)})\}. \quad (20.3)$$

An alternative, equivalent representation of this assignment of points to clusters is given by ‘responsibilities’, which are indicator variables  $r_k^{(n)}$ . In the assignment step, we set  $r_k^{(n)}$  to one if mean  $k$  is the closest mean to datapoint  $\mathbf{x}^{(n)}$ ; otherwise  $r_k^{(n)}$  is zero.

$$r_k^{(n)} = \begin{cases} 1 & \text{if } \hat{k}^{(n)} = k \\ 0 & \text{if } \hat{k}^{(n)} \neq k. \end{cases} \quad (20.4)$$

*What about ties?* – We don’t expect two means to be exactly the same distance from a data point, but if a tie does happen,  $\hat{k}^{(n)}$  is set to the smallest of the winning  $\{k\}$ .

**Update step.** The model parameters, the means, are adjusted to match the sample means of the data points that they are responsible for.

$$\mathbf{m}^{(k)} = \frac{\sum_n r_k^{(n)} \mathbf{x}^{(n)}}{R^{(k)}} \quad (20.5)$$

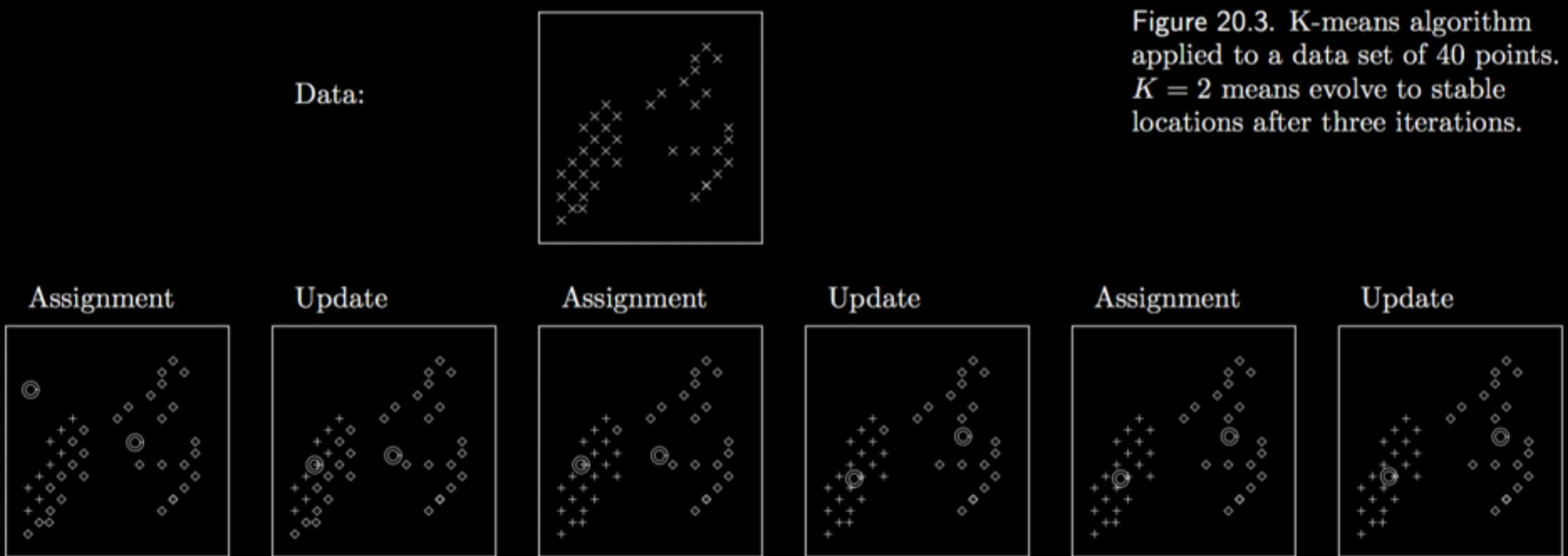
where  $R^{(k)}$  is the total responsibility of mean  $k$ ,

$$R^{(k)} = \sum_n r_k^{(n)}. \quad (20.6)$$

*What about means with no responsibilities?* – If  $R^{(k)} = 0$ , then we leave the mean  $\mathbf{m}^{(k)}$  where it is.

**Repeat the assignment step and update step until the assignments do not change.**

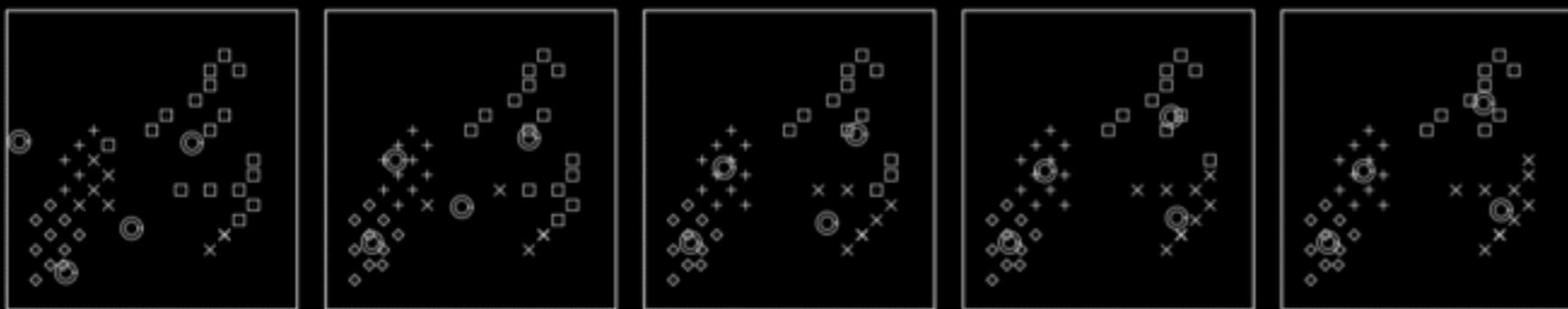
## K-MEANS ALGORITHM: IDENTIFY CLUSTER MEANS THAT HAVE MINIMAL DISTANCES TO ALL POINTS IN CLUSTER



K-MEANS ALGORITHM: IDENTIFY CLUSTER MEANS THAT HAVE MINIMAL DISTANCES TO ALL POINTS IN CLUSTER

## SOME CHALLENGES

Run 1



Run 2

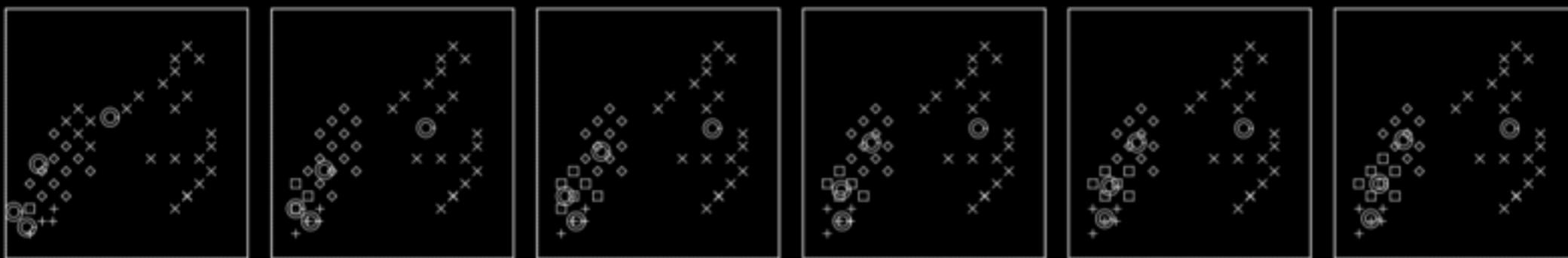
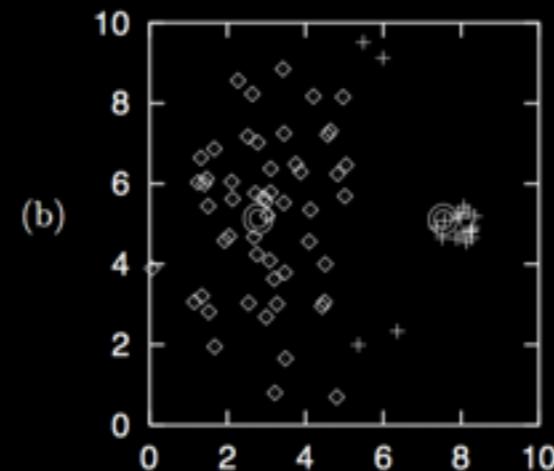
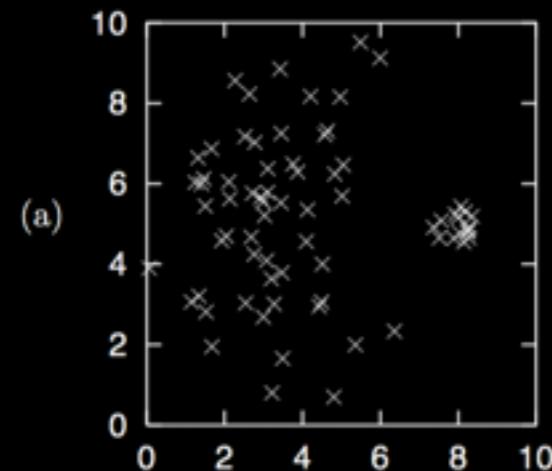
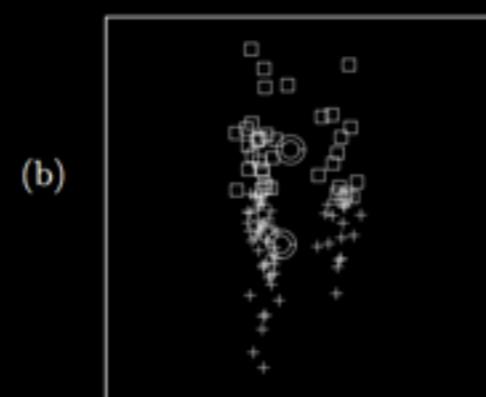
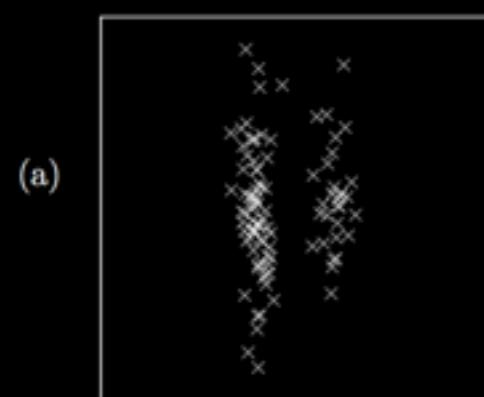


Figure 20.4. K-means algorithm applied to a data set of 40 points. Two separate runs, both with  $K = 4$  means, reach different solutions. Each frame shows a successive assignment step.

DEPENDANCE ON INITIALIZATION



**Figure 20.5.** K-means algorithm for a case with two dissimilar clusters. (a) The “little ‘n’ large” data. (b) A stable set of assignments and means. Note that four points belonging to the broad cluster have been incorrectly assigned to the narrower cluster. (Points assigned to the right-hand cluster are shown by plus signs.)



**Figure 20.6.** Two elongated clusters, and the stable solution found by the K-means algorithm.

## PROBLEMS WITH ‘HARD’ CLUSTERING APPROACHES

**Assignment step.** Each data point  $\mathbf{x}^{(n)}$  is given a soft ‘degree of assignment’ to each of the means. We call the degree to which  $\mathbf{x}^{(n)}$  is assigned to cluster  $k$  the *responsibility*  $r_k^{(n)}$  (the responsibility of cluster  $k$  for point  $n$ ).

$$r_k^{(n)} = \frac{\exp(-\beta d(\mathbf{m}^{(k)}, \mathbf{x}^{(n)}))}{\sum_{k'} \exp(-\beta d(\mathbf{m}^{(k')}, \mathbf{x}^{(n)}))}. \quad (20.7)$$

The sum of the  $K$  responsibilities for the  $n$ th point is 1.

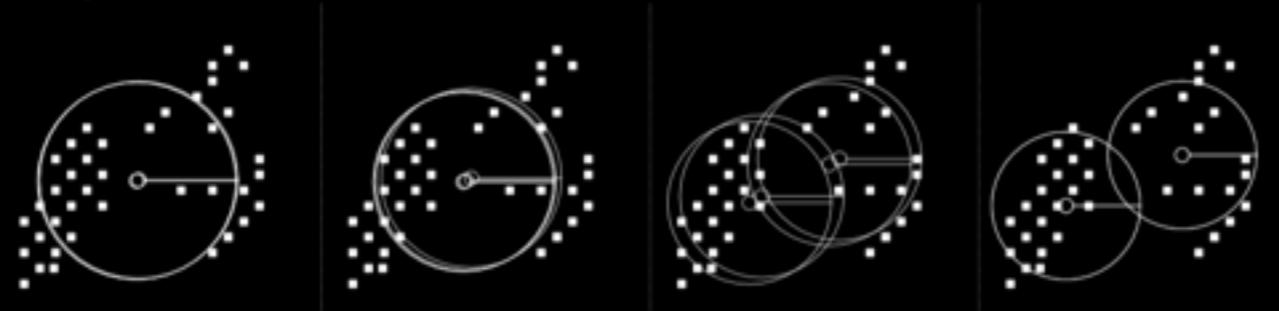
**Update step.** The model parameters, the means, are adjusted to match the sample means of the data points that they are responsible for.

$$\mathbf{m}^{(k)} = \frac{\sum r_k^{(n)} \mathbf{x}^{(n)}}{R^{(k)}} \quad (20.8)$$

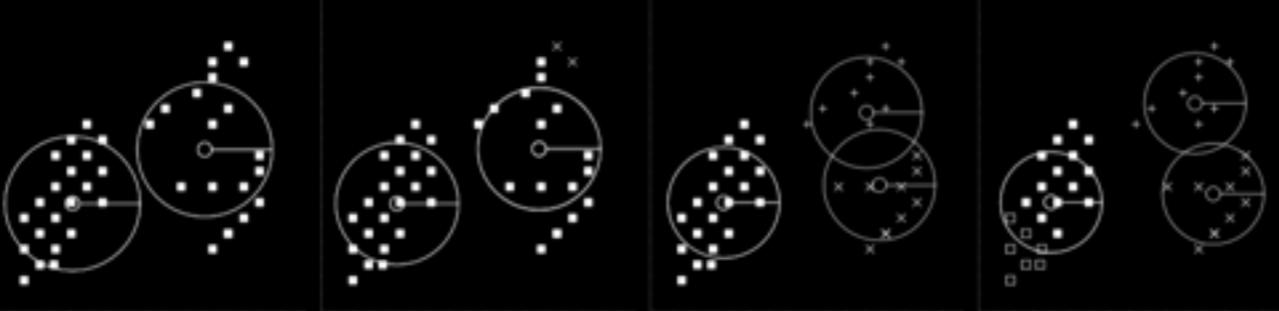
where  $R^{(k)}$  is the total responsibility of mean  $k$ ,

$$R^{(k)} = \sum_n r_k^{(n)}. \quad (20.9)$$

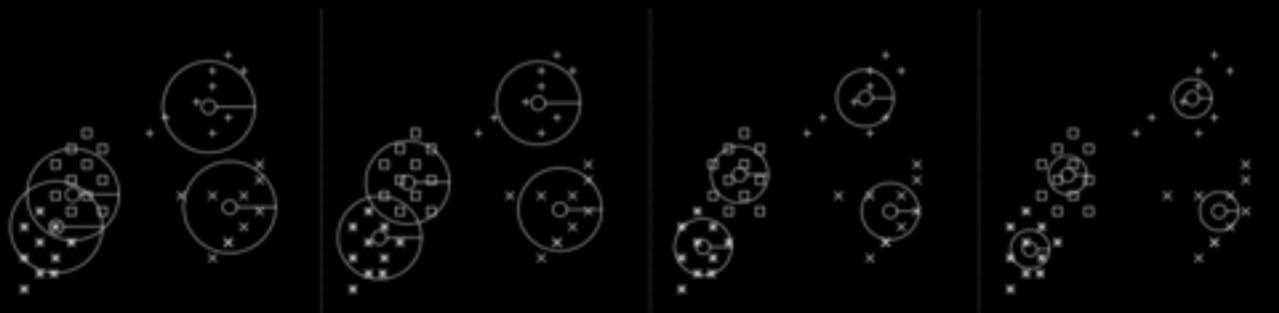
Large  $\sigma \dots$



...



... small  $\sigma$



**Figure 20.8.** Soft K-means algorithm, version 1, applied to a data set of 40 points.  $K = 4$ . Implicit lengthscale parameter  $\sigma = 1/\beta^{1/2}$  varied from a large to a small value. Each picture shows the state of all four means, with the implicit lengthscale shown by the radius of the four circles, after running the algorithm for several tens of iterations. At the largest lengthscales, all four means converge exactly to the data mean. Then the four means separate into two groups of two. At shorter lengthscales, each of these pairs itself bifurcates into subgroups.

## SOFT CLUSTERING SOLUTIONS

**Algorithm 22.2.** The soft K-means algorithm, version 2.

**Assignment step.** The responsibilities are

$$r_k^{(n)} = \frac{\pi_k \frac{1}{(\sqrt{2\pi}\sigma_k)^I} \exp\left(-\frac{1}{\sigma_k^2} d(\mathbf{m}^{(k)}, \mathbf{x}^{(n)})\right)}{\sum_{k'} \pi_k \frac{1}{(\sqrt{2\pi}\sigma_{k'})^I} \exp\left(-\frac{1}{\sigma_{k'}^2} d(\mathbf{m}^{(k')}, \mathbf{x}^{(n)})\right)} \quad (22.22)$$

where  $I$  is the dimensionality of  $\mathbf{x}$ .

**Update step.** Each cluster's parameters,  $\mathbf{m}^{(k)}$ ,  $\pi_k$ , and  $\sigma_k^2$ , are adjusted to match the data points that it is responsible for.

$$\mathbf{m}^{(k)} = \frac{\sum_n r_k^{(n)} \mathbf{x}^{(n)}}{R^{(k)}} \quad (22.23)$$

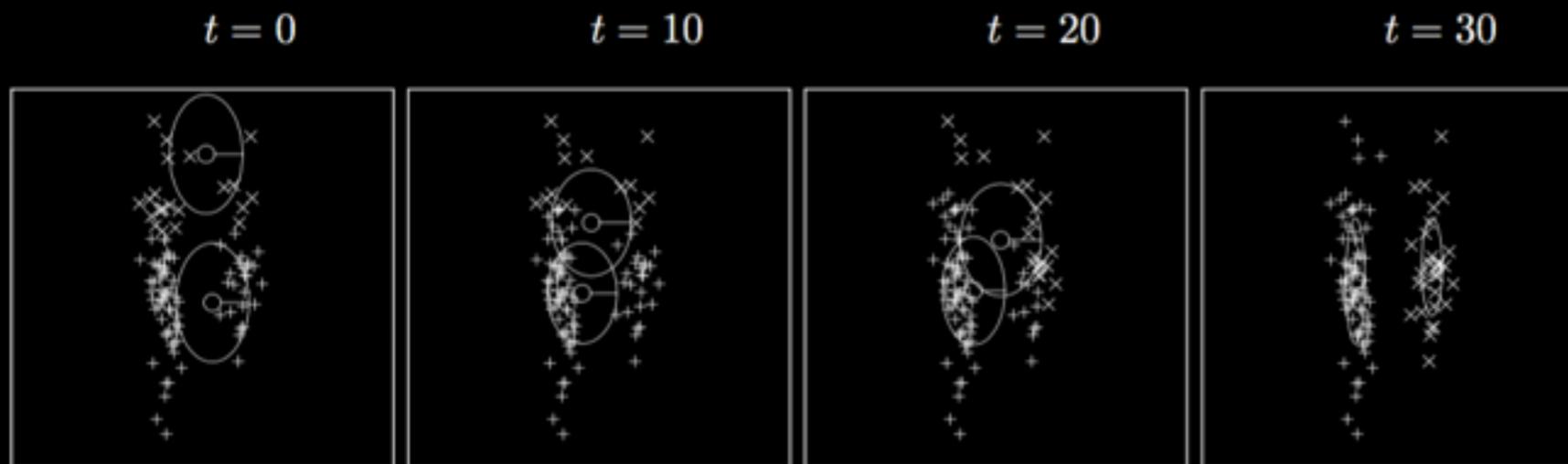
$$\sigma_k^2 = \frac{\sum_n r_k^{(n)} (\mathbf{x}^{(n)} - \mathbf{m}^{(k)})^2}{IR^{(k)}} \quad (22.24)$$

$$\pi_k = \frac{R^{(k)}}{\sum_k R^{(k)}} \quad (22.25)$$

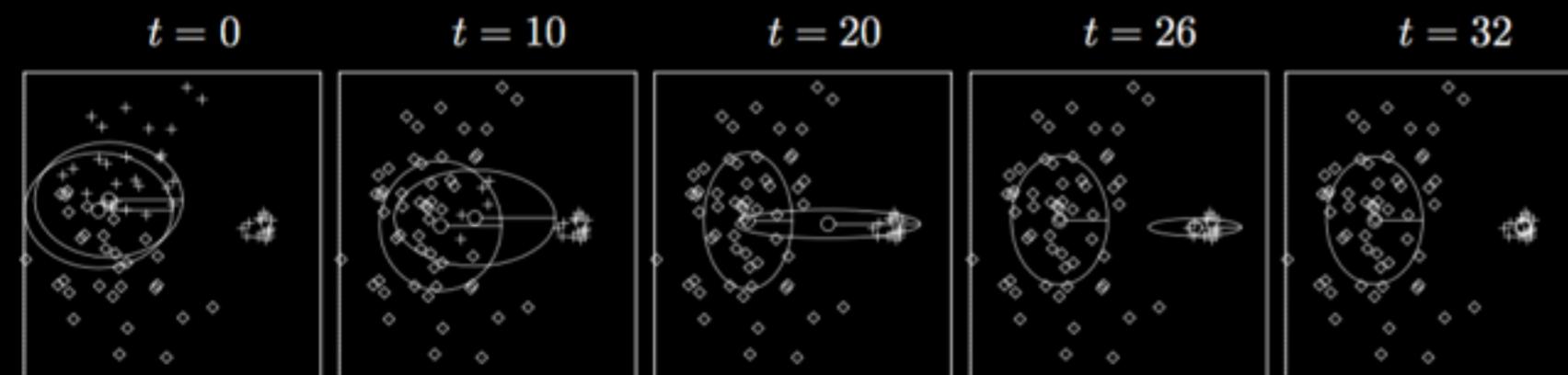
where  $R^{(k)}$  is the total responsibility of mean  $k$ ,

$$R^{(k)} = \sum_n r_k^{(n)}. \quad (22.26)$$

## SOFT CLUSTERING - MIXTURE OF GAUSSIANS



**Figure 22.5.** Soft K-means algorithm, version 3, applied to the data consisting of two cigar-shaped clusters.  $K = 2$  (cf. figure 20.6).

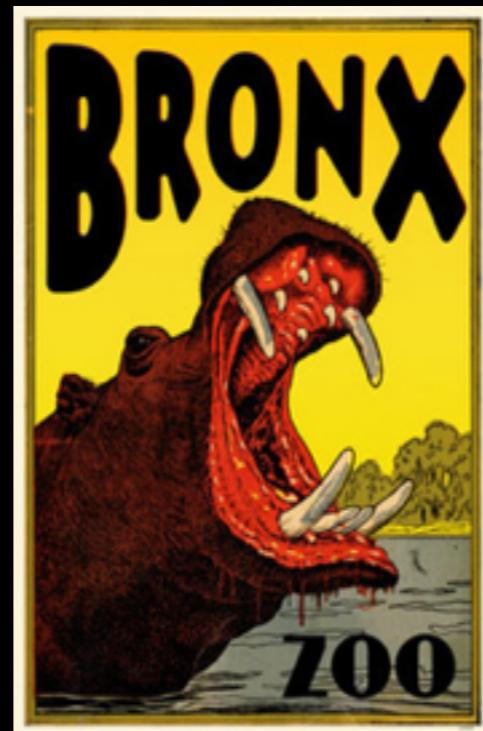


**Figure 22.6.** Soft K-means algorithm, version 3, applied to the little 'n' large data set.  $K = 2$ .

## SOFT CLUSTERING - MIXTURE OF GAUSSIANS

REQUIREMENTS FOR K-MEANS

DISTANCE MEASURE BETWEEN POINTS  
NUMBER OF CLUSTERS K  
INITIAL ASSIGNMENT



## A ZOO OF DISTANCE MEASURES

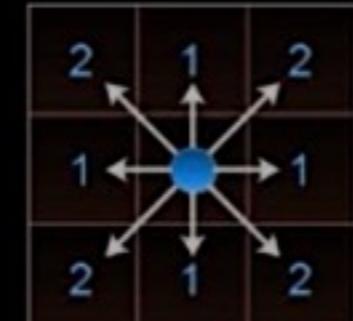
EUCLIDEAN, MANHATTAN, CHEBYCHEW  
MINKOWSKI, CANBERRA, MAXIMUM, COSINE

Euclidean Distance



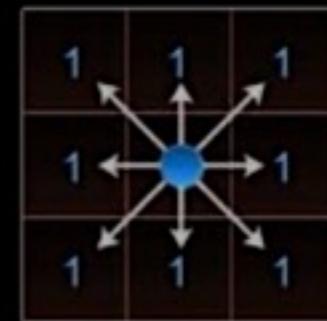
$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Manhattan Distance



$$|x_1 - x_2| + |y_1 - y_2|$$

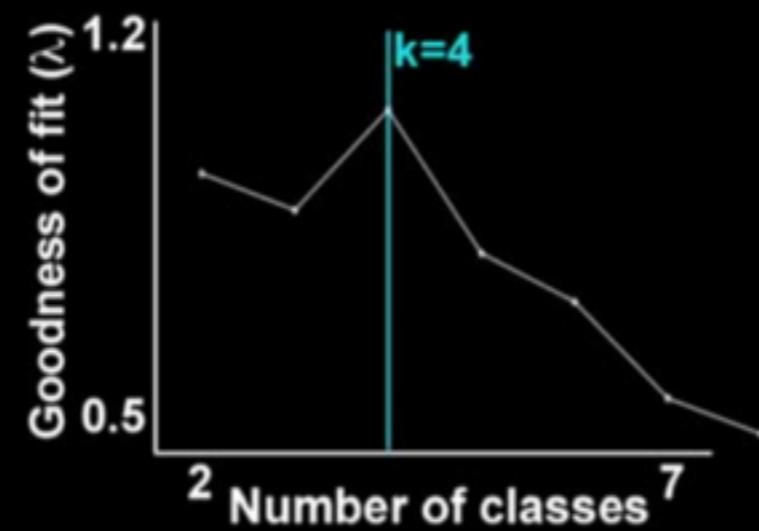
Chebyshev Distance



$$\max(|x_1 - x_2|, |y_1 - y_2|)$$

HOW MANY Ks

COMPACTNESS:  
MINIMIZE WITHIN-CLUSTER VARIANCE  
MAXIMIZE BETWEEN-CLUSTER VARIANCE



**NbClust Package for determining the best number of clusters****Description**

`NbClust` package provides 30 indices for determining the number of clusters and proposes to user the best clustering scheme from the different results obtained by varying all combinations of number of clusters, distance measures, and clustering methods.

**Usage**

```
NbClust(data, diss = NULL, distance = "euclidean", min.nc = 2, max.nc = 15,
method = "ward.D2", index = "all", alphaBeale = 0.1)
```

**Arguments**

- |                       |  |
|-----------------------|--|
| <code>data</code>     | matrix or dataset (the only mandatory argument)  |
| <code>diss</code>     | dissimilarity matrix to be used. By default, <code>diss=NULL</code> , but if it is replaced by a dissimilarity matrix, distance should be "NULL".  |
| <code>distance</code> | the distance measure to be used to compute the dissimilarity matrix. This must be one of: "euclidean", "maximum", "manhattan", "canberra", "binary", "minkowski" or "NULL". By default, <code>distance="euclidean"</code> . If the distance is "NULL", the dissimilarity matrix ( <code>diss</code> ) should be given by the user. If distance is not "NULL", the dissimilarity matrix should be "NULL". |
| <code>min.nc</code>   | minimal number of clusters, between 1 and (number of objects - 1)  |
| <code>max.nc</code>   | maximal number of clusters, between 2 and (number of objects - 1), greater or equal to <code>min.nc</code> . By default,   |

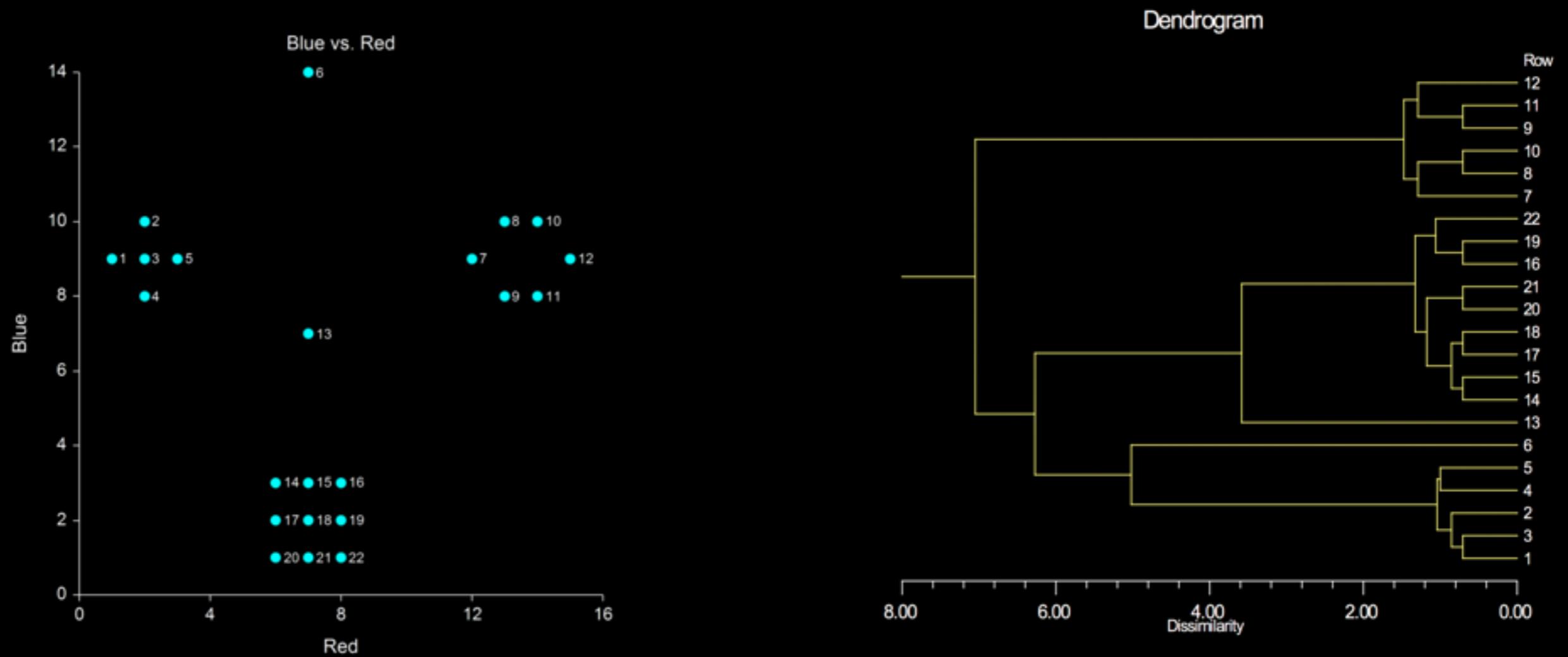
HIERARCHICAL AGGLOMERATIVE CLUSTERING:  
ONLY NEEDS SIMILARITY MEASURE BETWEEN GROUPS OF POINTS

HIERARCHICAL AGGLOMERATIVE CLUSTERING:  
ONLY NEEDS SIMILARITY MEASURE BETWEEN GROUPS OF POINTS

ALGORITHM:

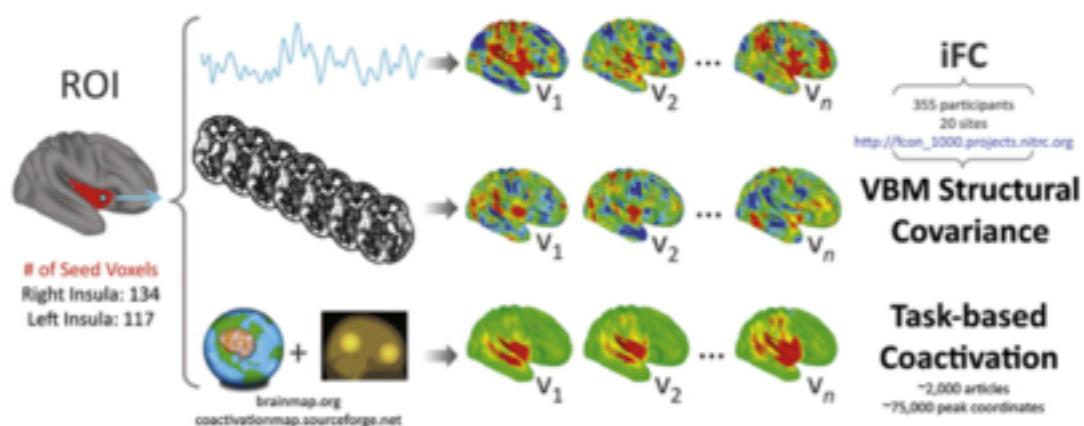
- 1) PLACE EACH DATA POINT INTO A SINGLE GROUP
- 2) MERGE THE TWO MOST SIMILAR GROUPS, REPEAT
- 3) STOP WHEN ONLY ONE CLUSTER LEFT

## HIERARCHICAL AGGLOMERATIVE CLUSTERING: ONLY NEEDS SIMILARITY MEASURE BETWEEN GROUPS OF POINTS

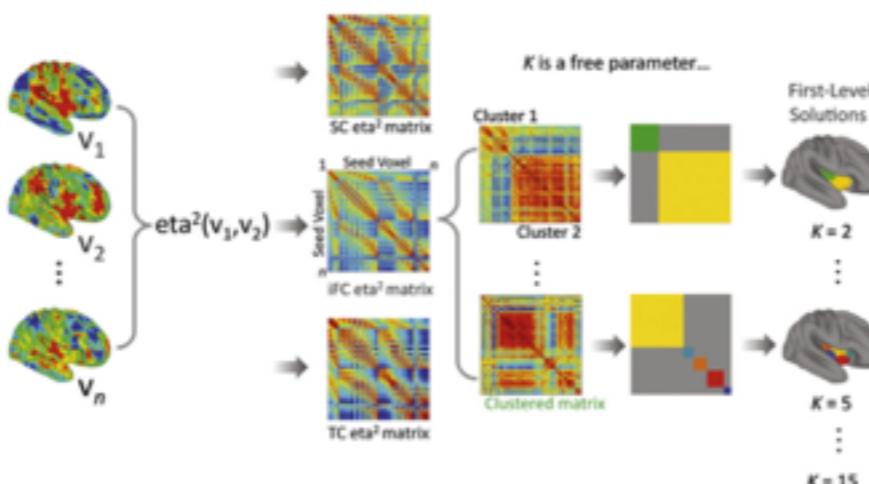


KELLY

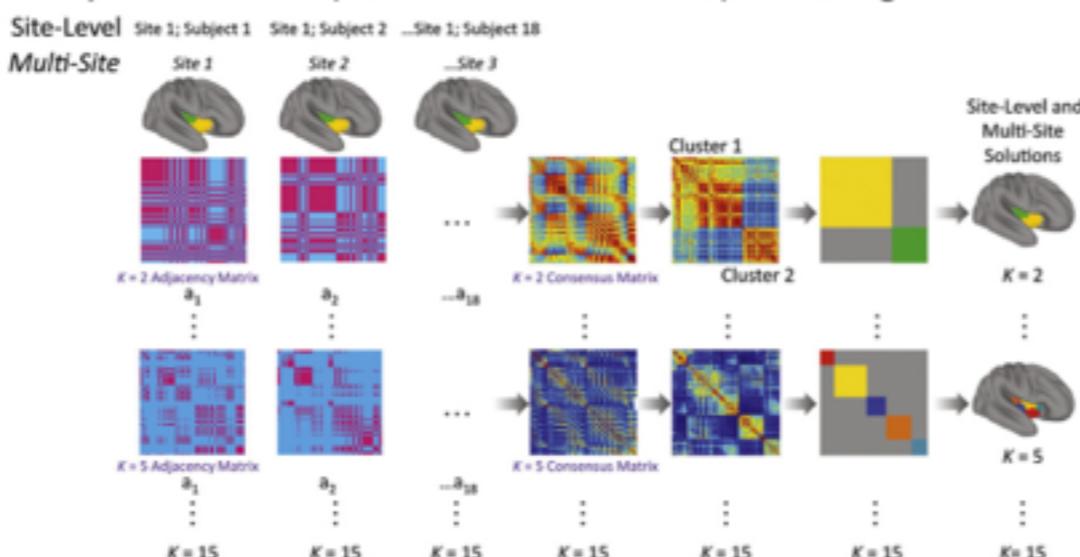
## Step 1: Covariance-Based Measures



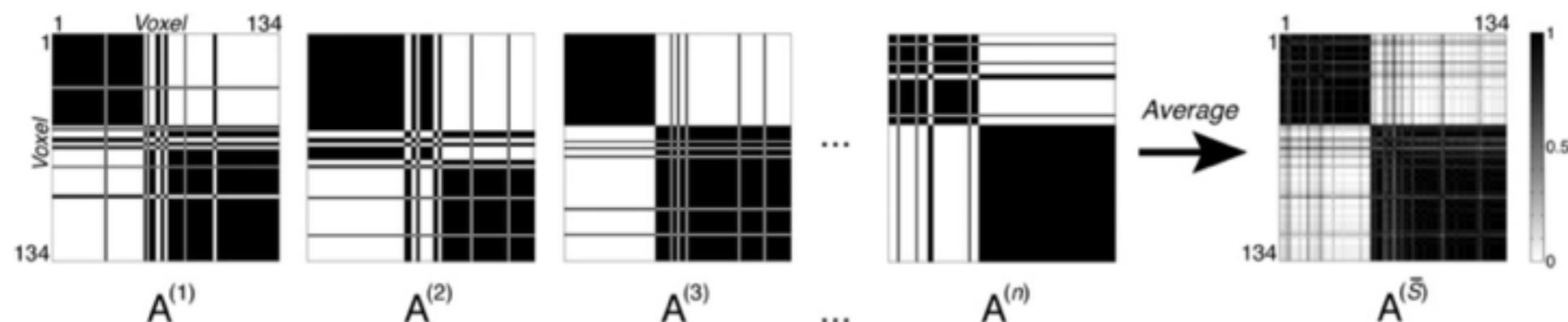
## Step 2: $\eta^2$ and First-Level Clustering



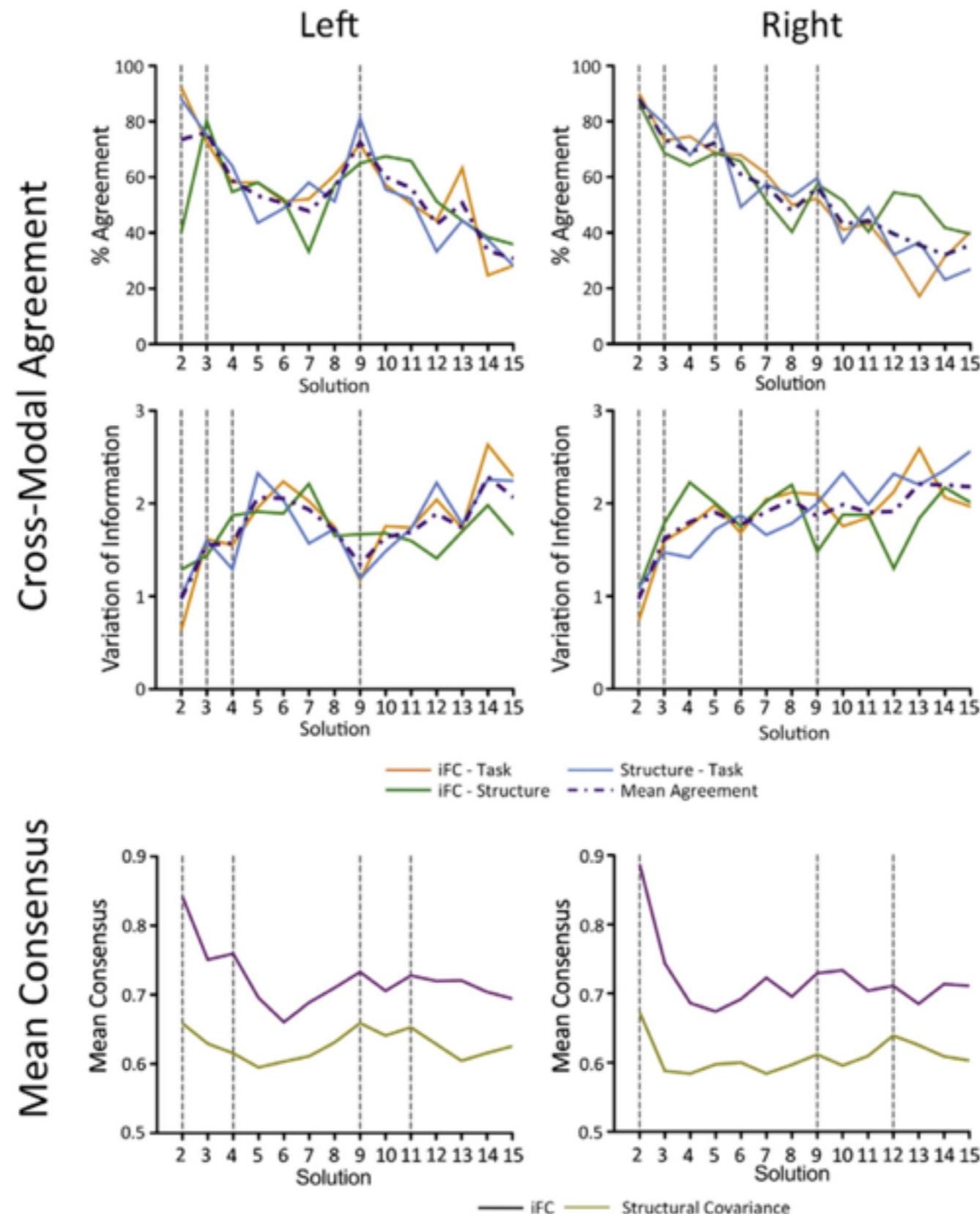
## Step 3: Consensus (Site-Level and Multi-Site) Clustering



**Fig. 1.** Analysis schematic. Step 1: Covariance-based measures. Starting with ROIs comprising the right (134 voxels) and left (117 voxels) insula, per the Harvard–Oxford atlas that accompanies FSL (25% probability, resampled to  $4 \times 4 \times 4$  mm voxels), we computed whole-brain intrinsic functional connectivity (iFC), gray-matter structural covariance, and task-based coactivation, for each voxel ( $v_1, \dots, v_n$ ) within each ROI ( $n = 134$  for the right insula;  $n = 117$  for the left). Step 2:  $\eta^2$  and First-Level Clustering. For each covariance measure we used  $\eta^2$  to quantify the similarity between every pair of covariance maps, producing a set of  $134 \times 134$  matrices containing 8911 unique  $\eta^2$  values ranging between 0 and 1 for the right hemisphere; and a set of  $117 \times 117$  matrices containing 6786 unique  $\eta^2$  values for the left. We then used spectral clustering (Meila and Shi, 2001) to partition the insula into clusters of voxels maximizing intra-cluster/minimizing inter-cluster similarity. This produced three sets of clustering solutions for  $K = 2, \dots, 15$ , one for each covariance measure. Step 3: Consensus (Site-Level and Multi-Site) Clustering. Where more than one set of clustering solutions is available, it is possible to apply a stability or consensus clustering approach to derive clustering solutions that are stable across instances. For the iFC data, two levels of consensus clustering were possible — one that derived 20 sets of site-level clustering solutions that were stable across individuals within each site, and a second that derived a single set of clustering solutions that were stable across data collection sites. For the structural covariance data, only one level of consensus clustering was possible — the one that captured stable solutions across data collection sites. Because the task coactivation data produced only a single set of clustering solutions at the first level, no consensus clustering was possible for those data. All the scripts used to perform the multisite consensus clustering analysis are available via [http://fcon\\_1000.projects.nitrc.org](http://fcon_1000.projects.nitrc.org).



**Fig. 2.** Consensus clustering schematic for  $K=2$ . The schematic illustrates the consensus clustering process. For each scale  $K$ , each clustering instance contributes an adjacency matrix  $A^{(s)}$ , each element  $a_{ij}^{(s)}$  of which contains a value of 1 if voxels  $i$  and  $j$  are assigned to the same cluster  $k$ , and 0 otherwise. In this example, let each instance be a data collection site, so  $A^{(1)}$  is contributed by Bangor;  $A^{(2)}$  is contributed by Berlin, etc. A consensus matrix  $A^{(\bar{S})}$  is derived by averaging across adjacency matrices. Each element of the consensus matrix thus contains a number between 0 and 1, corresponding to the proportion of times a given pair of voxels appeared in the same cluster, across instances (here, data collection sites). The spectral clustering algorithm can then be applied to identify the most stable pattern of cluster assignments across instances, using the same scale  $K$  that was used to generate the consensus matrix (here,  $K=2$ ).



**Fig. 3.** Clustering metrics. We identified optimal clustering solutions on the basis of their cross-modal correspondence. For each  $K$ , we compared the cluster solutions generated for each modality using Percent Agreement (top two panels) and Variation of Information (VI; middle two panels). We selected the values of  $K$  showing the highest mean Percent Agreement or lowest mean VI (purple dashed lines), across the three different pairings of the three modalities. In addition to strong agreement for the trivial  $K=2$  solutions, these metrics indicated that strong agreement was obtained for  $K=2, 3, 4$  and  $9$  (left hemisphere), and  $3, 5, 6, 7$  and  $9$  (right hemisphere). The bottom two panels show, for iFC and structural covariance, the mean intra-cluster consensus (stability) values for each  $K$ . The locations of peaks in the Mean Consensus ( $K=2, 4, 9$ , and  $11$ , left hemisphere;  $K=2, 9$ , and  $12$ , right hemisphere) plots show good correspondence with the locations of peaks in the Agreement plots, suggesting that the solutions showing the best agreement across modalities also exhibited the strongest stability, across data collection sites.

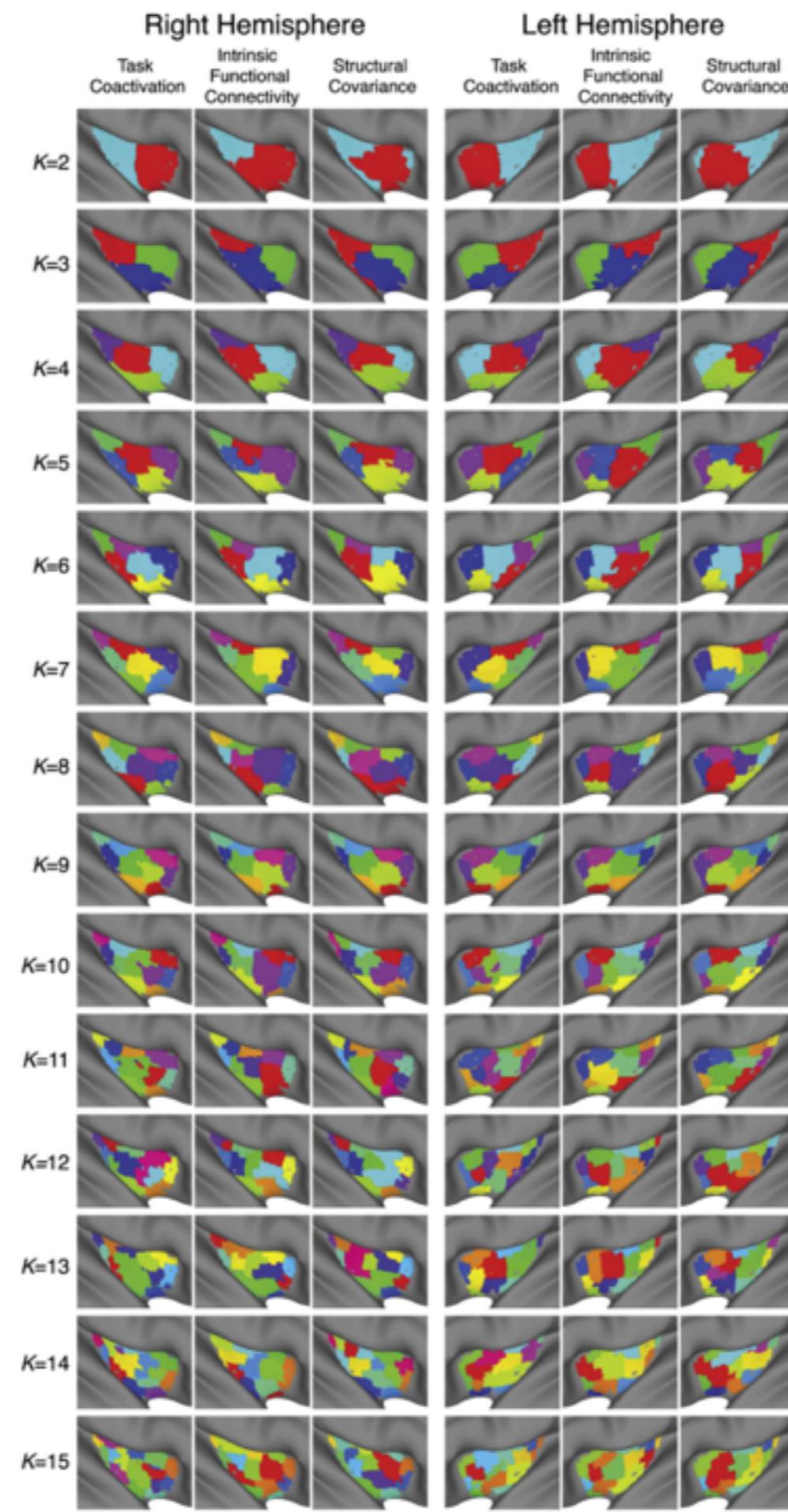
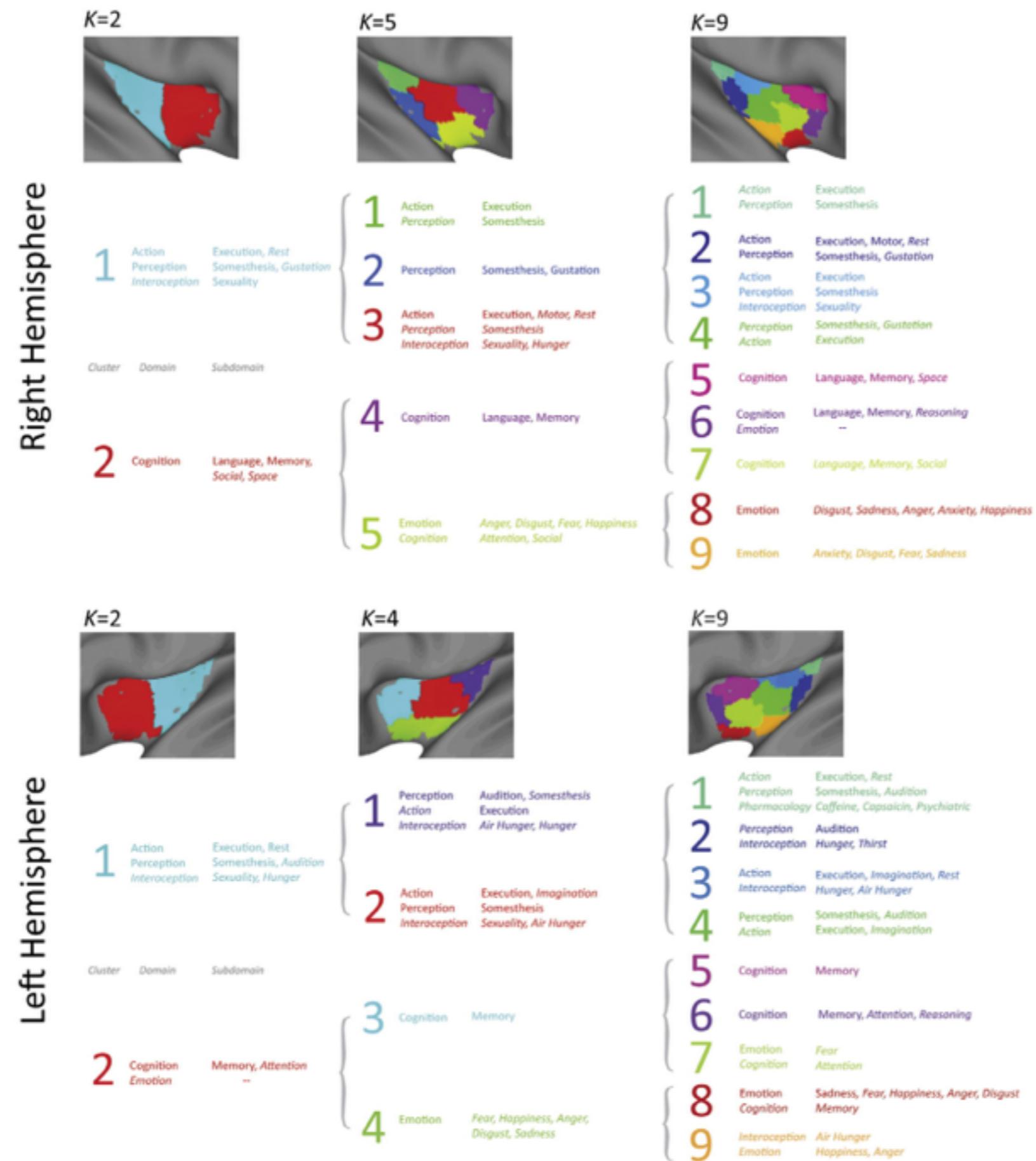
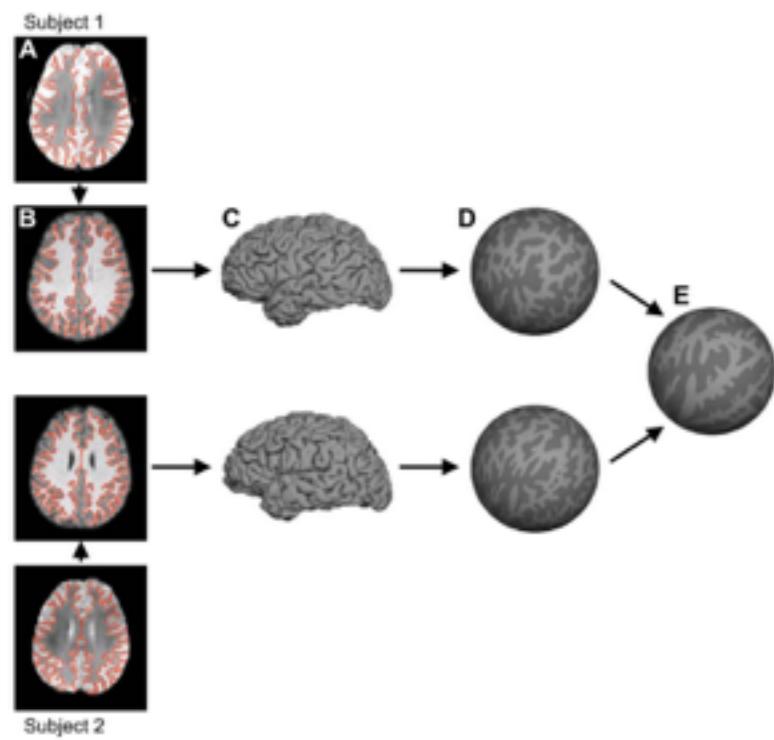


Fig. 4. Final cluster solutions. Consensus/multi-site cluster solutions for  $K = 2, \dots, 15$  for the structural covariance, intrinsic functional connectivity (IFC) and task-based coactivation data.

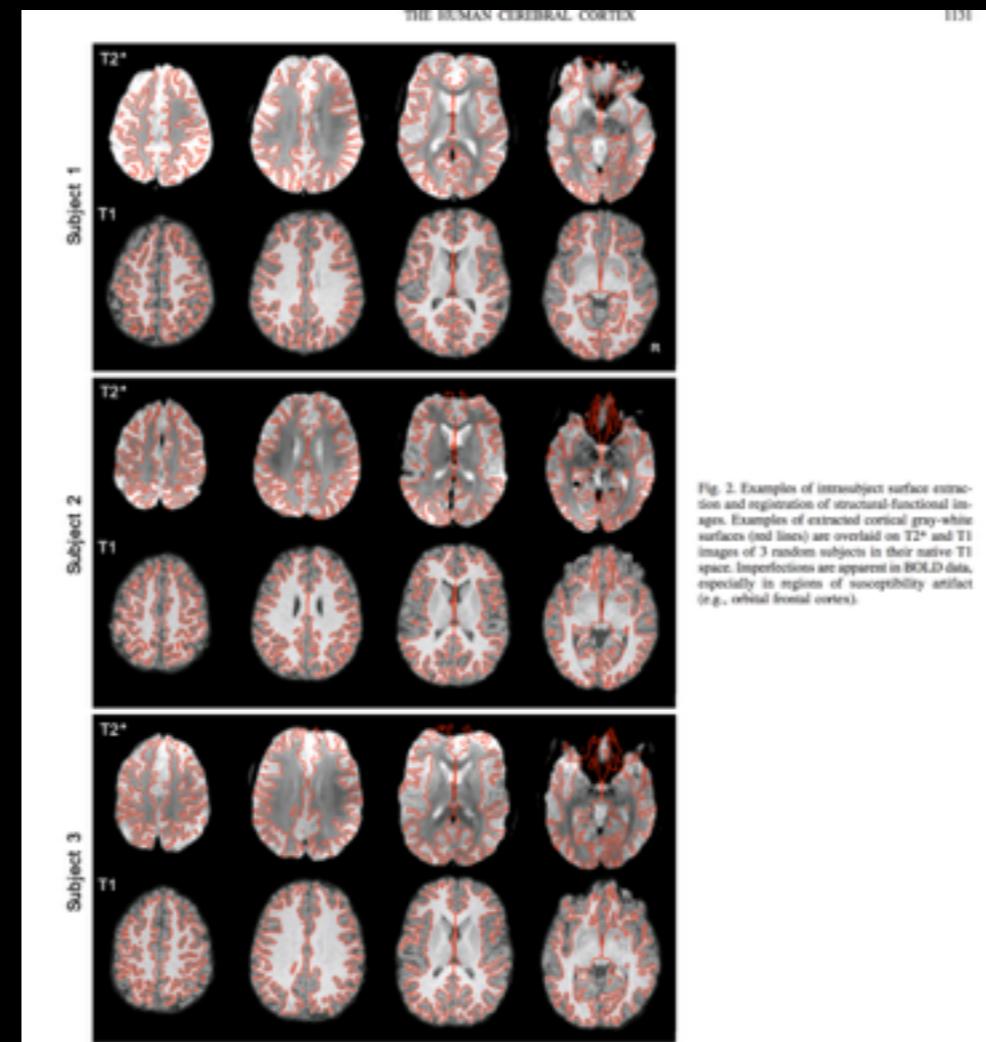


**Fig. 5.** Behavioral domain analysis. Significant behavioral domains (and subdomains) identified for K = 2, 4 and 9 (left hemisphere), and 2, 5 and 9 (right hemisphere). Domains and subdomains exceeding an uncorrected threshold of  $p < 0.05$  but not the Bonferroni threshold are in italics. Distributions for each domain and cluster are shown in Supplemental Figs. S13–S18.

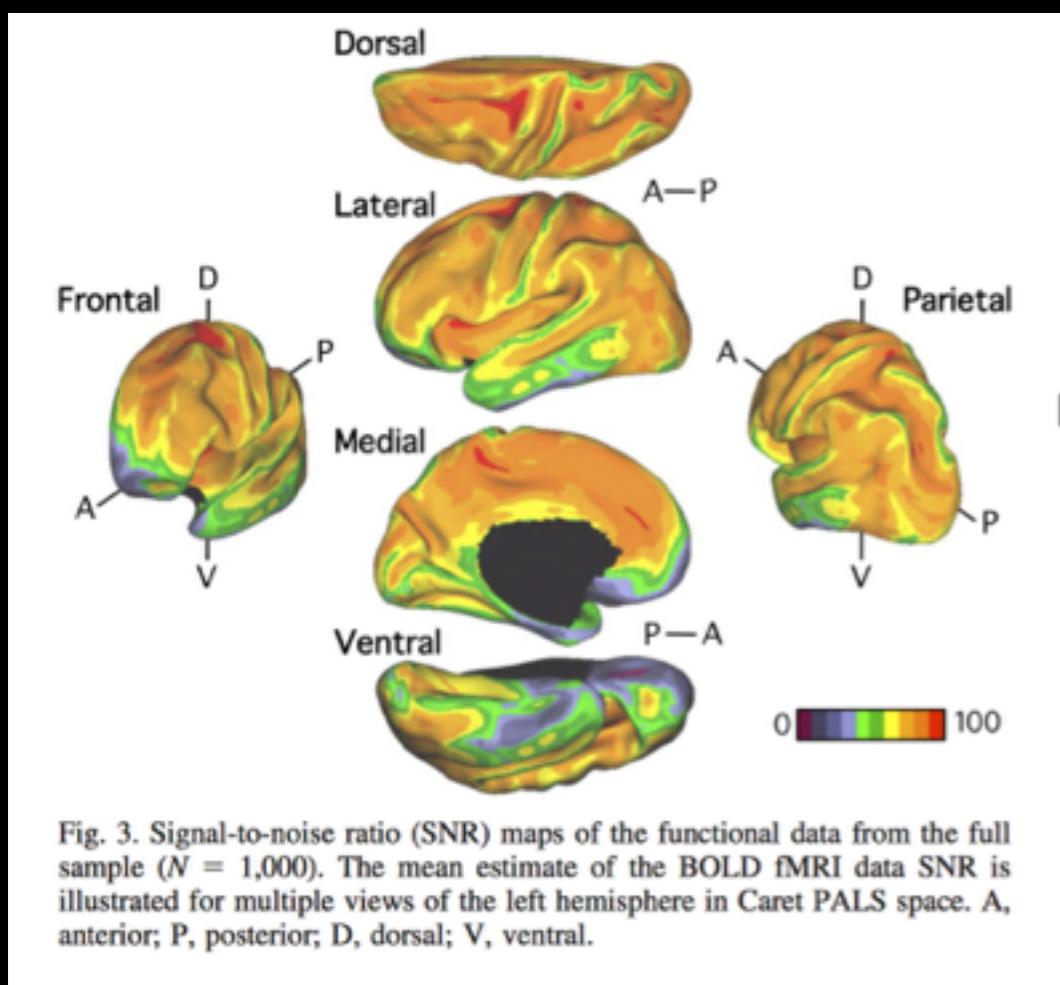
YEO



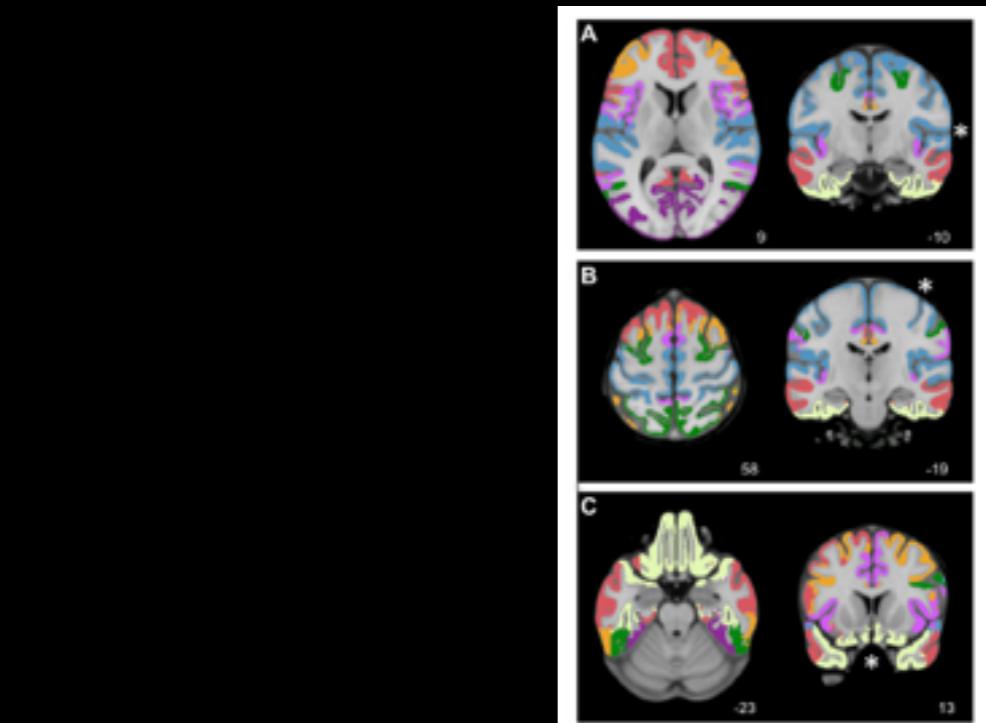
**Fig. 1.** Surface coordinate system for functional magnetic resonance imaging (fMRI) analysis. For each subject, the  $T2^*$  images yielding blood oxygenation level-dependent (BOLD) contrast fMRI data (**A**) were registered to the  $T1$ -weighted structural data (**B**). The cortical gray-white and pial surfaces were estimated from the structural data. The red lines show the estimated gray-white surface (**A** and **B**). Pial surface is shown in **C**. The gray-white surface was inflated into a sphere (**D**). The inflated spheres were then aligned across subjects using surface-based registration of the cortical folding pattern, resulting in a common spherical coordinate system (**E**). BOLD data of individual subjects (**A**) can then be projected onto the spherical coordinate system (**E**) in a single transformation step to reduce artifacts due to multiple interpolations.



**Fig. 2.** Examples of intrasubject surface extraction and registration of structural-functional images. Examples of extracted cortical gray-white surfaces (red lines) are overlaid on  $T2^*$  and  $T1$  images of 3 random subjects in their native  $T1$  space. Imperfections are apparent in BOLD data, especially in regions of susceptibility artifact (e.g., orbital frontal cortex).



**Fig. 3.** Signal-to-noise ratio (SNR) maps of the functional data from the full sample ( $N = 1,000$ ). The mean estimate of the BOLD fMRI data SNR is illustrated for multiple views of the left hemisphere in Caret PALS space. A, anterior; P, posterior; D, dorsal; V, ventral.



**Fig. 4.** Uncertain observations due to limited data resolution and MR susceptibility. When the clustering results are interpreted, potential artifacts and uncertainties must be considered. Because of the close proximity of the somatomotor and auditory cortices (**A**) and the close proximity of the pre- and postcentral gyri (**B**), we are unable to resolve whether the clustering of the somatosensory and primary motor cortices (**A**) and the clustering of the primary somatosensory and auditory cortices (**B**) are due to the result of fMRI blurring across sulci or a true, coupled network of distributed areas as predicted by macaque tracing studies. **C**: the orbitofrontal-temporal-polar network (green color) consists of temporo-polar and orbital frontal regions that are affected by MR susceptibility. Since MR susceptibility spatially distorts the MR signal and reduces SNR, there is uncertainty in the exact boundary of the orbitofrontal-temporal-polar network, and the true extent of the network is probably underestimated.

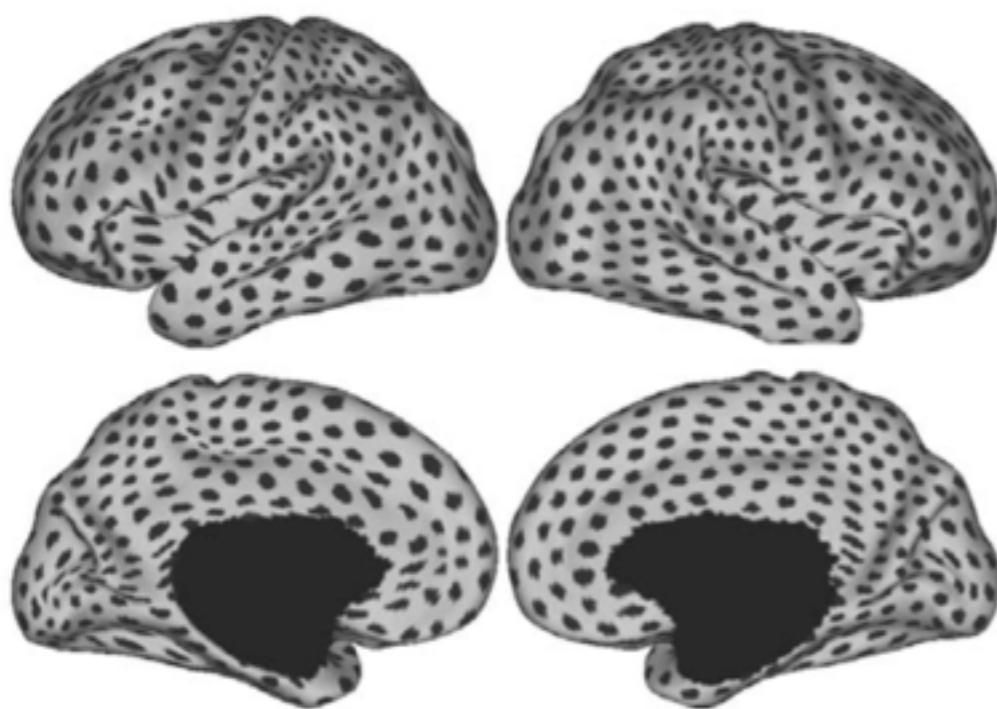


Fig. 4. Cortical regions utilized in constructing functional connectivity profiles. A total of 1,175 regions were sampled uniformly on the surface-based representations of the left and right hemispheres within the FreeSurfer surface coordinate system and are shown in Caret PALS space, where each dark patch represents the location of a single regional vertex. Each vertex in the surface coordinate system is characterized by its profile of functional connectivity to the 1,175 regions. The visually nonuniform distribution of the regions in Caret PALS space is due to the nonlinear deformation from FreeSurfer space to Caret PALS space. This image thus also serves to illustrate the subtle differences between the 2 surface coordinate systems.

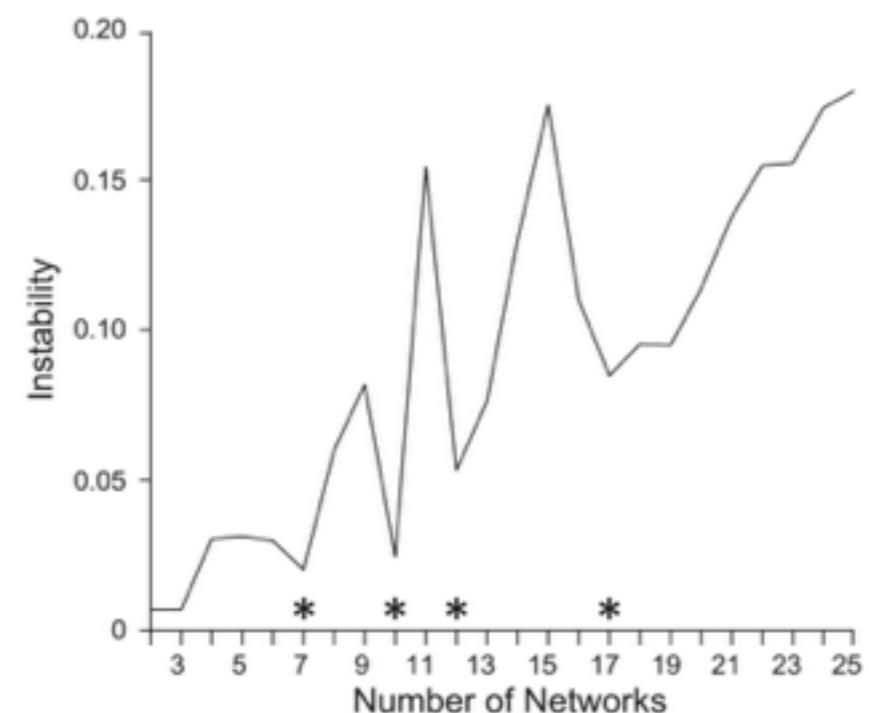
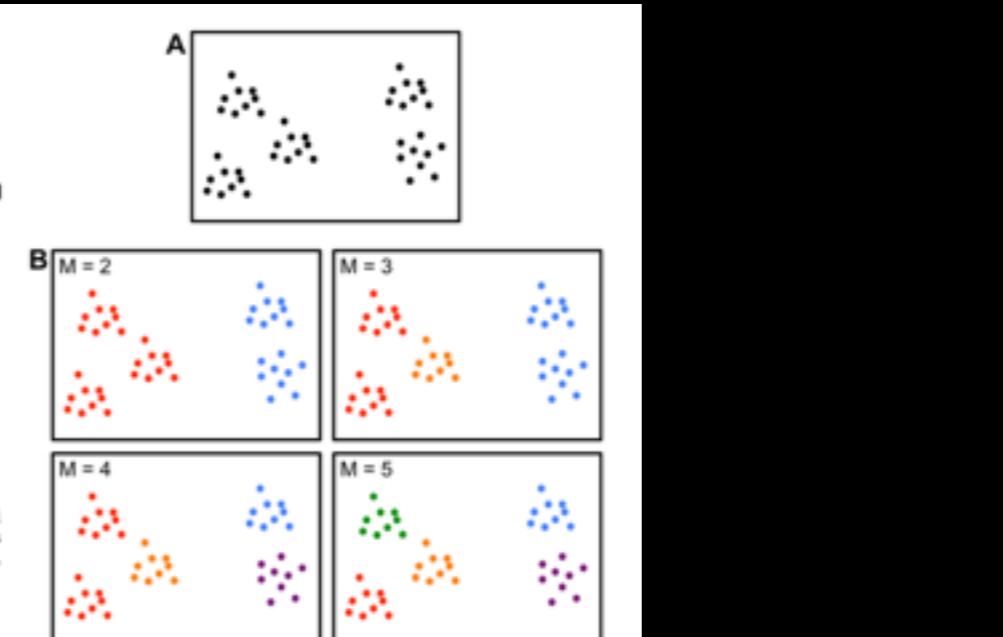
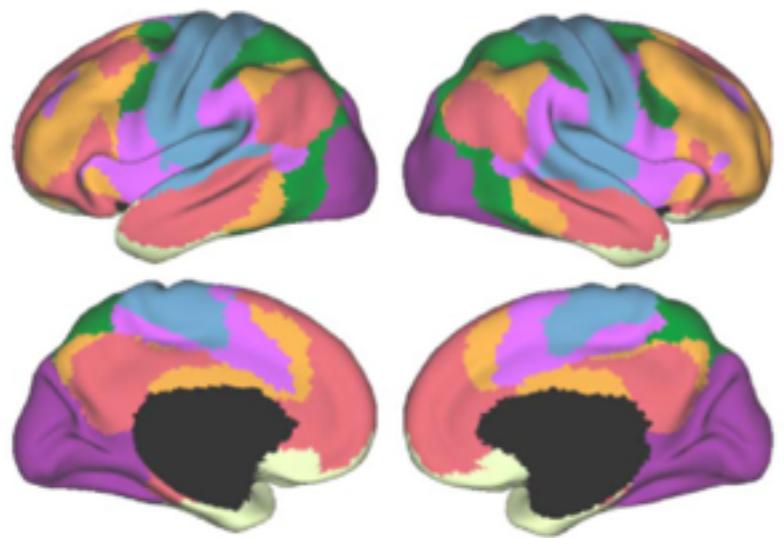


Fig. 6. Seven and 17 networks can be stably estimated. Instability of the clustering algorithm is plotted as a function of the number of estimated networks for the vertex-resampling variant of the stability analysis applied to 1,000 subjects. The clustering algorithm is less stable with increasing number of estimated networks, which is an expected property, since the number of estimated networks enlarges the solution space (and thus complexity) of the clustering problem. The local minima of the graphs (marked with asterisks) indicate the number of networks that can be stably estimated by the clustering algorithm. The stability analysis suggests that 7, 10, 12, or 17 networks can be stably estimated. Resampling the regions of interest yields almost identical results and is not shown. In this study we focus on the 7- and 17-network estimates to provide a broad survey of the solution space.

Discovery Sample ( $n = 500$ )



Replication Sample ( $n = 500$ )

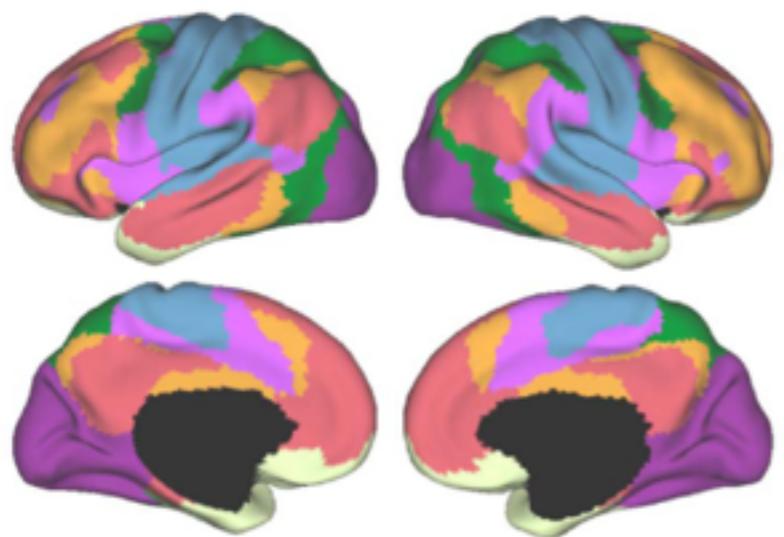


Fig. 7. Discovery and replication of a 7-network cortical parcellation. The 7-network estimates are highly consistent across the discovery ( $n = 500$ ) and replication ( $n = 500$ ) data sets. A total of 97.4% of the vertices were assigned to the same network.

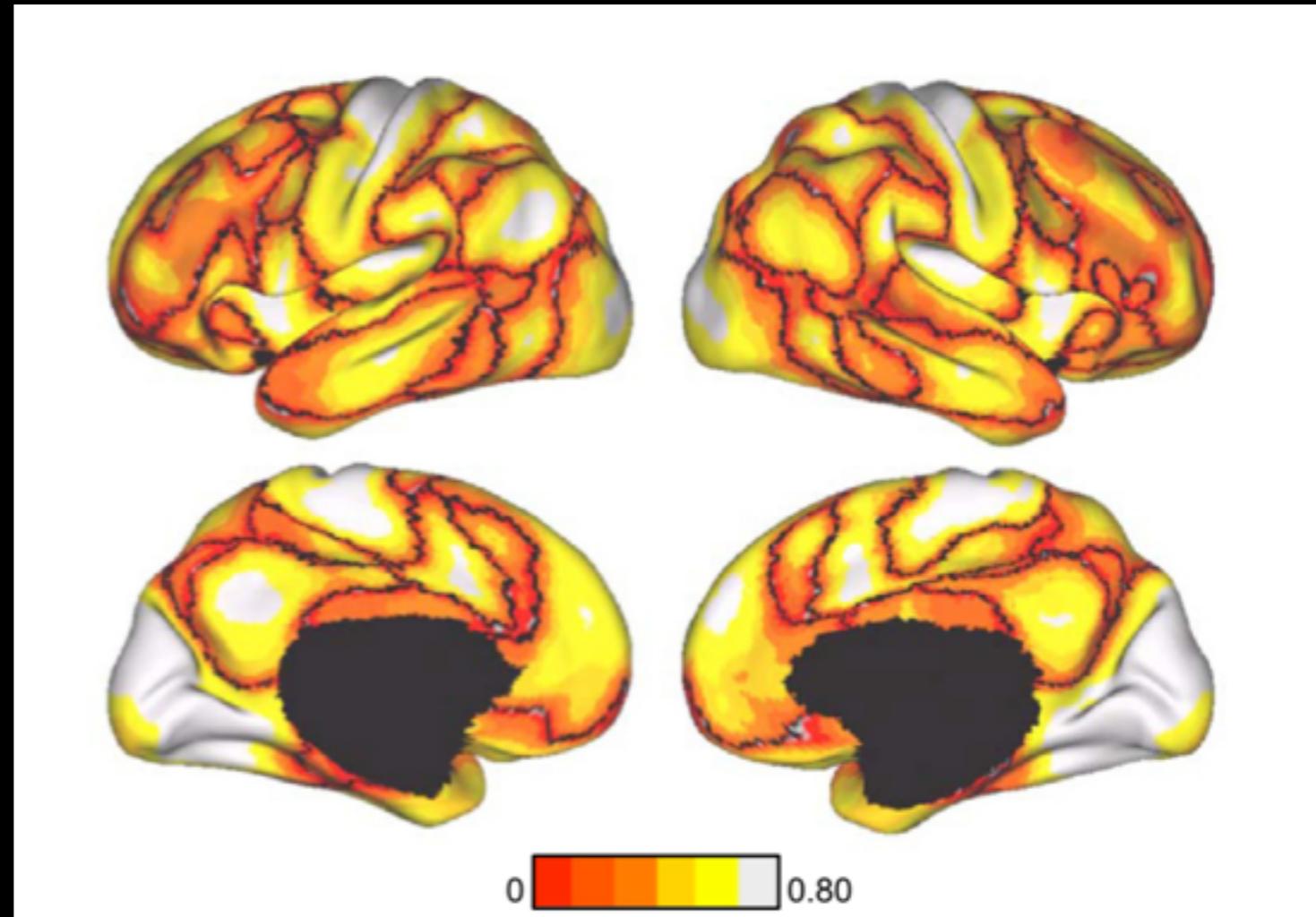
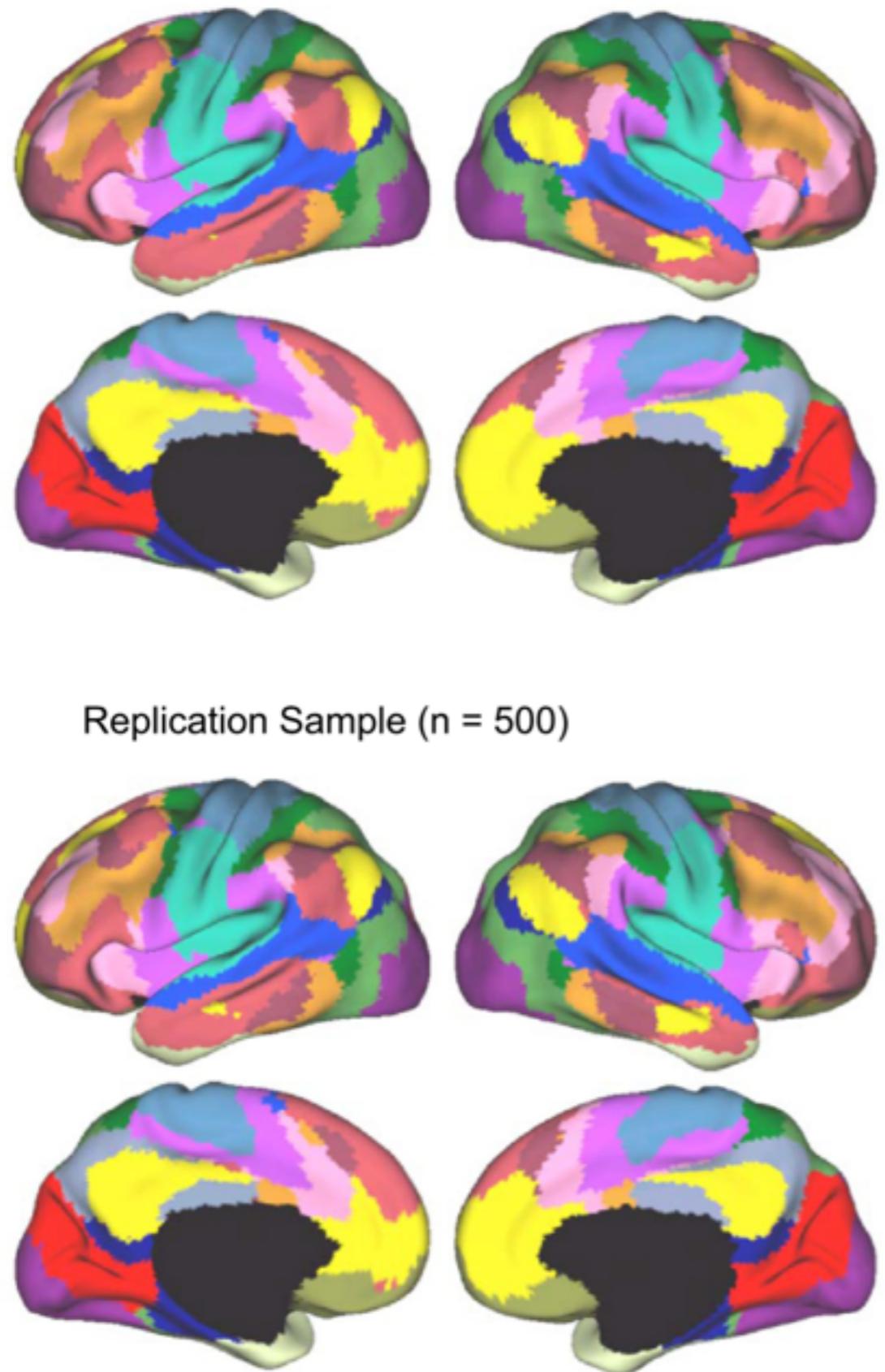


Fig. 8. Confidence of the 7-network estimate in the discovery data set. Confidence (silhouette) value for each vertex with respect to its assigned network is shown for the discovery data set. Regions close to the boundaries between networks were less confident of their assignment, although we also observed structured spatial variation within individual components of the estimated networks, such as lateral prefrontal cortex, which foreshadows its division in the 17-network estimate (see Fig. 9).

Discovery Sample ( $n = 500$ )



Replication Sample ( $n = 500$ )

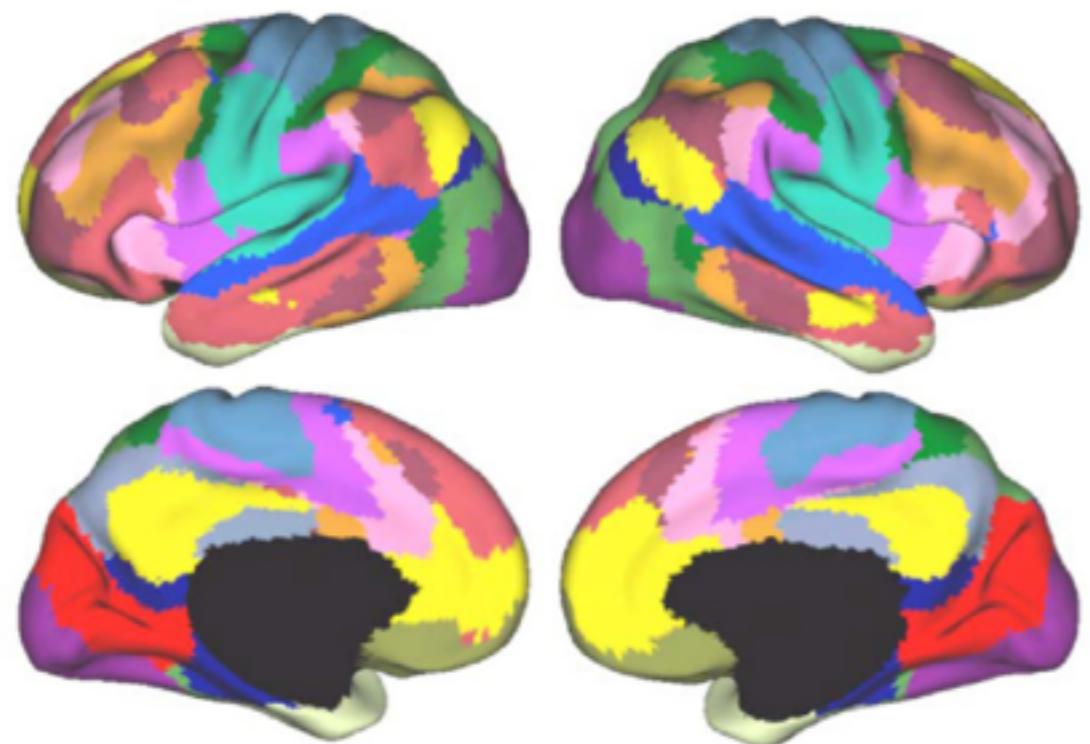


Fig. 9. Discovery and replication of a 17-network cortical parcellation. The 17-network estimates are highly consistent across the discovery ( $n = 500$ ) and replication ( $n = 500$ ) data sets. A total of 97.0% of the vertices were assigned to the same network.

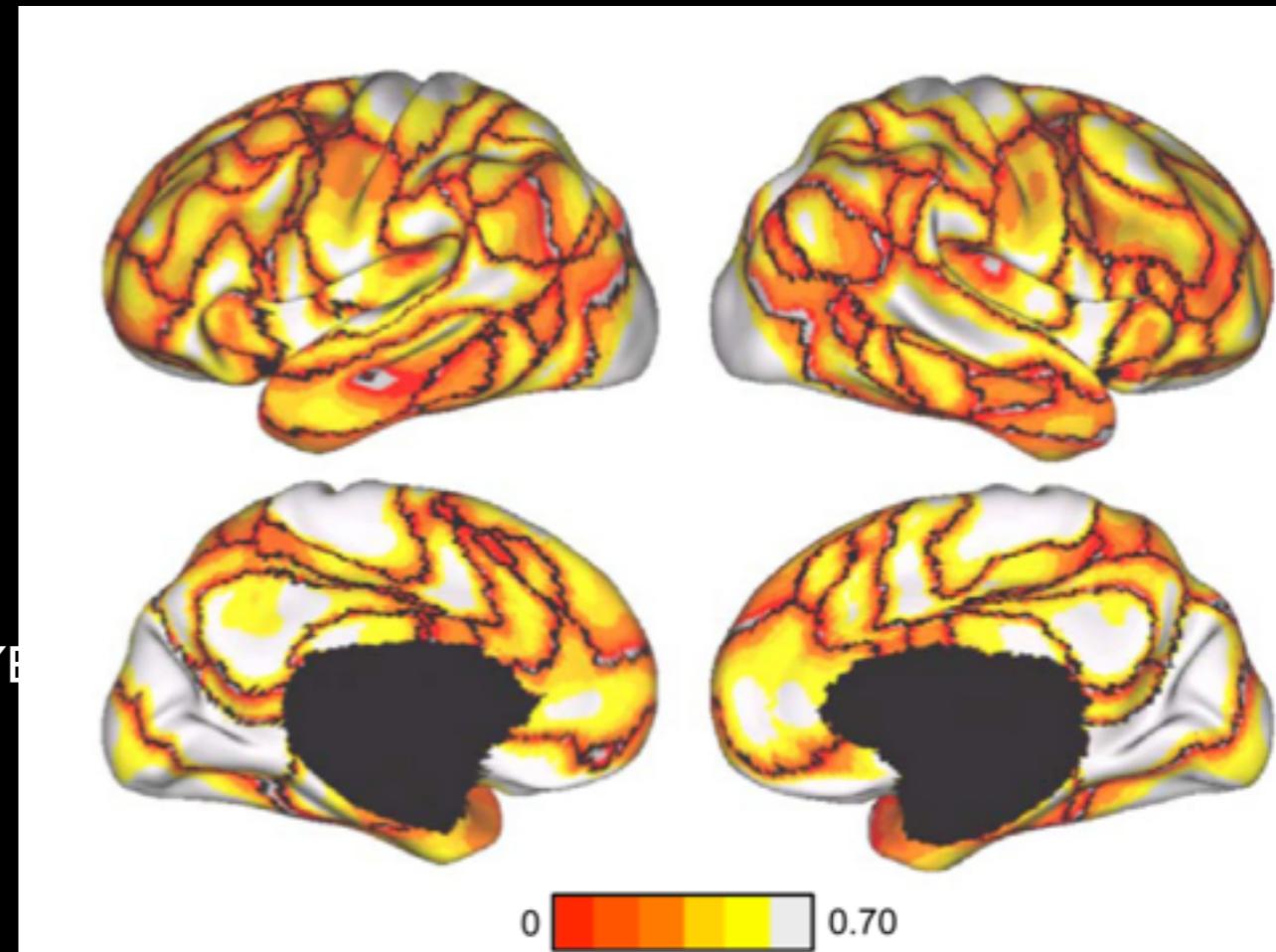


Fig. 10. Confidence of 17-network estimate in the discovery data set. Confidence (silhouette) value for each vertex with respect to its assigned network is shown for the discovery data set. Again, regions close to the boundaries between networks were less confident of their assignment.

## 7-Network Parcellation (N=1000)

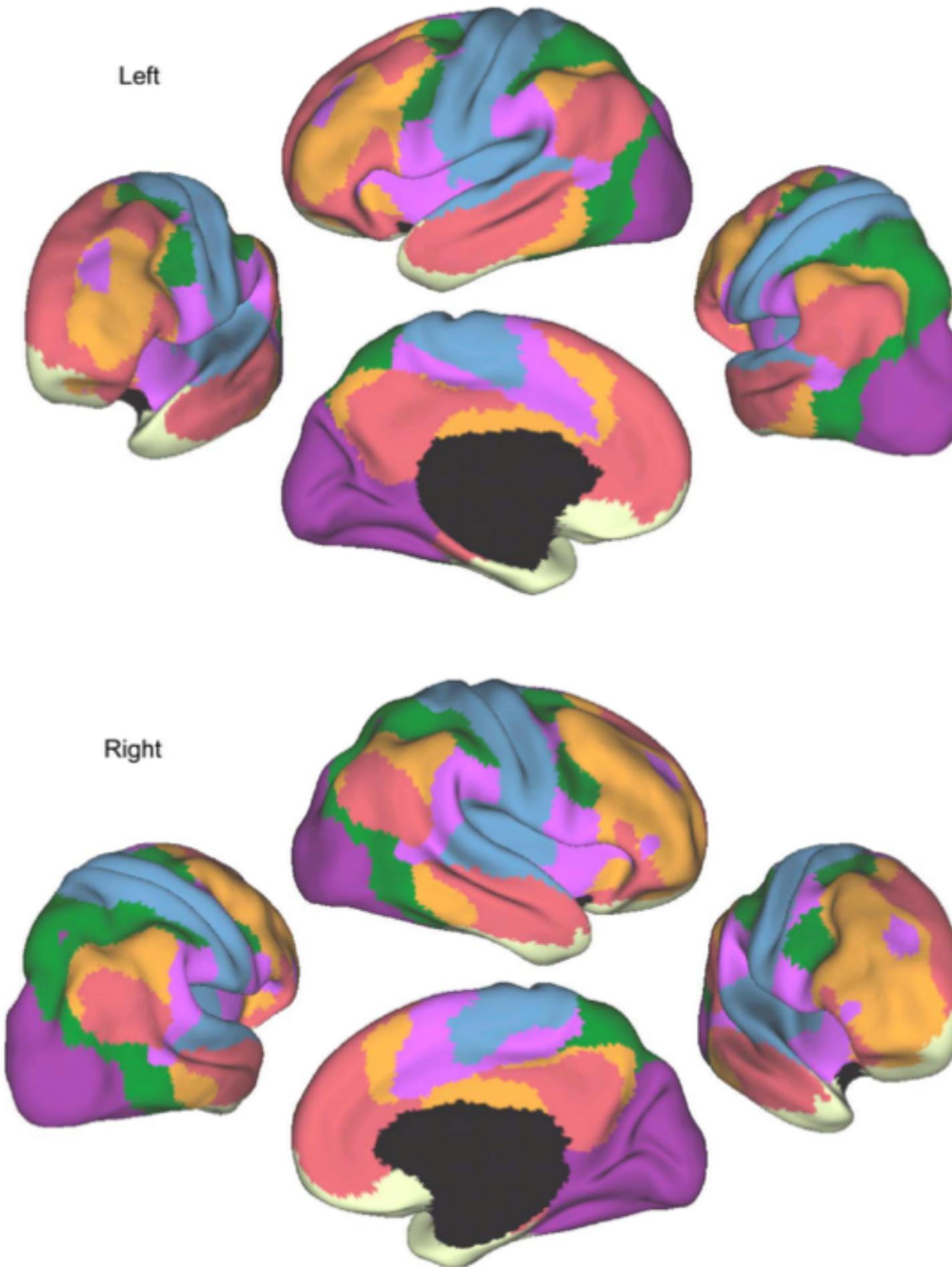


Fig. 11. A coarse (7-network) parcellation of the human cerebral cortex based on 1,000 subjects. To provide the best estimates of the 7 cortical networks, clustering was performed on the fMRI data of the full 1,000 subjects. A salient feature is the separation of the early sensory and late motor cortices (blue and purple) from the association cortex. The association networks converged and extended on networks previously described in the resting-state literature, including the dorsal attention, ventral attention, frontoparietal control, and default networks.

- Purple (Visual)
- Blue (Somatomotor)
- Green (Dorsal Attention)
- Violet (Ventral Attention)
- Cream (Limbic)
- Orange (Frontoparietal)
- Red (Default)

### 17-Network Parcellation (N=1000)

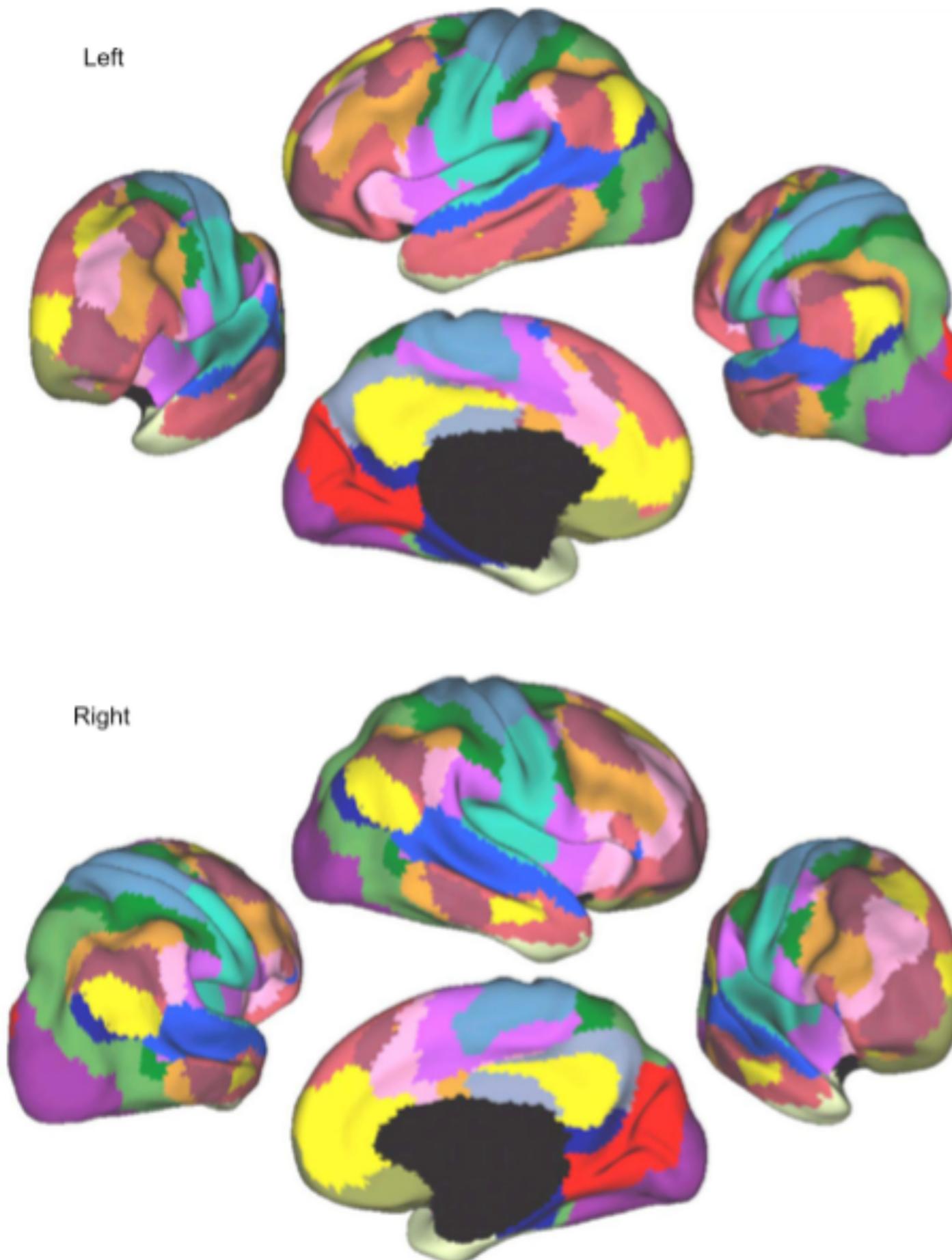


Fig. 13. A fine-resolution (17-network) parcellation of the human cerebral cortex based on 1,000 subjects. To provide the best estimates of the 17 cortical networks, clustering was performed on the fMRI data of the full 1,000 subjects. The 17-network estimate fractionated the 7-network into smaller networks. Some aspects of the fractionations have been previously noted in other studies.

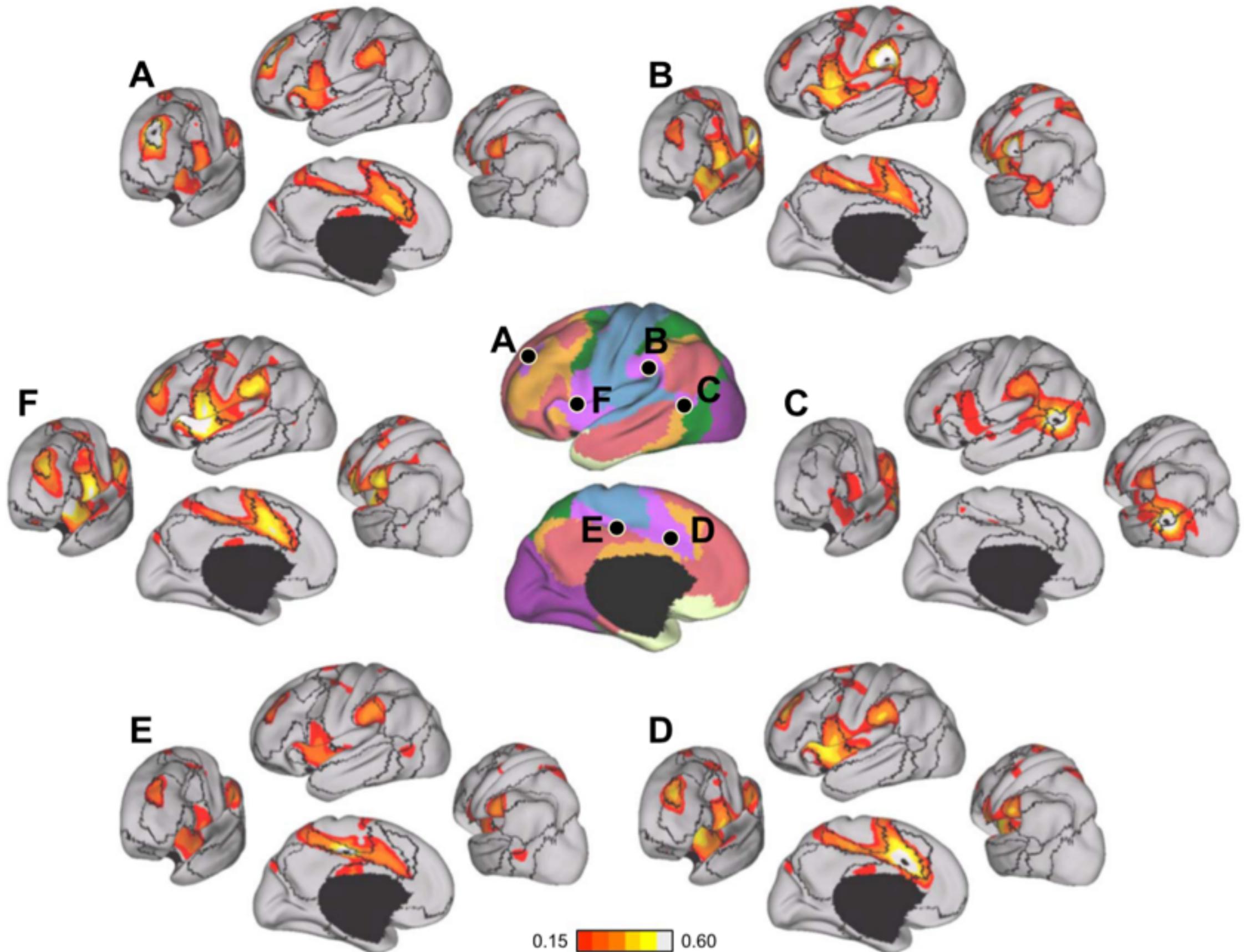


Fig. 33. Functional connectivity for distributed regions within a second large-scale association network. This network is often called the ventral attention network (but see Fig. 12 legend). The format and plotting are the same as for Fig. 32, and coordinate locations are reported in Table 5. Each seed region is functionally coupled mostly with regions within the same network, revealing that each component of the network recapitulates the others. There is a general absence of cross talk between networks except for local correlation around the seed regions.

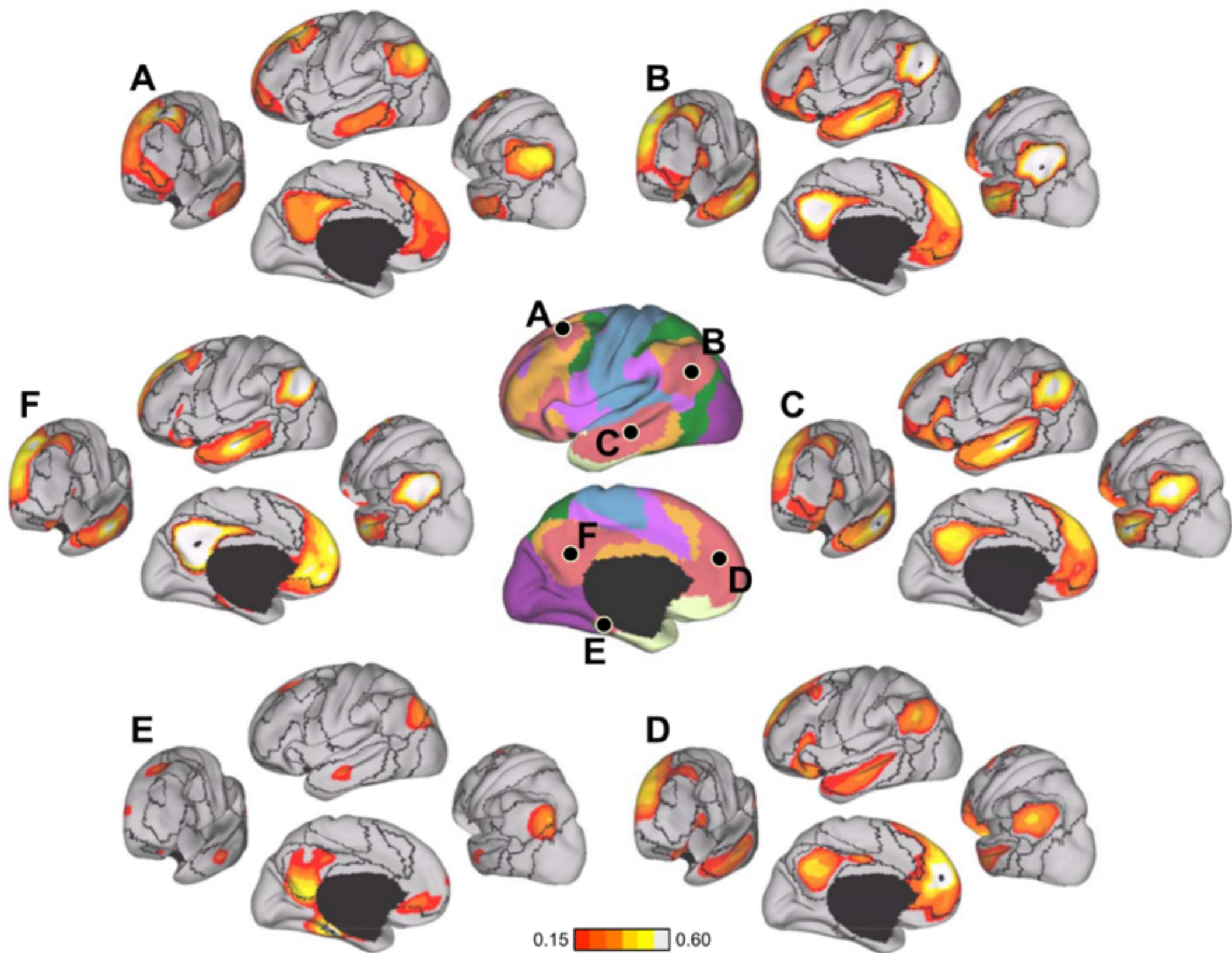


Fig. 35. Functional connectivity for distributed regions within a fourth large-scale association network. This network makes up the prominent components of the network often called the default network. The format and plotting are the same as for Fig. 32, and coordinate locations are reported in Table 5. Each seed region is functionally coupled mostly with regions within the same network, again revealing that each component of the network recapitulates components of the remaining network, with some exceptions. For example, the seed region in the parahippocampal cortex (*E*) shows functional connectivity with the retrosplenial cortex, ventral medial prefrontal cortex, and a specific region with the caudal IPL. These patterns of functional connectivity lead to the fractionation into subnetworks as illustrated in Fig. 13. The fractionation of this particular network is largely to subdivide the broader network, rather than to display functional coupling with regions in distinct networks.

TOPOGRAPHICALLY SPECIFIC CONNECTIVITY

## VISUAL

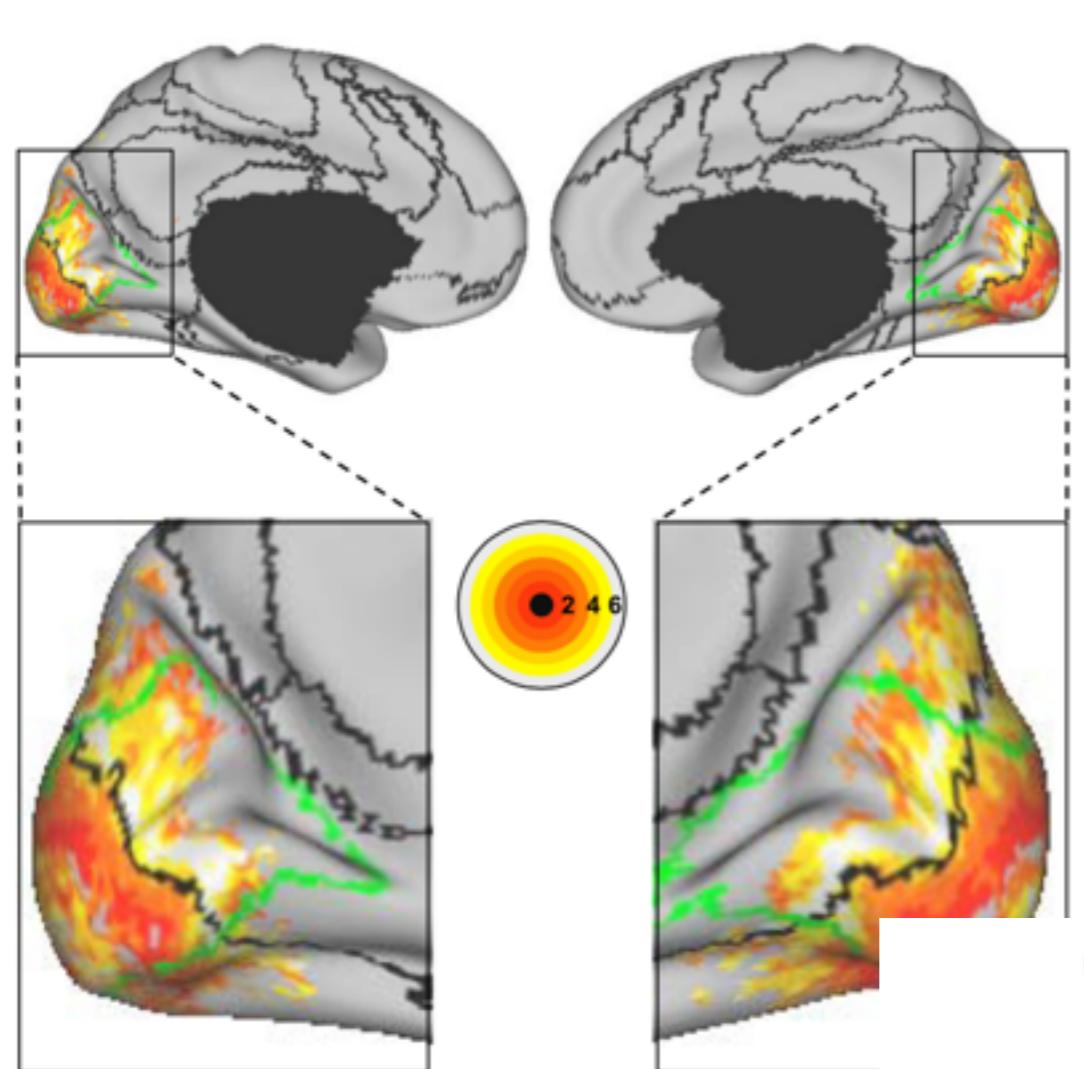


Fig. 15. Eccentricity estimates quantify the division of the early visual system into central and peripheral systems. Eccentricity estimates in the areas of 4 subjects were averaged and overlaid on the boundaries of the 17-network estimate. The boundary between areas 18 and 19 from the histological data set is overlaid in green. The 17-net divides the early visual areas along an isoeccentricity line of  $\sim 4^\circ$ . Eccentricity estimates are not reliable outside the V1-V3 complex.

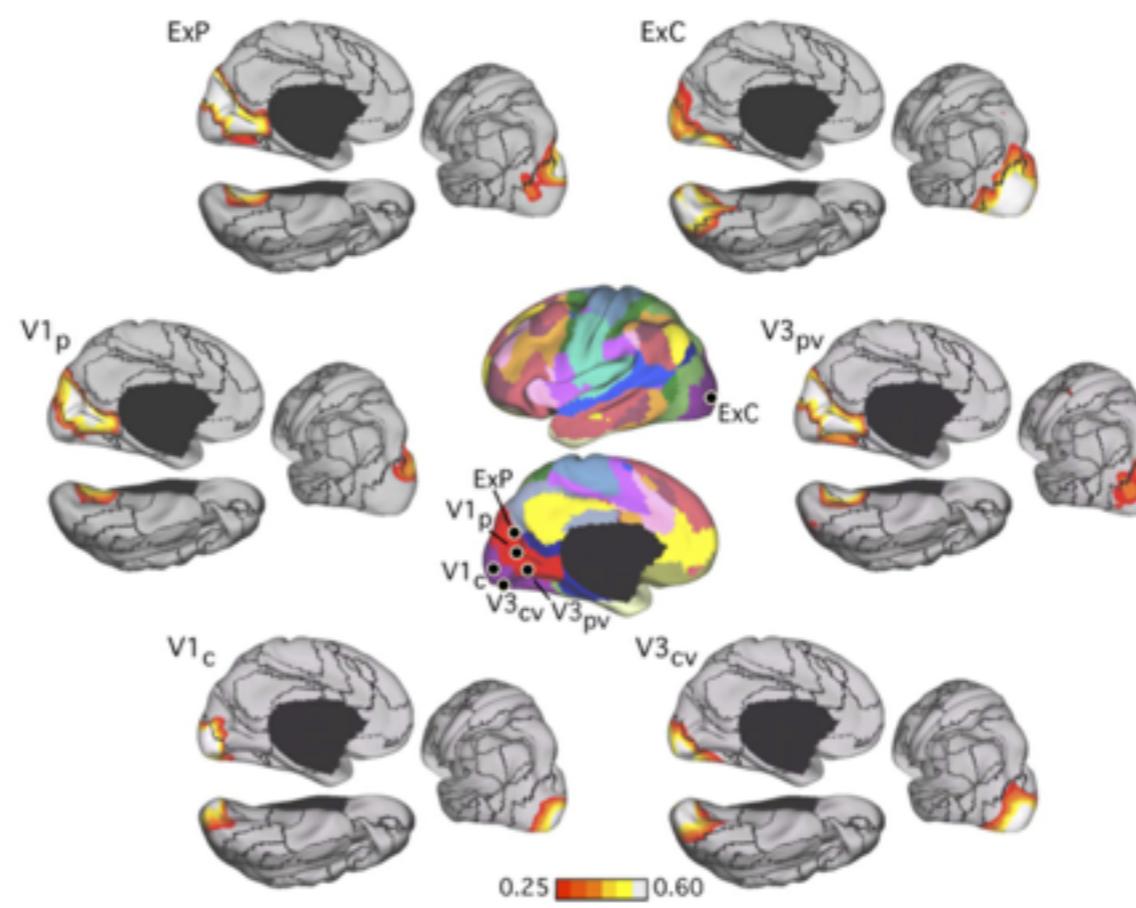
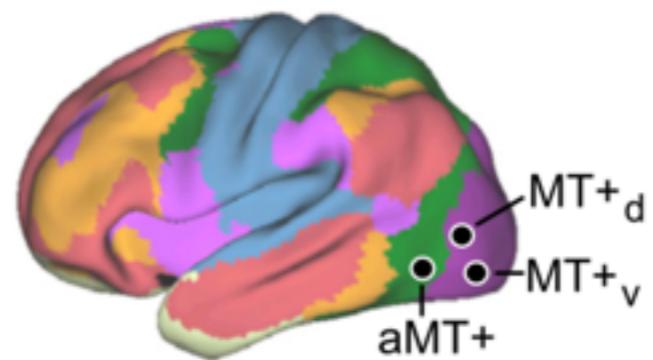
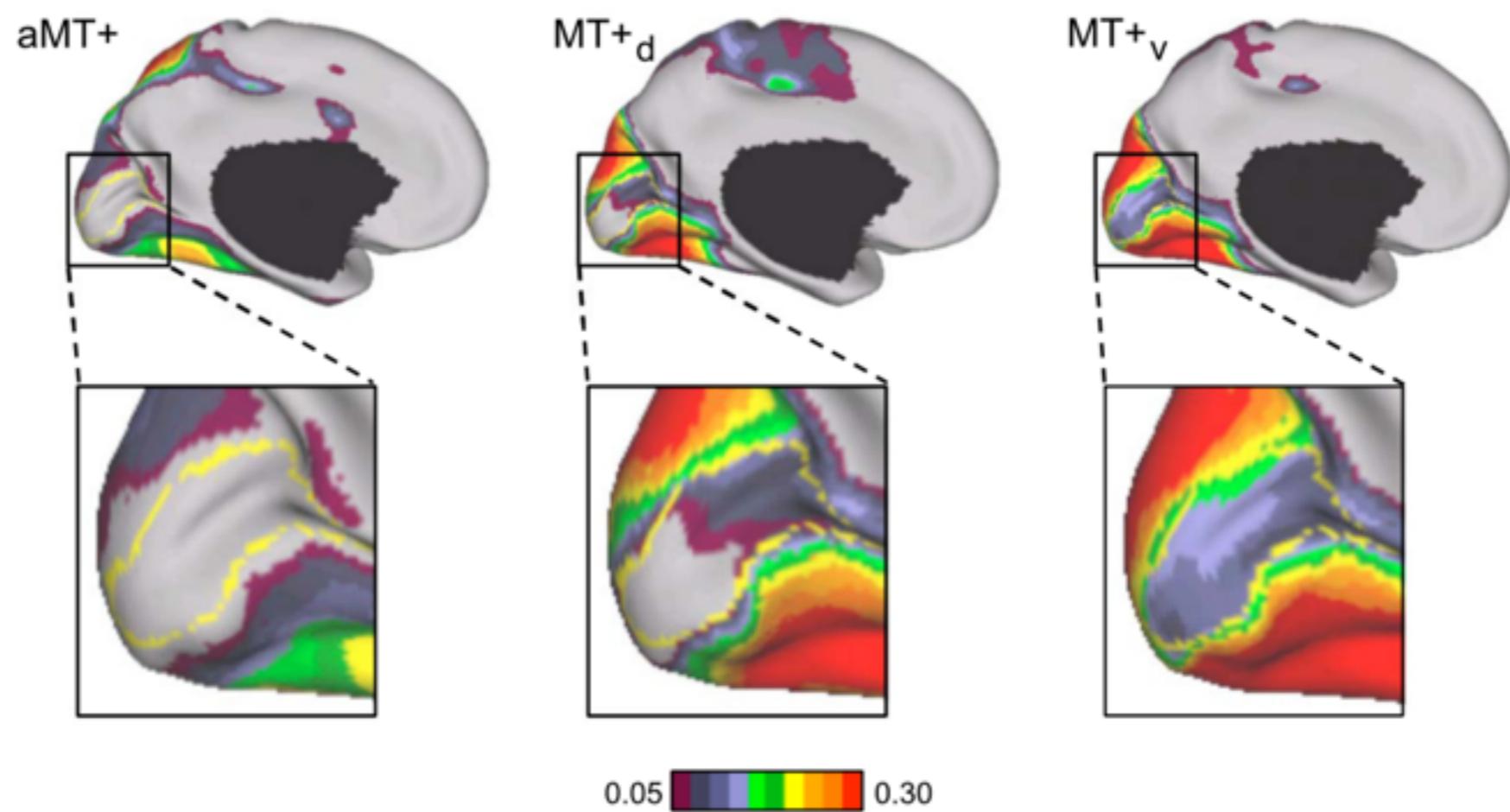


Fig. 16. Evidence that the fractionation of the visual system reflects functional connectivity MRI (fcMRI) topography within the visual cortex. Six left hemisphere seed regions were picked from the discovery dataset: V1<sub>c</sub> and V1<sub>p</sub> correspond to central and peripheral visual field representation within V1, respectively; V3<sub>cv</sub> and V3<sub>pv</sub> correspond to central and peripheral V3v, respectively; ExC and ExP correspond to 2 seed regions within the extrastriate visual cortex in the estimated locations of the central and peripheral visual fields (purple and bright red at center). The 6 seed regions are illustrated at center, and their coordinate locations are reported in Table 1. Their left hemisphere fcMRI maps were computed using the replication data set and arranged around the center images. Note that the central visual seed regions are selectively correlated with the central visual representation, whereas the peripheral visual seed regions are selectively correlated with the peripheral visual representation.



**Fig. 23.** Functional connectivity maps of MT+ reveal topographic organization. Functional connectivity maps of aMT+, MT+<sub>v</sub>, and MT+<sub>d</sub> were computed using the replication data set and are shown with views focusing on V1. V1 shows little or no correlation to aMT+ but strong correlations with both MT+ seeds. In both MT+ fcMRI maps, there is stronger correlation with dorsal V1 (lower visual field) than ventral V1 (upper visual field). There is also increasing correlation with central V1 as we proceed from MT+<sub>d</sub> to MT+<sub>v</sub>. The yellow line denotes the areal boundary of V1.



## SOMATOMOTOR

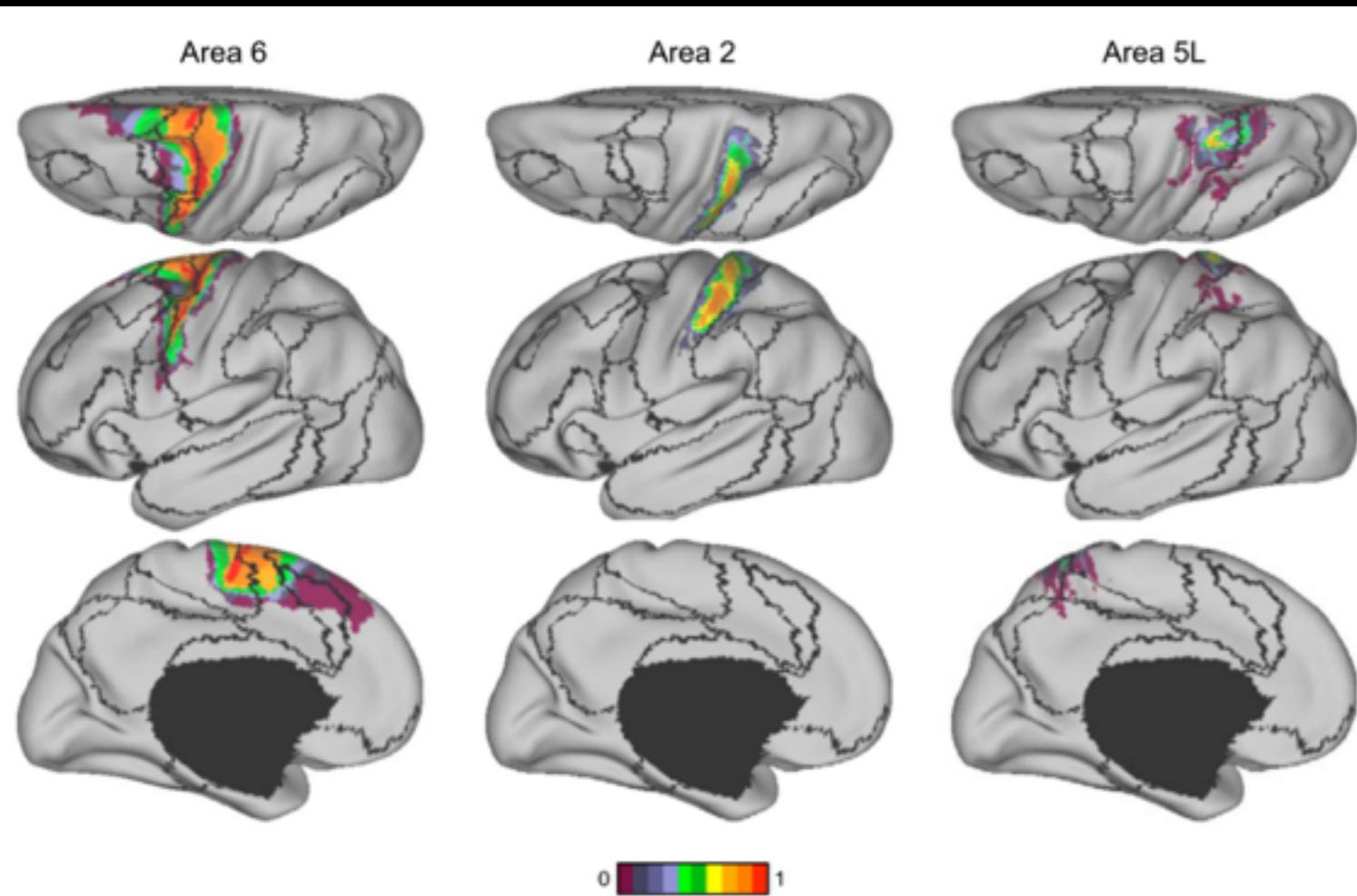
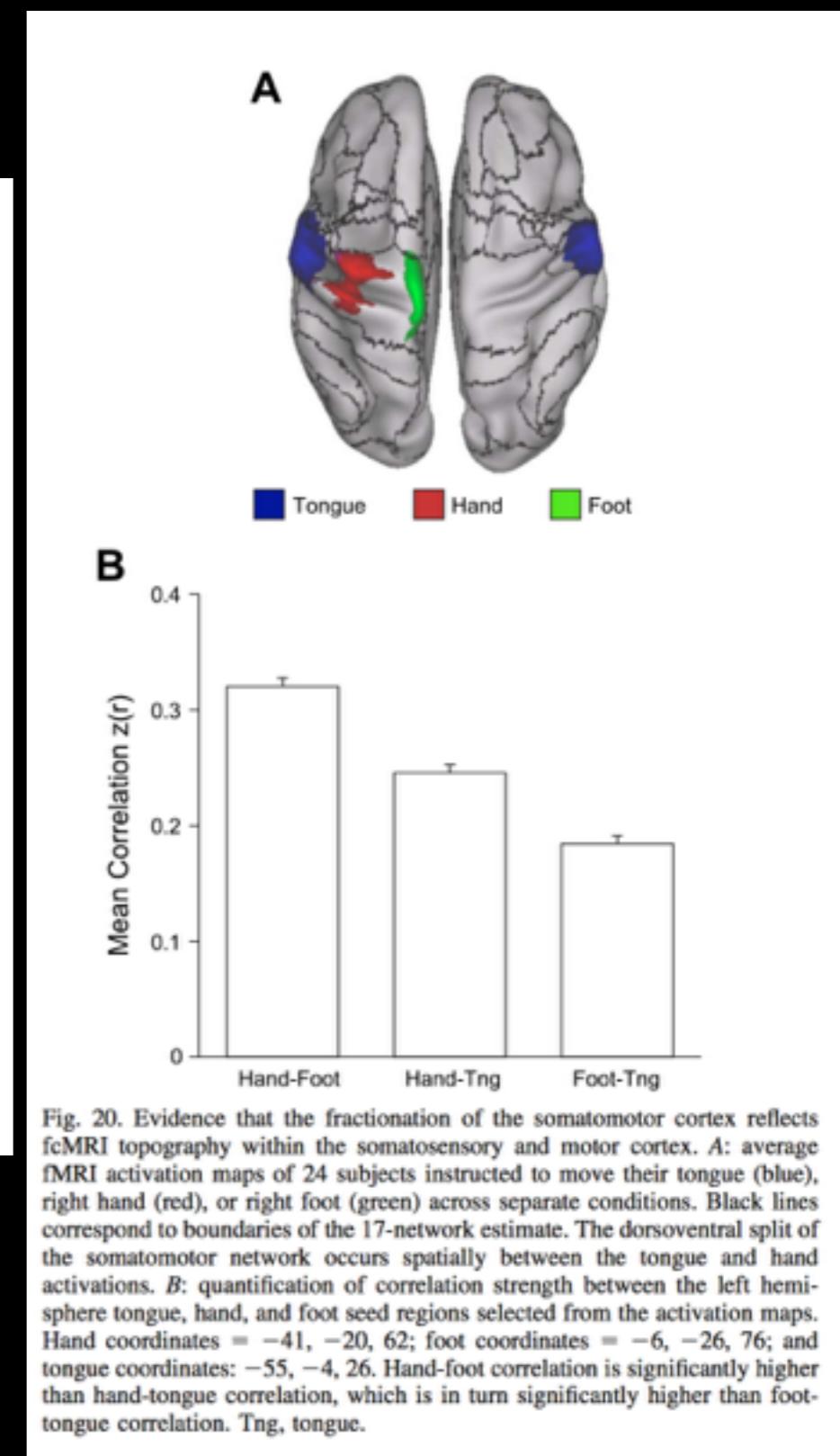


Fig. 19. Seven-network boundaries on probabilistic maps of areas 6, 2, and 5L. Boundaries of 7-network estimate based on the full sample of 1,000 subjects are overlaid on the surface-based probabilistic histological maps of areas 6, 2, and 5L. The somatomotor network includes most, if not all, of areas 2 and 5L, but only the caudal half of area 6.



## HIGHER ORDER SYSTEMS

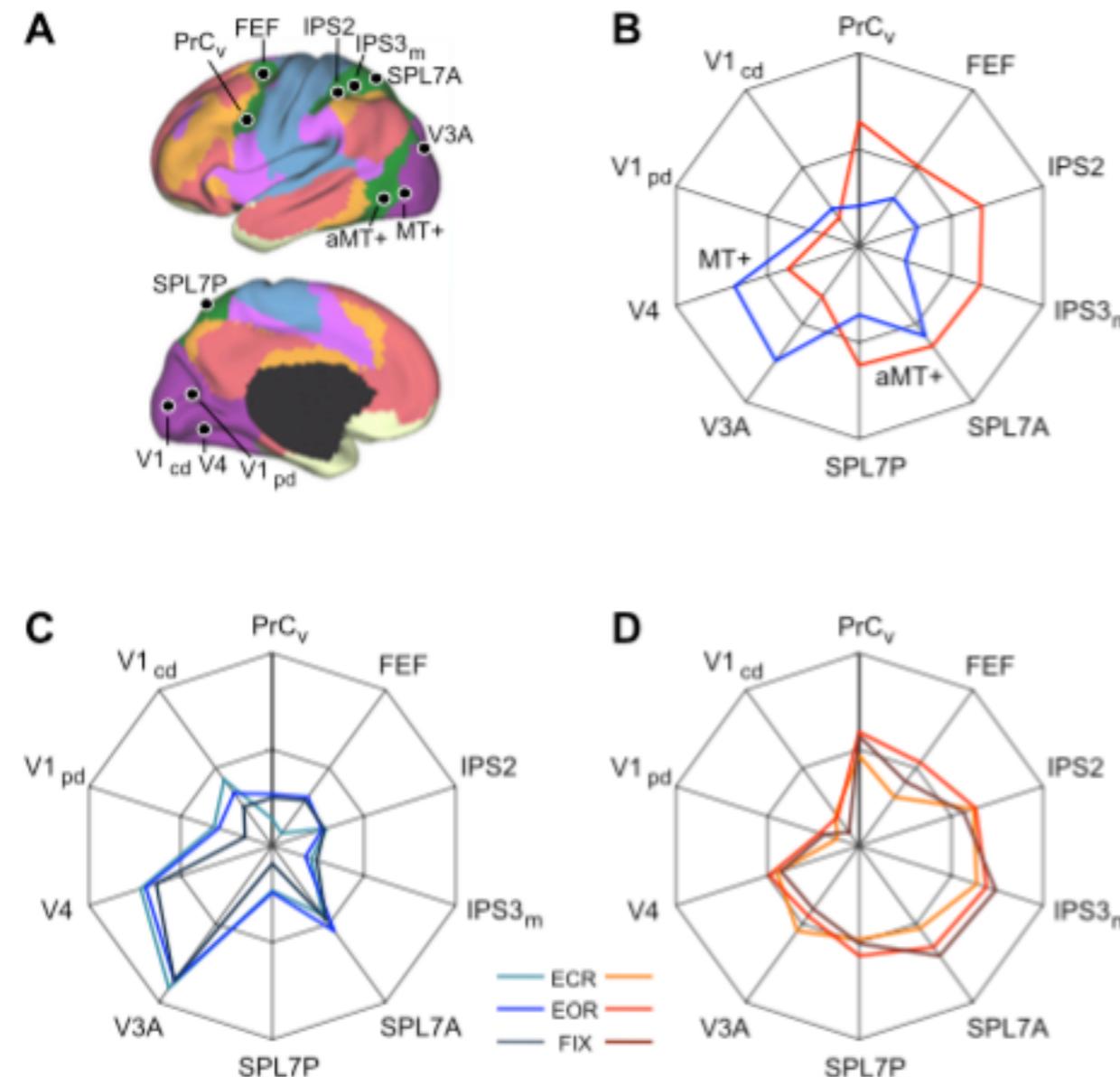
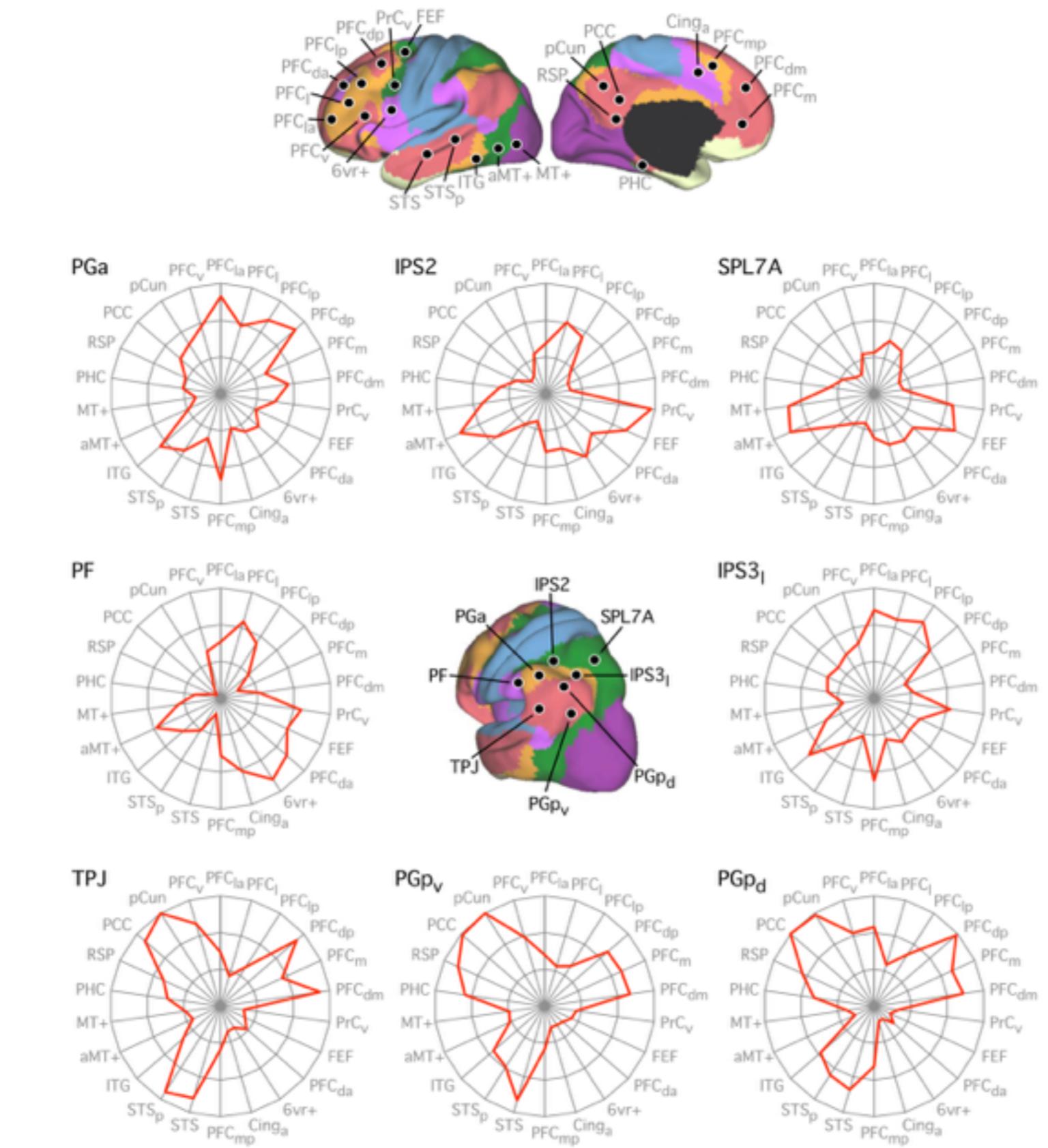


Fig. 25. aMT+ and MT+ functional connectivity patterns generalize across task conditions. *A*: 4 visual, 4 parietal, and 2 frontal seed regions were used to quantify the functional coupling of aMT+ and MT+ to distributed cortical regions. Coordinate locations of the seed regions are reported in Table 2 and were chosen using either the discovery data set or meta-analysis of fMRI studies (Table 3). *B*: polar plots of MT+ (blue) and aMT+ (red) connectivity with the visual, parietal, and frontal seed regions were computed using the replication data set. MT+ is more strongly correlated with visual cortex compared with parietal and frontal cortices. The converse is true for aMT+. *C* and *D*: polar plots of MT+ (blue) and aMT+ (red) connectivity replicated in the task effects data set demonstrate that the functional coupling differences generalize across multiple data acquisition conditions. The polar scales range from  $r = -0.1$  (center) to  $r = 0.6$  (outer boundary) in 0.35-step increments.



**Fig. 30.** Adjacent parietal regions exhibit distinct functional connectivity fingerprints. Correlations of 8 parietal seed regions (center) with 22 cortical target regions (top) from the replication data set, displayed as polar plots. Colors represent the 7-network segmentation (from Fig. 11). The coordinate locations are reported in Table 4. Parietal seed regions that belong to the same network (e.g., TPJ, PGp<sub>v</sub>, and PGp<sub>d</sub>) have generally similar functional connectivity fingerprints that are distinct from other parietal seed regions. Close inspection of the polar plots reveals distinct connectivity fingerprints even for parietal regions within the same network, some of which anticipate the further fractionation of the parietal cortex in the 17-network estimate (Fig. 13). Note that the cortical targets from anterior cingulate cortex to pCun on the left side of the polar plots are the same as that of the frontal polar plots (see Fig. 31) to allow for comparison across the 2 sets of polar plots. The remaining cortical targets are different across the 2 sets of polar plots but are arranged so that cortical targets at the same location in the polar plots belong to the same network in the 7-network estimate. The polar scales range from  $r = -0.4$  (center) to  $r = 0.5$  (outer boundary) in 0.3-step increments.

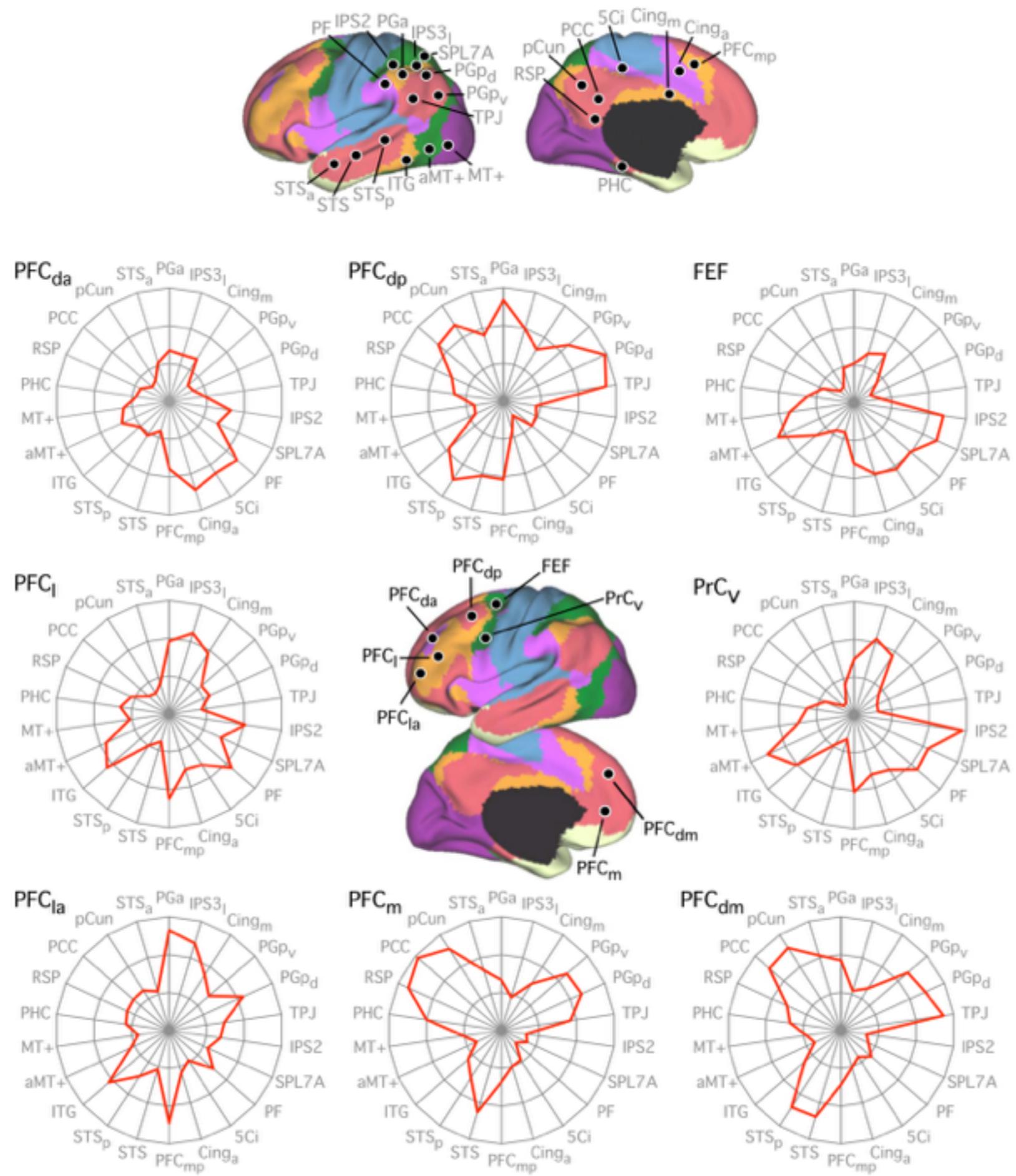


Fig. 31. Adjacent frontal regions exhibit distinct functional connectivity fingerprints. The format and plotting are the same as for Fig. 30 with regions tailored for exploration of frontal cortex. The coordinate locations are reported in Table 4. The polar scales range from  $r = -0.4$  (center) to  $r = 0.5$  (outer boundary) in 0.3-step increments.