# How to Perform a Co-activation Based Parcellation.

# Table of Contents

# 1. Creating a seed VOI

To search the BrainMap database, you will need a seed volume of interest (VOI). This VOI must be saved in ~/MACM7/VOIs. While the final VOI must be in MNI152 space, there are many options for creating this seed:

a. Native patient space: create the seed in native patient space (manually, via FIRST, via FreeSurfer) and then transform this to MNI152 space (SPM, FSL…). When using freesurfer, be sure to use the current volume representation of your anatomical scan for spatial normalization.
b. FSL-bundled labels: use atlases contained within the FSL software distribution (MNI152, Jülich Cytoarchitecture Talairach, Harvard-Oxford, etc) to create a tissue-specific ROI.
c. JuBrain Cytoarchitectonic Atlas: Be sure to output your seeds in MNI (rather than anatomical MNI space). In version 1.x of the Anatomy Toolbox, data is in MNI single subject space, make sure this matches reasonably with MNI153
d. Free-Surfer label: create FreeSurfer cortical segmentations (recon-all) of the MNI152 atlas. This atlas is distributed in the FSL software bundle. Recall that because FreeSurfer performs segmentation in FreeSurfer space, these labels must be converted from FreeSurfer space back into the original MNI152 space before being used in CBP.
e. Output from another analysis: output from a coordinate-based meta-analysis (or any study, for that matter) can be used as a VOI. Make sure the final VOI is in MNI (ideally MNI152) space.

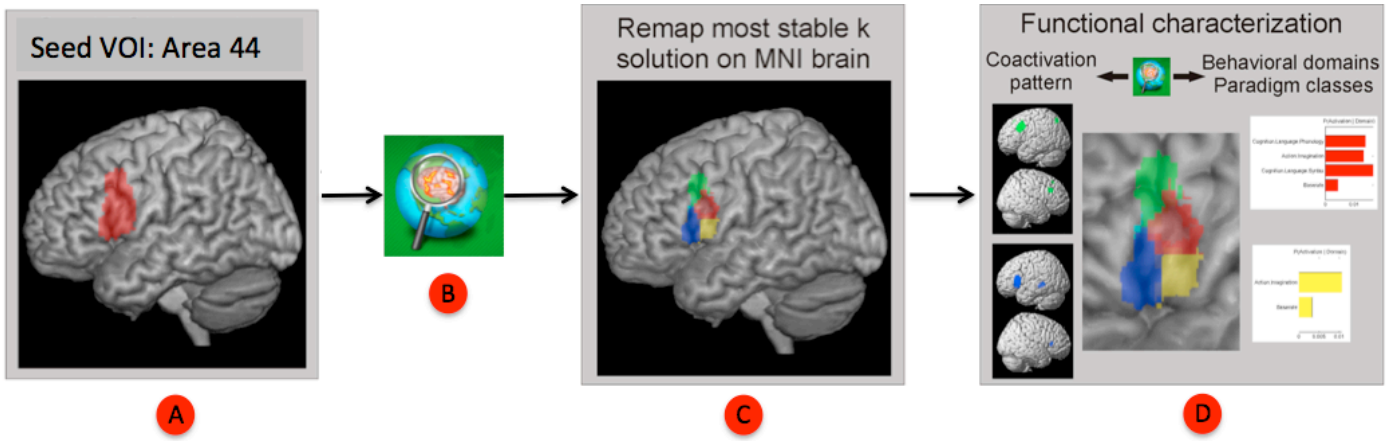



Figure 1. Overview of CBP Pipeline. The co-activation profile of a defined VOI (A) is calculated by seeding this VOI into the BrainMap database (B). The seed VOI is then parcellated using a non-hierarchical clustering, producing clusters with unique co-activation profiles (C). Each cluster can then be functionally characterized by creating a meta-analytic connectivity model for that cluster (D).

## 2. Set Up MatLab Environment

Within MatLab, 3 things must be properly set up for CBP to work:

1. "Current Folder" must be set to "~/MACM7," (or whatever version is used)
2. Statistics Toolbox is purchased (from MatLab) and installed
3. SPM8 is installed and properly pathwayed within the MatLab environment.
4. If this is the case, type "se_SpecifyCBP" in the MatLab workspace and the SPM8 project window shown in Figure 2 will pop-up.

At this point, the operations are performed sequentially by clicking each tab and specifying certain settings, described below.

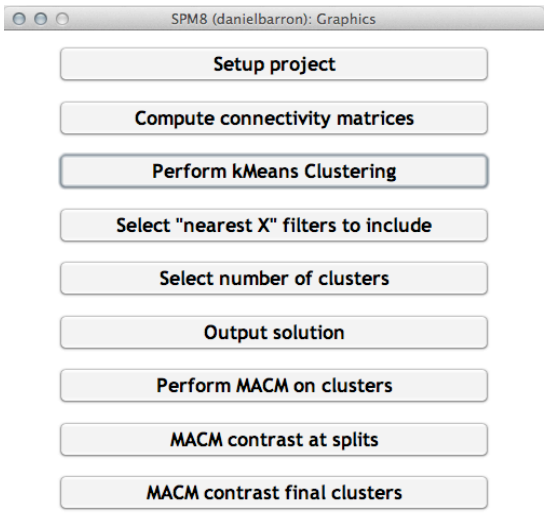



Figure 2. Graphic User Interface for CBP procedure. Steps are performed sequentially, top→bottom.

## 3. Setup Project

Navigate to ~/MAM7/VOIs and select yourvoi.nii. This will create yourvoi_CBP.mat file that will be use in subsequent processing. This .mat file will be stored in ~/MACM7/CBP and needs to be selected as the first input for the step 4-8.

## 4. Compute connectivity matrices

Navigate to ~/MACM7/CBP and select yourvoi_CBP.mat Once selected, the SPM8 command window will pop up. Select the following

**a.** Exclude VOI as target: Yes/No: Yes only considers connectivity with the rest of the brain, no also uses the local connectivity information. "No" is hence more biased towards continuous clusters, Large VOIs, where within-VOI (but not necessarily local) connectivity may be relevant "No" should be selected.

**b.** Median-filter VOI: Yes/No: 'Yes' should be chosen when VOI is very frayed to smooth the borders of VOI. Obviously, however, the VOI does not exactly match your original input any more afterwards

**c.** Use GM Mask: Yes/No: Uses the ICBM grey matter template to include only those voxels where the probability for GM is > 10%. May make the VOI more specific but at the danger of loosing meaningful voxels. Not recommended for subcortical VOIs, where GM probabilities are heterogeneous.

**d.** Specify 'nearest X:' 20:5:200 (this preset value will select the closest 20, 25, 30,…, 200 experiments for each voxel ). Tighter spacing samples the distribution-structure better but

takes longer. For smaller regions, settings like 20:2:100 may be useful, however, as the largest filters will be too coarse here.

This computes a co-activation map for each voxel within yourvoi.nii to each target voxel based on a portion of the BrainMap database. This portion of BrainMap includes only fMRI and PET experiments reporting foci in healthy adults; i.e. all experiments involving pathological populations or children were excluded. The brain-wide co-activation pattern for each voxel in yourvoi.nii is computed by meta-analyzing (via ALE) the BrainMap experiments containing co-activation. Because there is usually a variable and low number of foci related to a particular voxel in yourvoi.nii, this procedure is performed at different degrees of association (the filter sizes specified in 'nearest X') for each voxel within yourvoi.nii.

The brain-wide co-activation pattern for each voxel in yourvoi.nii is then computed by a meta-analysis over the experiments associated with that voxel at a given filter size. The preset nearest X value above thus yields 37 co-activation profiles (corresponding to analyses based on the 20, 25, 30… nearest experiments) for each seed voxel which are recorded in 37 $N_S$ x $N_T$ co-activation matrices where $N_S$ denotes the number of seed voxels and $N_T$ the number of target voxels in the reference brain. These co-activation matrices can be found under ~/MACM7/CBP/yourvoi/DATA/yourvoi_TopX.mat. That is, the co-activation matrix based on the 20 closest experiments for each seed voxel of yourvoi is recorded in yourvoi_Top20.mat.

Note that the mean distance [in mm] between each seed voxel and the included activation foci for a given filter size can be found under ~/MACM7/CBP/yourvoi.mat in the MeanD variable.

This analysis will take at least 30 min (potentially longer) to complete, depending on the size of yourvoi.nii as well as the selected spacing and range of the filters.

## 5. Perform K-Means Clustering:

Navigate to ~MAM7/CBP and select yourvoi_CBP.mat Once selected, the same SPM8 command window will pop up with new options. Enter the following:

a. Max # of clusters: default is 9, this will create 8 parcellations of the seed region into 2, 3, …, 9 clusters). More is rarely necessary. If your seed region is rather small, 7 may be sufficient. Less is not recommended.
b. Replicates: e.g., 10. This will repeat the K-means clustering X times using randomly selected initial cluster centers, with the best-fitting version being retained. Probably it would be better to go for 50, 100 or even more. But this takes much longer.

This will now assign each seed voxel of yourvoi.nii to a one of K clusters based on its connectivity profile to every other brain voxel as determined by the BrainMap database using K-means with K varying from 2 to 9. K-means clustering is a non-hierarchical clustering method that uses an iterative algorithm to separate the seed region into the previously selected number of K non-overlapping clusters. K-means aims at minimizing the variance within clusters (that is, the co-activation profiles of voxels assigned to the same cluster should be as similar as possible) and

maximizing the variance between clusters (that is, the co-activation profiles of voxels assigned to different clusters should be as dissimilar as possible) by first computing the centroid of each cluster and subsequently reassigning voxels to the clusters such that their difference from the centroid is minimal. Note that the initial centroids are chosen at random for each new replication (so more replications provide a better confidence in finding the "true" optimal solution). As default, one minus the correlation between the connectivity patterns of the individual seed voxels is used as the distance measure (correlation distance). For our previous input of nearest X = 20:5:200 and Max # of clusters = 9, this parcellation will result in 8 (K number of clusters) x 37 (filter size) independent cluster solutions. For each of the 8x37 parcellations the best solutions from the (per default 10) replications with randomly placed initial centroids are recorded under ~/MACM7/CBP/yourvoi/DATA/Correspondenz_K.mat in the variable 'T'.

This analysis will take at least 3 hrs to complete (potentially days, weeks), depending on the size of yourvoi.nii, your choice of the maximum number of clusters K and the number of replications – and of course your computer.

## 6. Select 'nearest X' filters to include

The previous 'nearest X' filter and K-means cluster steps calculated parcellations for a wide range of filter and cluster values. In the next two steps, we must decide which values are most appropriate for the current dataset, that is, the 37x8 parcellations should be combined into a single 'best' cluster solution. In a first step, the optimal filter range is determined. Rather than averaging across all filter sizes, the following process is used to determine which filter range is most stable.

Selecting this tab will pull up a new SPM8 window similar to Figure 3.

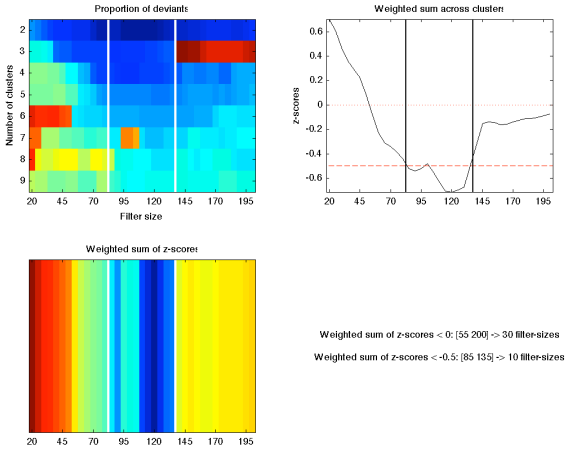



Figure 3. Selection of 'nearest X' filter.

These graphs represent how stable a particular filter size (20:5:200, chosen in Step 4) is at a given K-means cluster number (2-9 clusters, chosen in Step 5) based on the consistency of the cluster assignment for the individual voxels across the different filter sizes. The red⇔blue spectrum represents less⇔more stable results, i.e. less⇔more voxels that were assigned differently as compared to the solution from the majority of filters (deviants). When these results are displayed another SPM8 window will pop up with a suggested filter size. The present graph shows 85-135 (the white line), which was the most stable solution for *this particular analysis.*

This is based on the normalized z-scores of the number of deviants, with higher weighting given to the smaller K-solutions (graph on the top right). The analysis for yourvoi.nii might be different so consult the suggested solution for your particular analysis. Usually the suggested values work fine, but be sure to inspect your graphs. Enter these selected values and proceed to K-means cluster selection.

In all subsequent steps the analysis is then restricted to K parcellations based on co-activation computed for the nearest (in this particular case) 85 to 135 experiments (note that this filter range information is recorded in ~/MACM7/CBP/yourvoi_CBP.mat in the second row of the 'box' variable). That is, once you selected your filter-size, all other filter sizes are discarded from further analysis.

## 7. Select number of clusters

Similar to selecting the filter size, you must now select the most stable number of K-means clusters for your dataset. There are 8 total criteria that can help you determine which cluster solution is optimal. These criteria are based on topological, information-theoretic and cluster separation characteristics of each cluster solution.

Based on the sum of scores obtained across all criteria, the cluster solution with the maximum score will automatically be suggested as the most stable solution. However, you should NOT simply accept this prompt!

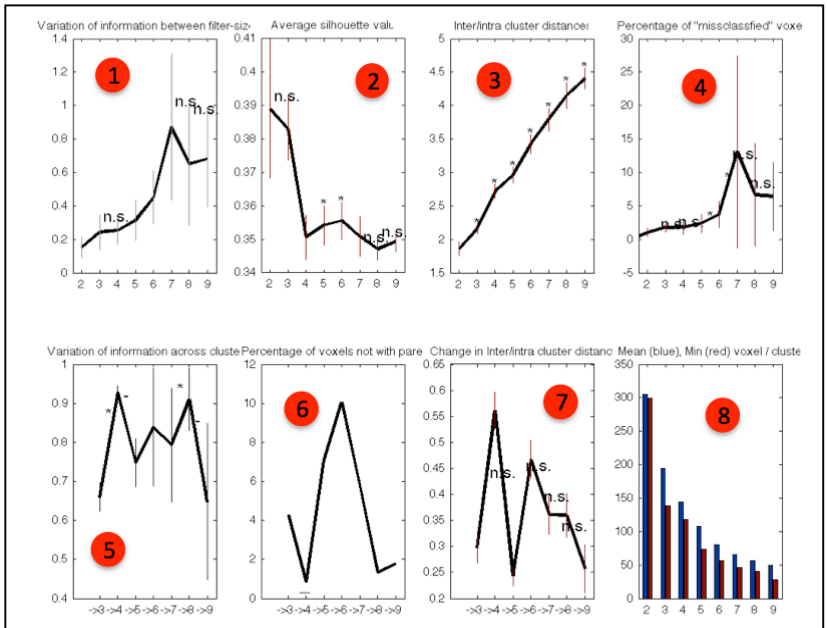



Figure 4. Data used to determine an appropriate cluster solution.

Rather, it is usually worthwhile to take a closer look at the performance of the cluster solutions with regard to the different criteria. If you can support a particular cluster solution with one of each of these (topological, information theoretic and cluster separation characteristics), this is considered a very strong argument for selecting that cluster solution. These criteria are explained below:

a. Topological criteria:
**Number 4**, Figure 4 reports the percentage of misclassified voxels (deviants). This criterion indirectly reflects the amount of noise and potentially local effects in the clustering. In particular, the criterion addresses the across-filter stability, that is, the average percentage of voxels for each filter-size that were assigned to a different cluster compared to the most

frequent (mode) assignment of these voxels across all filter sizes. *Good solutions* are considered those K parcellations where percentages of deviants (presumably reflecting noise and local variance) are not significantly increased compared to the K-1 solution, in particular if the subsequent K+1 solution leads to a significant higher percentage of deviants (this is for example the case for the 5-cluster solution in Figure 4).

**Number 6**, Figure 4 reports the percentage of voxels not related to the dominant parent cluster compared to the *K*-1 solution. This measure is related to the hierarchy-index (Kahnt et al., 2012) and corresponds to the percentage voxels that are not present in hierarchy, K, compared to the previous K-1 solution. That is, voxels assigned e.g. to cluster 3 in the $K = 3$ solution stemming from a subset of voxels previously assigned to cluster 2 (in the $K = 2$ solution) would be excluded if the majority of cluster 3 voxels actually stemmed from cluster 1 (in the $K = 2$ solution*). Good solutions* for a given K cluster parcellation were those wherein the percentage of lost voxels was below the median across all possible solutions (cluster parcellations 2-9), where the respective clustering step resulted in a local minimum and/or the following clustering-step featured a maximum in the percentage of lost (hierarchically inconsistent) voxels. For example, Figure 4 shows that the 4-cluster solution is a good solution because (1) the percentage of lost voxels in the 4-cluster solution is lower than the median value (across all K solutions), and (2) the 4-cluster solution comes before the cluster 5 and 6 solutions which report the maximum of lost voxels.

**Number 8,** Figure 4 shows the third topological criterion. This concerns the number of consistent voxels per cluster, i.e. the sizes of the individual cluster after removal of hierarchically inconsistent voxel (previous criterion). *K* parcellations are evaluated by considering the proportion of the minimum cluster size (red) to the mean cluster size (blue) provided by a given *K* solution. *Good solutions* may be considered to be those where the size of the minimum cluster size was more than half of the average cluster size within the K solution. For example, in Figure 4 all 8 cluster solutions satisfied this condition as the ratio between the minimum cluster size (red bar) and the average cluster size (blue bar) was more than 0.5 for all 2 to 9 K solutions. In particular, however, solutions in which the smallest cluster becomes zero would have been disregarded, as these indicate that at least one cluster did not contain any hierarchically consistent voxel any more.

b. Information-theoretic criteria:

**Number 1**, Figure 4 shows the variation of information between the filter sizes you selected in step 6. The variation of information (VI) metric has been previously used to determine the optimal *K*-means parcellation of a given brain region by Kelly et al. (2010) and by Kahnt et al. (2012). This VI metric was computed for each *K* solution between all (unique) combinations of the eleven (85, 90, 95, … 135, cf. above) filter sizes selected in step 6. *Good/stable solutions* should not show an increase in VI between filter sizes compared to the previous K-1 solution and/or show a significant increase for the K+1 solution. For example, the 4-cluster solution is a good/stable solution because it does not show a significant increase in VI from the 3- cluster solution.

**Number 5**, Figure 4 shows the VI metric between a solution, K, and subsequent neighbouring solutions, K+1. *Good/stable solutions* show a significant increase in VI between

the subsequent set of solutions (primary criterion) or a significant decrease from the previous to the current clustering step (secondary criterion). For example, the 3-cluster and the 7-cluster solution satisfy the primary criterion and the 5-cluster solution satisfies the secondary criterion in Figure 4. In general, however, this is a rather weak criterion.

| Criteria | A good solution (K)… |
|---|---|
| 1 | Should not show an increase in VI between filter sizes compared to the previous K-1 solution and/or show a significant increase for the K+1 solution. |
| 2 | Should show a significantly higher silhouette value compared to the K-1 solution (primary criterion) or whose silhouette value is at least not significantly decreased compared to the previous K-1 solution (secondary criterion). |
| 3 | Does not have a significantly larger subsequent (K+1) increase in intercluster to intracluster distance. |
| 4 | Is one where percentages of deviants (presumably reflecting noise and local variance) is not significantly increased compared to the K-1 solution, in particular if the subsequent K+1 solution leads to a significant higher percentage of deviants. |
| 5 | Shows a significant increase in VI between the subsequent (K+1) set of solutions (primary criterion) or a significant decrease from the previous to the current clustering step (secondary criterion). |
| 6 | Shows a percentage of lost voxels that is below the median across all possible solutions (cluster parcellations 2-9), where the respective clustering step resulted in a local minimum and/or the following clustering-step featured a maximum in the percentage of lost (hierarchically inconsistent) voxels. |
| 7 | 7 is the first derivative of 3, information from criteria 3 also applies. |
| 8 | Is where the size of the minimum cluster size was more than half of the average cluster size within the K solution. |

c. Cluster Separation Criteria:

**Number 3 and 7,** Figure 4 report information on cluster separation. This is computed as the intercluster to intracluster distance ratio (Chang et al. 2012) for the filter size you selected in Step 6. That is, the ratio between the average distance of a voxel to its cluster centre and the average distance between the cluster centres. A significant increased ratio compared to the *K*-1 solution would indicate a better separation of the obtained clusters. For example, all cluster solutions in Figure 4 (except for the 2-cluster solution which cannot be compared to a K-1 solution) are good solutions with regard to criterion 3. However, because of the monotonous increase usually observed with this ratio, criterion 7 shows the first derivative of criterion 3 which can be used to evaluate the change in this ratio to the previous *K*-1 solution. *Good solutions* are those where the subsequent K+1 solution does NOT show a significantly larger increase in intercluster to intracluster distance. For example, the 4-cluster solution is considered a good solution because moving from the 4- to the 5-cluster solution results in a

comparable or even lower change in this ratio.

**Number 2**, Figure 4 reports the silhouette value averaged across voxels for your selected filter size. The silhouette value is a measure of how similar that voxel is to voxels in its own cluster compared to voxels in other clusters. This value ranges from -1 to +1. *Good solutions* are those with a significantly higher silhouette value compared to the K-1 solution (primary criterion) or whose silhouette value is at least not significantly decreased compared to the previous K-1 solution (secondary criterion). For example, the 5-cluster solution satisfied the primary criterion because its silhouette value is significantly higher compared to the 4-cluster solution. The 3-cluster solution satisfies the secondary criterion because its silhouette value is not significantly lower compared to the 2-cluster solution.

Once the best cluster solution is chosen and entered in the prompt, this information is recorded in ~/MACM7/CBP/yourvoi_CBP.mat in the first row of the 'box' variable and you can proceed with the 'Output Solution' step.

## 8. Output Solution

Having selected the filter size and number of clusters most justified by the data, the final solution is extracted from the .mat file. After clicking "Output Solution" (see Figure 2), select ~/MACM7/CBP/yourvoi_CBP.mat to compute this final solution, which will be saved as a separate folder within ~/MACM7/CBP/yourvoi/yourvoi_4Clusters_85-135 (assuming you chose a 4 cluster solution with filter size 85-135).

Within ~/MACM7/CBP/yourvoi/yourvoi_4Clusters_85-135 (or whatever reflects the selected filter sizes and cluster-number) there are a number of files. Figure 5 shows Clustering_Results.png. This figure shows how the data were clustered and how well the clusters are separated. The left upper part of Figure 5 shows the visualization of the 4-cluster solution by multidimensional scaling (MDS). MDS allows the visualization of signals residing in an N-dimensional 'functional space' in 2D. Higher proximity between points (voxels) indicates more similar co-activation patterns of these voxels. For example, it can be seen that the red and

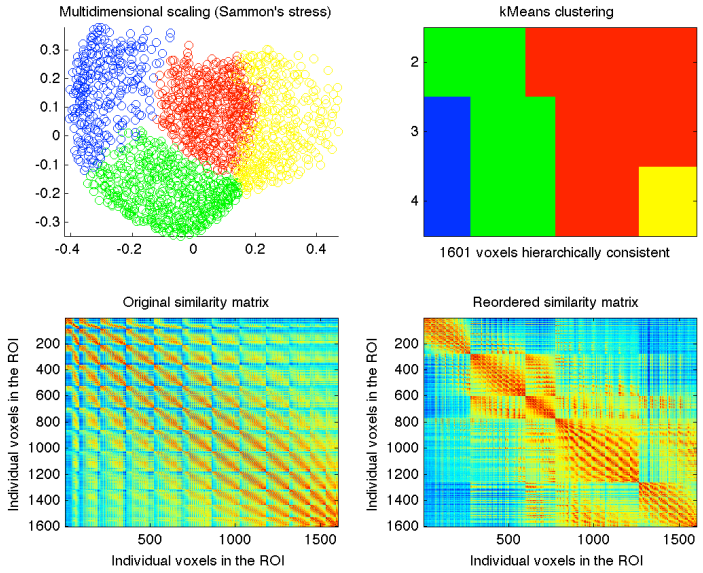



Figure 5. Clustering_Results.png

the yellow cluster are more similar to each other than to the green and blue clusters, as the points (voxels) of the red and yellow cluster are relatively close together and more clearly separate from the green and blue cluster points. The right upper part shows the pattern of cluster assignment and splitting of clusters across levels of K (pseudo-hierarchical clustering, note that K-means is per se NON hierarchical). It basically shows you the 2- and 3-cluster solutions and their relationship to the 4-cluster solution. For example, you can see here that the blue cluster stems from a subset of the green cluster and the yellow cluster stems from a subset of the red cluster splitting off at K = 3 and K = 4, respectively. Note that the pattern of separation observed in MDS corresponds well to this hierarchical splitting pattern, as the red and yellow cluster which looked most similar in MDS are also splitting up last. The left lower part of Figure 5 is the similarity matrix of the seed voxels (the $N_S$ x $N_S$ distance matrix) and the right lower part represents the similarity matrix of the seed voxels reordered according to the splitting scheme derived from K-means clustering illustrated above.

Another folder, 'Clusters,' was created in ~/MACM7/CBP/yourvoi/yourvoi_4Clusters_85-135. It contains the nifti files of the each cluster which can be used to visualize the location of the cluster using brain viewers such as mricron, spm, fslview or Mango. Mrparcellation.nii contains all clusters in a single nifti-file, with the values for each voxel reflecting the cluster-number it is assigned to. This can be useful for visualization as well. The locations of the clusters are also illustrated in the .png files located in the folder 'location' contained in the 'Clusters' folder. The summary .png, gives an overview of all clusters arranged in a rostral to caudal sequence.

## 9. Perform MACM on clusters

The purpose of this step is to characterize each CBP-derived cluster based on their co-activation profile. In this context, "clusters" refer to sets of voxels within the seed region that were identified by the co-activation-based parcellation outlined above as having similar co-activation patterns to each other but distinct ones to the rest of the seed voxels. Note, that this analysis is done at the selected level of K, in the current example K=4. The co-activation profiles of the different clusters are obtained by first identifying all experiments in the BrainMap database that feature at least one focus of activation in a particular CBP-derived cluster. Next, an ALE meta-analysis is performed on these experiments. In contrast to the MACM underlying the co-activation-based parcellation, where ALE maps were not thresholded to retain the complete pattern of co-activation likelihoods, statistical inference is now performed.

To establish which regions were significantly co-activated with a given cluster, ALE scores for the MACM analysis of this cluster are compared to a null-distribution reflecting a random spatial association between experiments with a fixed within-experiment distribution of foci (Eickhoff et al. 2009, 2012). This random-effects inference assesses above-chance convergence between experiments, not clustering of foci within a particular experiment. The observed ALE

scores from the actual meta-analysis of experiments activating within a particular cluster were then tested against the ALE scores obtained under this null-distribution yielding a P-value based on the proportion of equal or higher random values. The resulting non-parametric P-values were transformed into Z-scores and thresholded at FWE-corrected threshold of $P < 0.05$. The output of this post-hoc MACM analysis can be found under ~/MACM7/MACM/yourvoi_4Clusters_85-135. The most important results for now are found in the folder 'Results' and 'ResultsFWE' where the nifti files of the significant co-activations (FWE-corrected at cluster level ('Results') and voxel level ('ResultsFWE)) for each cluster are saved. For example, the file 'Cluster1_cFWE05_001_279.nii' would contain the co-activations of cluster 1 significant at 0.05 (FWE-corrected at cluster level, cluster height threshold at 0.001 – the last number 279 indicates the size of the smallest cluster that passed significance thresholding) activating within this cluster. Corresponding illustrations of these co-activations can be found under 'Images' and 'ImagesFWE'.

Secondly, this step also performs the functional characterization of the CBP-derived clusters in terms of behavioral domains (BD) and paradigm classes (PC) using forward and reverse inference. In the forward inference approach, a cluster's functional profile is determined by identifying taxonomic labels, for which the probability of finding activation in the respective cluster was significantly higher than the overall chance (across the entire database) of finding activation in that particular cluster. Significance is established using a binomial test. That is, we test whether the conditional probability of activation given a particular label [P(Activation|Task)] is higher than the baseline probability of activating the region in question per se [P(Activation)]. In the reverse inference approach, a cluster's functional profile was determined by identifying the most likely behavioral domains and paradigm classes given activation in a particular cluster. This likelihood P(Task|Activation) can be derived from P(Activation|Task) as well as P(Task) and P(Activation) using Bayes rule. Significance was then assessed by means of a chi-squared test. The output of this analysis can be found in the 'BDPC' folder which contains png files for each cluster displaying the results. For example, 'Cluster1_FDR05.png' shows the forward and reverse inference for cluster 1 significant at 0.05 (FDR-corrected) and 'Cluster1_uc05.png' shows the forward and reverse inference for cluster 1 significant at 0.05 (uncorrected). Associated ps files of this analysis which can be processed with image processing software such as Corel Draw can be found in the folder 'BDPCprint' (under ~/MACM7/MACM/yourvoi_4Clusters_85-135).

Details on the values displayed in these png files can be found in the corresponding text files 'Cluster1_BD' and 'Cluster1_PC. These also list the number of experiments that activate each cluster.

## 10. MACM at splits

The purpose of this step is to assess the differences and similarities in co-activation patterns as well as in BDPC between *splitting* clusters. Thus, this step computes contrasts and conjunctions between the newly emerged child cluster and its remaining parent cluster at the same level of

K (cf. Figure 5 the right upper pattern of hierarchical cluster splitting). In the current example, we would thus compare the entire green and the entire red cluster at K = 2, the emerging blue and the remaining green cluster at K = 3 and the emerging yellow and the remaining red cluster at K = 4.

Firstly, differences in co-activation patterns between the respective clusters are tested by performing MACM separately on the experiments associated with either cluster and *computing the voxel-wise difference between the ensuing ALE maps*. All experiments contributing to either analysis are then pooled and randomly divided into 2 groups of the same size as the 2 original sets of experiments defined by activation in the first or second cluster (Eickhoff, Bzdok et al. 2011). ALE-scores for these 2 randomly assembled groups are calculated and the difference between these ALE-scores is recorded for each voxel in the brain. Repeating this process 10000 times then yields a null-distribution of differences in ALE scores between the MACM analyses of the 2 clusters. The 'true' difference in ALE scores was then tested against this null-distribution yielding a P-value for the difference at each voxel based on the proportion of equal or higher random differences. The resulting non-parametric P-values are transformed into Z-scores, thresholded at $P < 0.001$ and inclusively masked by the respective main effects, that is, the significant effects in the MACM for the particular cluster. Conjunctions in turn are computed as the intersection between both cluster-level FWE corrected analyses.

The most important output of this MACM analysis can be found under ~/MACM7/MACM/yourvoi_4Clusters_85-135, in the 'Contrasts' and 'Conjunctions' folders, respectively. For example, the file '2_Cluster1--2_Cluster2_P95.nii' in the 'Contrasts' folder contains the color-coded contrast between Cluster 1 and Cluster 2 at K=2 thesholded at $P > 0.95$. Associated .png files of the contrasts and conjunctions can be found in the respective subfolder 'Images'.

Secondly, this step computes the conjunctions and the contrasts of the BDPC between the splitting clusters. For each comparison of the splitting cluster, the analysis was constrained to all BrainMap experiments activating either cluster. From this pool of experiments, the baserate is the a priori probability of any focus to lie in either of the two compared clusters. Forward inference here compared the activation probabilities of the clusters given a task compared to the a priori baserate by means of a binomial test. In the likewise performed reverse inference approach, we compared the occurrence probabilities of the tasks given activation in the one cluster (rather than in the other cluster) and assessed them by means of a chi-squared test. Again, the conjunction lists only those BDs and PCs that were significant in both individual analyses. The output can be found under ~/MACM7/MACM/yourvoi_4Clusters_85-135, in the 'BDPC-Contrasts' and 'BDPC-Conjunctions' folders. For example, 2_Cluster1--Cluster2 specifies the contrast between the splitting clusters 1 and 2 at the level of K = 2.

## 11. MACM contrast final clusters

This last step computes the co-activation contrasts and conjunctions between the final, CBP-

derived clusters (in our example on the level of K = 4) and between their functional profiles (BDPC). These results can again be found under ~/MACM7/MACM/yourvoi_4Clusters_85-135, in the 'Contrasts' and Conjunctions' and in the 'BDPC-Contrasts' and 'BDPC-Conjunctions'. For example, Cluster1--Cluster2 specifies the contrast between final cluster 1 and final cluster 2.

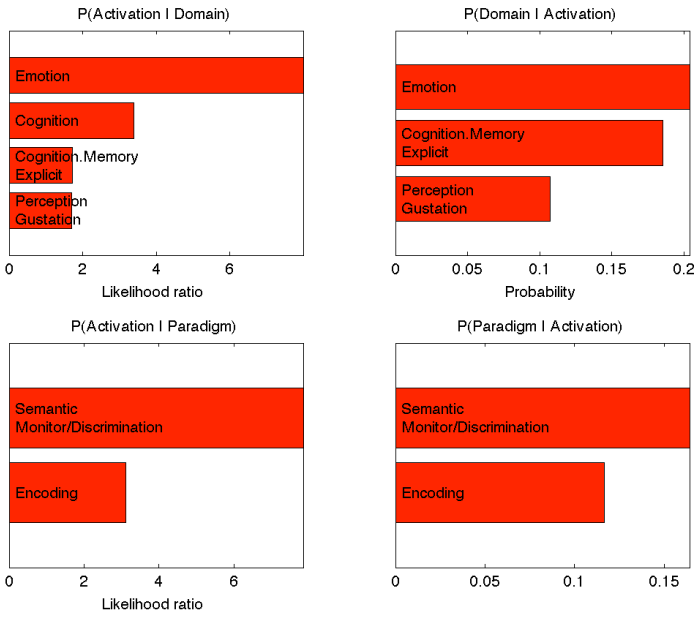
Furthermore, also the specific co-activation of each final CBP-derived cluster is computed, that is, brain regions significantly more co-activated with a given cluster than with any of the other clusters. This is achieved by performing a conjunction analysis over the differences between a given cluster and the other four ones. The output can be found in the 'Specifics' folder.

## 12. General information: MACM Output

There are many files produced by the MACM steps. Below is a look at these files based on the same folder/file organization as your output.

a. ALEvolumes: this folder contains the ALE scores for each MACM analysis
b. ALEvolumesZ: this folder contains the Z-scores for each MACM analysis
c. BDPC: this folder contains Behavioral Domain and Paradigm Class (BDPC) information
d. BDPC-Conjunctions: this folder contains Behavioral Domain and Paradigm Class (BDPC) information for conjunctions of different cluster pairs, i.e. where two clusters show the same BDPC information. The most common behavioral domains are displayed on the top row; most common paradigm classes are displayed on the bottom row. The left column shows likelihood ration that activity in Cluster X will produce a particular behavioral domain or paradigm class; the right column shows probability that whenever a particular behavioral domain or paradigm class is engaged, there will be activation within Cluster X.

**Figure 6:** Behavior Domain and Paradigm Class (BDPC) graphs. See 12.d for explanation of layout.



e. BDPC-Contrasts: this folder contains Behavioral Domain and Paradigm Class (BDPC) information for contrasting cluster pairs, i.e. where two clusters show different BDPC information.
f. BPPDprint: Same as above but in postscript format, which is easy to import into graphic-programs such as Corel or Adobe
g. Conjunctions: this folder contains conjoined MACM results for cluster pairs, i.e. where two clusters show the same MACM distribution.
h. Contrasts: this folder contains MACM results cluster pairs, i.e. where two clusters show the same MACM distribution.
i. Foci: this folder contains .nii files of all foci included in the different MACM analyses.
j. Images: this folder contains cortical projections of the ALE analyses. The file extension, e.g. Cluster1_cFWE05_001_105_LTR.png, indicates ClusterNumber_statistical level_clusterFormingThreshold_SizeOfSmallestSignificantCluster
k. ImagesFWE: Voxel-level FWE-corrected analyses
l. NullDistributions: this file contains the .mat files that were used to create the null distribution used in each statistical test.
m. Results: The thresholded (cluster-level FWE) Z-scores.
n. ResultsFWE: The thresholded (voxel-level FWE) Z-scores
o. Specifics: this folder contains areas of MACM-reported connectivity that are *unique* to each cluster.
p. Workspaces: For import into Sleuth.

# 13. References


Chang SE, Kenney MK, Loucks TM, Poletto CJ, Ludlow CL. 2009. Common neural substrates support speech and non-speech vocal tract gestures. Neuroimage. 47:314-325.

Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. 2012. Activation likelihood estimation meta-analysis revisited. Neuroimage. 59:2349-2361.

Eickhoff SB, Bzdok D, Laird AR, Roski C, Caspers S, Zilles K, Fox PT. 2011. Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage. 57:938-949.

Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. 2009b. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum.Brain Mapp. 30:2907-2926.

Kahnt T, Chang LJ, Park SQ, Heinzle J, Haynes JD. 2012. Connectivity-based parcellation of the human orbitofrontal cortex. J.Neurosci. 32:6240-6250.

Kelly C, Uddin LQ, Shehzad Z, Margulies DS, Castellanos FX, Milham MP, Petrides M. 2010. Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. Eur.J.Neurosci. 32:383-398.