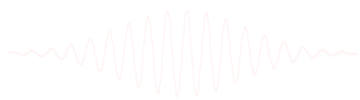
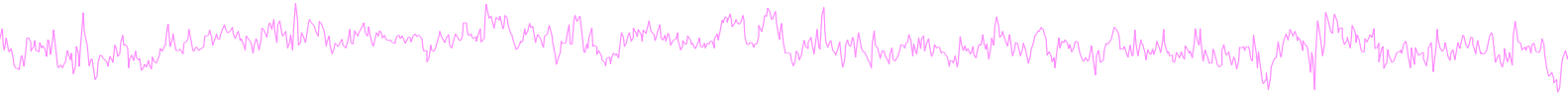
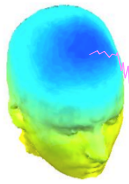


Clustering of ICA components

Arnaud Delorme

(with Julie Onton, Romain Grandchamp, Nima Bigdely Shamlo, Scott Makeig)



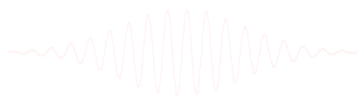


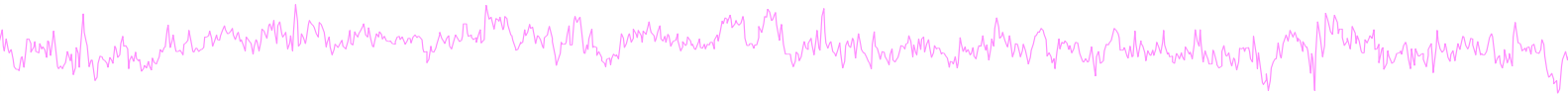
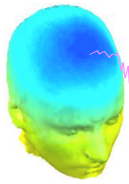
Outline

- ICA clusters and reliability within subjects

REF: Grandchamp, Makeig, Delorme, IEEE, 2012

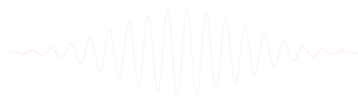
- Clustering in EEGLAB theory & Practice

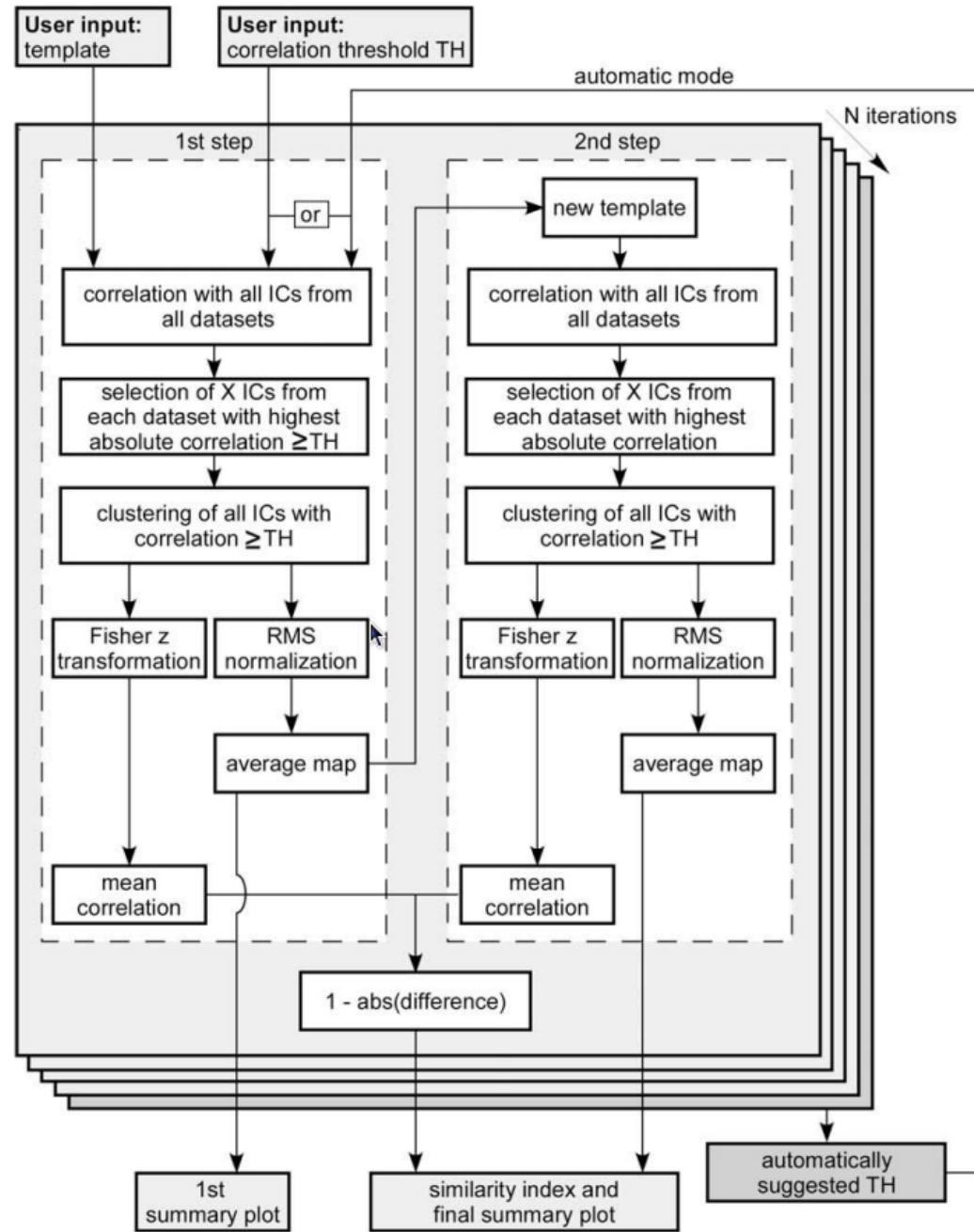
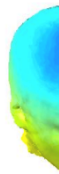




ICA decomposition of multiple data sets from the same individuals

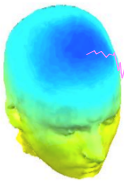
- Experimental protocol
 - Mind wandering experiment
 - 2 subjects
 - 11 x 30 min. sessions
 - 2 sessions per week
 - EEG from Biosemi 64 channels
 - $F_s=1024$ Hz





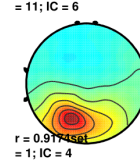
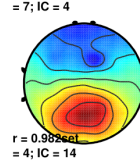
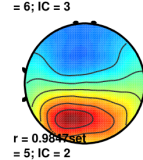
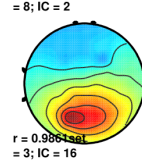
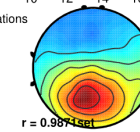
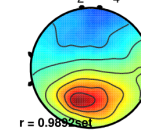
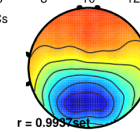
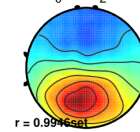
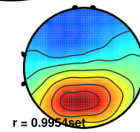
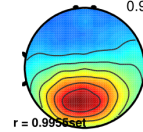
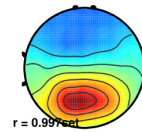
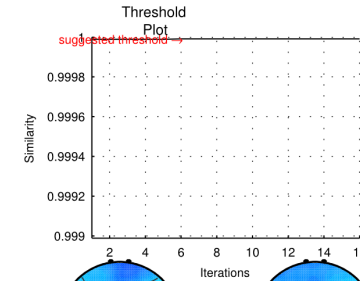
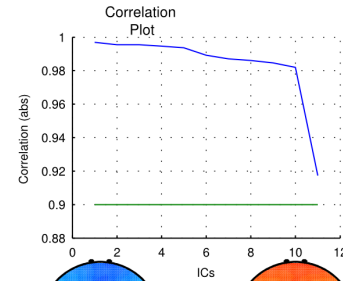
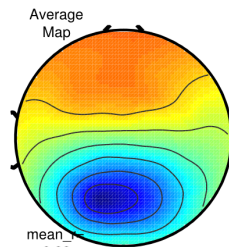
F. Campos Viola et al., "Semi-automatic identification of independent components representing EEG artifact," Clinical Neurophysiology 120, no. 5 (2009): 868–877.

Results (Cluster 1)

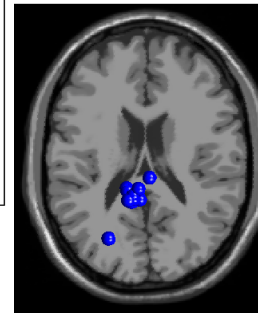
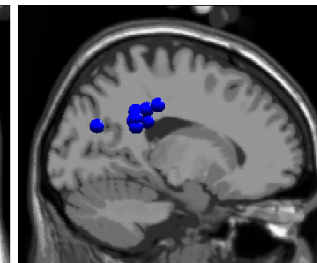
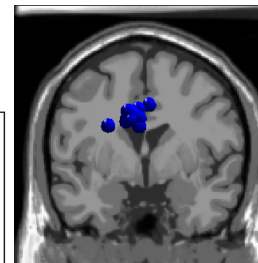
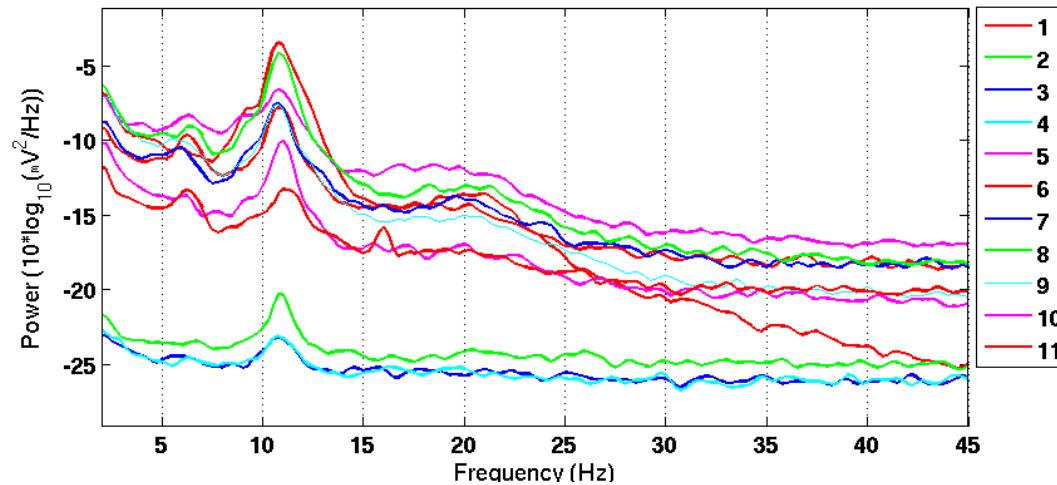


100 % Sessions contribute

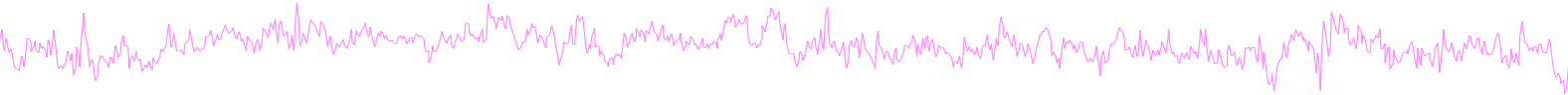
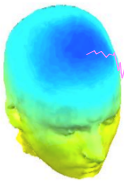
INFO:
Template: CB Session 7 PREPROC:STEP 2; Set 7; IC 3;
Number of datasets: 11
Correlation threshold: 0.9 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.
Similarity = 1.0000



Cls 3 Spectrum



Results (Cluster 2)



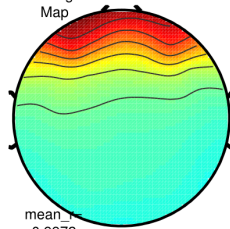
100 % Sessions contribute

INFO:

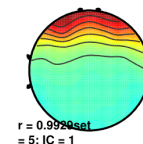
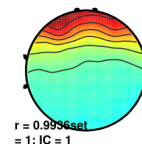
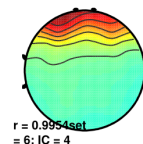
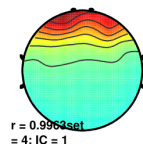
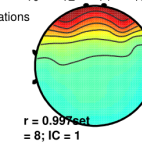
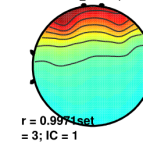
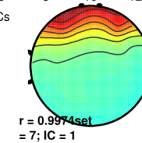
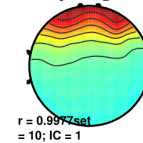
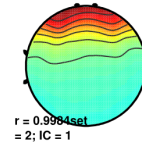
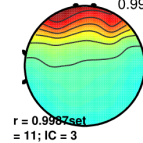
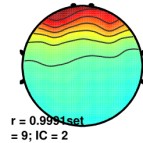
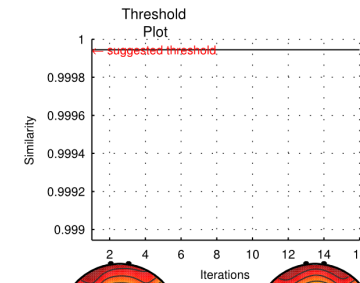
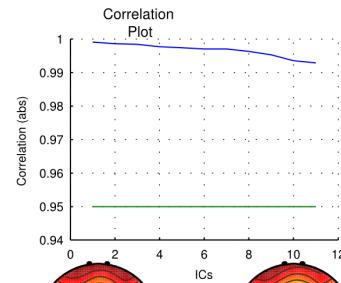
Template: CB Session 5 PREPROC:STEP 2; Set 5; IC 1;
Number of datasets: 11
Correlation threshold: 0.95 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.

Similarity = 0.9999

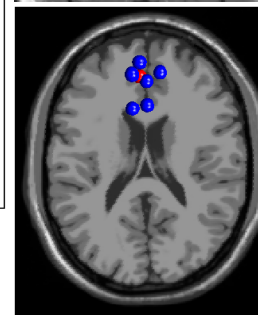
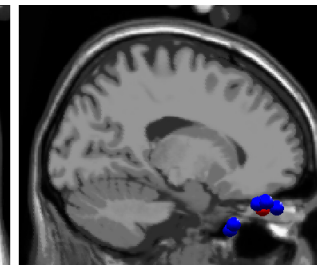
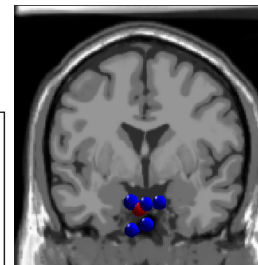
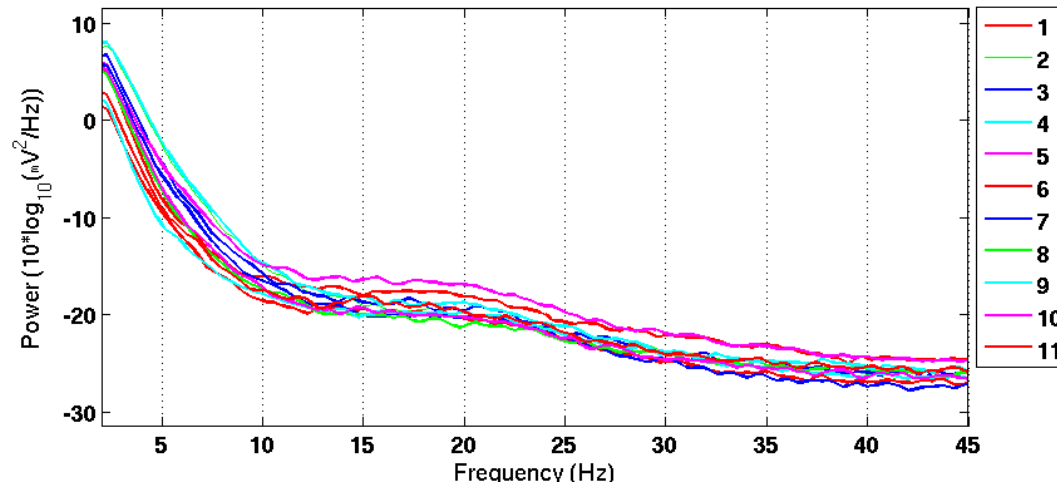
Average



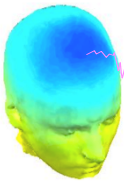
mean
0.9972



Cls 4 Spectrum

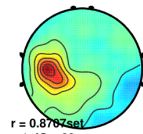
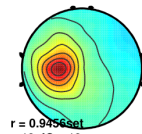
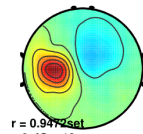
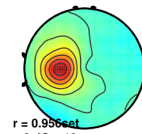
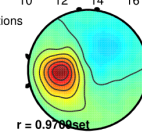
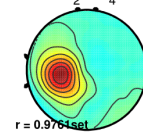
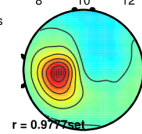
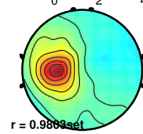
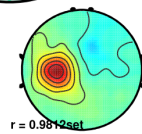
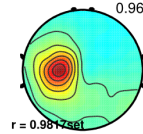
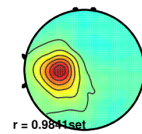
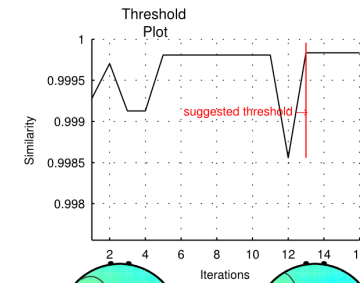
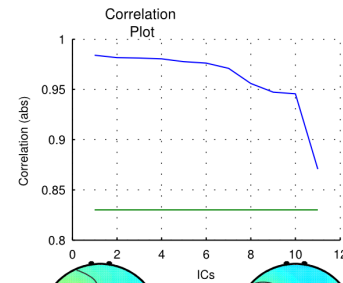
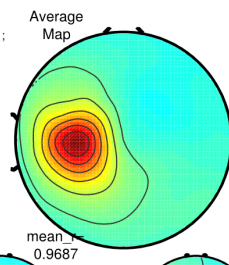


Results (Cluster 8)

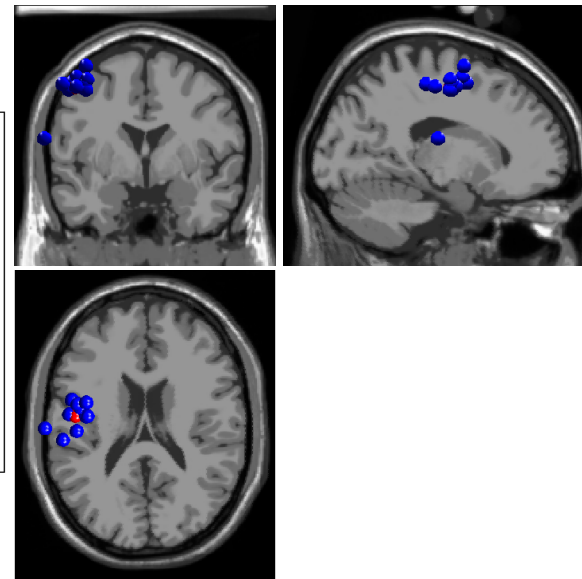
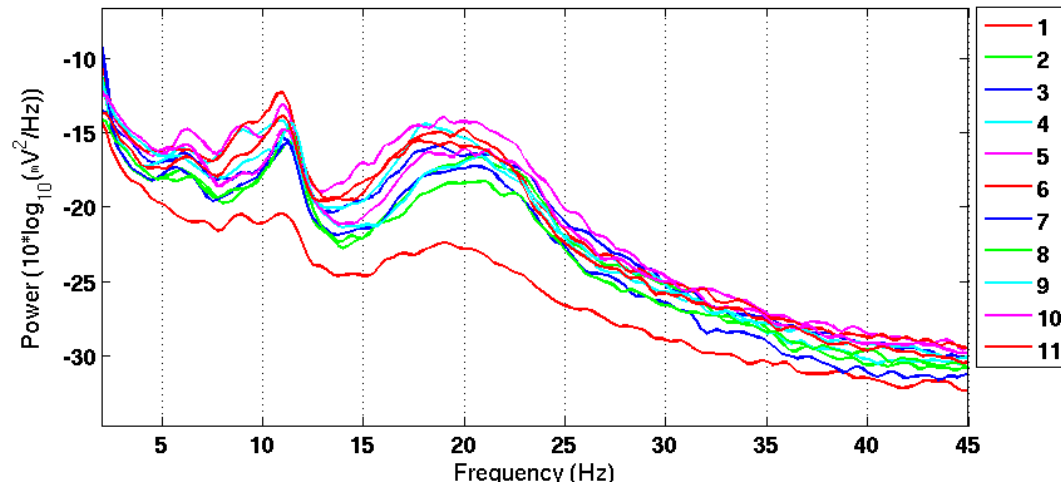


100 % Sessions contribute

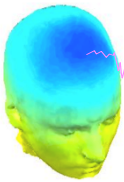
INFO:
Template: CB Session 7 PREPROC:STEP 2; Set 7; IC 11;
Number of datasets: 11
Correlation threshold: 0.83 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.
Similarity = 0.9998



Cls 8 Spectrum

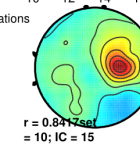
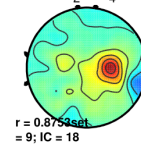
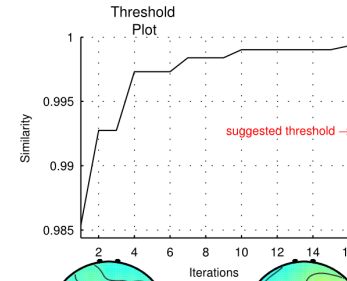
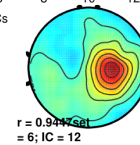
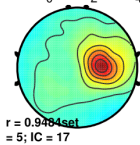
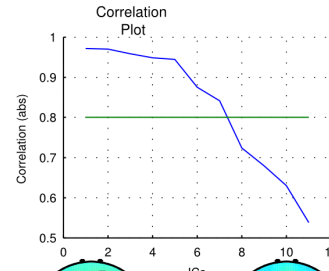
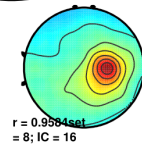
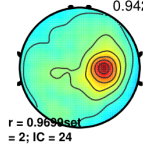
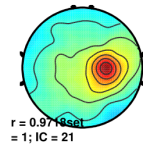
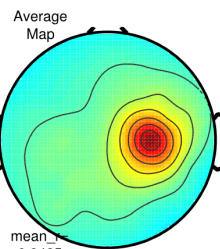


Results (Cluster 13)

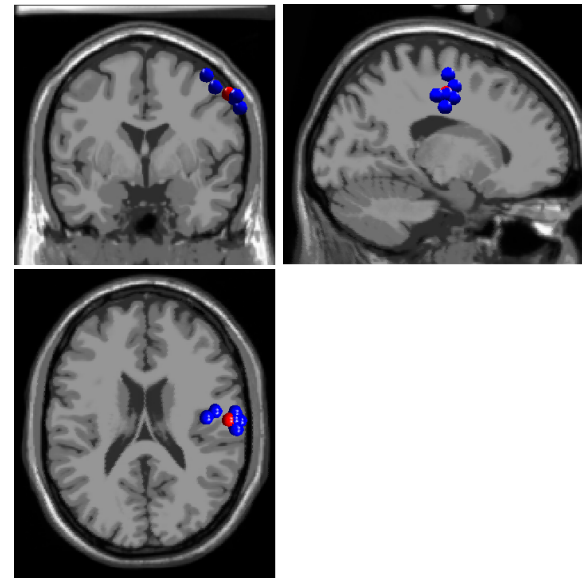
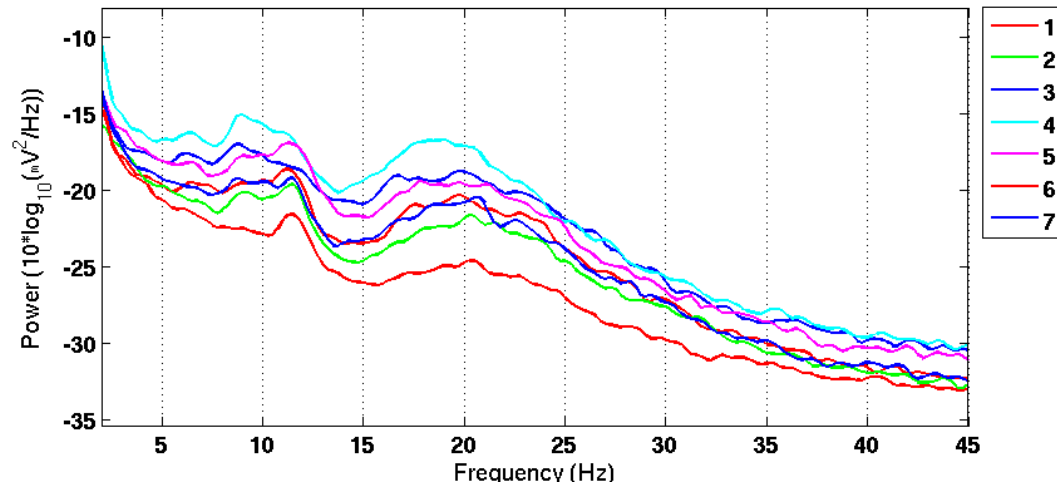


63.64% Sessions contribute

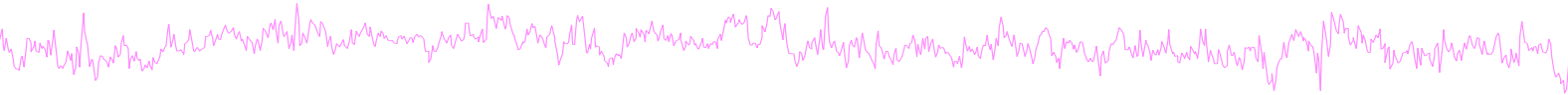
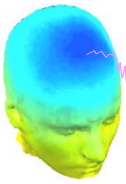
INFO:
 Template: CB Session 2 PREPROC:STEP 2; Set 2; IC 24;
 Number of datasets: 11
 Correlation threshold: 0.8 (green line)
 Max ICs from each dataset: 1
 Cluster: 7 ICs from 7 sets
 Sets not contributing:
 #3; #4; #7; #11;
 Similarity = 0.9993



Cl 13 Spectrum

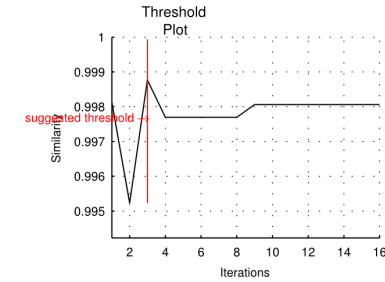
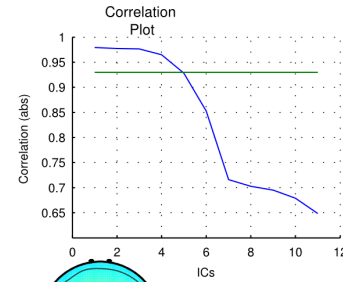
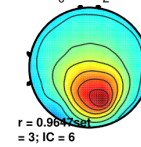
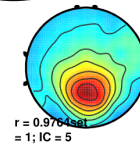
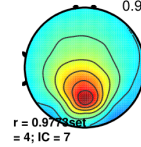
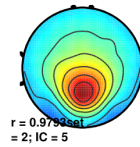
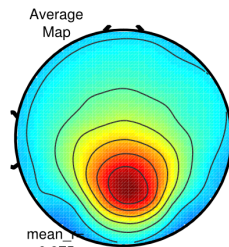


Results (Cluster 14)

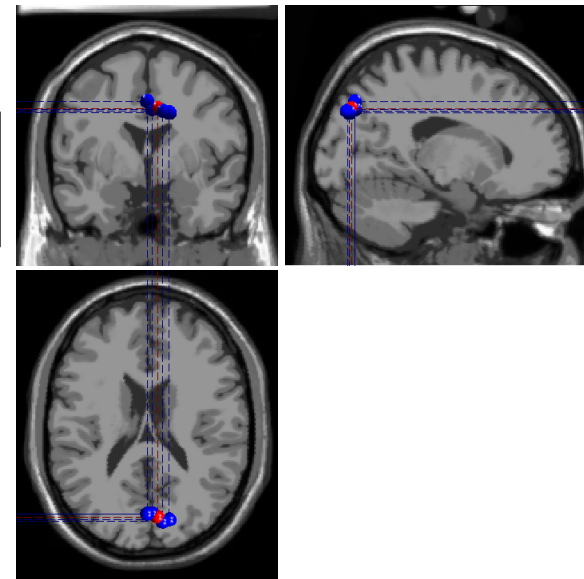
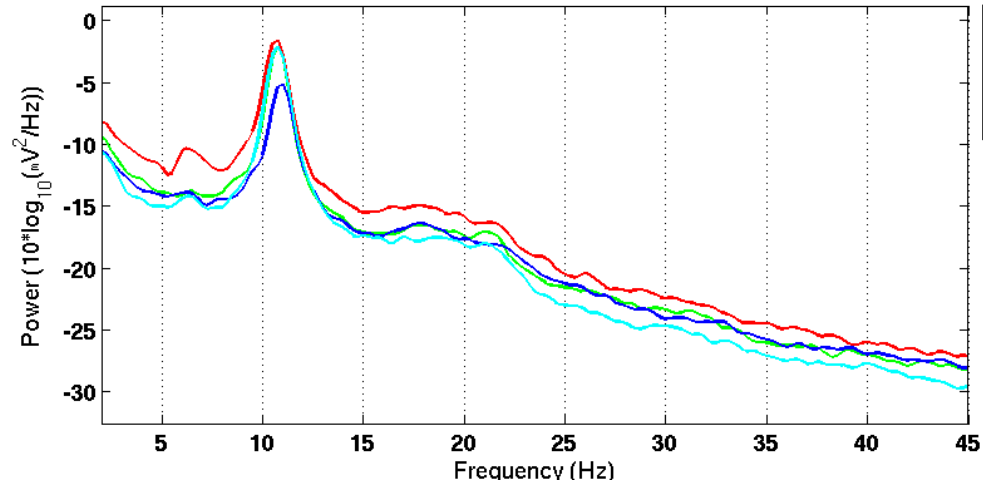


36.36% Sessions contribute

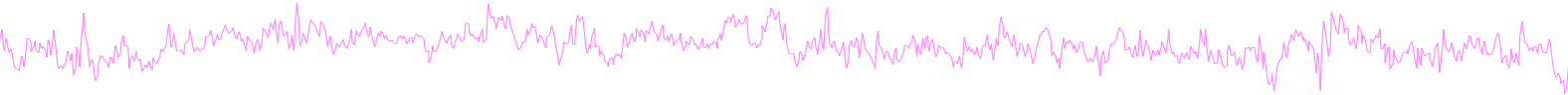
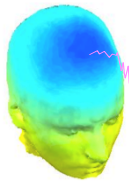
INFO:
Template: CB Session 1 PREPROC:STEP 2; Set 1; IC 5;
Number of datasets: 11
Correlation threshold: 0.93 (green line)
Max ICs from each dataset: 1
Cluster: 4 ICs from 4 sets
Sets not contributing:
#5; #6; #7; #8; #9; #10#11;
Similarity = 0.9988



Cl 14 Spectrum

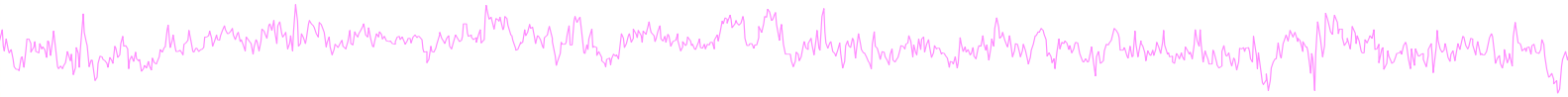
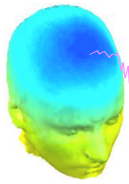


Inter iteration Cluster Consistency



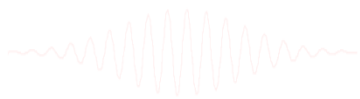
Iterations

Clusters		1	2	3	4	5	6	7	8	9	10	Mean
	3	100	100	100	100	100	100	100	100	100	100	100
	4	100	100	100	100	100	100	90	100	100	100	99
	5	90	40	10	90	90	60	100	10	60	90	64
	6	60	0	100	60	100	90	60	60	90	60	68
	7	90	100	90	90	60	90	90	100	90	90	89
	8	80	80	60	80	40	80	80	80	80	100	76
	9	60	90	50	60	80	60	0	10	60	50	52
	10	40	90	10	40	0	50	50	0	50	60	39
	11	60	20	0	0	10	60	10	90	60	60	37
	12	100	50	50	100	50	100	100	50	100	50	75
	13	50	10	20	50	90	50	50	10	50	20	40
	14	20	10	10	20	20	30	20	20	30	30	21

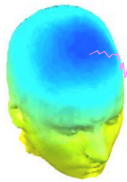


Outline

- ICA clusters and reliability within subjects
- Clustering in EEGLAB theory & Practice



Edit dataset info



pop_study(): Pre-select components

Enter maximum residual (topo map – dipole proj.) var. (in %)
NOTE: This will delete any existing component clusters!

15

☒ Keep only in-brain dipoles.

Cancel Help Ok

Create a new STUDY set -- pop_study()

Edit STUDY set information - remember to save changes

STUDY set name: Sternberg

STUDY set task name: Sternberg

STUDY set notes:

	dataset filename	browse	subject	session	condition	group	Select by r.v.	
1	C:\Users\julie\Documents\Wor	...	S01		memorize		Comp.: 3 5 ...	Clear
2	C:\Users\julie\Documents\Wor	...	S01		ignore		Comp.: 3 5 ...	Clear
3	C:\Users\julie\Documents\Wor	...	S01		probe		Comp.: 3 5 ...	Clear
4	C:\Users\julie\Documents\Wor	...	S02		memorize		Comp.: 5 6 ...	Clear
5	C:\Users\julie\Documents\Wor	...	S02		ignore		Comp.: 5 6 ...	Clear
6	C:\Users\julie\Documents\Wor	...	S02		probe		Comp.: 5 6 ...	Clear
7	C:\Users\julie\Documents\Wor	...	S03		memorize		Comp.: 6 7 ...	Clear
8	C:\Users\julie\Documents\Wor	...	S03		ignore		Comp.: 6 7 ...	Clear
9	C:\Users\julie\Documents\Wor	...	S03		probe		Comp.: 6 7 ...	Clear
10	C:\Users\julie\Documents\Wor	...	S04		memorize		Comp.: 1 2 ...	Clear

Important note: Removed datasets will not be saved before being deleted from EEGLAB memory

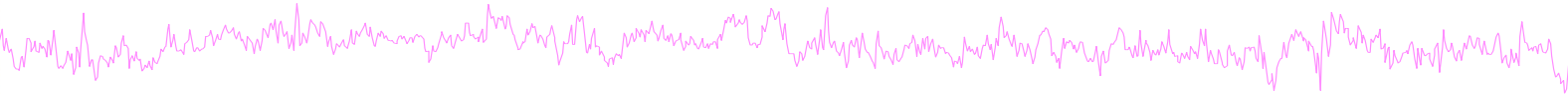
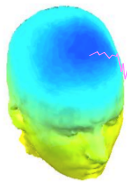
< Page 1 >

☐ Dataset info (condition, group, ...) differs from study info. [set] = Overwrite dataset info.

☐ Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

Help Cancel Ok

ICs to cluster



Create a new STUDY set -- pop_study()

Edit STUDY set information - remember to save changes

STUDY set name: Sternberg
STUDY set task name: Sternberg
STUDY set notes:

	dataset filename	condition	group	Select by r.v.	
1	C:\Users\julieW\...	memorize		Comp.: 3 5 ...	Clear
2	C:\Users\julieW\...	ignore		Comp.: 3 5 ...	Clear
3	C:\Users\julieW\...	probe		Comp.: 3 5 ...	Clear
4	C:\Users\julieW\...	memorize		Comp.: 5 6 ...	Clear
5	C:\Users\julieW\...	ignore		Comp.: 5 6 ...	Clear
6	C:\Users\julieW\...	probe		Comp.: 5 6 ...	Clear
7	C:\Users\julieW\...	memorize		Comp.: 6 7 ...	Clear
8	C:\Users\julieW\...	ignore		Comp.: 6 7 ...	Clear
9	C:\Users\julieW\...	probe		Comp.: 6 7 ...	Clear
10	C:\Users\julieW\...	memorize		Comp.: 1 2 ...	Clear

Important note: Removed datasets will not be saved before being deleted from EEGLAB memory

< Page 1 >

☐ Dataset info (condition, group, ...) differs from study info. [set] = Overwrite dataset info.
☐ Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

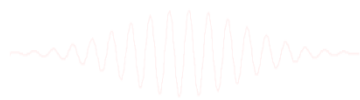
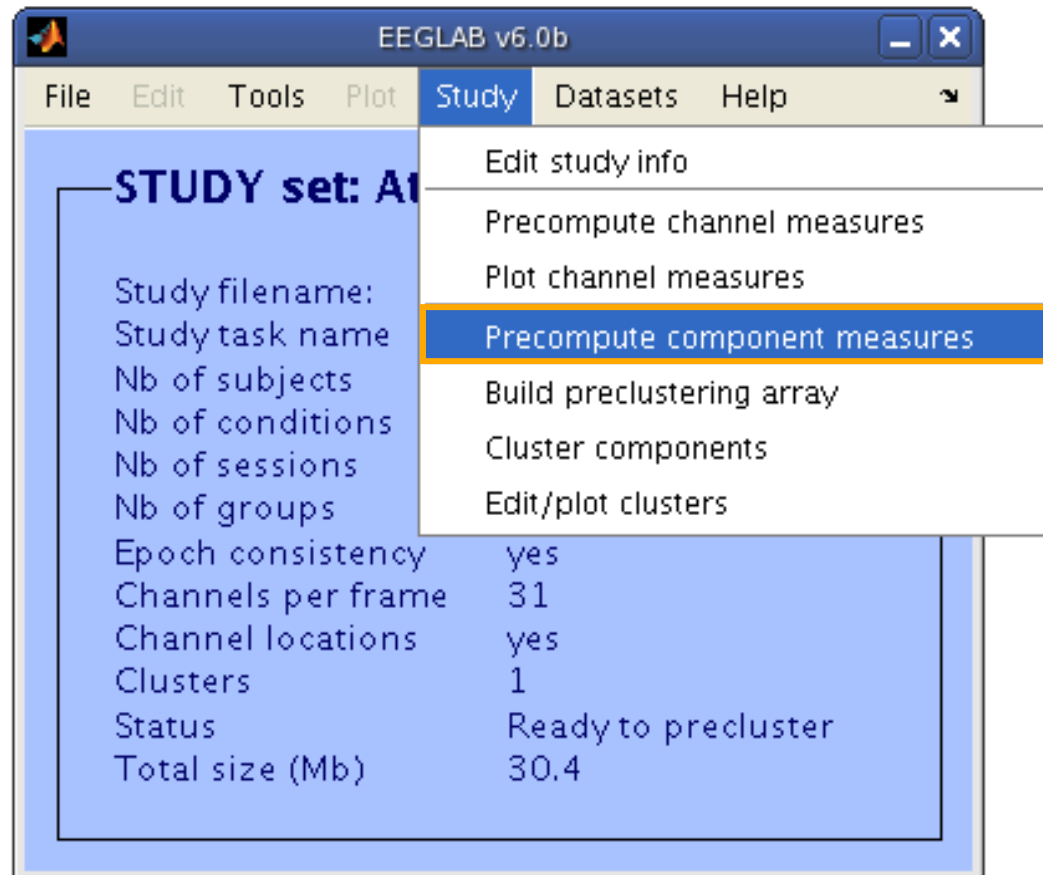
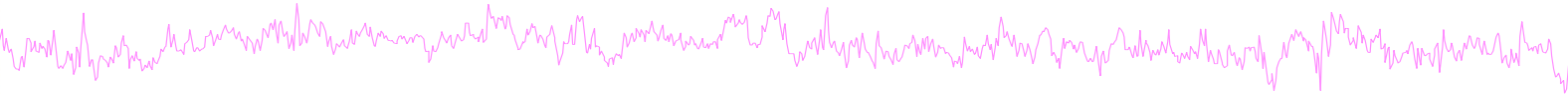
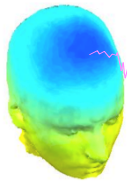
Help Cancel Ok

select components

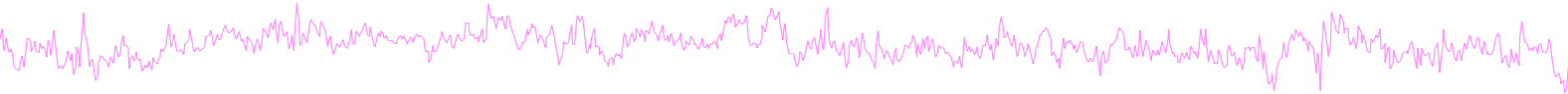
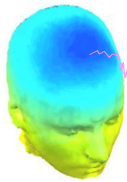
- ic 19
- ic 20
- ic 21
- ic 22
- ic 23
- ic 24
- ic 25
- ic 26
- ic 27
- ic 28
- ic 29
- ic 30
- ic 31
- ic 32

Cancel Ok

Precompute data measures



Precompute data measures



TIP: Compute all measures so you can test different combinations for clustering

Select and compute component measures for later clustering -- pop_precomp()

Pre-compute component measures for STUDY 'Sternberg'

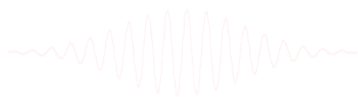
☒ Compute ERP/spectrum/ERSP only for components selected by RV (set) or for all components (unset)

List of measures to precompute

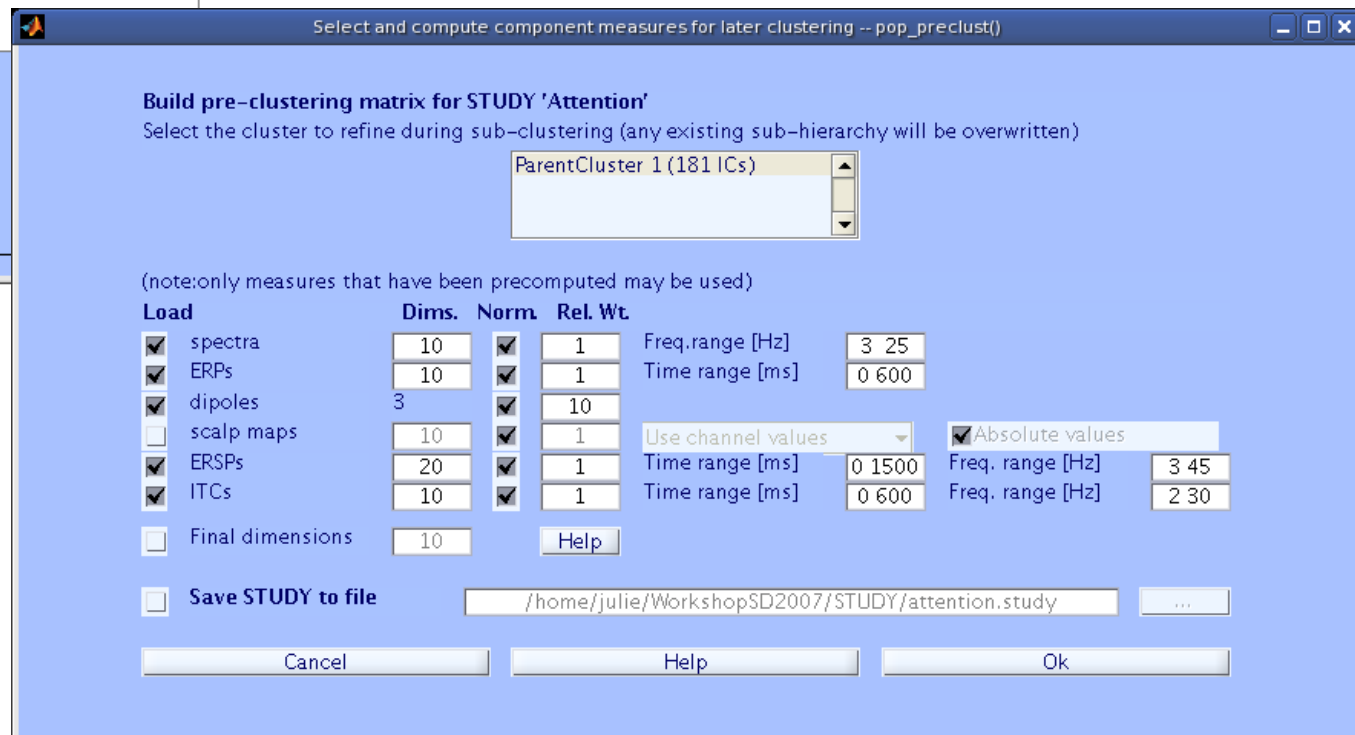
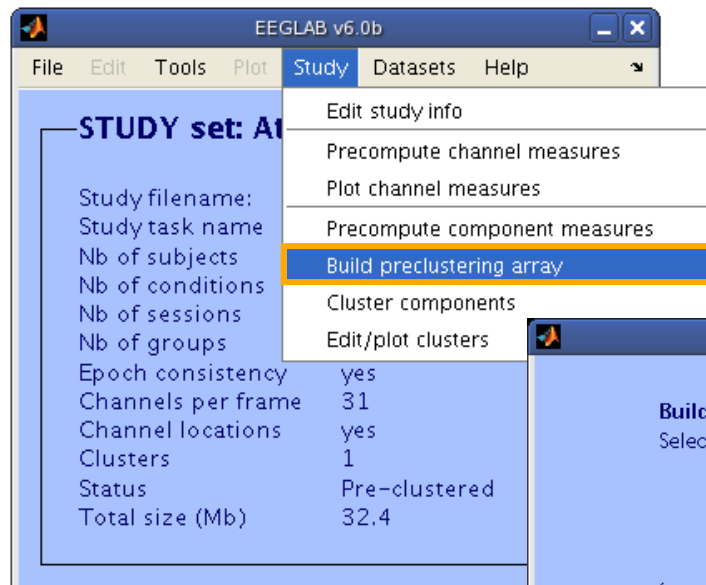
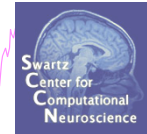
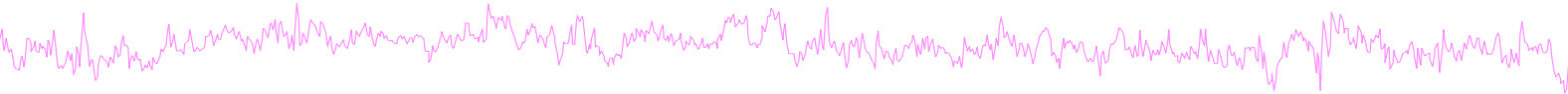
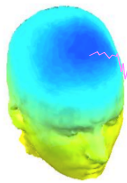
<input checked="" type="checkbox"/> ERPs	Baseline ([min max] in ms)	<input type="text" value="[-200 0]"/>	
<input checked="" type="checkbox"/> Power spectrum	Spectopo parameters	<input type="text"/>	<input type="button" value="Test"/>
<input checked="" type="checkbox"/> ERSPs	Time/freq. parameters	<input type="text" value="'cycles', [3 0.5], 'nfreqs', 100"/>	<input type="button" value="Test"/>
<input checked="" type="checkbox"/> ITCs			
<input checked="" type="checkbox"/> Scalp maps			

☐ Recompute even if present on disk

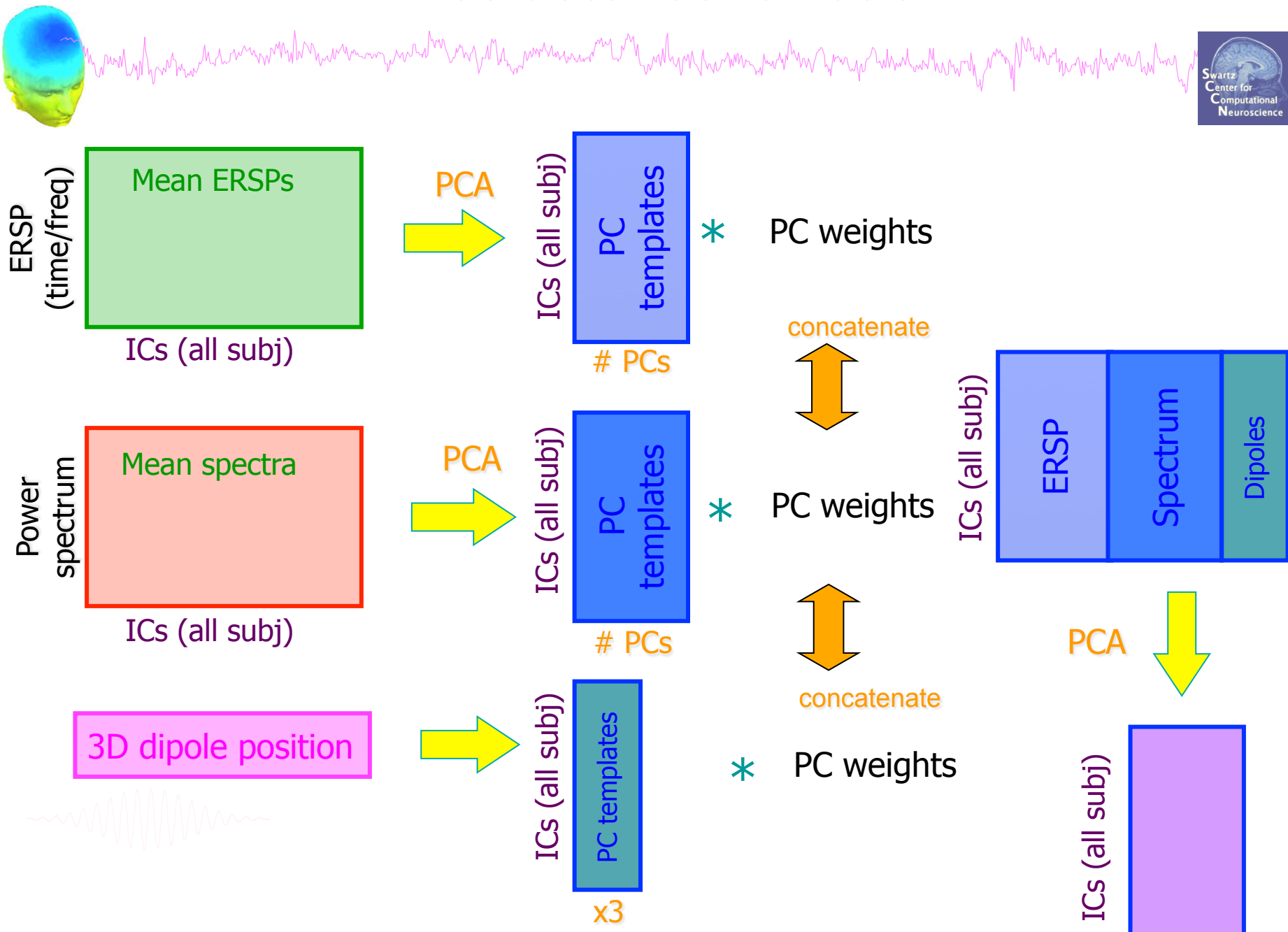
Time-frequency options



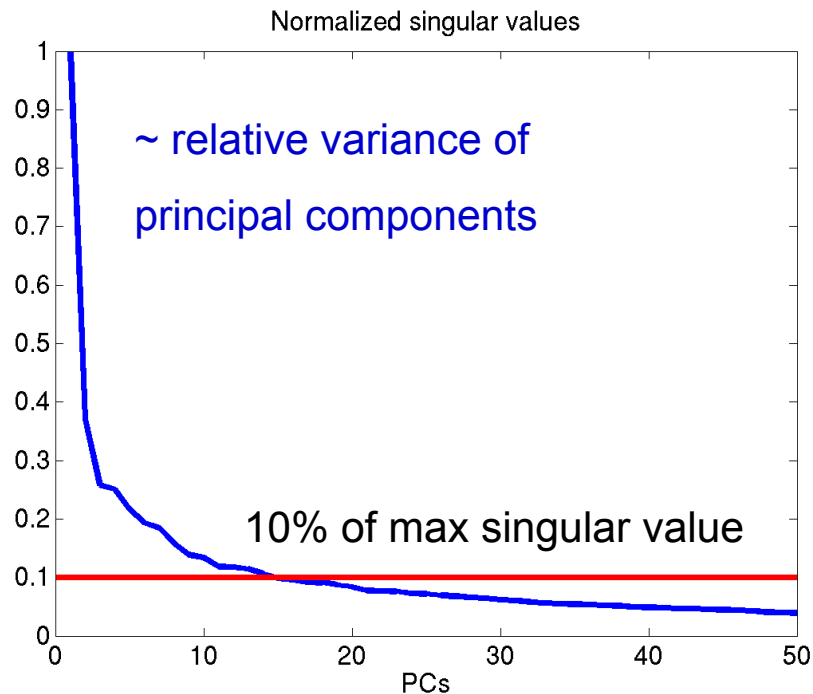
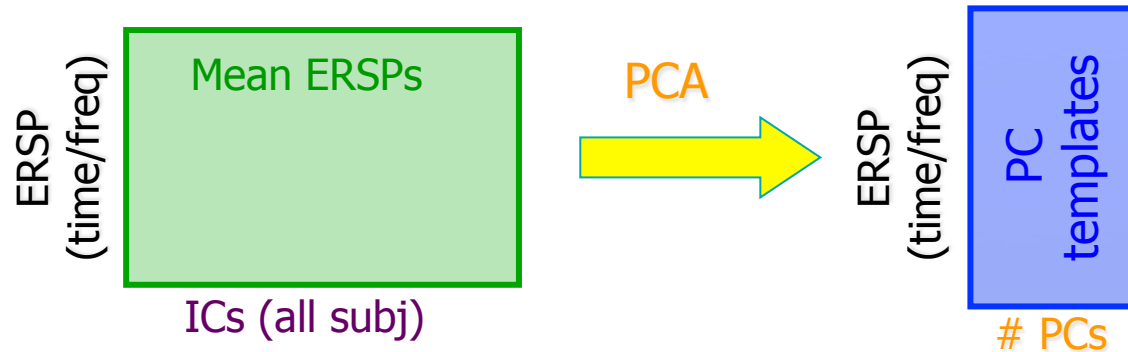
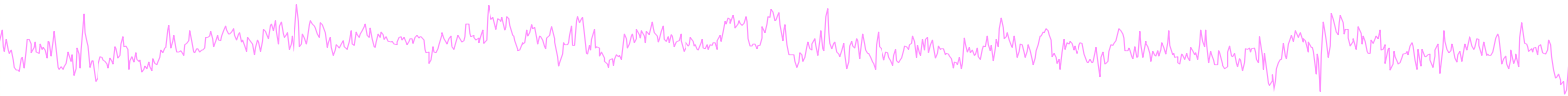
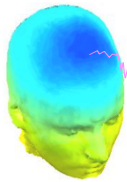
3. Cluster components



Precluster schematic

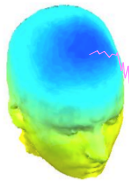


Precluster: Use singular values from PCA

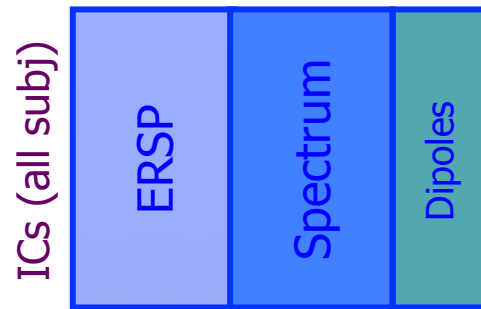


Credit: Julie Onton

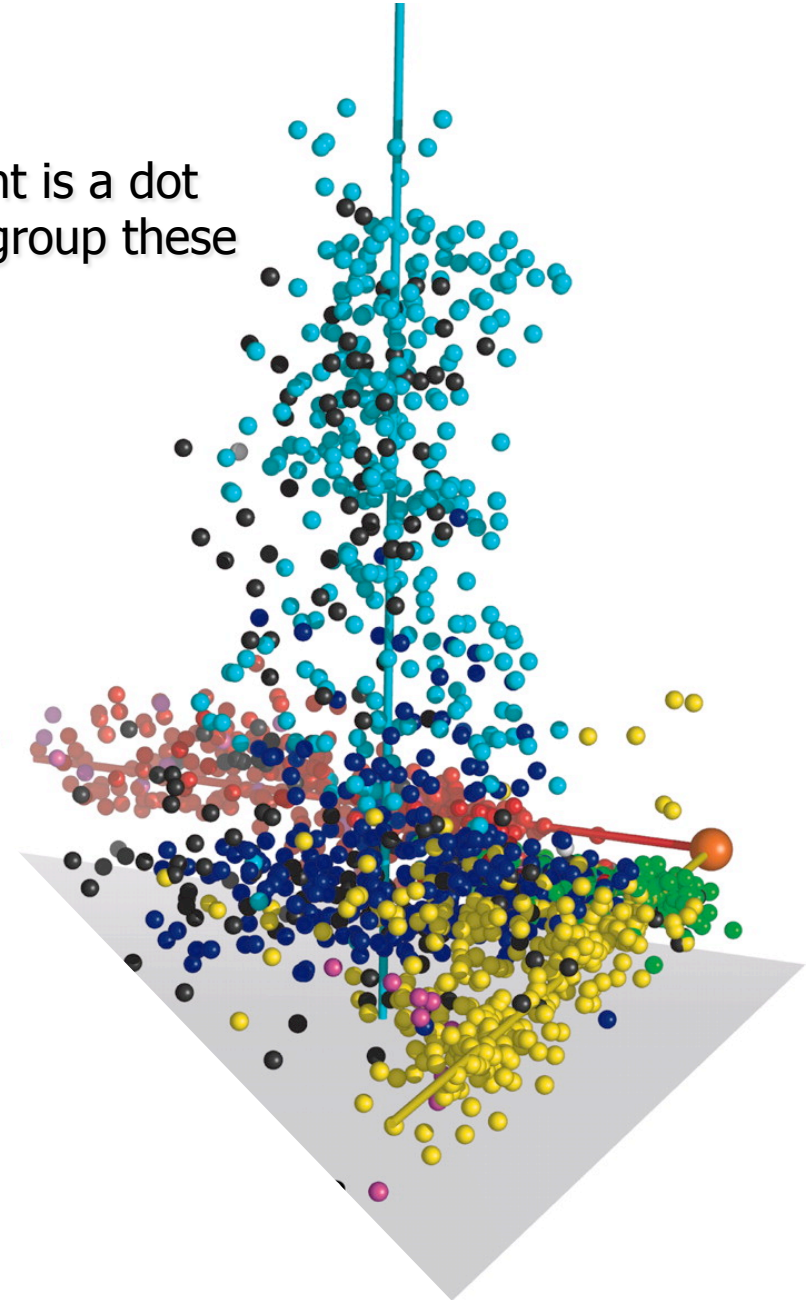
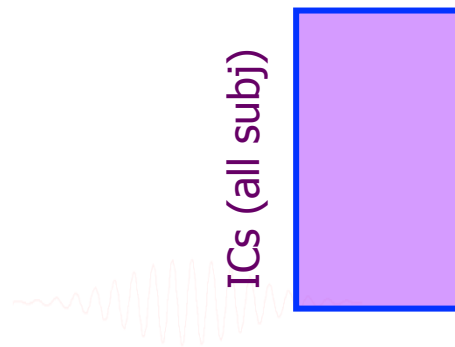
Precluster schematic



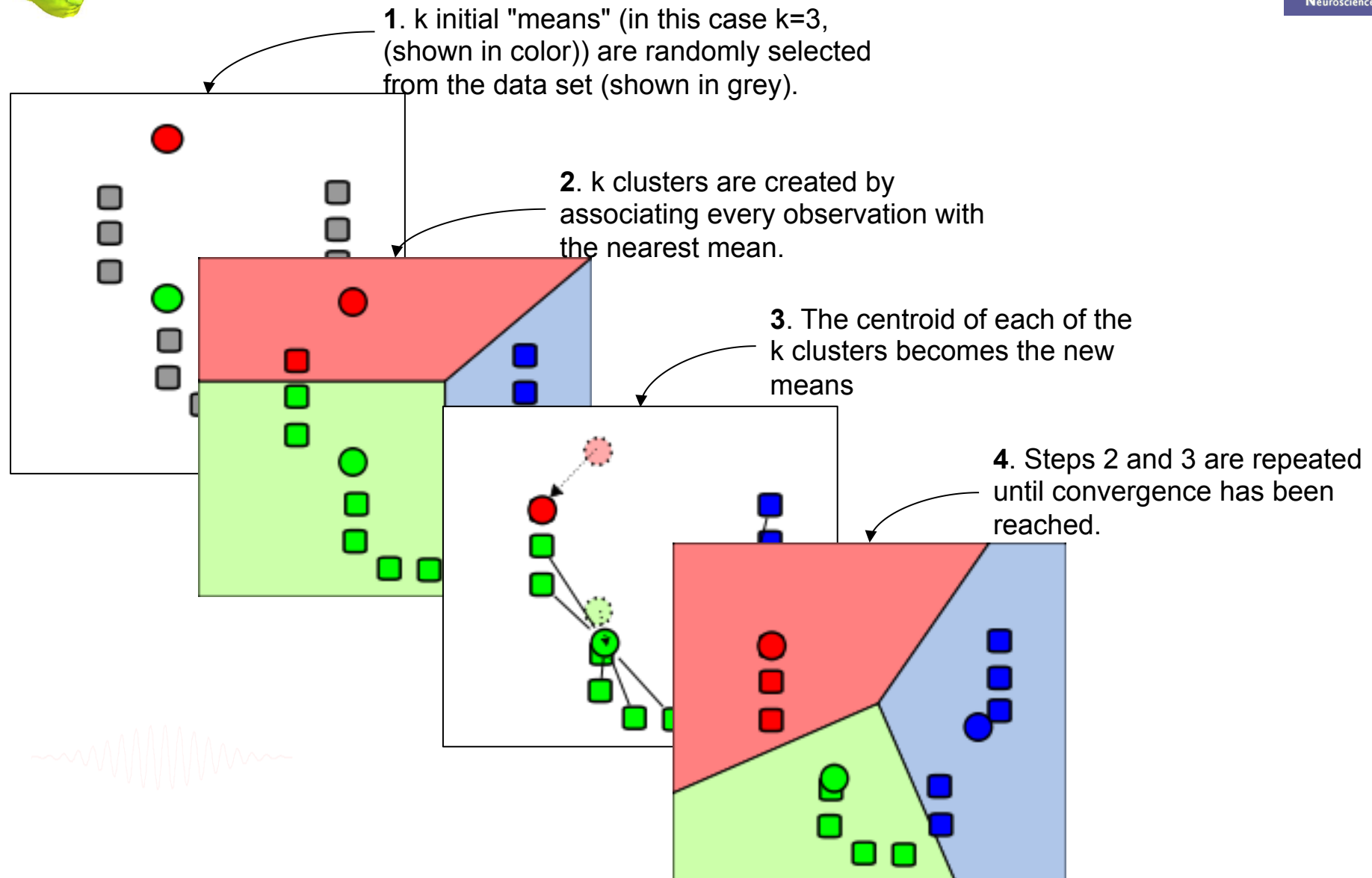
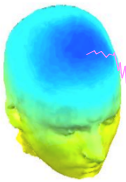
Each component is a dot
Clustering will group these dots

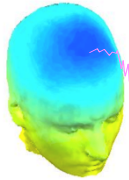


OR



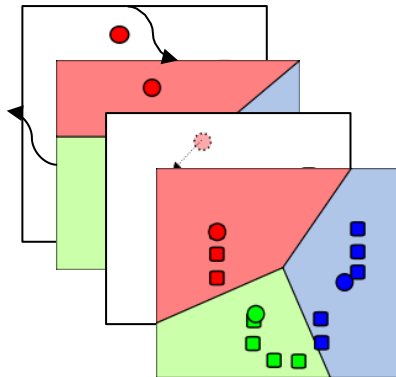
Classical KMean





Customized KMean

(no more than 1 session per cluster)



1. A first KMean solution is computed for N clusters

2. Select the cluster with minimum residual distance to centroid

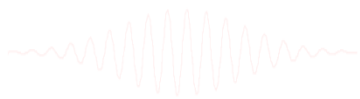
3. Keeps at most one component per session
(min dist. to centroid)

4. Store the resulting cluster

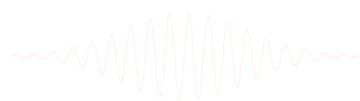
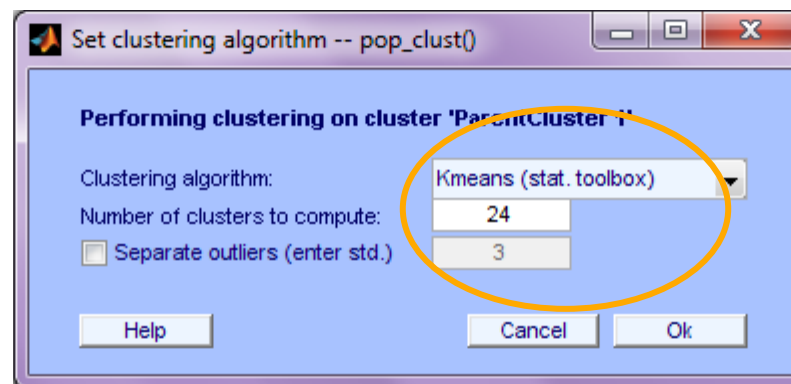
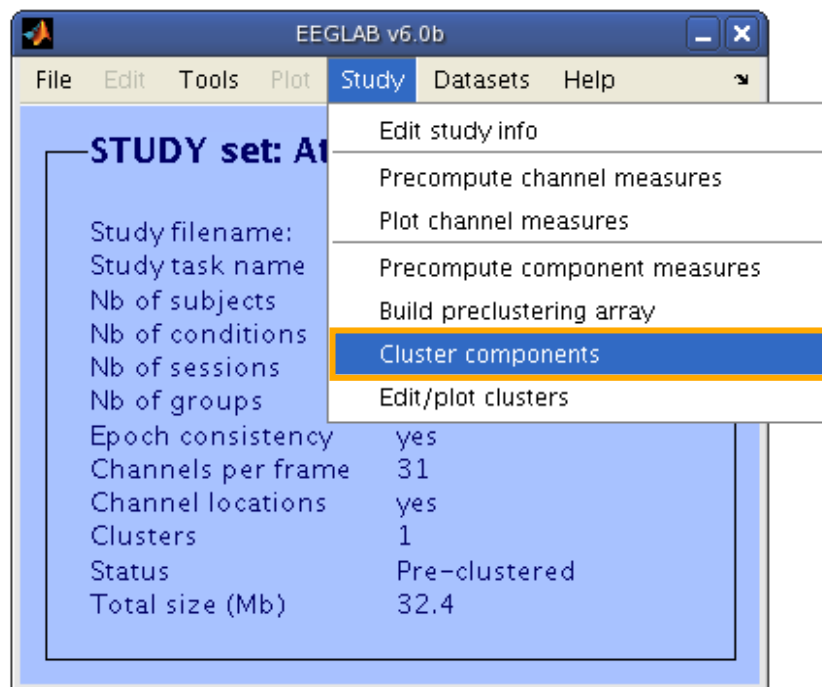
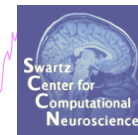
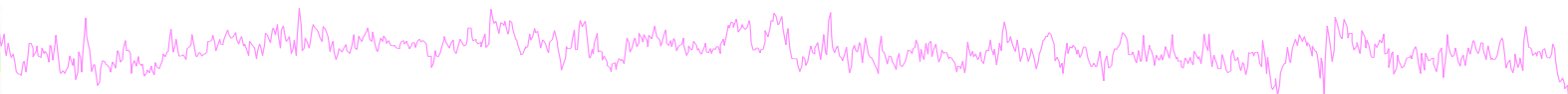
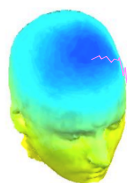
5. Remove the cluster's ICs from the pool of all ICs

6. Compute a new KMean solution for N-1 clusters on the new pool

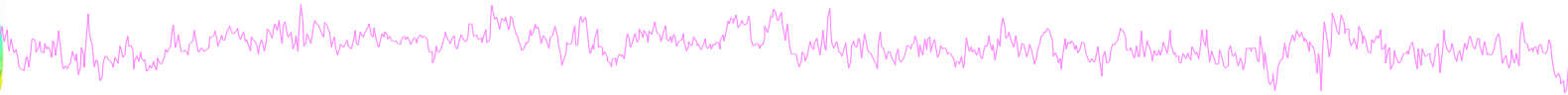
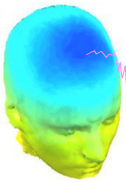
7. Loop until the desired number of selected clusters is reached



Cluster components



Choosing data measures



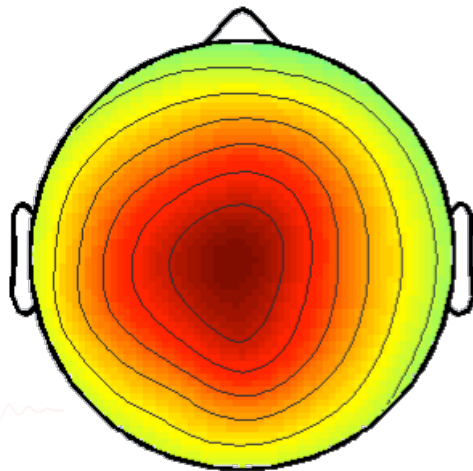
What measure(s) should you use?

- It depends on your final cluster criteria...
 - If for example, your priority is dipole location, then cluster only based on dipole location...

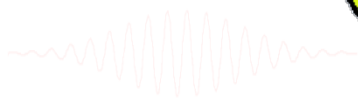
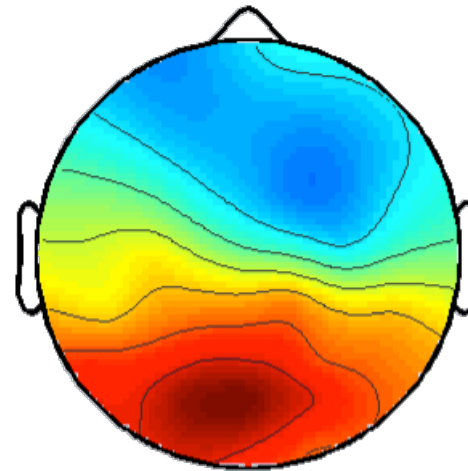
But consider:

- What is the difference between these two components?

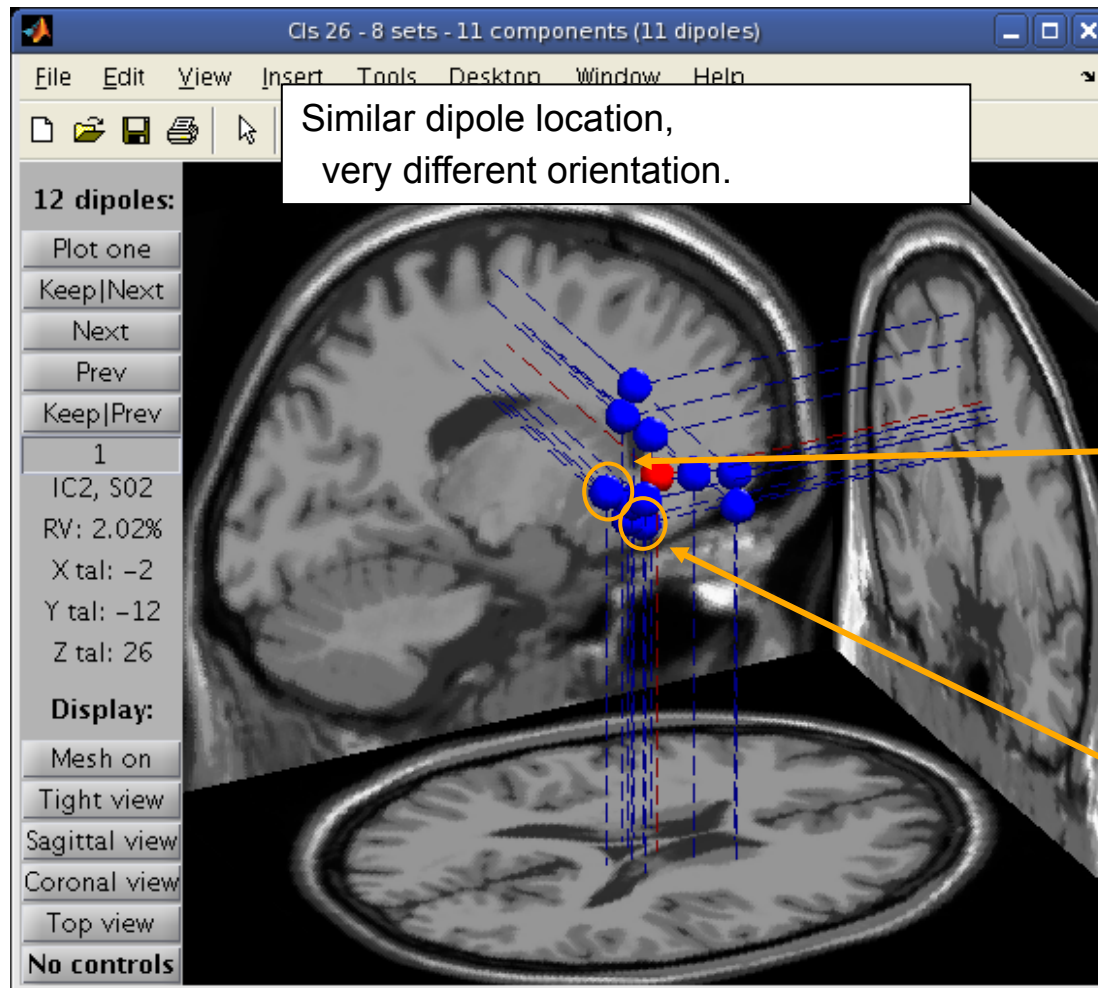
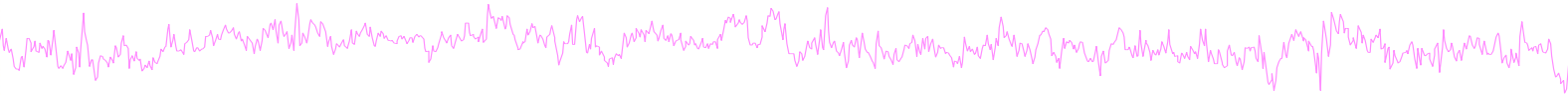
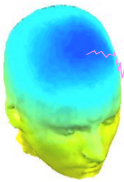
IC2 / S02, CIs 26



IC5 / S05, CIs 26

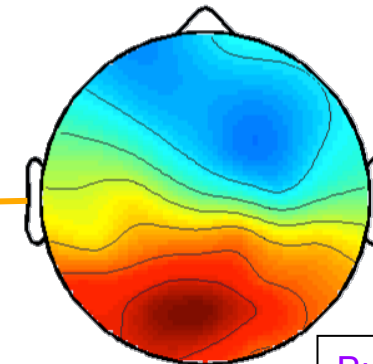


Choosing data measures

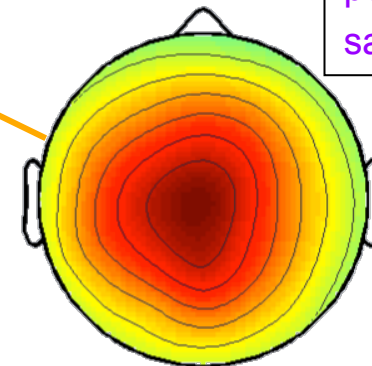


Obvious dramatic effect on
scalp map topography:

IC5 / S05, Clis 26

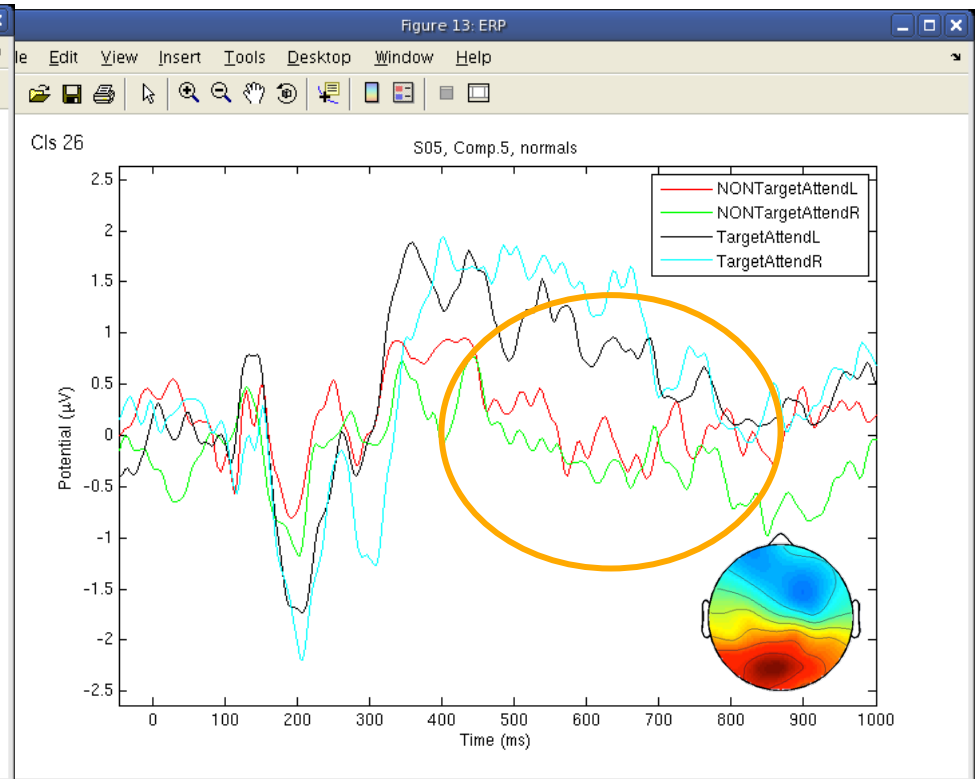
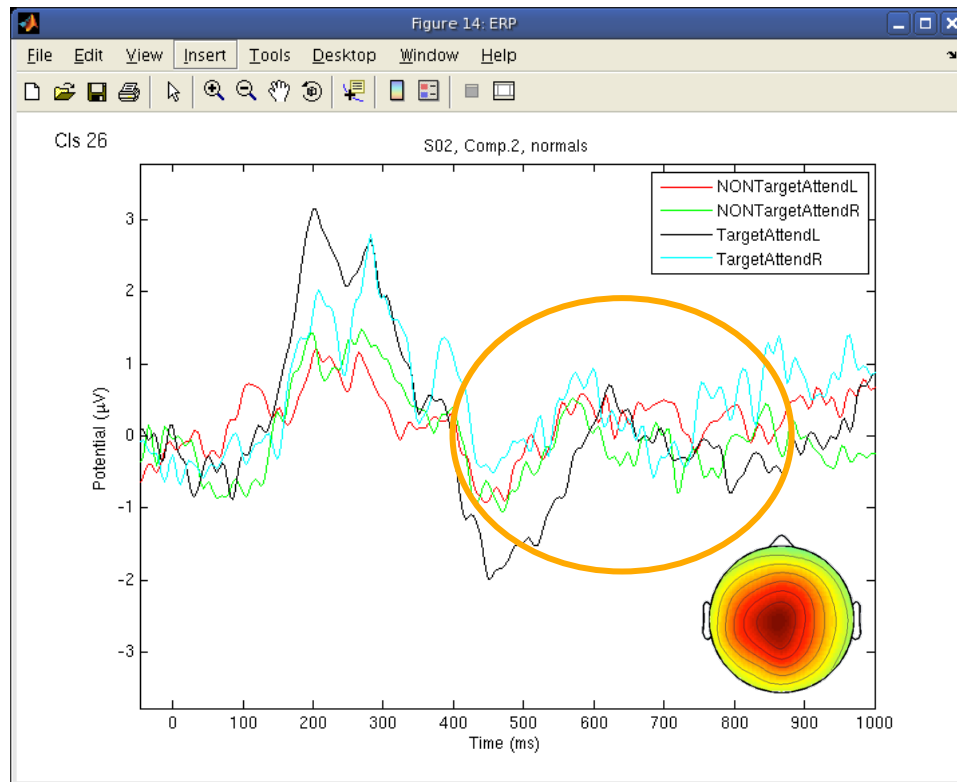
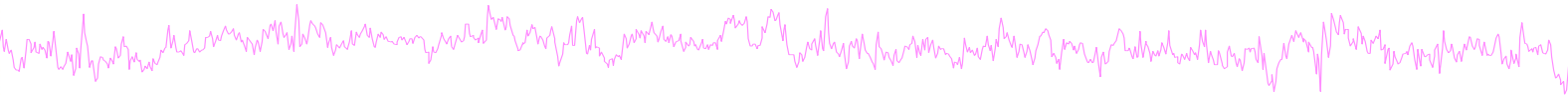
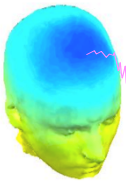


IC2 / S02, Clis 26

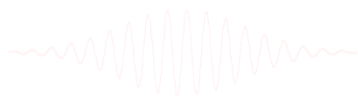


But, do they
perform the
same functions?

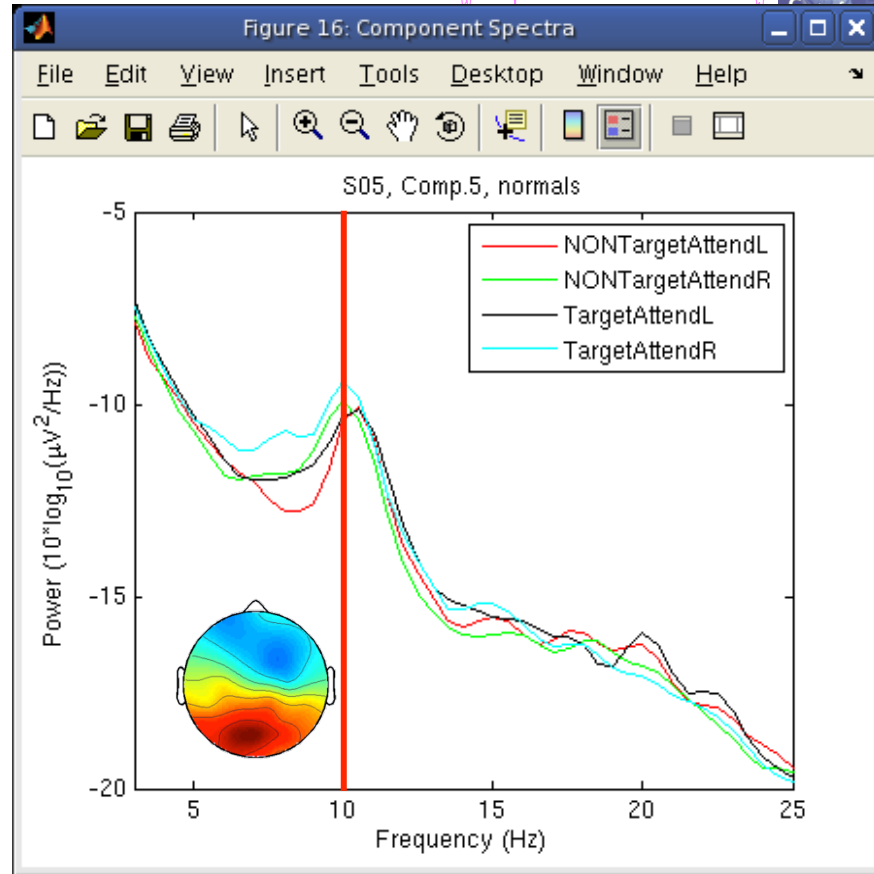
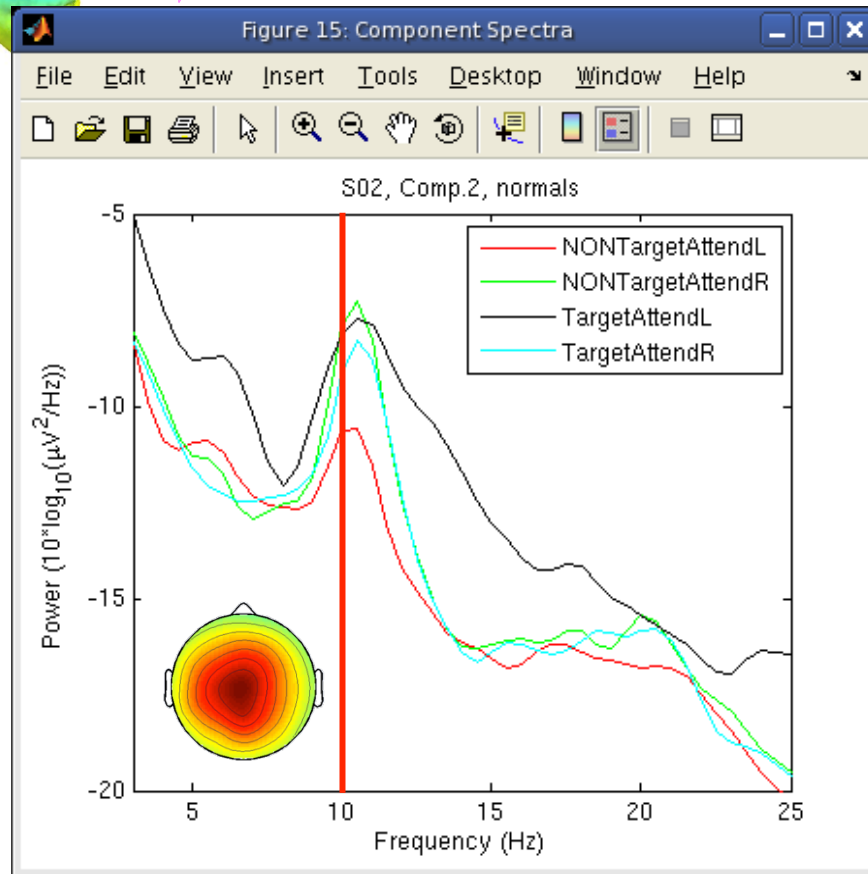
Choosing data measures



ERPs seem different...

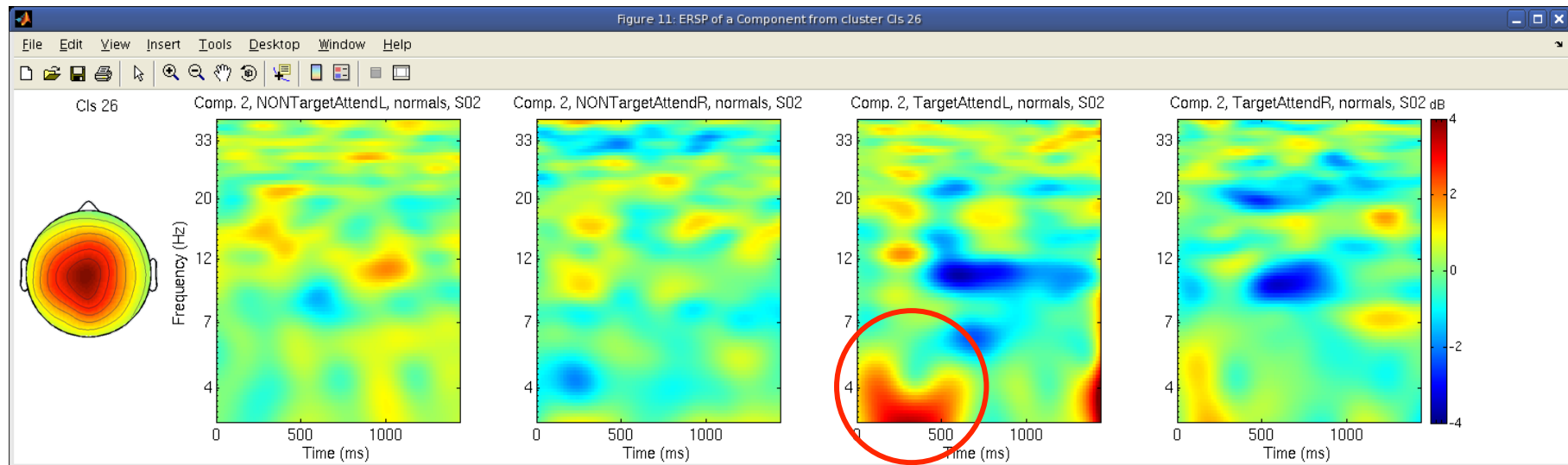
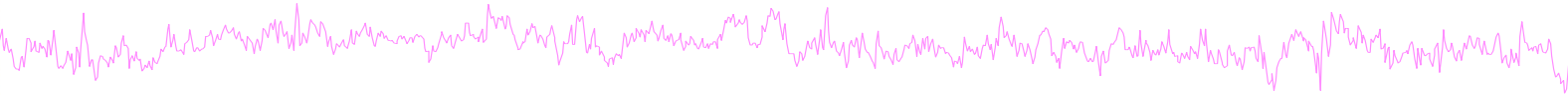
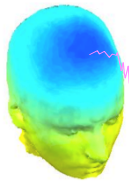


Choosing data measures

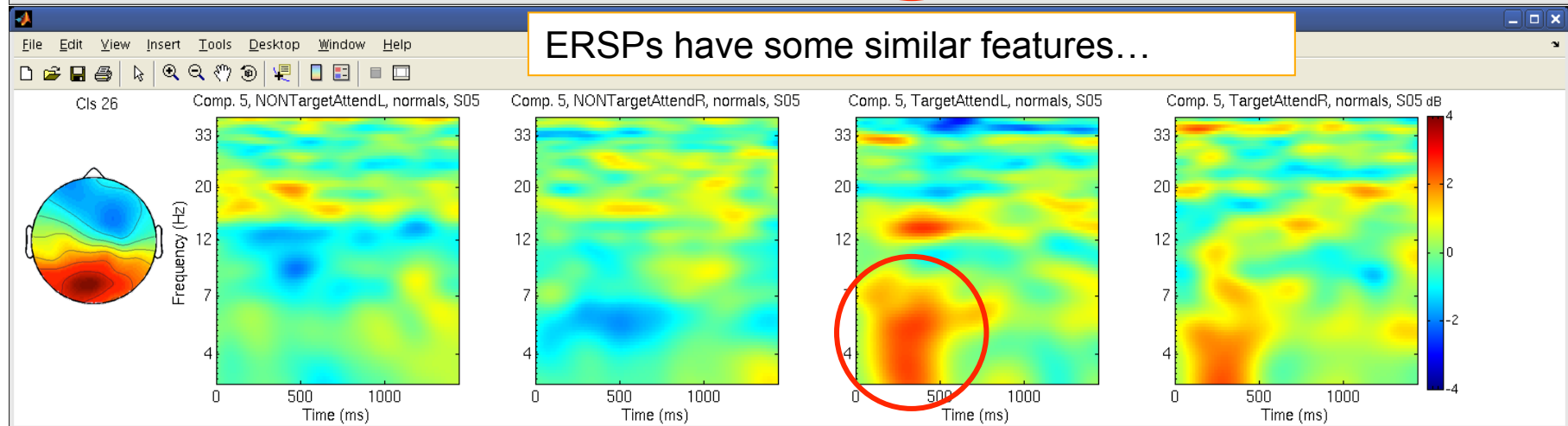


Spectra are similar, but they have
variable responses to different conditions...

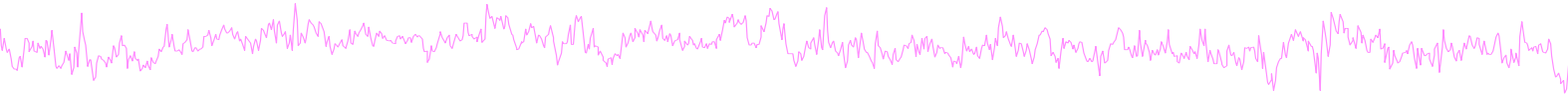
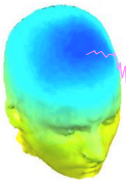
Choosing data measures



ERSPs have some similar features...



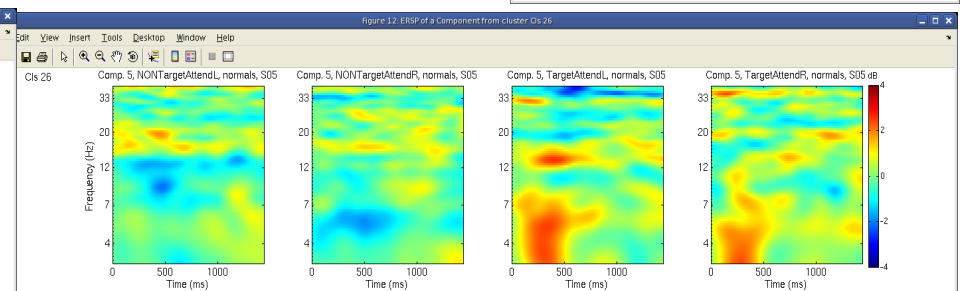
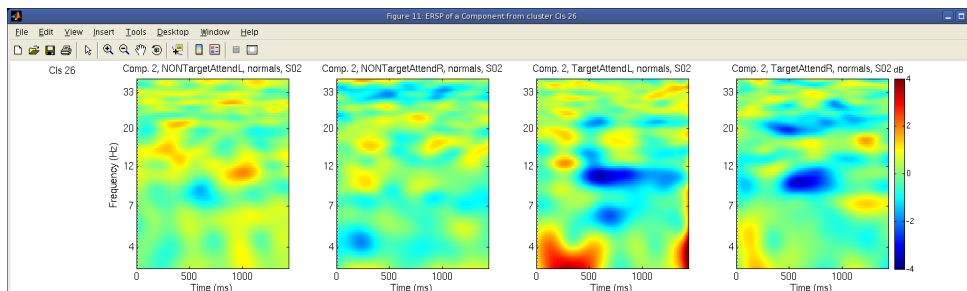
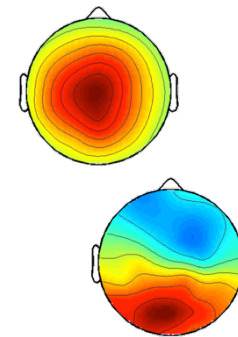
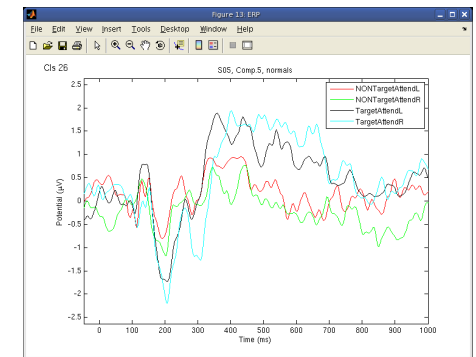
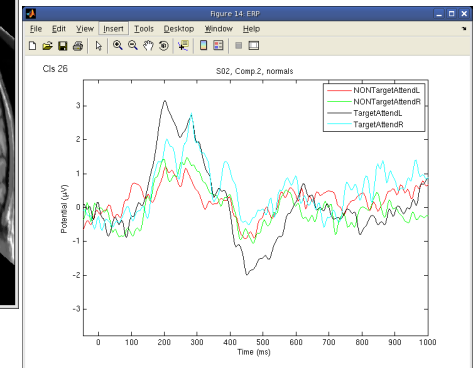
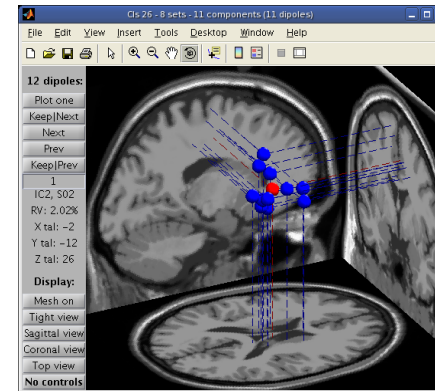
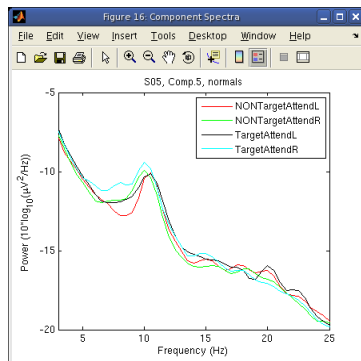
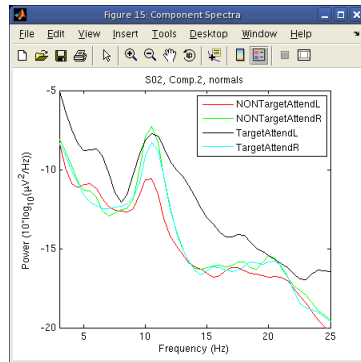
Choosing data measures



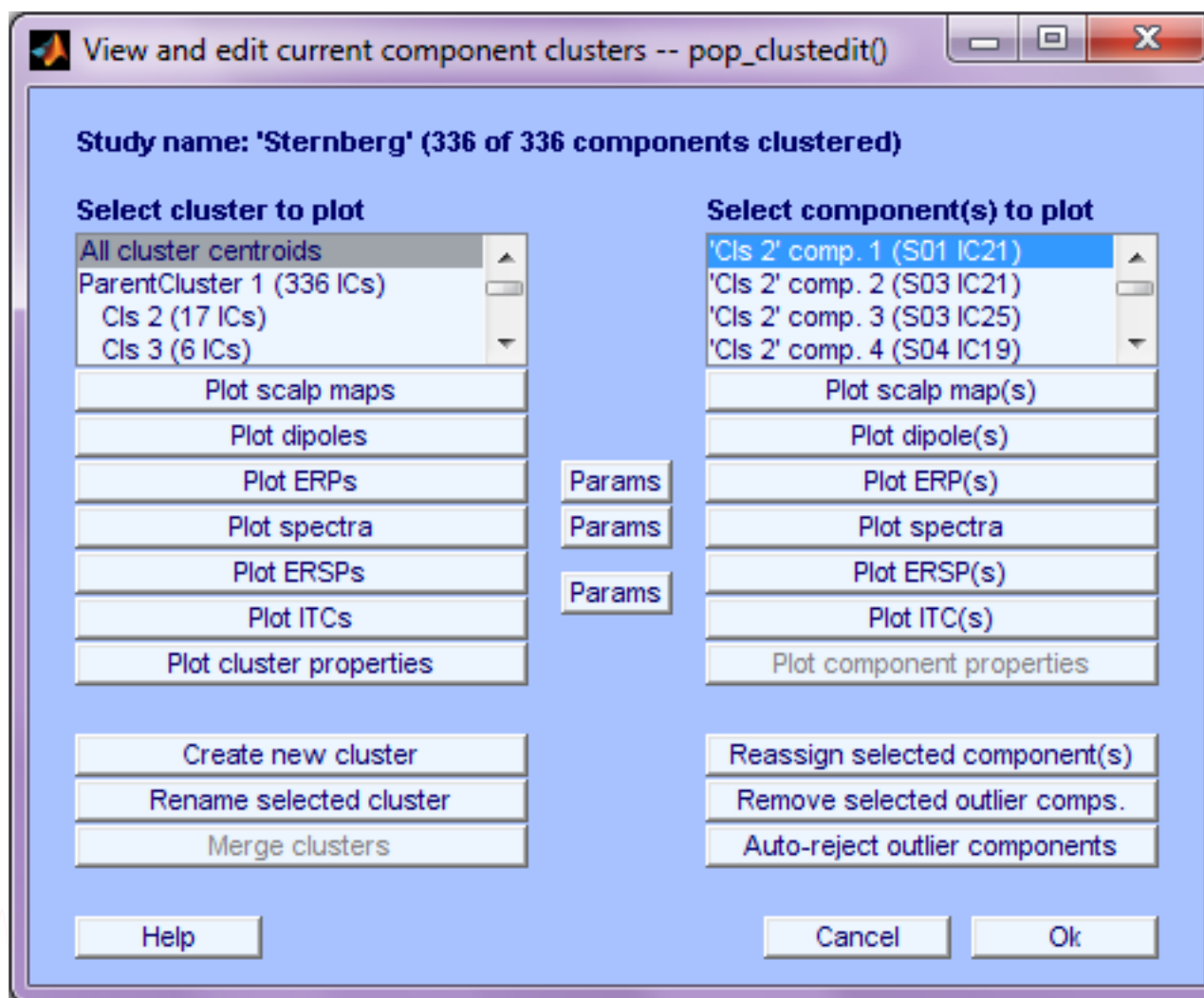
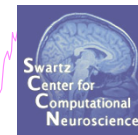
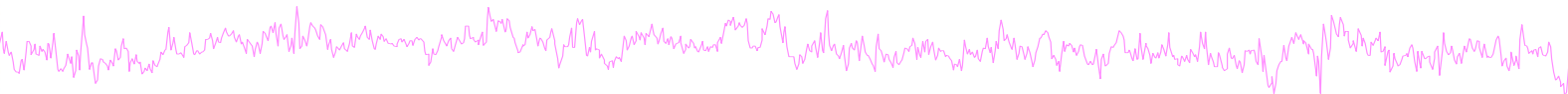
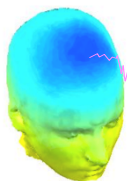
What data measures
should you use?

It depends...

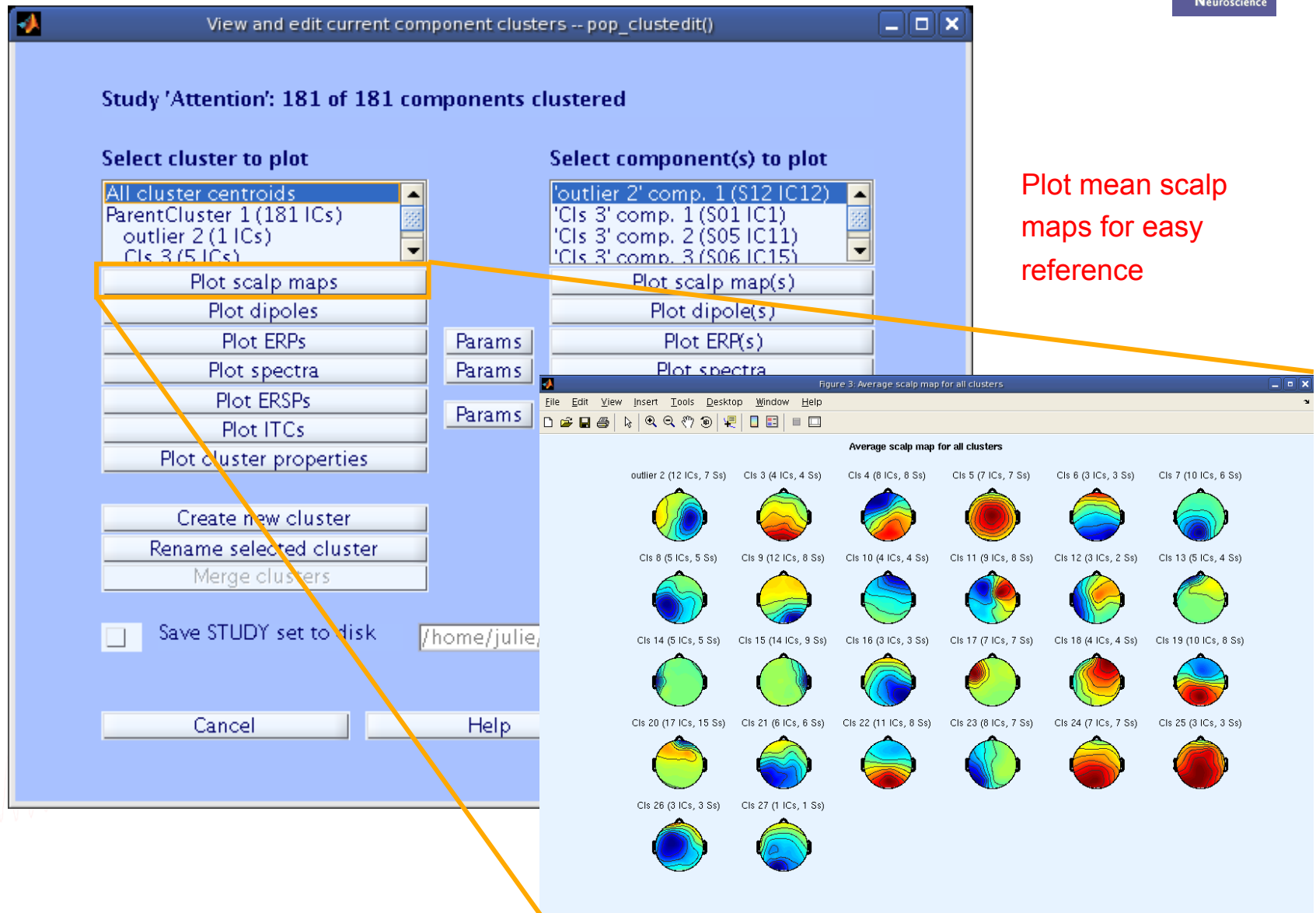
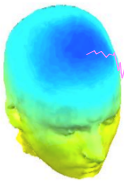
- broadly-matched ICs: use many/all of the measures.
- specifically-matched ICs: use one/few of the measures.



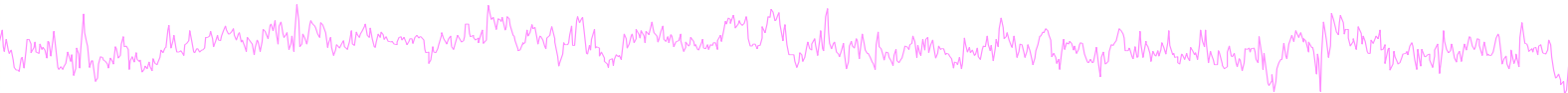
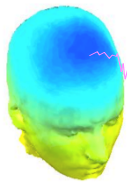
Plot/edit clusters



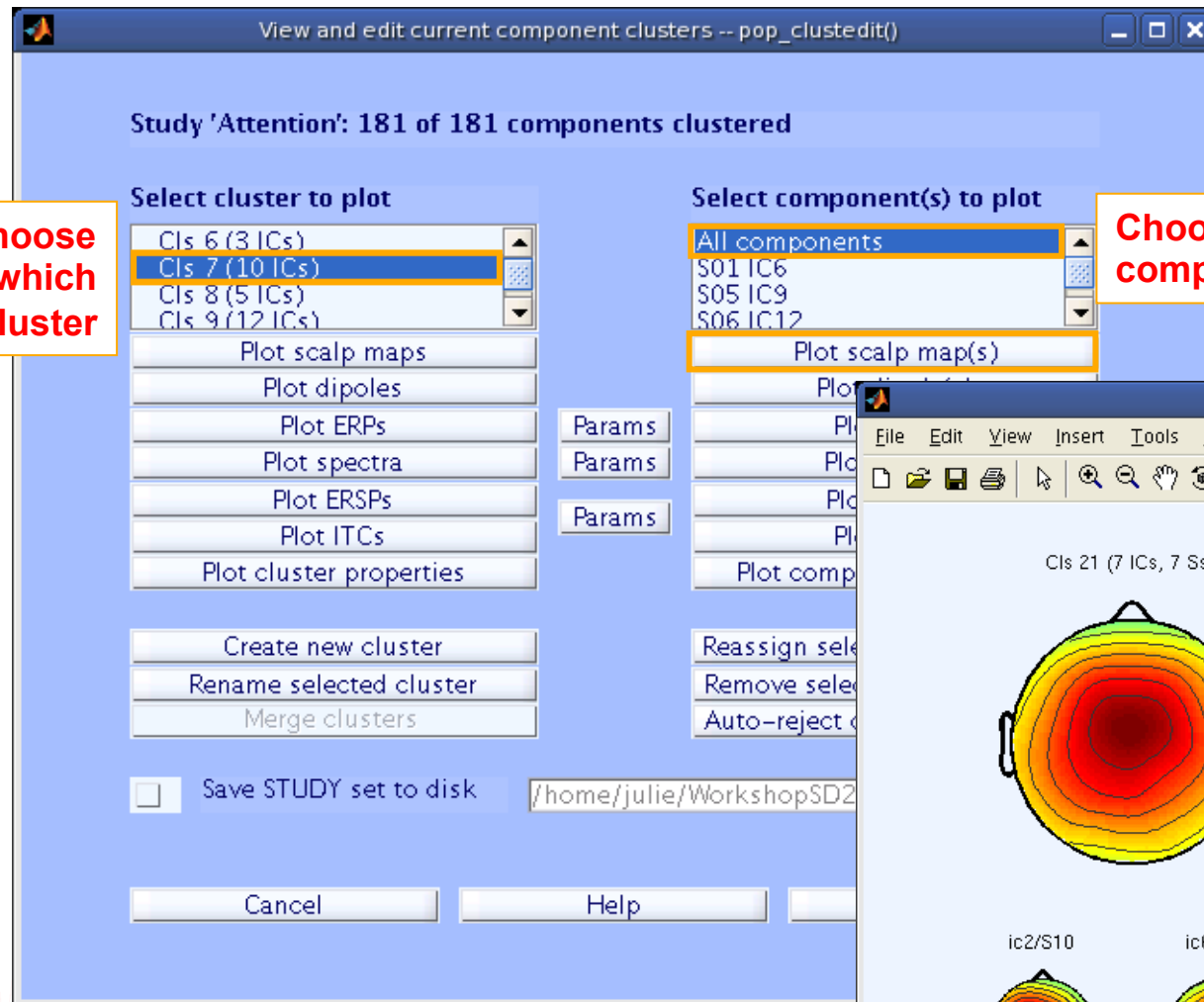
Plot cluster data



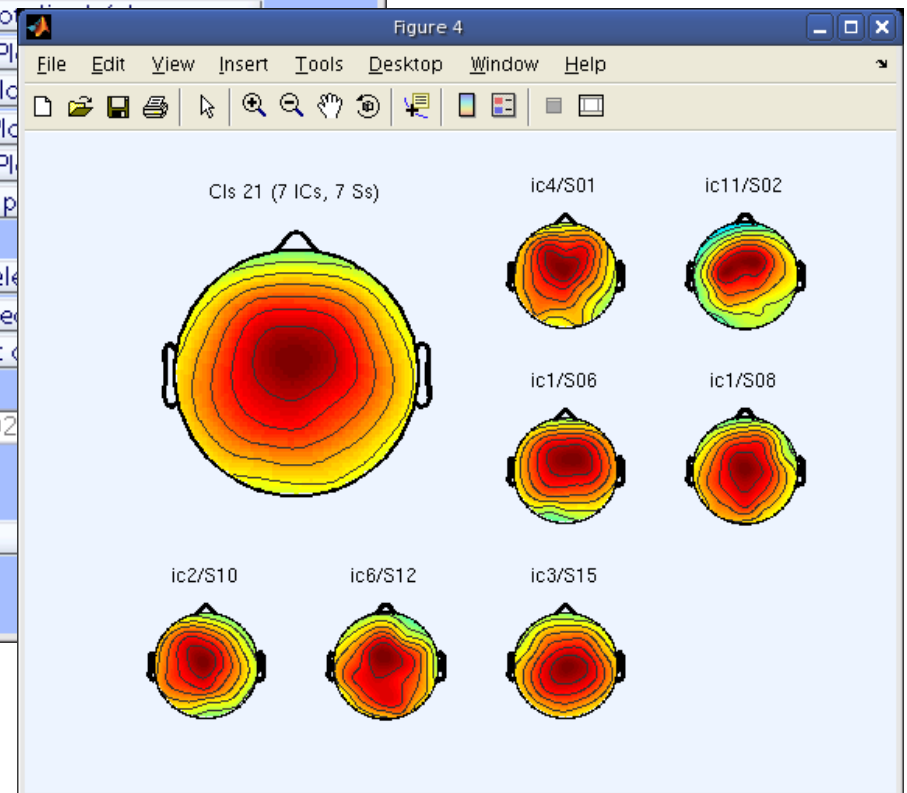
Plot cluster data



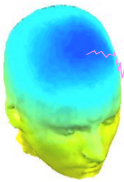
Choose which cluster



Choose which components



Plot cluster data



Clis 19 - 5 sets - 14 components (14 dipoles)

File Edit View Insert Tools Desktop Window Help

15 dipoles:
Plot one
Keep|Next
Next
Prev
Keep|Prev
1
IC3, S02
RV: 2.62%
X tal: -6
Y tal: -13
Z tal: 21
Display:
Mesh on
Tight view
Sagittal view
Coronal view
Top view
No controls

Plot scalp maps
Plot dipoles
Plot ERPs
Plot spectra
Plot ERSPs
Plot ITCs
Plot cluster properties
Create new cluster
Rename selected cluster
Merge clusters
☐ Save STUDY set to disk
Cancel Help Ok

component clusters -- pop_clustedit()

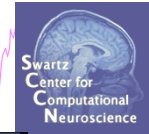
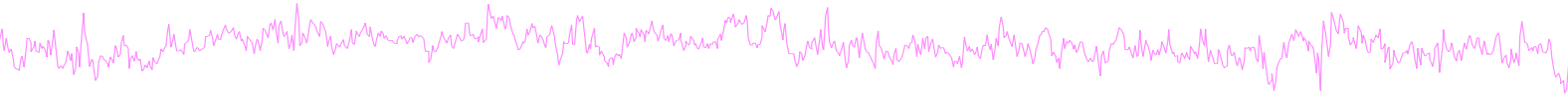
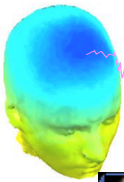
Select component(s) to plot

All components
S02 IC3
S02 IC11
S02 IC12
S02 IC17

Plot scalp map(s)
Plot dipole(s)
Plot ERP(s)
Plot spectra
Plot ERSP(s)
Plot ITC(s)
Plot component properties
Reassign selected component(s)
Remove selected outlier comps.
Auto-reject outlier components
Params
Params
Params

/home/julie/workshop06/5subjects/WSstudy.study

Plot cluster data



View and edit current component clusters -- pop_clustedit()

Study #: 151 of 151 components clustered

Select cluster to plot

- Cls 15 (8 ICs)
- Cls 16 (6 ICs)
- Cls 17 (4 ICs)
- Cls 18 (14 ICs)
- Cls 19 (14 ICs)

Plot scalp maps

Plot dipoles

Plot ERPs

Plot spectra

Params

Params

Select component(s) to plot

- All components
- S02 IC3
- S02 IC11
- S02 IC12
- S02 IC17

Plot scalp map(s)

Plot dipole(s)

Plot ERP(s)

Plot spectra

Set ERP plotting parameters -- pop_erpparams()

Time range in ms [low high]

Plot limits in uV [low high]

Plot scalp map at latency [ms]

Display filter in Hz [high]

☐ Plot conditions on the same panel

☐ Plot groups on the same panel

Statistical method to use

Statistical threshold (p<)

☐ Compute condition statistics

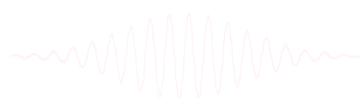
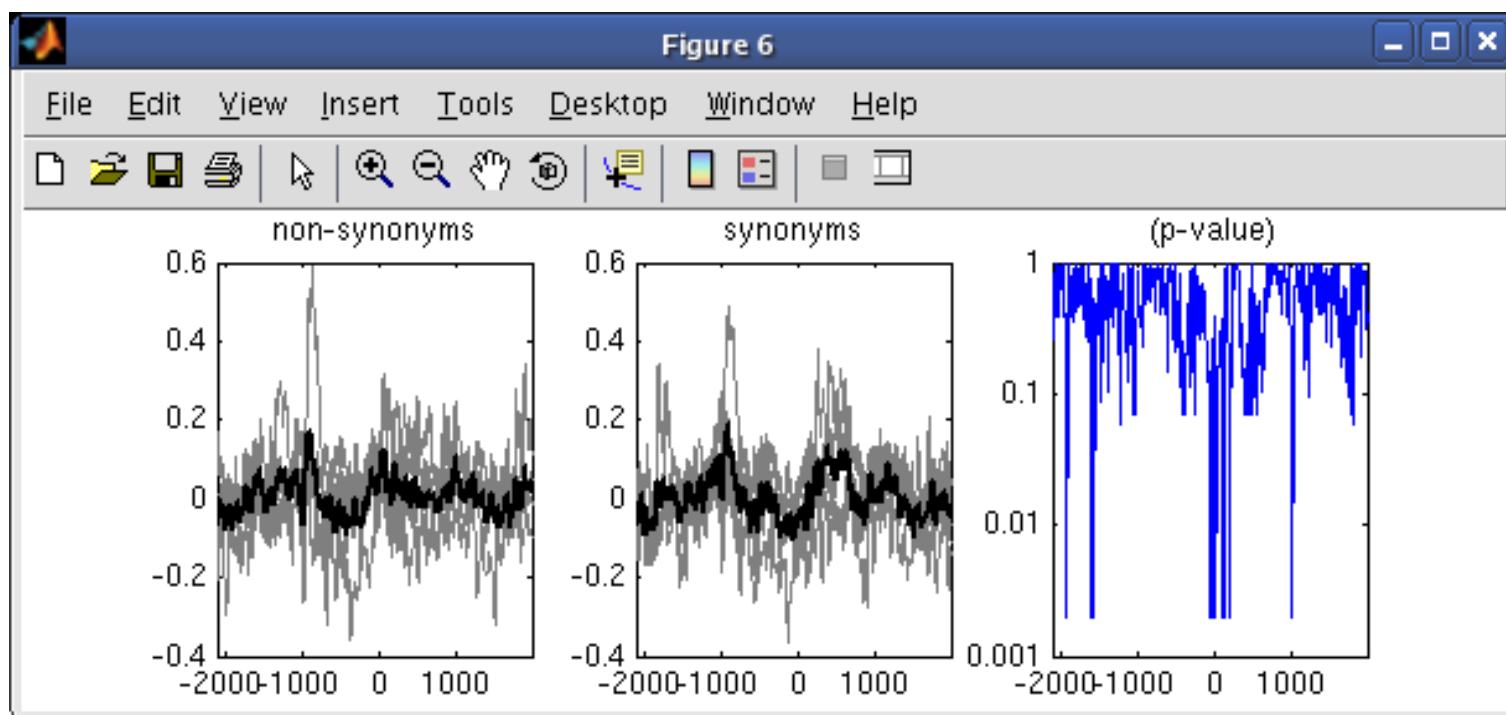
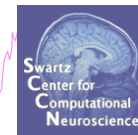
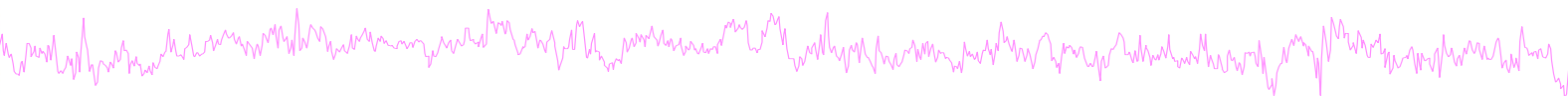
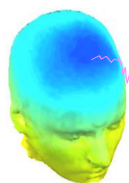
☐ Compute group statistics

☐ Use single trials (when available)

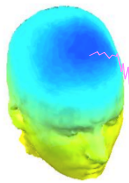
☐ Use False Discovery Rate to correct for multiple comparisons

Help Cancel Ok

Plot cluster ERP



Other plotting options...



Set ERP plotting parameters -- pop_erpparams()

Time range in ms [low high] Plot limits in uV [low high]

Plot scalp map at latency [ms] Display filter in Hz [high]

☐ Plot conditions on the same panel

☐ Plot groups on the same panel

Statistical method to use Statistical threshold (p<)

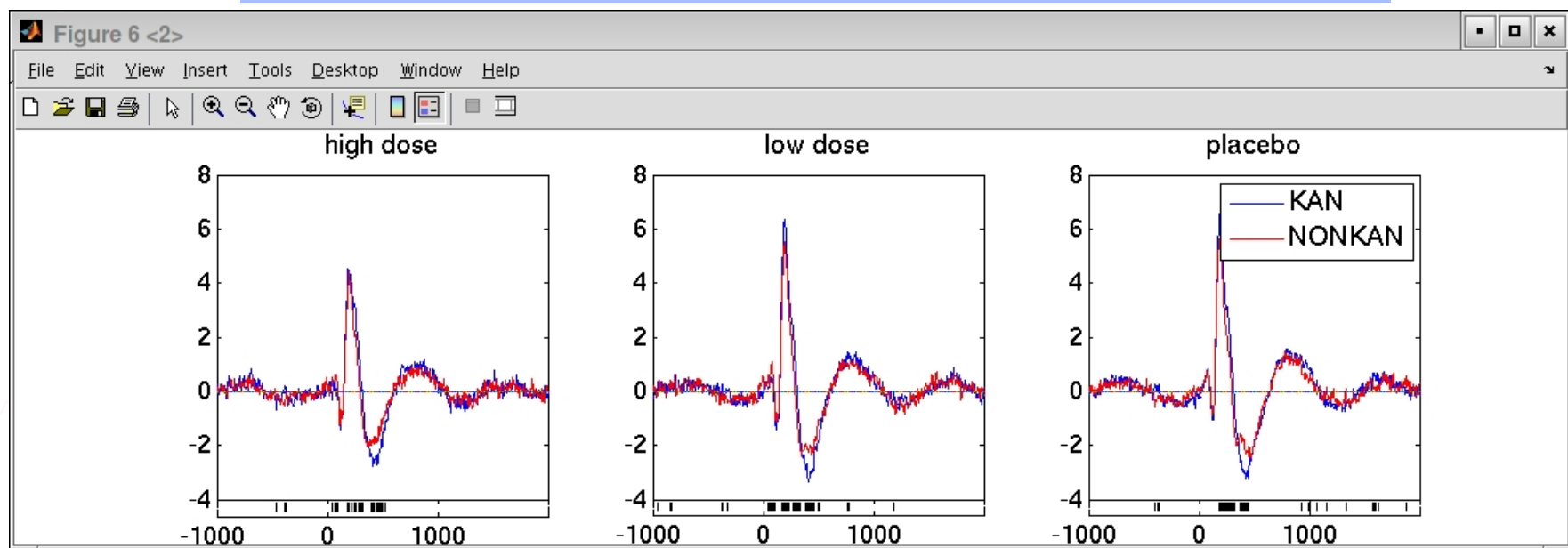
☐ Compute condition statistics

☐ Compute group statistics

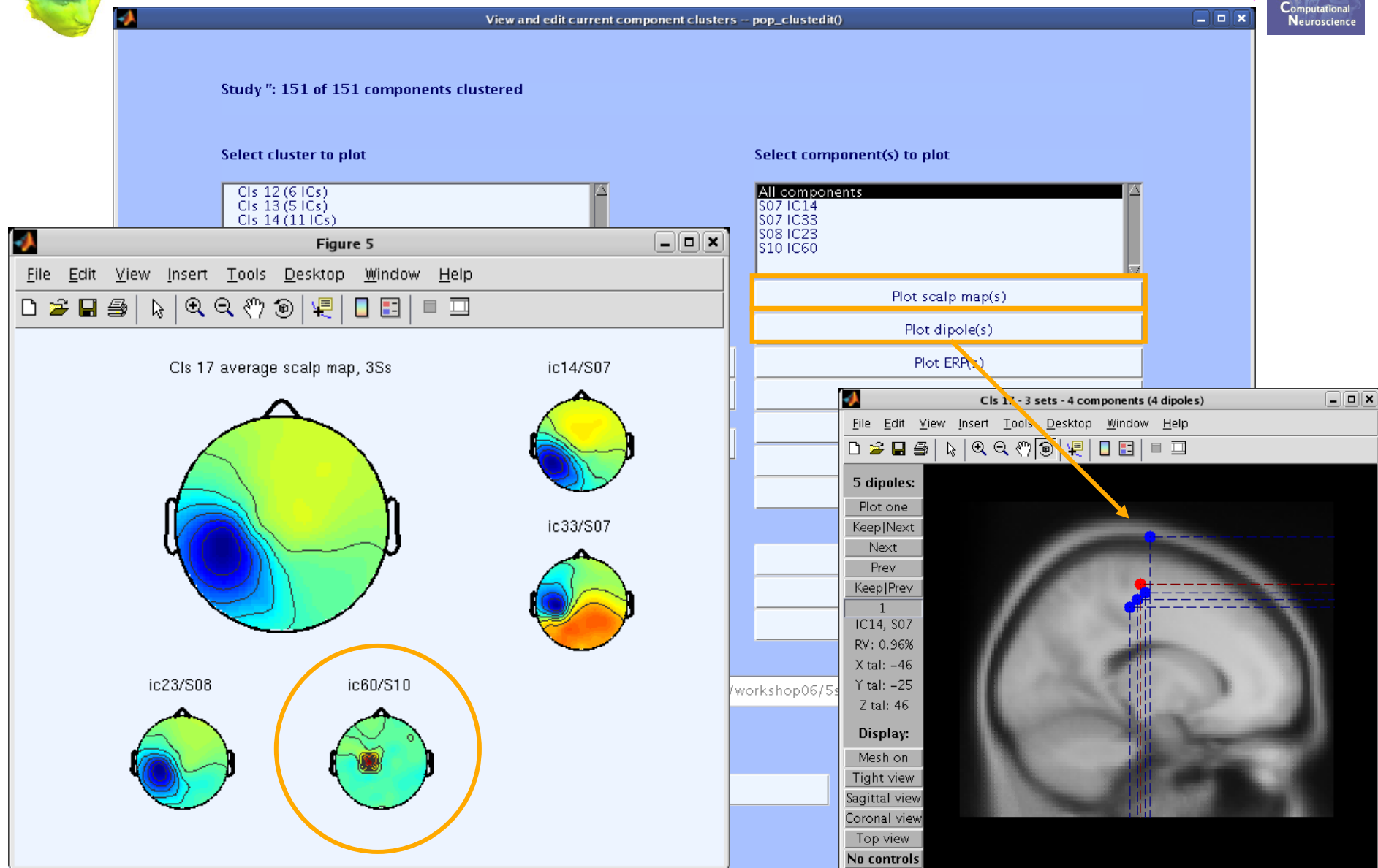
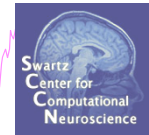
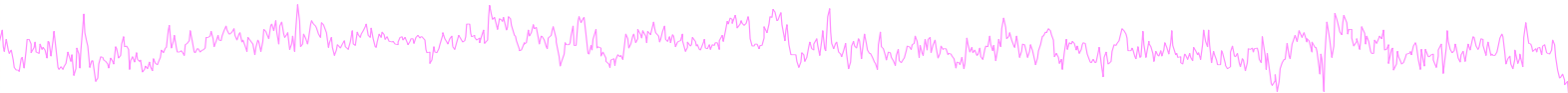
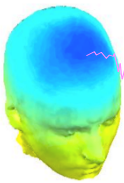
☐ Use single trials (when available)

☐ Use False Discovery Rate to correct for multiple comparisons

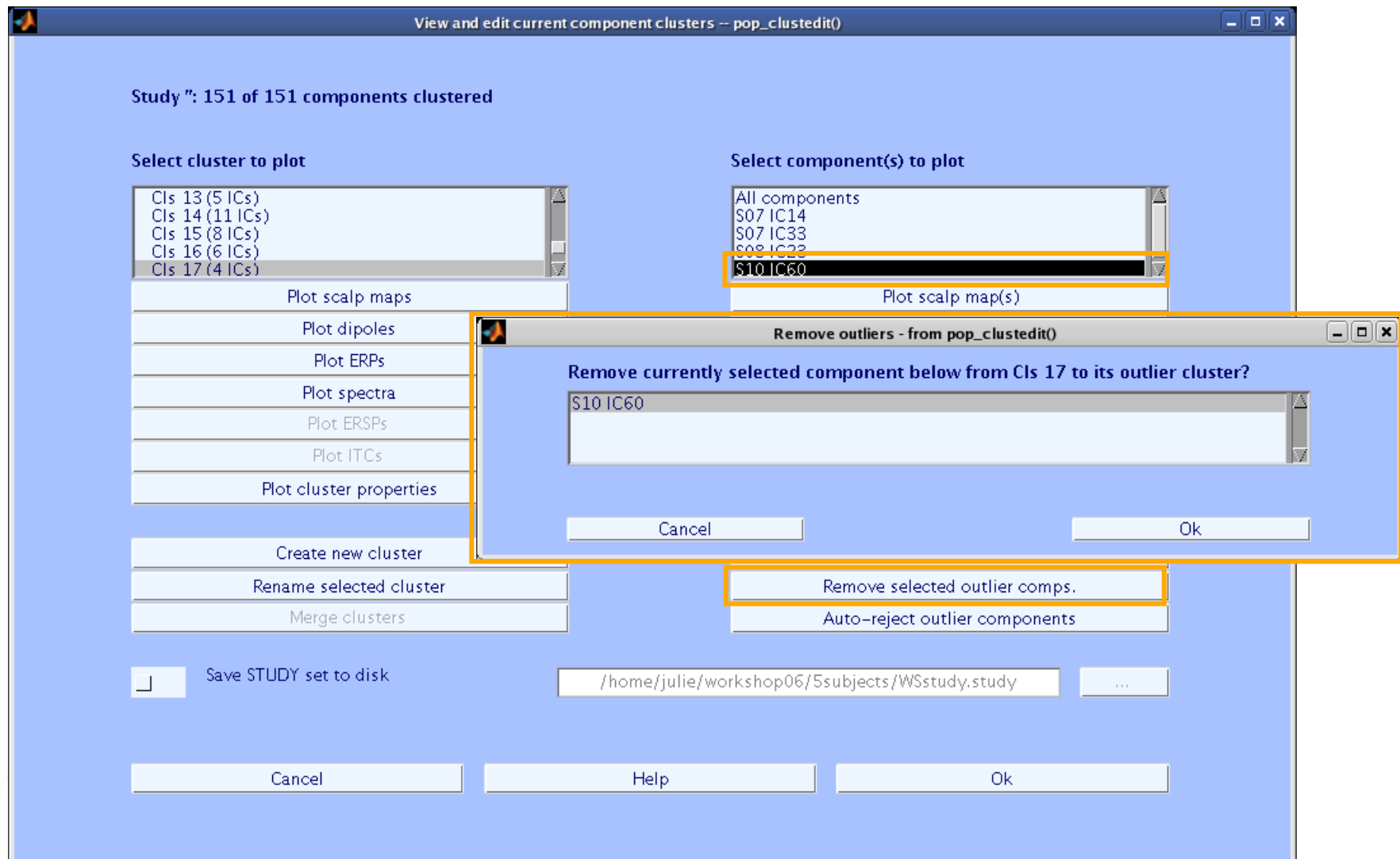
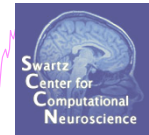
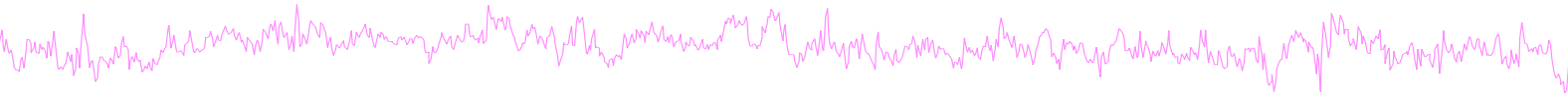
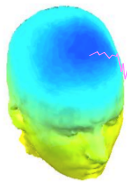
Help Cancel Ok



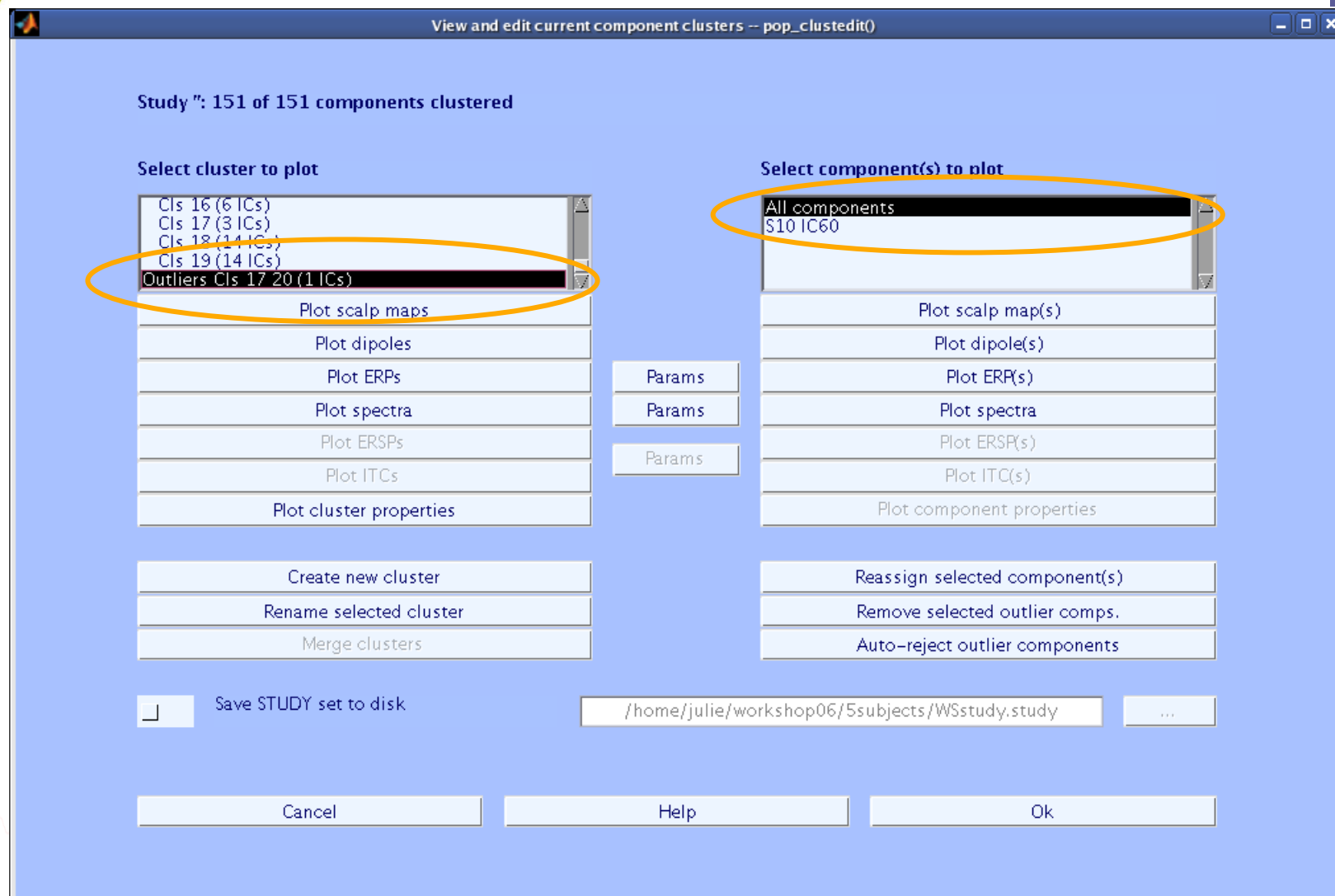
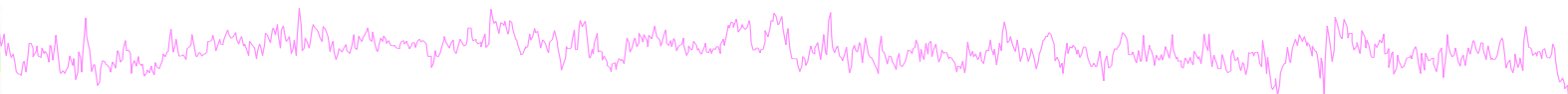
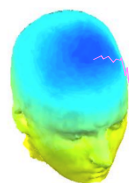
Reassigning components



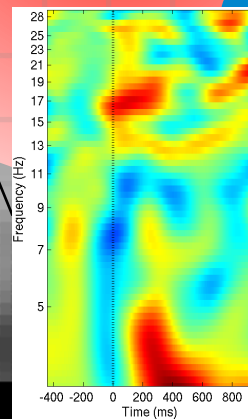
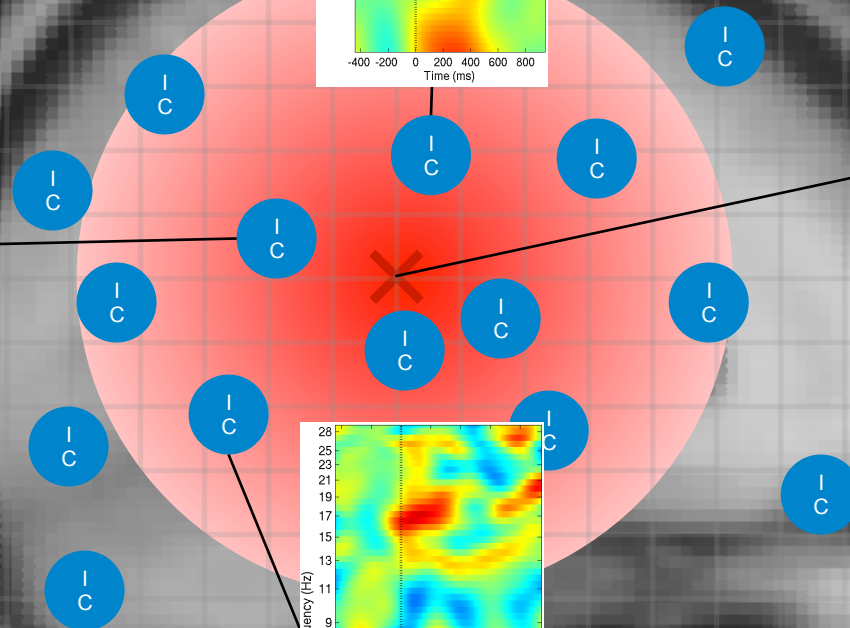
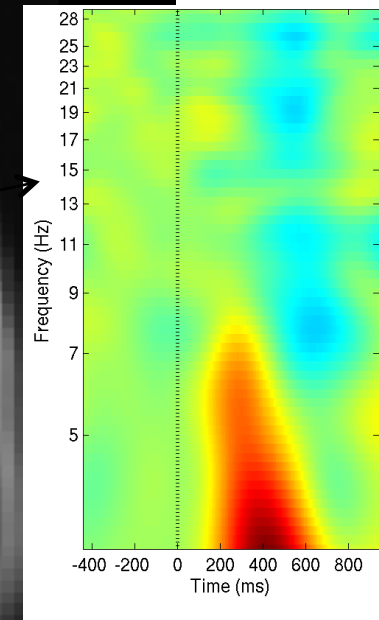
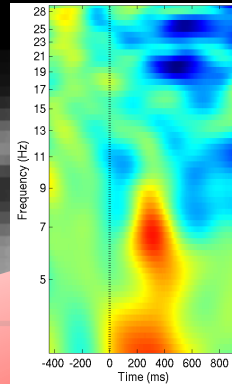
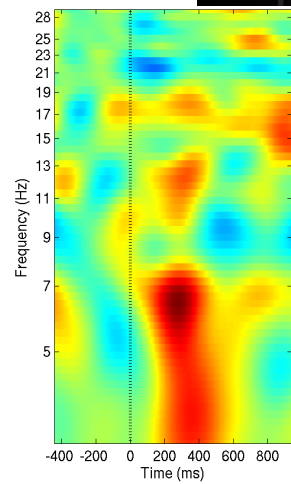
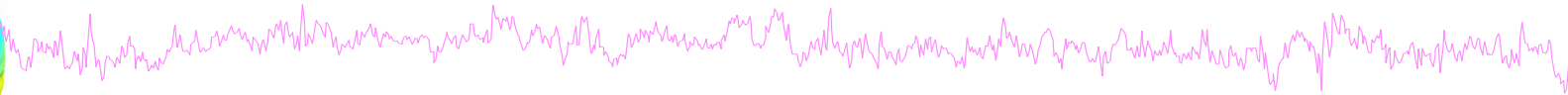
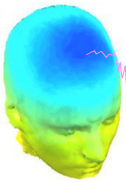
Reassigning components



Outlier cluster reassignment

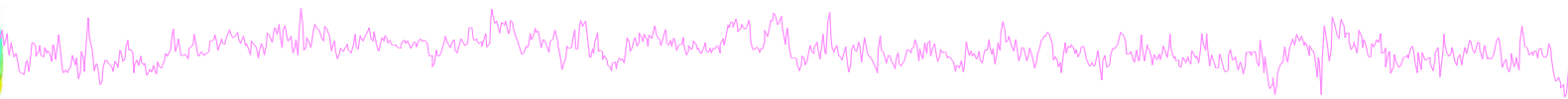
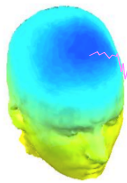


Measure Projection Toolbox

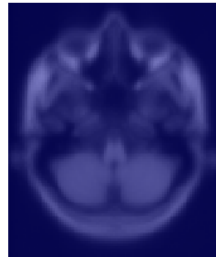


Projected ERSP

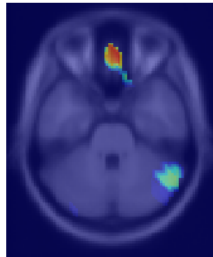
Measure Projection Toolbox



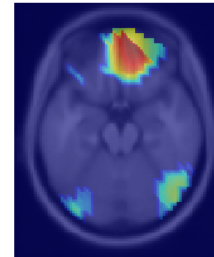
-50 mm



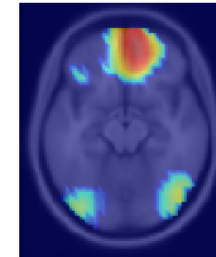
-30 mm



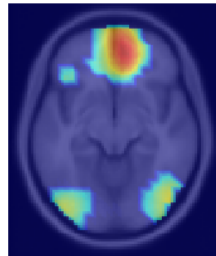
-20 mm



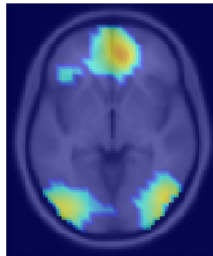
-15 mm



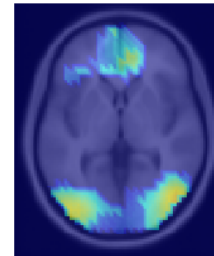
-10 mm



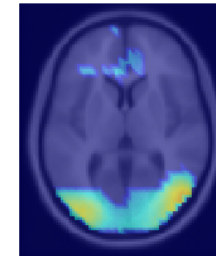
-5 mm



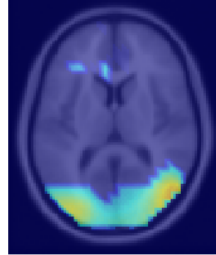
0 mm



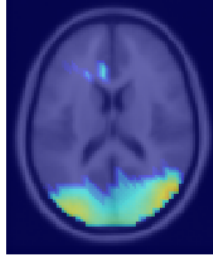
5 mm



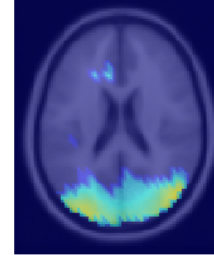
10 mm



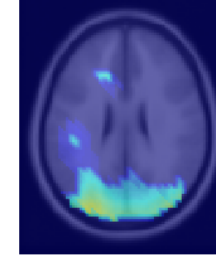
15 mm



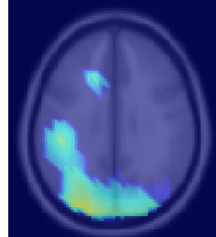
20 mm



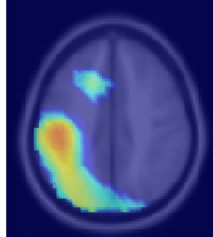
25 mm



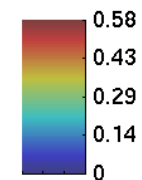
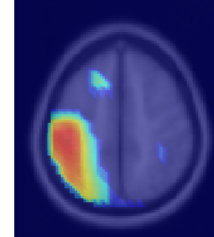
30 mm

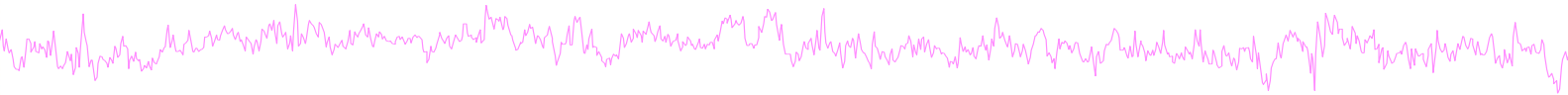
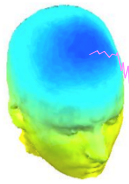


40 mm



50 mm





Exercise

Load the Stern STUDY (STUDY folder)

Precluster (pre-computation already done) and cluster components using measures of your choice. Experiment with different measures.

