

# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

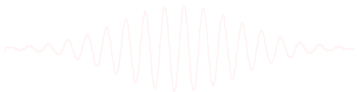
## STEP 3

Precluster the data

## STEP 4

Cluster the data

**Exercise...**



# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

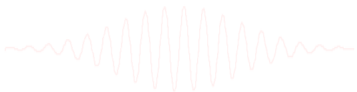
## STEP 3

Precluster the data

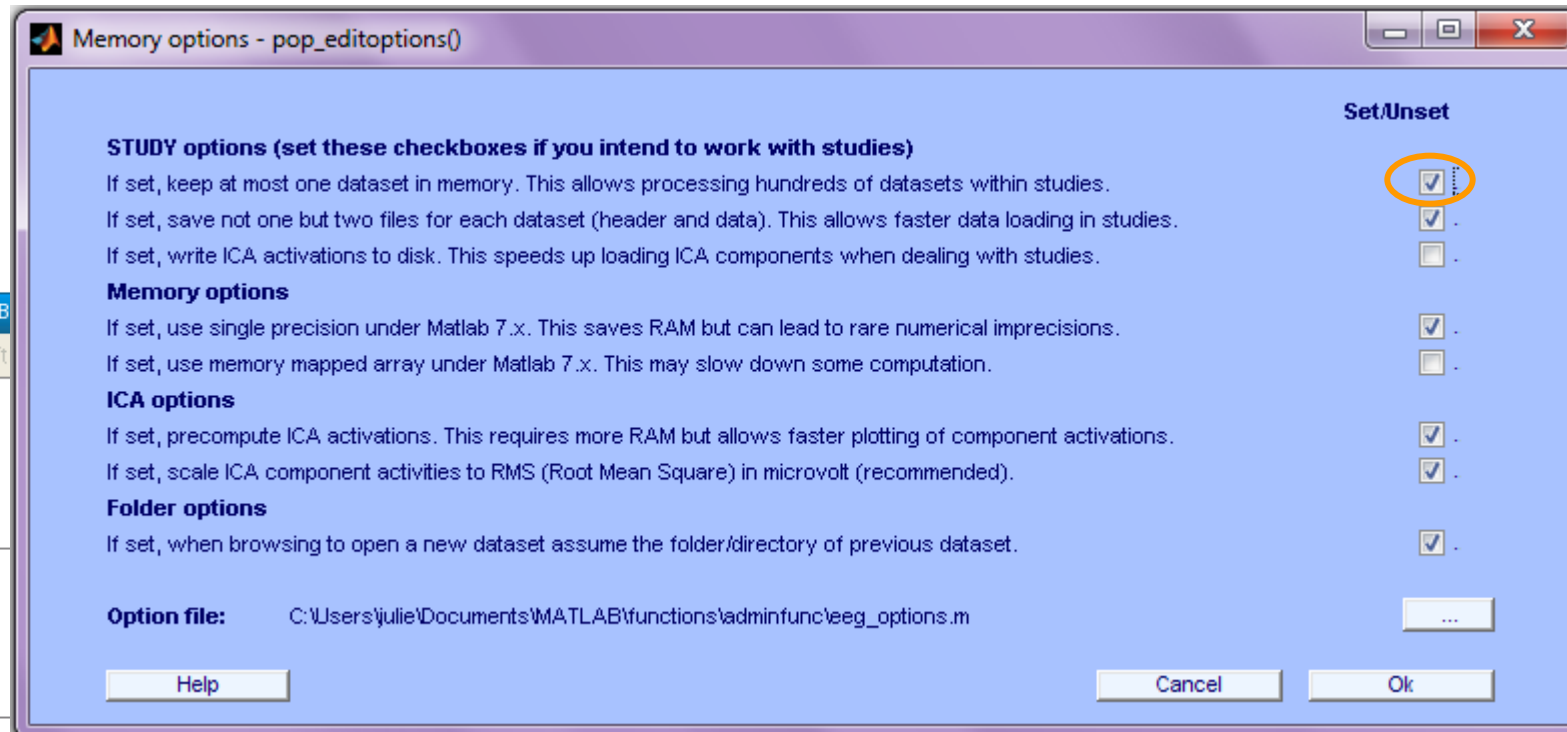
## STEP 4

Cluster the data

**Exercise...**

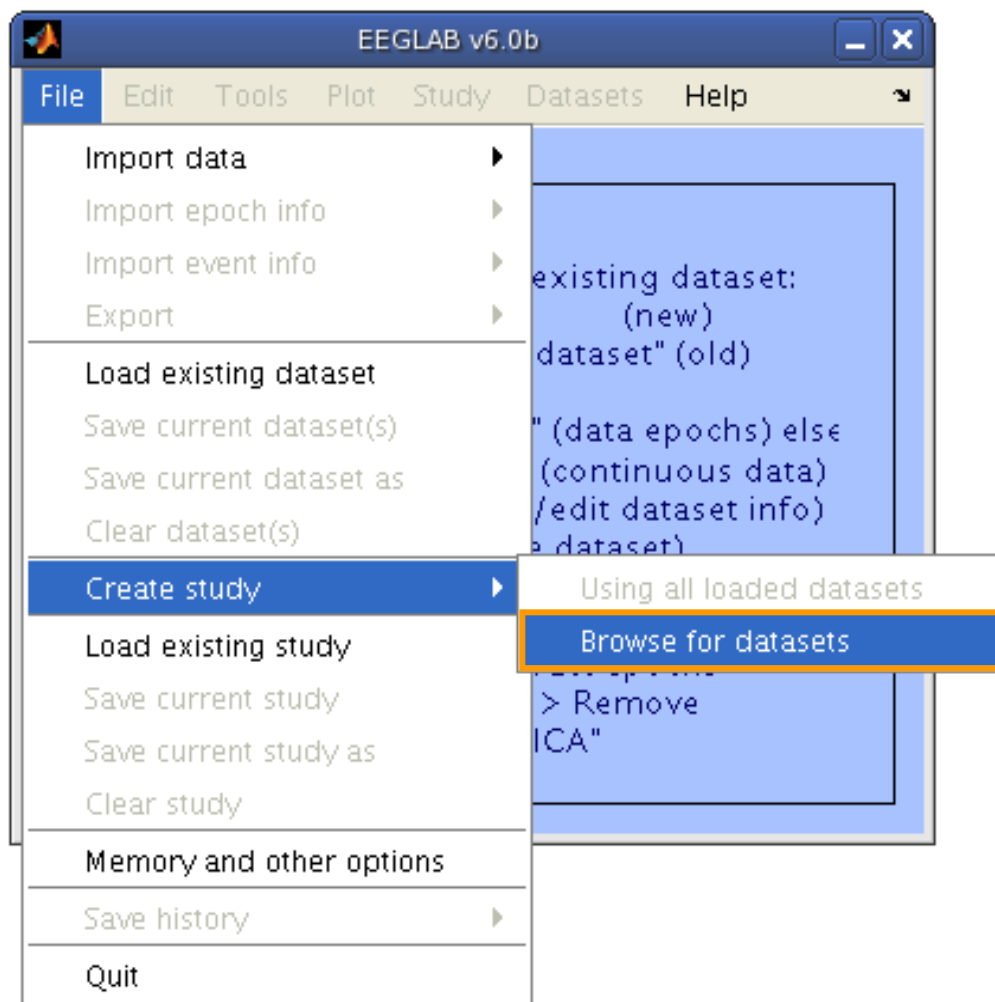


# Memory options

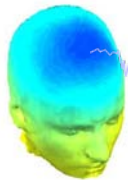


**Memory options should change  
when using STUDY vs single dataset**

# Build a STUDY



# Build a STUDY, cont'd



Create a new STUDY set -- pop\_study()

Create a new STUDY set

STUDY set name:

STUDY set task name:

STUDY set notes:

	dataset filename	browse	subject	session	condition	group
1	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
6	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
7	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
8	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
9	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
10	<input type="text"/>	...	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Select by r.v.

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

< Page 1 >

☒ Update dataset info - datasets stored on disk will be overwritten (unset = Keep study info set)

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

Help

Choose dataset to add to STUDY -- pop\_study()

Look in: S01

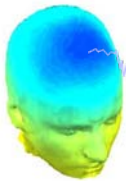
Name	Date modified	Type
Ignore.set	11/8/2009 7:06 PM	SET File
Memorize.set	11/8/2009 7:06 PM	SET File
Probe.set	11/12/2009 10:02 ...	SET File

File name:

Files of type: (\*.set, \*.SET)

Open Cancel

# 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

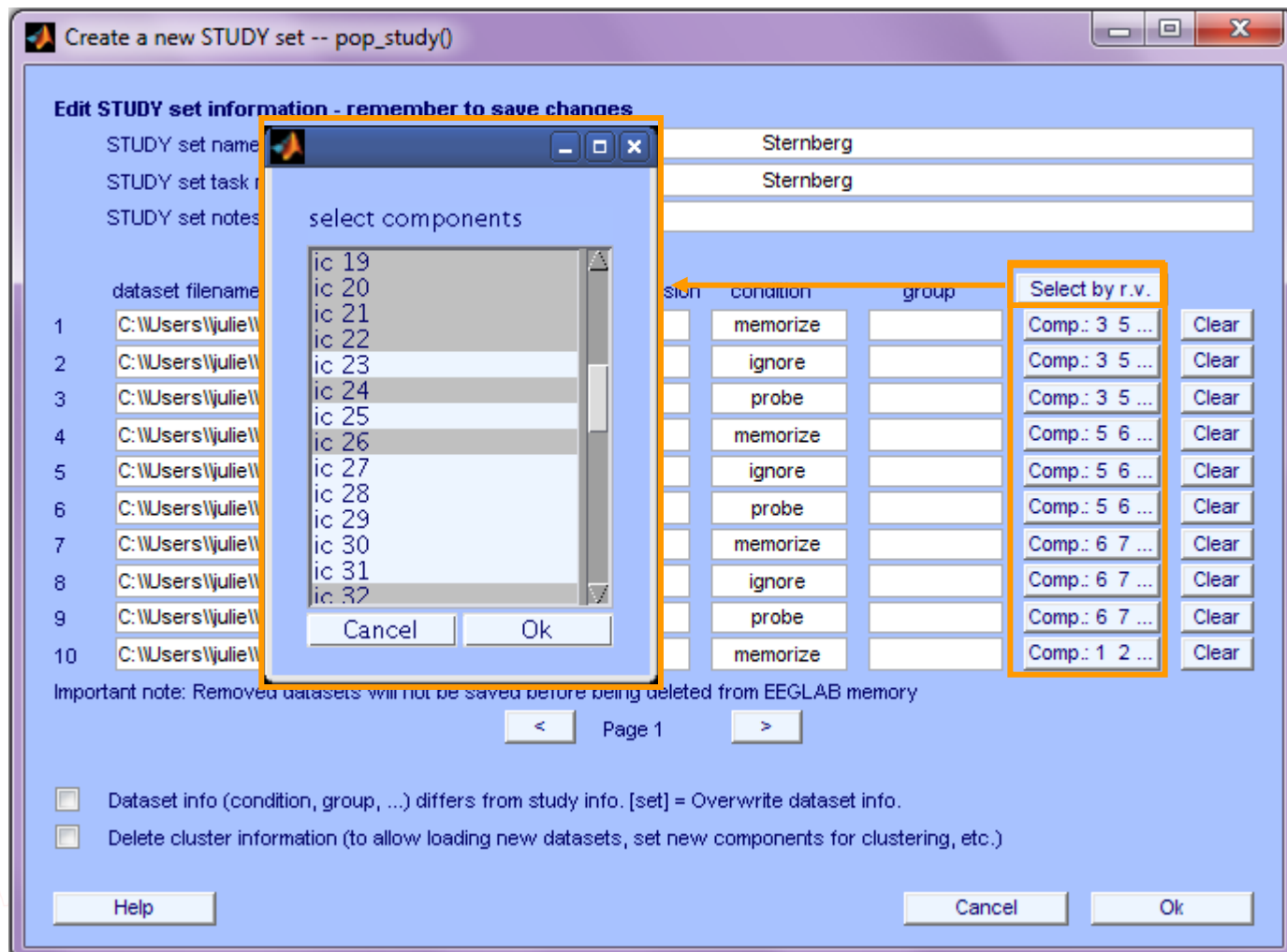
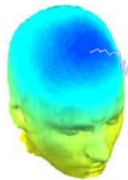
< Page 1 >

☐ 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



# Build a STUDY



## Most important option:

- Allows only one dataset to be loaded at once.
- Most STUDYs are too big to have all data loaded at once.

% Open eeglab:

```
[ALLEEG EEG CURRENTSET ALLCOM] = eeglab;
```

% Set memory options:

```
pop_editoptions( 'option_storedisk', 1, 'option_savetwofiles', 1,...  
    'option_saveica', 1, 'option_single', 0, 'option_memmapdata',...  
    0, 'option_computeica', 1, 'option_scaleicarms', 0,...  
    'option_rememberfolder', 1);
```

% saves a file 'eeg\_options.m' to your current working directory

% Initialize EEGLAB/STUDY variables:

```
STUDY = []; CURRENTSTUDY = 0; ALLEEG=[]; EEG=[]; CURRENTSET=[];
```



# Build a STUDY



% Faster alternative to building a STUDY manually

% Example STUDY: 13 subjects, 3 conditions

% Define variables:

```
basedir = 'C:\...\EEGLAB_WORKSHOP\STUDY\';
```

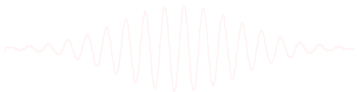
```
setnames = {'Memorize.set','Ignore.set','Probe.set'};
```

```
subjs = {'S01','S02','S03','S04','S05','S06','S07',...  
'S08','S09','S10','S11','S12','S13'};
```

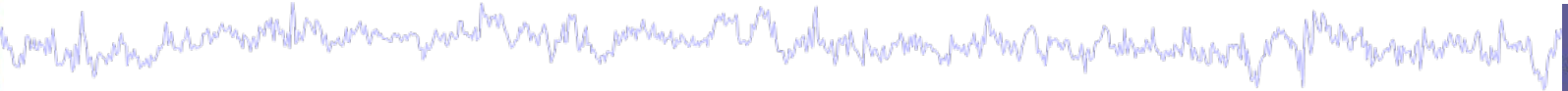
```
studyname = 'Sternberg';
```

```
taskname = 'Sternberg';
```

```
savename = 'stern.study';
```



# Define variables

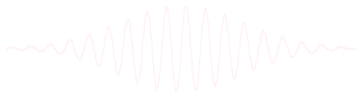


`% concatenate string variables:`

`[] % strings inside brackets will be concatenated`

```
dataset = [basedir,subjs{subj},'\',setnames{cond}];
```

```
C:\EEGLAB_Workshop\STUDY\ S01\ Memorize.set
```



# Load dataset info from commandline



```
%Now loop through subjects and add to the STUDY:
index = 1; %initialize STUDY index
for subj = 1:length(subjs) %for each subject
    for cond = 1:length(setnames) %for each condition
        dataset = [basedir,subjs{subj},'\',setnames{cond}];
        [STUDY ALLEEG] = std_editset( STUDY, ALLEEG,...
            'name', studyname, 'task', taskname,...
            'commands',{{'index',index,'load',dataset},...
                {'dipselect',0.15},{ 'subject',subjs{subj}},...
                {'condition',conds{cond}}},...
            'inbrain','on','updatedat','off',...
            'savedat', 'off','filename', [basedir, savename]);
        index = index + 1;
        CURRENTSTUDY = 1; EEG = ALLEEG; CURRENTSET = [1:length(EEG)];
        [STUDY, ALLEEG] = std_checkset(STUDY, ALLEEG);
    end;
end;
eeglab redraw
```

# STUDY structure

STUDY =

```
    name: 'Sternberg'
    task: 'Sternberg'
datasetinfo: [1x39 struct]
    notes: ''
filename: 'stern.study'
filepath: 'C:\Users\julie\Documents\Workshops\Finland\STUDY'
history: [1x7332 char]
subject: {1x13 cell}
    group: {''}
session: []
condition: {'ignore' 'memorize' 'probe'}
    setind: [3x13 double]
        etc: [1x1 struct]
preclust: [1x1 struct]
    cluster: [1x1 struct]
changrp: [1x71 struct]
    saved: 'yes'
```

>>

# Subject info in STUDY structure



**Gives information  
for each dataset  
of each subject**

```
>> STUDY.datasetinfo
```

```
ans =
```

```
1x39 struct array with fields:
```

```
    filepath
```

```
    filename
```

```
    subject
```

```
    session
```

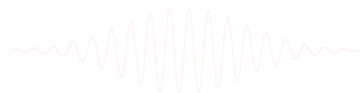
```
    condition
```

```
    group
```

```
    index
```

```
    comps
```

```
>>
```



# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

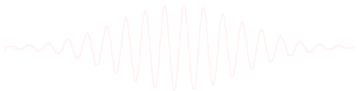
## STEP 3

Precluster the data

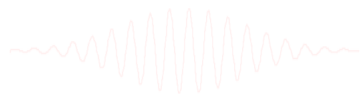
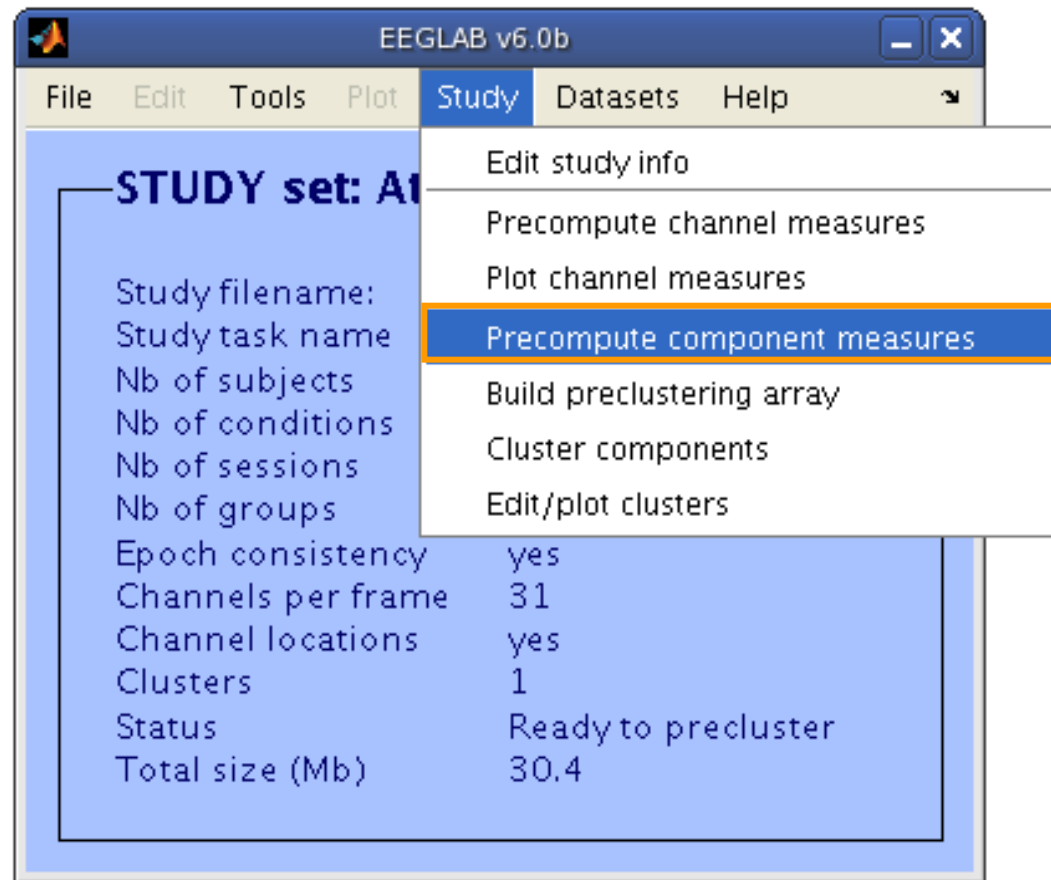
## STEP 4

Cluster the data

**Exercise...**



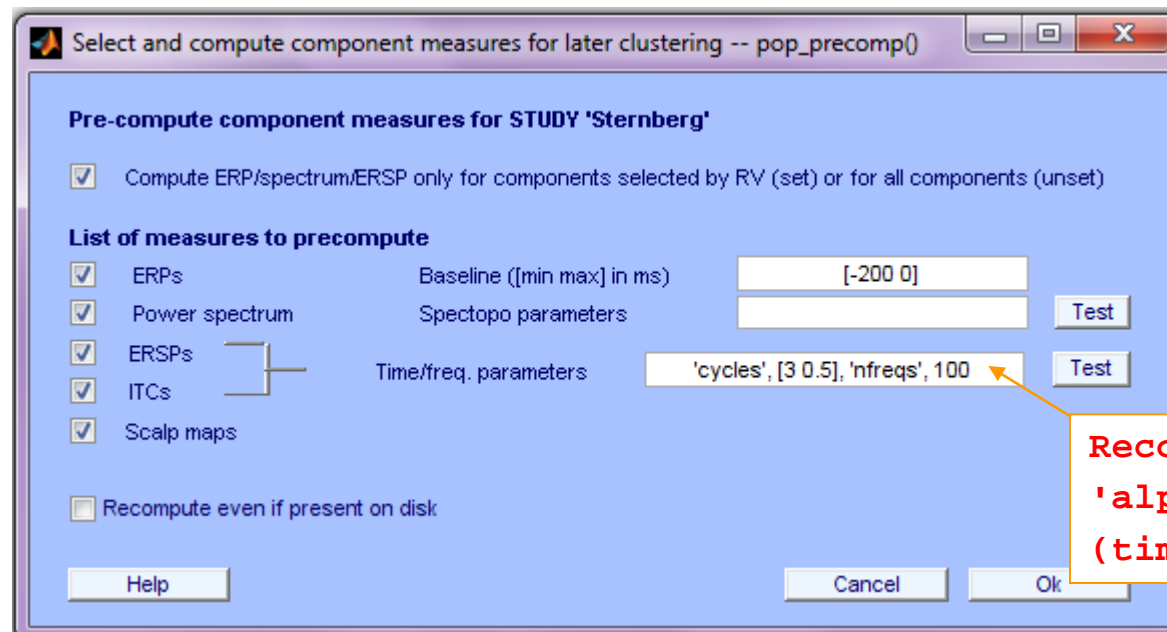
# Precompute data measures



# Precompute data measures



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



**Recommend:**  
**'alpha', .01**  
**(time-consuming)**

```
[STUDY ALLEEG] = std_precomp(STUDY, ALLEEG, 'components', ...  
'erp', 'on', 'rmbase', [-200 0], 'scalp', 'on', 'spec', ...  
'on', 'specparams', {}, 'ersp', 'on', 'erspparams', ...  
{ 'cycles', [3 0.5], 'nfreqs', 100, 'freqs', [3 70], ...  
'alpha', 0.01 }, 'itc', 'on');
```



# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

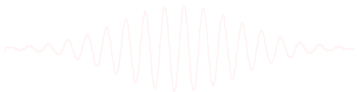
## STEP 3

Precluster the data

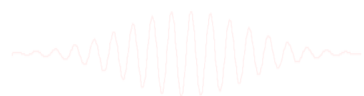
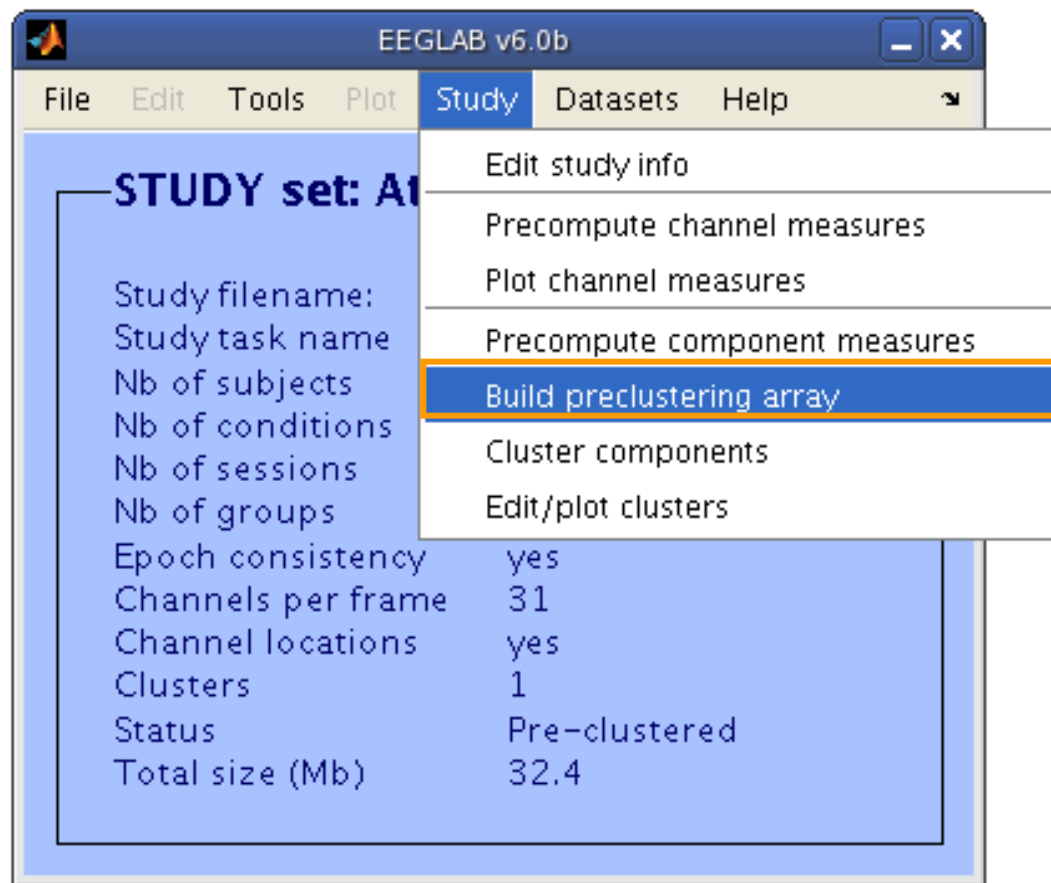
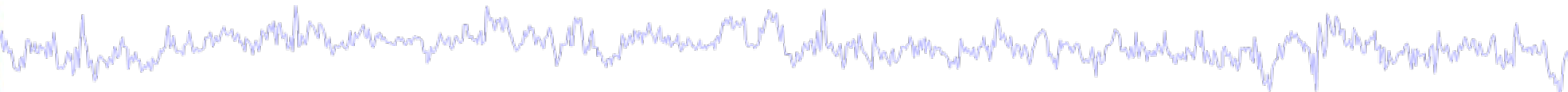
## STEP 4

Cluster the data

**Exercise...**



# Precluster the data



# Precluster the data



Select and compute component measures for later clustering -- pop\_precl...

**Build pre-clustering matrix for STUDY set: Sternberg**

Select the cluster to refine by sub-clustering (any existing sub-hierarchy will be overwritten)

ParentCluster 1 (336 ICs)

Note: Only measures that have been precomputed may be used for clustering.

Measures	Dims.	Norm.	Rel. Wt.
<input checked="" type="checkbox"/> spectra	10	<input checked="" type="checkbox"/> 1	
<input type="checkbox"/> ERPs	10	<input checked="" type="checkbox"/> 1	
<input checked="" type="checkbox"/> dipoles	3	<input checked="" type="checkbox"/> 10	
<input type="checkbox"/> scalp maps	10	<input checked="" type="checkbox"/> 1	
<input checked="" type="checkbox"/> ERSPs	10	<input checked="" type="checkbox"/> 1	
<input type="checkbox"/> ITCs	10	<input checked="" type="checkbox"/> 1	
<input type="checkbox"/> Final dimensions	10		

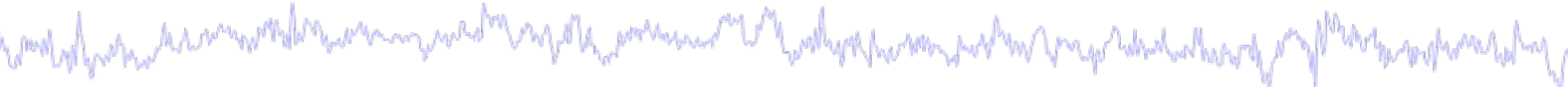
☐ Use Measure Product clustering

Help

Cancel Ok

```
parentclust = 1; % cluster 1 is always full parent cluster
[STUDY ALLEEG] = std_preclust(STUDY, ALLEEG, parentclust,{'spec','npca',5,...
    'norm',1,'weight',1,'freqrange',[3 25]},{ 'erp','npca',6,'norm',1,...
    'weight',1, 'timewindow',[0 400]},{ 'scalp','npca',10,'norm',1,'weight',1,...
    'abso',1},{ 'dipoles','norm',1,'weight',10},{ 'ersp','npca',20,...
    'freqrange',[3 30] , 'timewindow',[0 600], 'norm',1,'weight',1},{ 'itc',...
    'npca',6,'freqrange',[3 30], 'timewindow',[0 400] , 'norm',1, 'weight',1});
```

# 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...

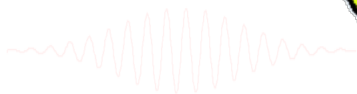
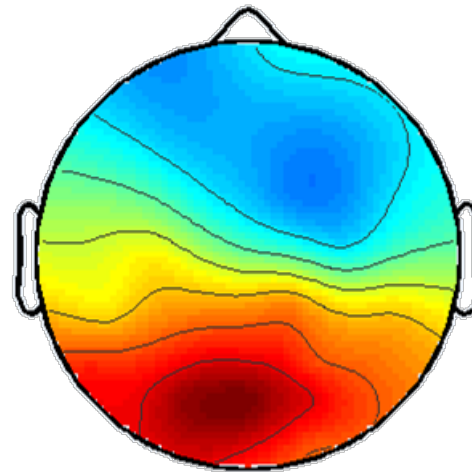
But consider:

- What is the difference between these two components?

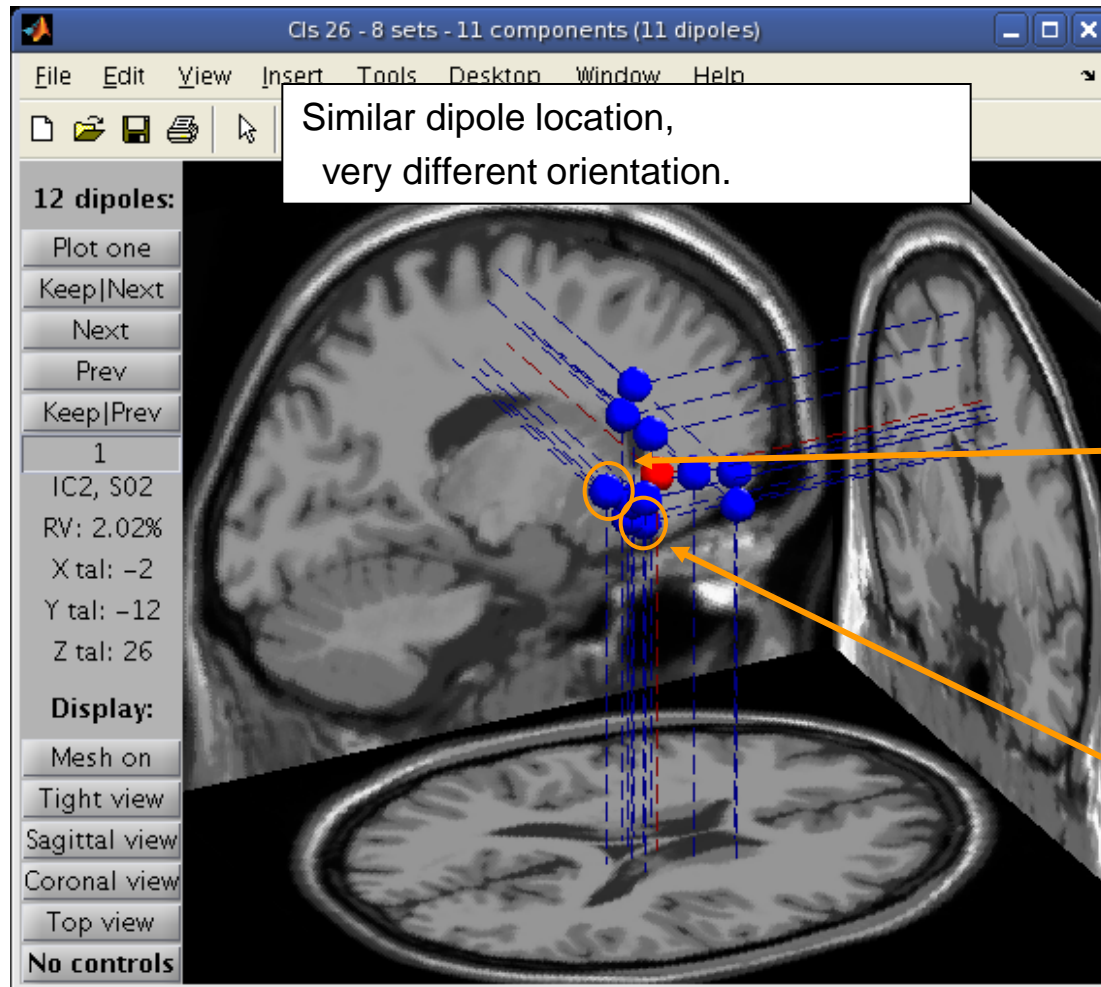
IC2 / S02, Cls 26



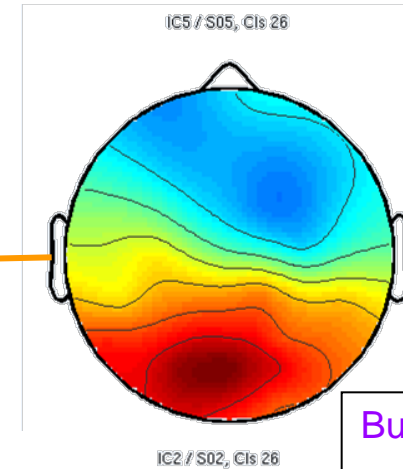
IC5 / S05, Cls 26



# Choosing data measures

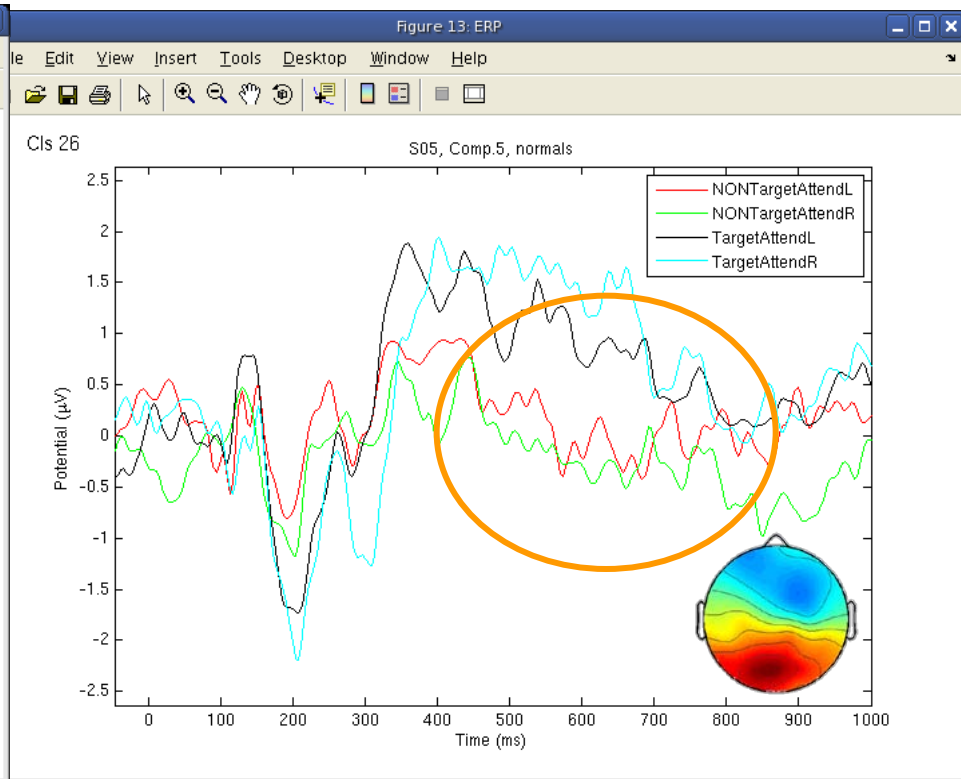
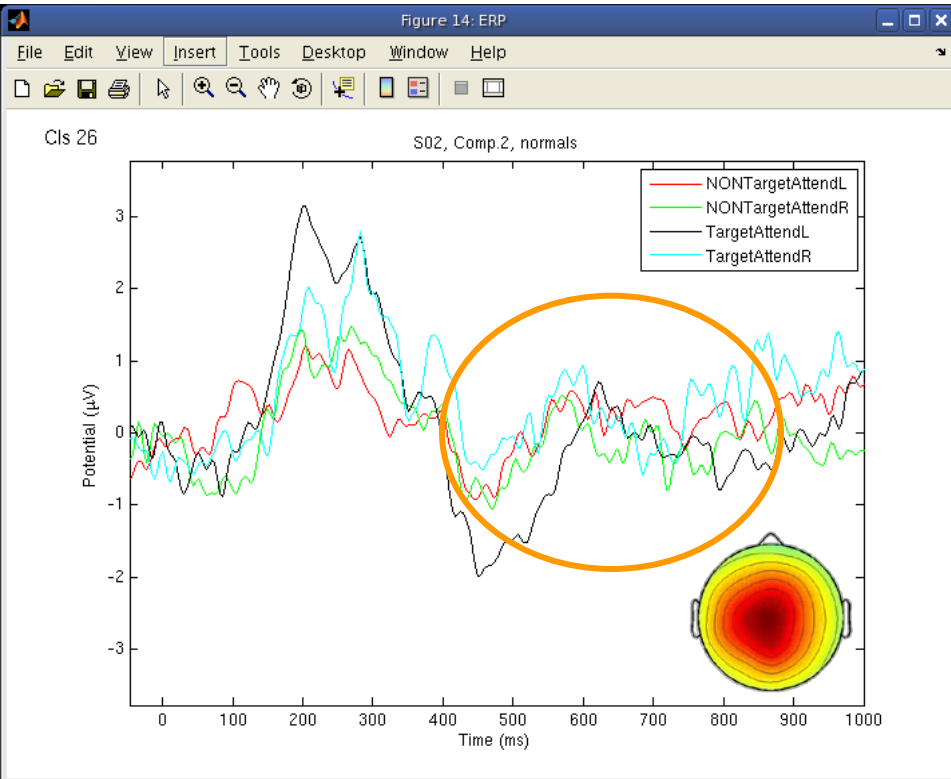
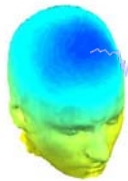


Obvious dramatic effect on scalp map topography:

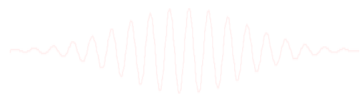


But, do they perform the same functions?

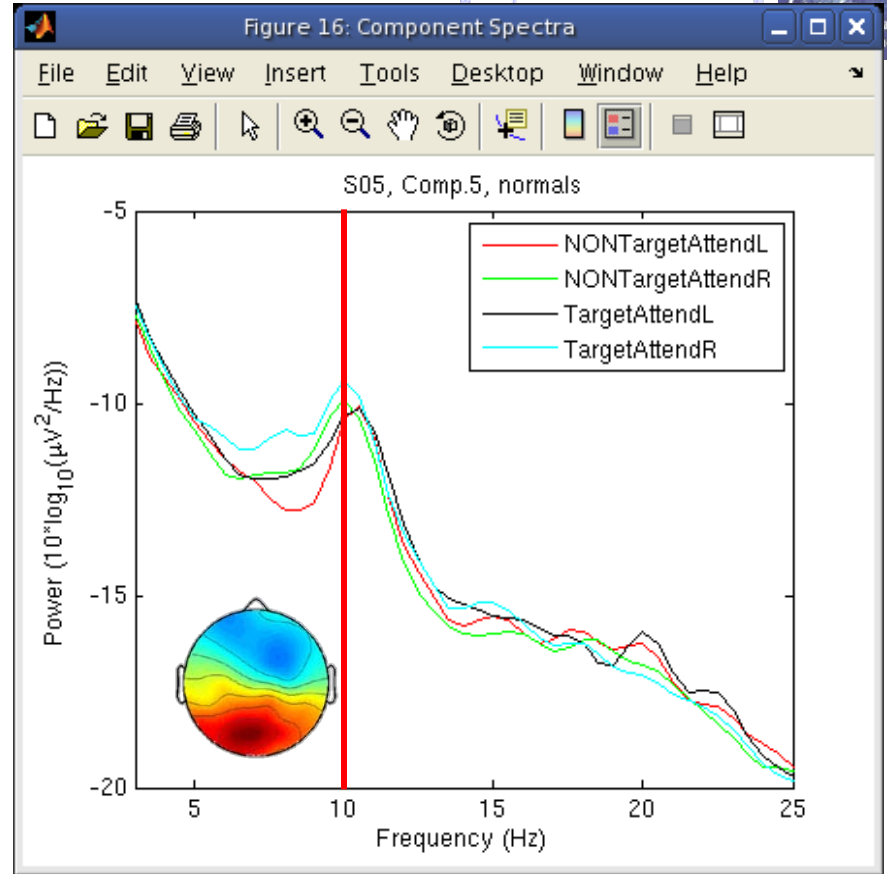
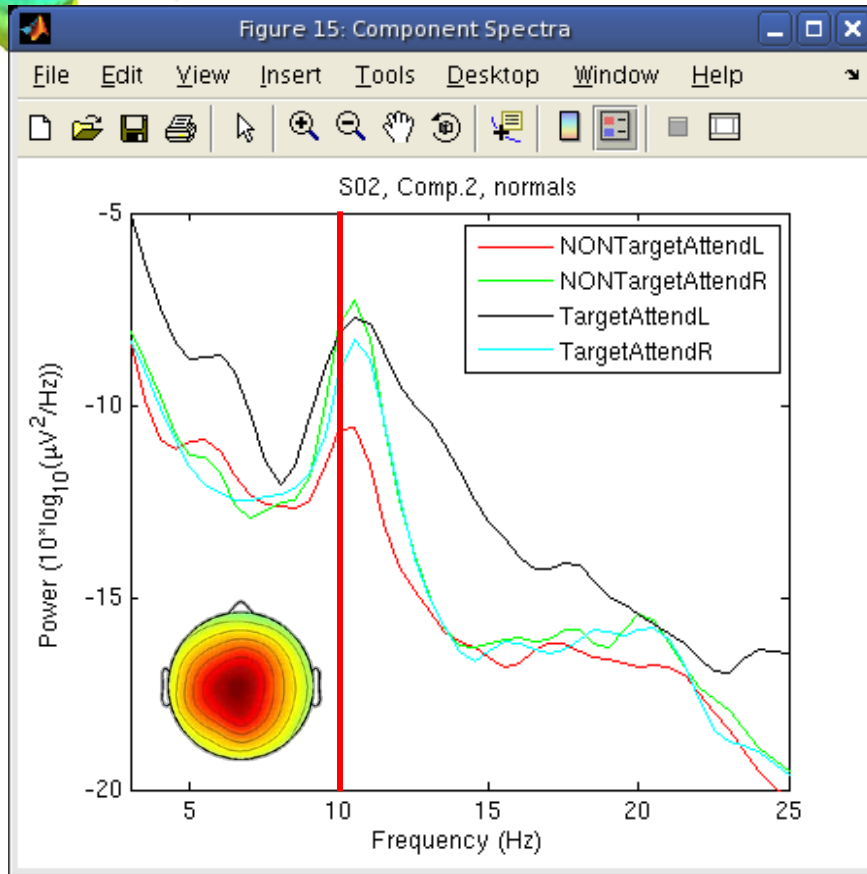
# Choosing data measures



ERPs seem different...



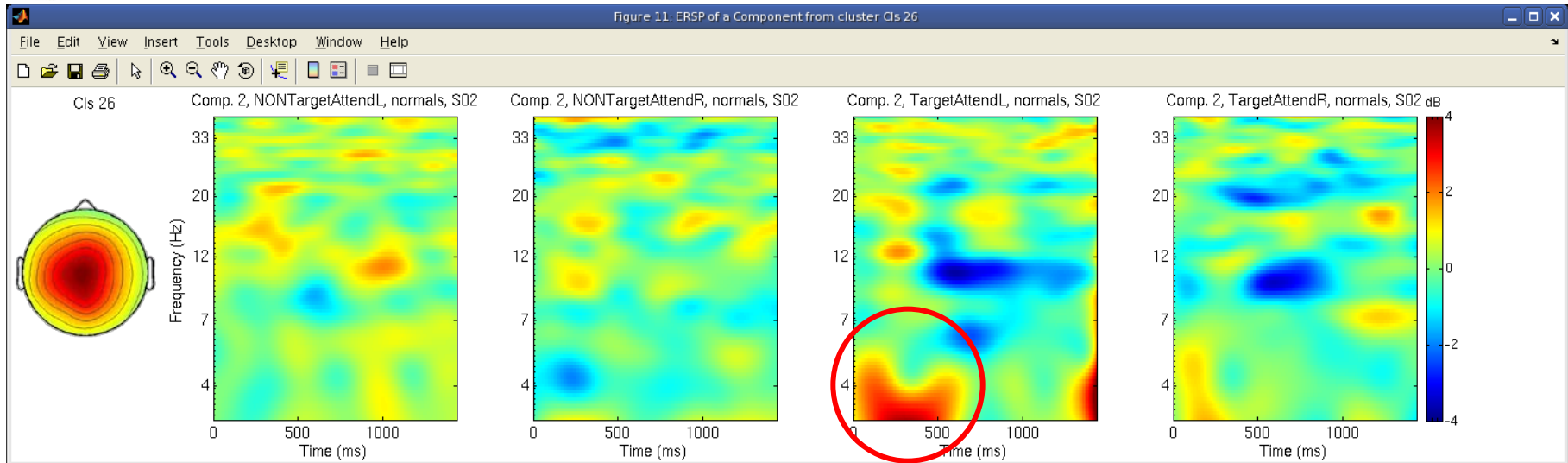
# Choosing data measures



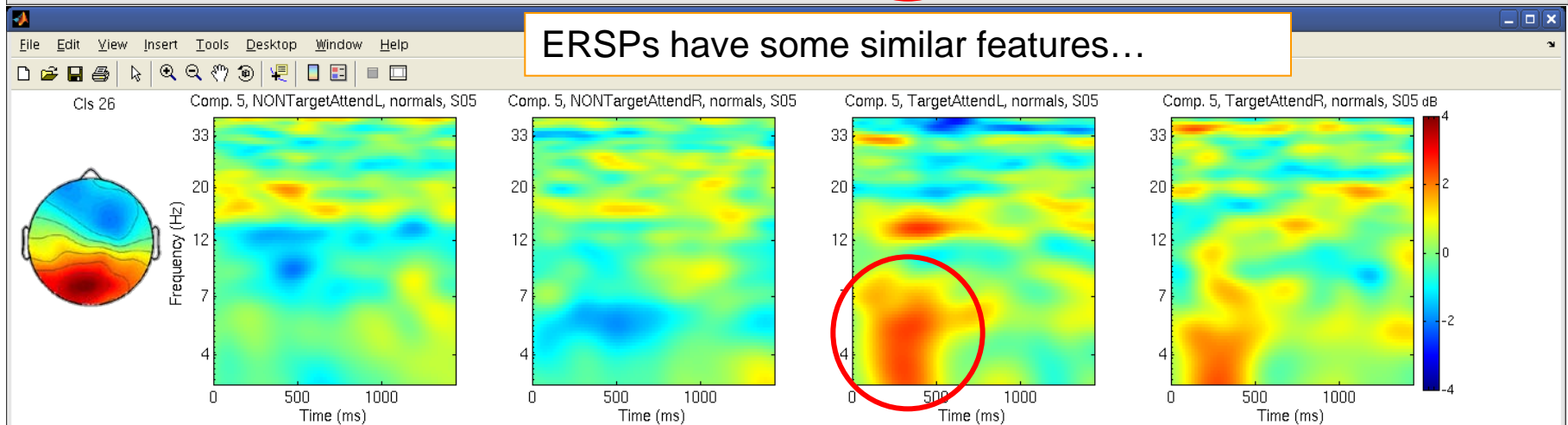
Spectra are similar, but they have  
variable responses to different conditions...



# Choosing data measures



ERSPs have some similar features...





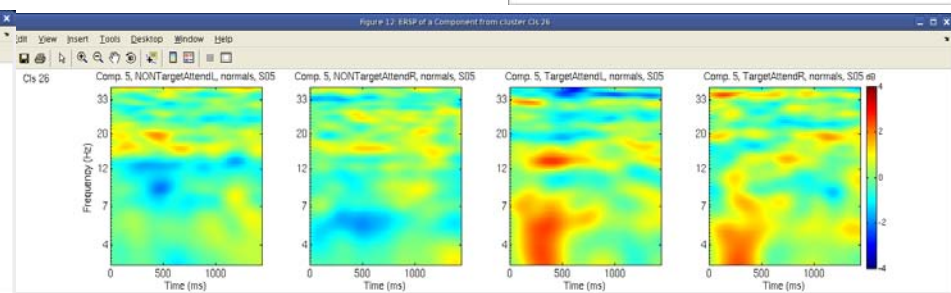
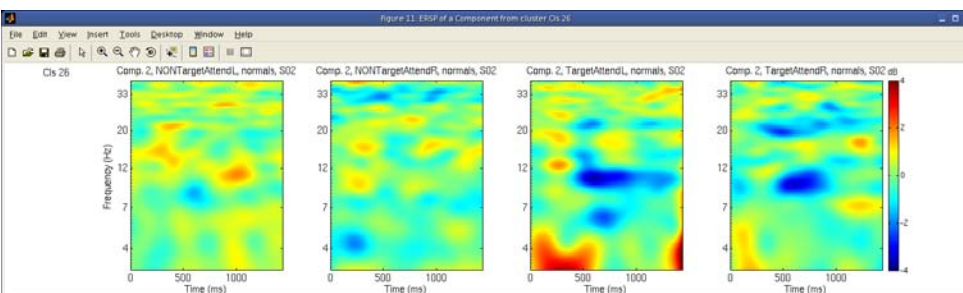
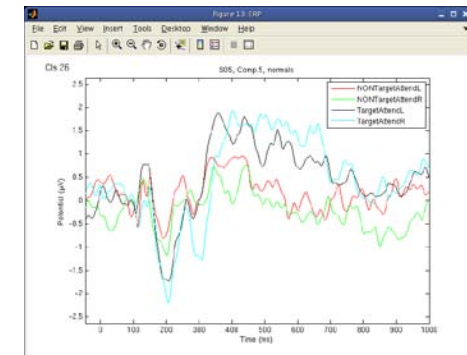
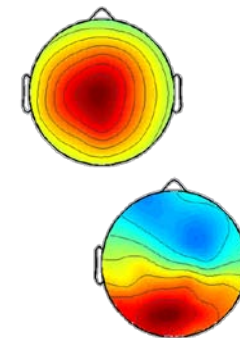
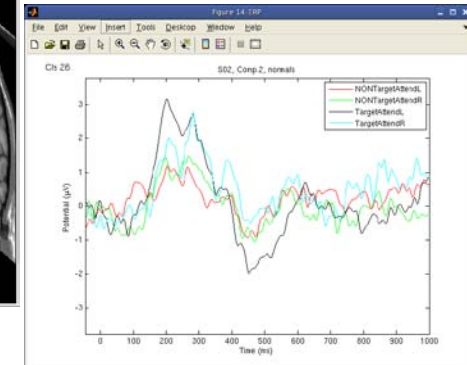
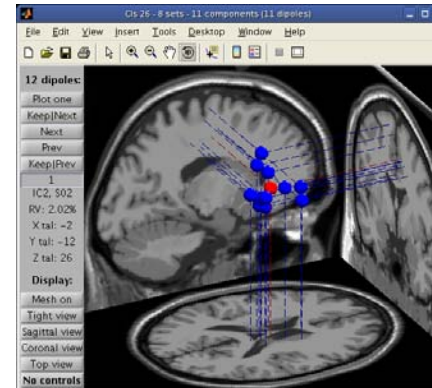
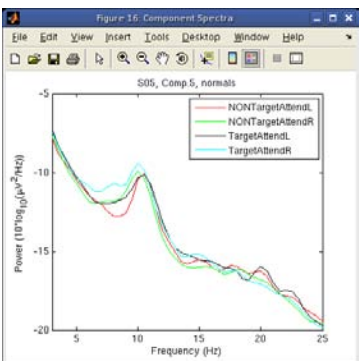
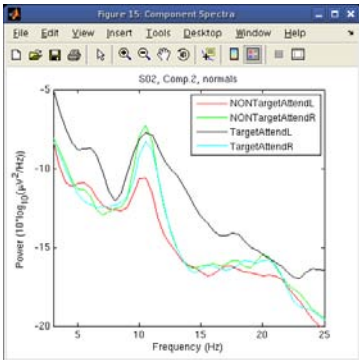
# Choosing data measures



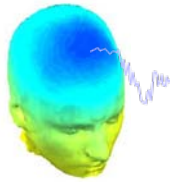
What data measures should you use?

It depends...

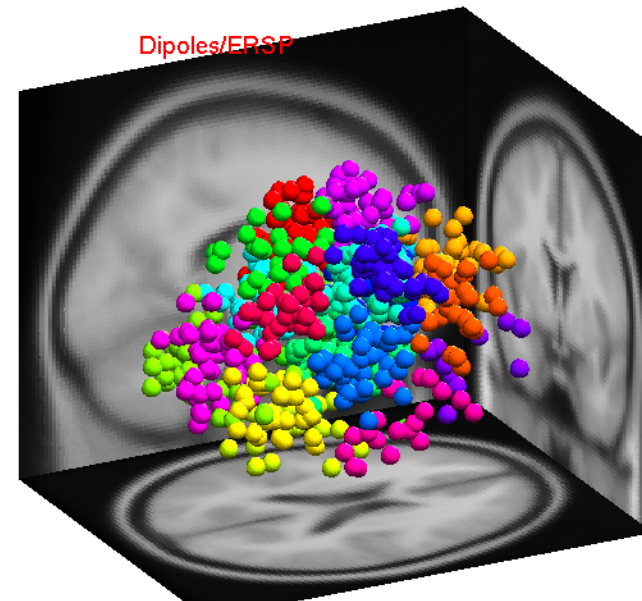
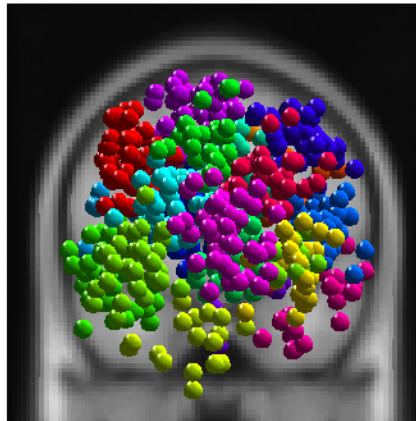
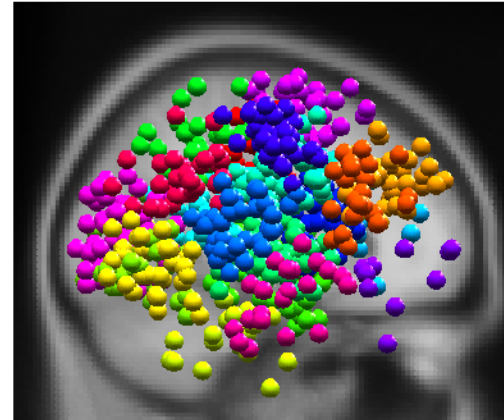
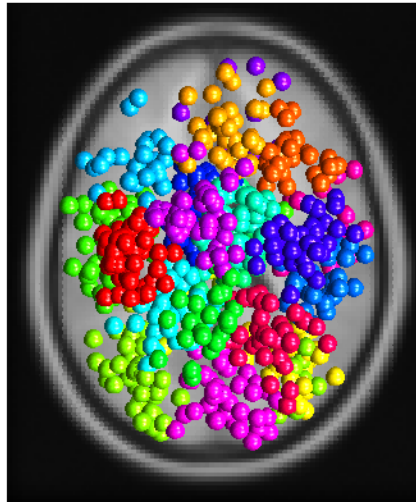
- broadly-matched ICs: use many/all of the measures.
- specifically-matched ICs: use one/few of the measures.



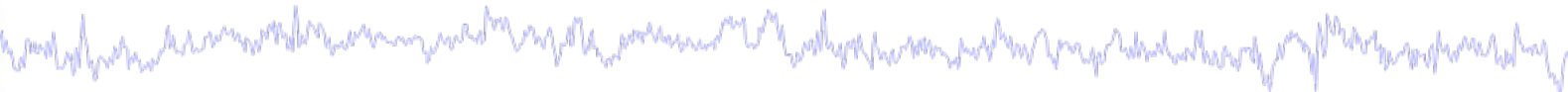
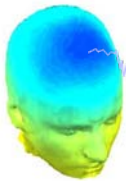
# What should clusters look like?



ICs clustered  
by dipole  
location and  
ERSP activity



# Plot STUDY dipoles



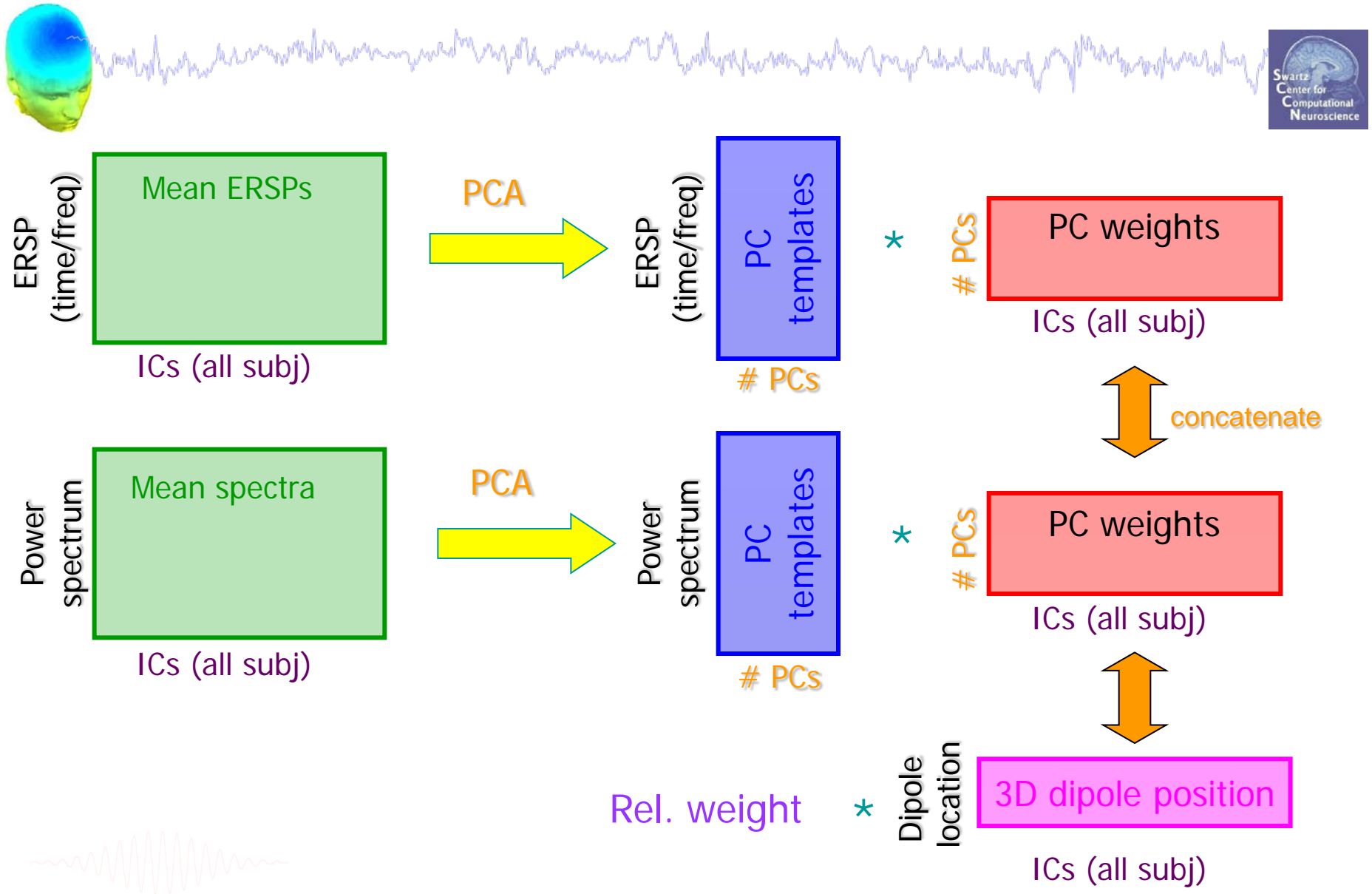
```
% std_dipoleclusters() variables:
```

```
clusters = [3:length(STUDY.cluster)]; % clusters to plot  
title = 'Cluster Dipoles'; % figure title  
plot_params = [2,2,1]; % [nrows,ncols,subplot]  
views = [1,2,3,4]; % 1=top,2=side,3=rear,4=oblique  
cols = hsv(length(clusters));
```

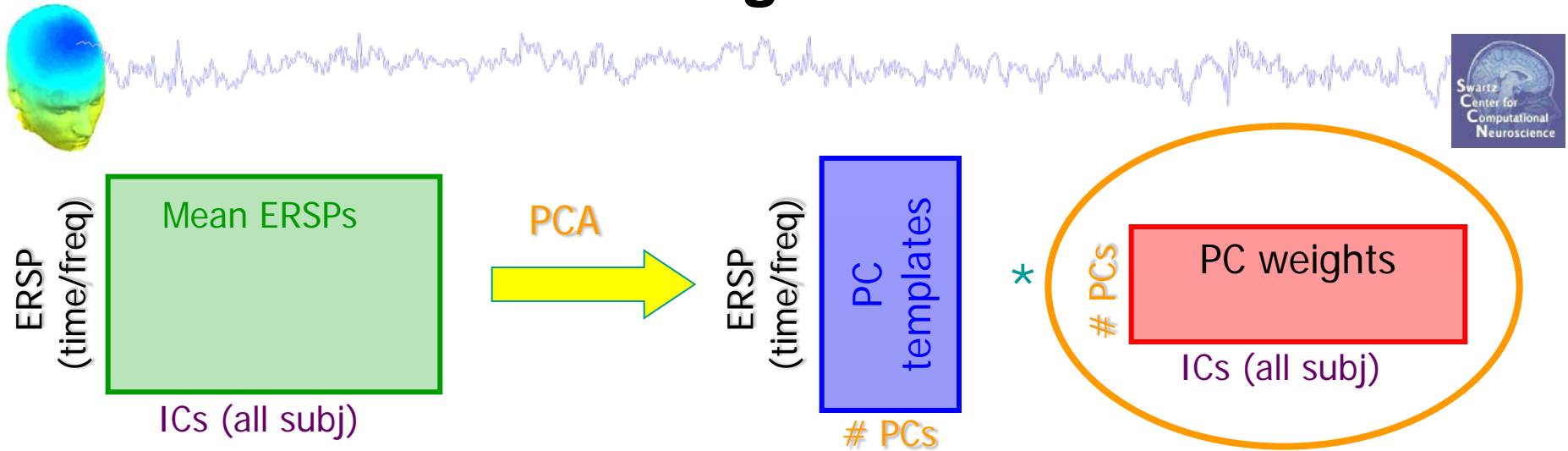
```
% std_dipoleclusters function call:
```

```
std_dipoleclusters(STUDY,ALLEEG,'clusters',clusters,...  
'title',title,'viewnum',views,'rowcolplace',plot_param,...  
'centroid','off','colors',cols);
```

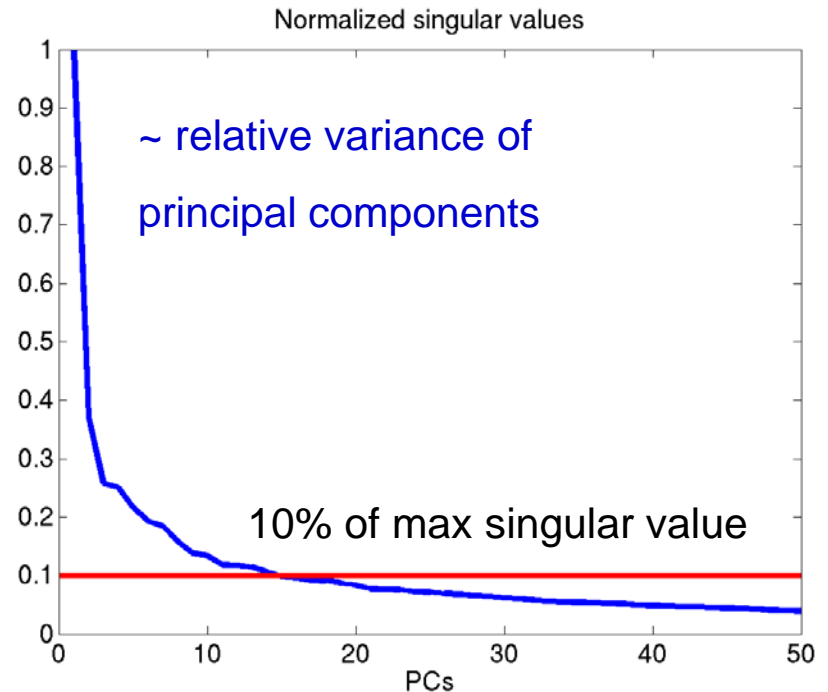
# Precluster schematic



# Precluster: Use singular values from PCA



%% Do it yourself:  
%% Load all ERSP data  
%% decompose with PCA  
%% plot singular values  
(See code in 'Tutorial\_8\_BuildSTUDY.m')



# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

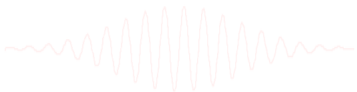
## STEP 3

Precluster the data

## STEP 4

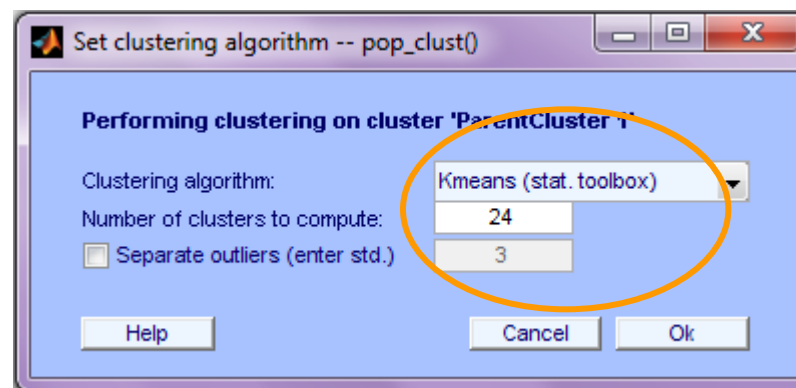
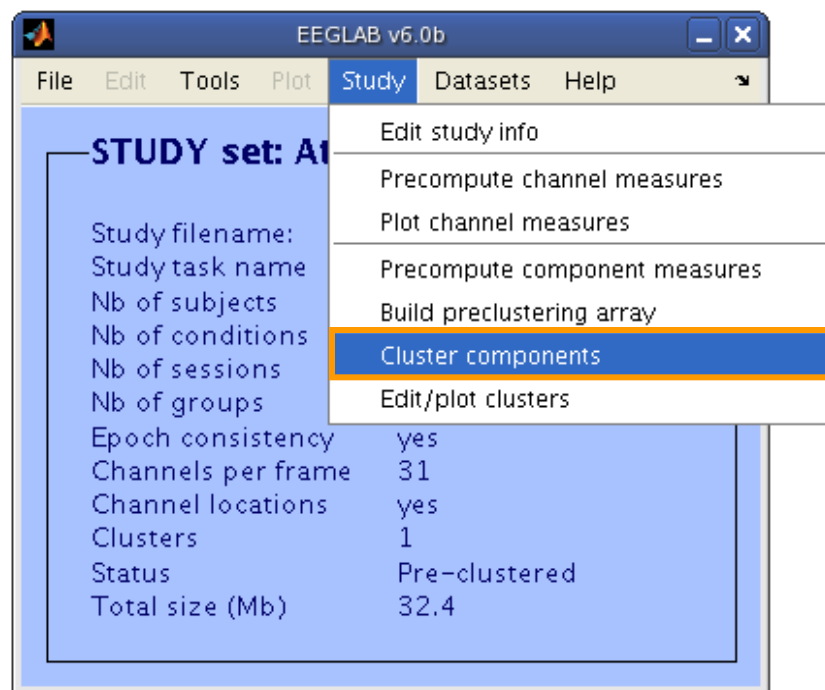
Cluster the data

**Exercise...**





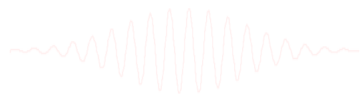
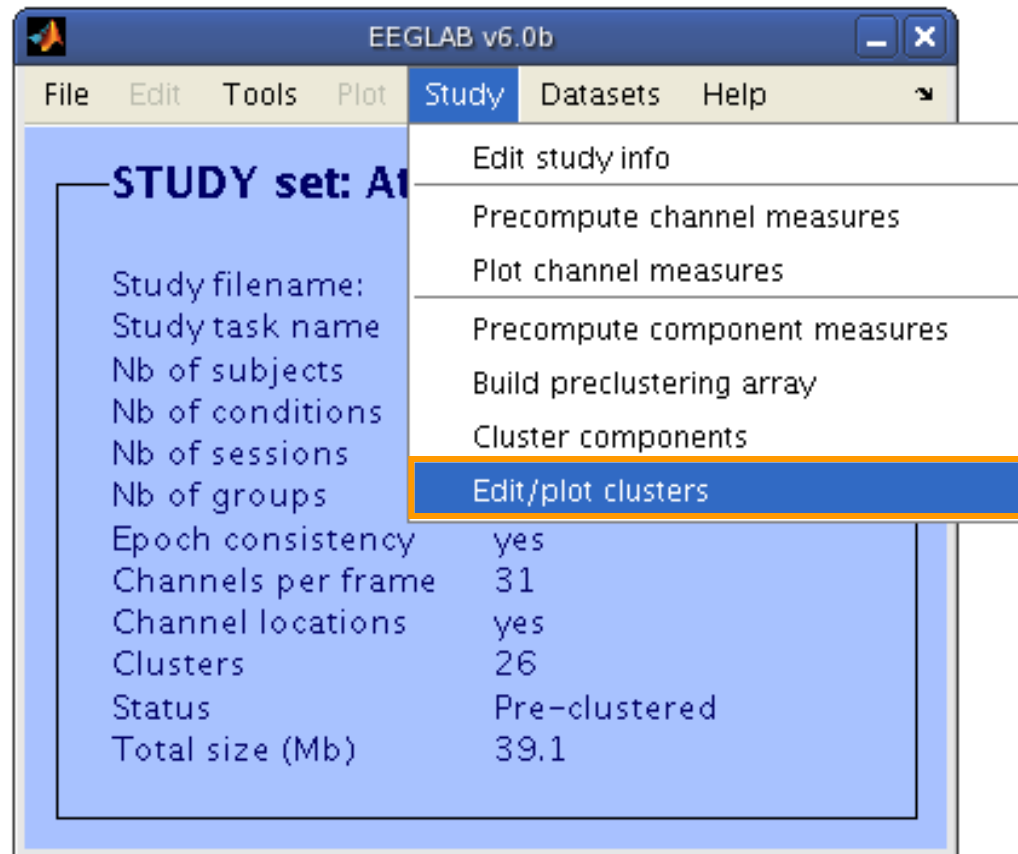
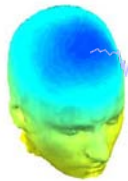
# Cluster components



```
nclusts = 25; % choose # of clusters to create
```

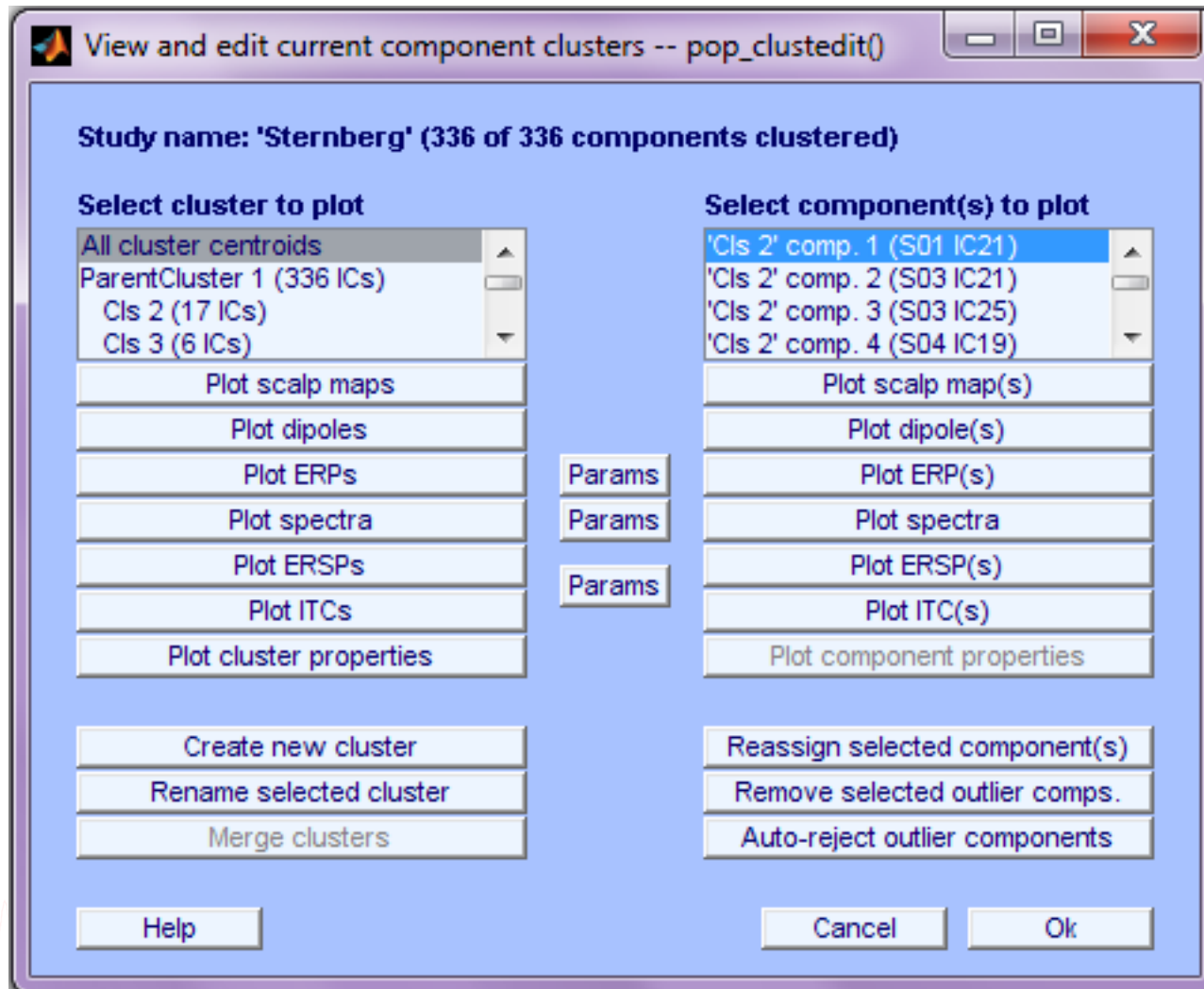
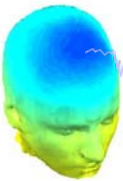
```
[STUDY] = pop_clust(STUDY, ALLEEG, 'algorithm', 'kmeans', 'clus_num', nclusts);
```

# View and edit clusters

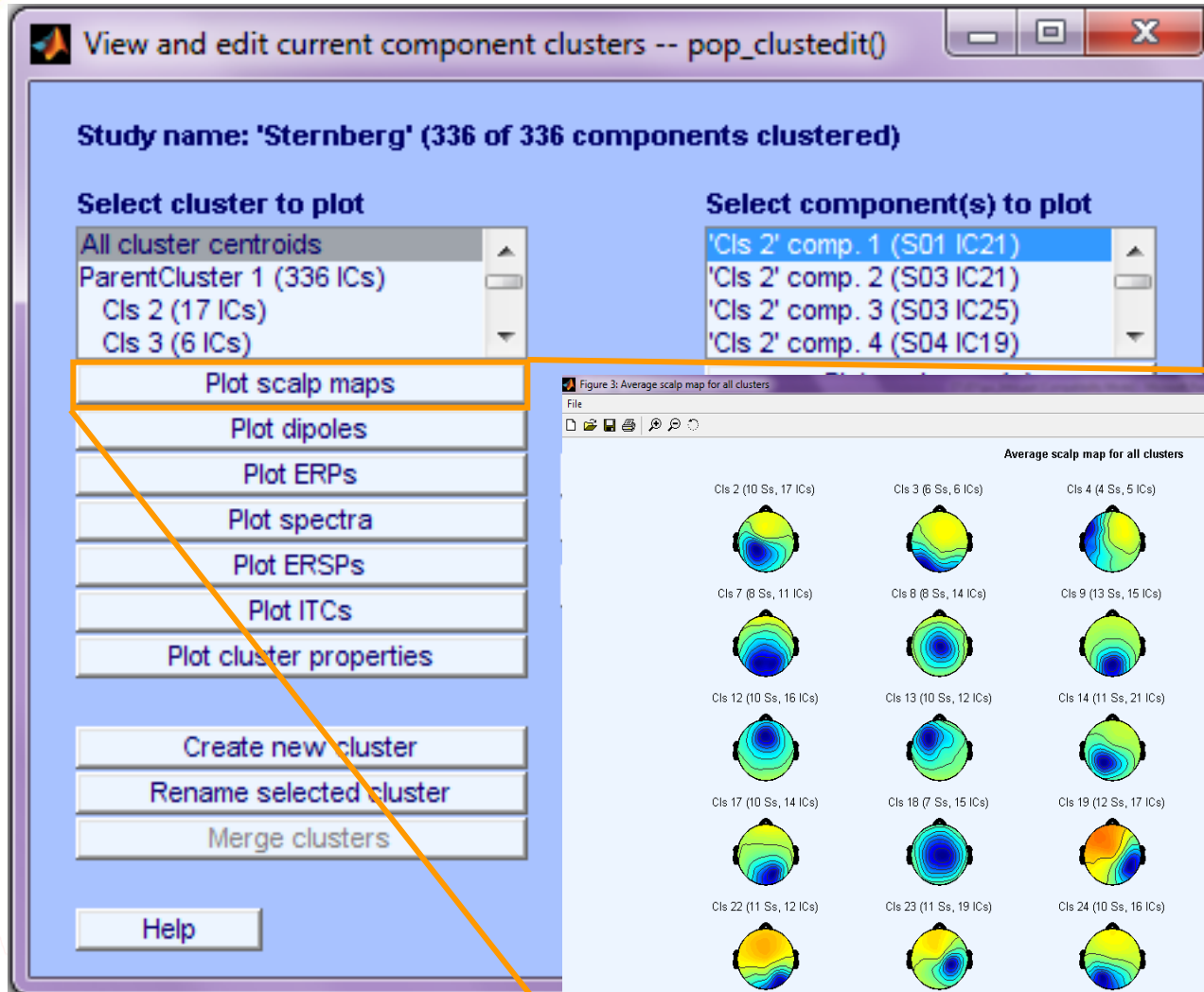




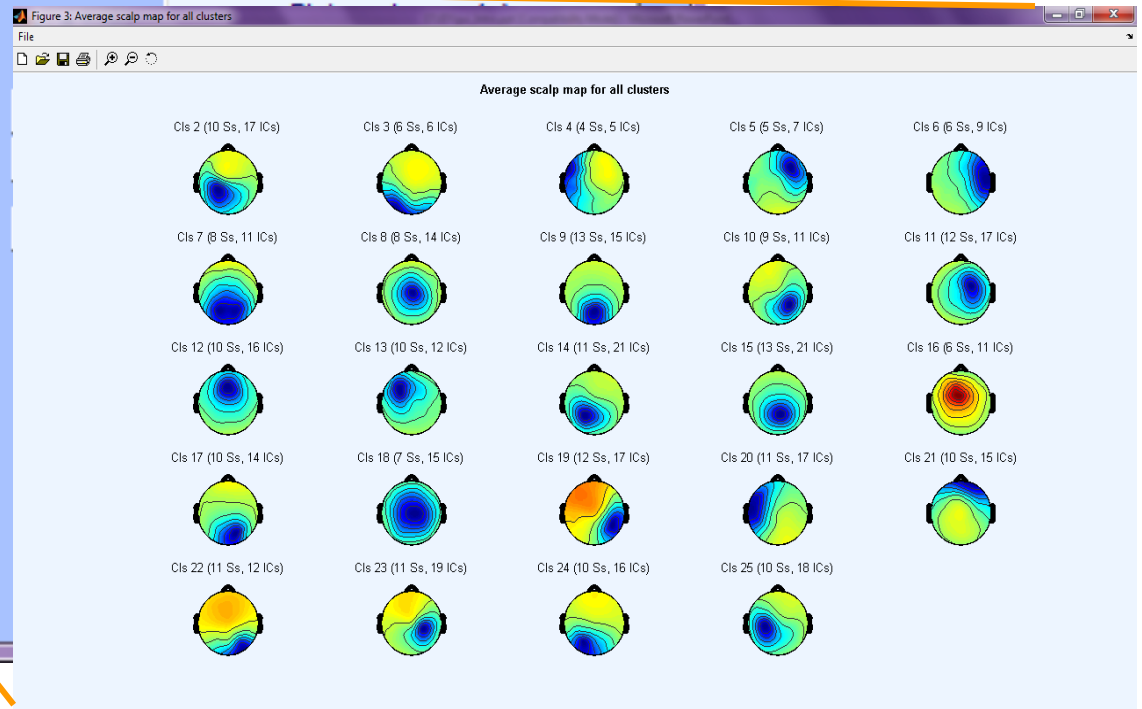
# Plot/edit clusters



# Plot cluster data



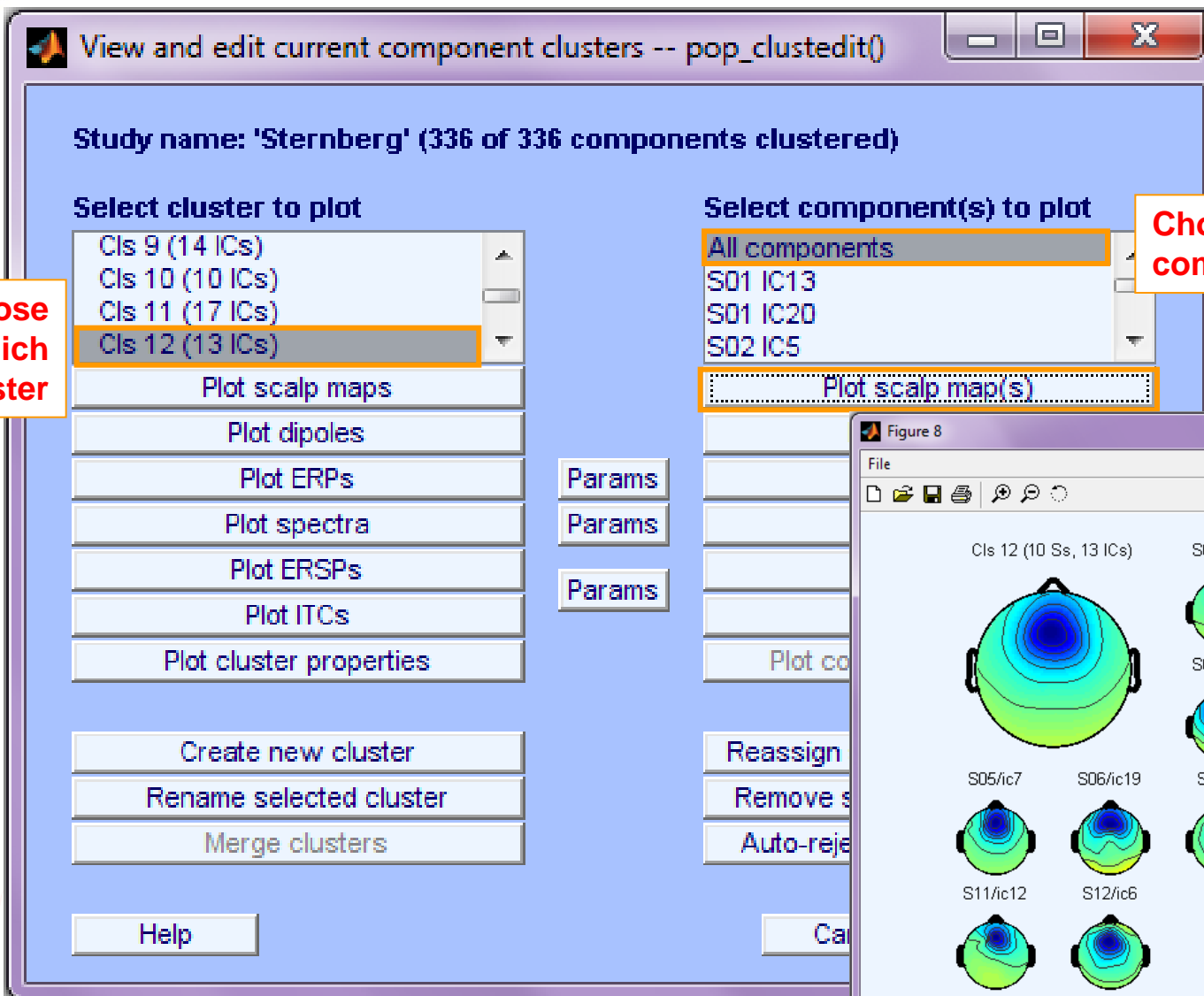
Plot mean scalp maps for easy reference



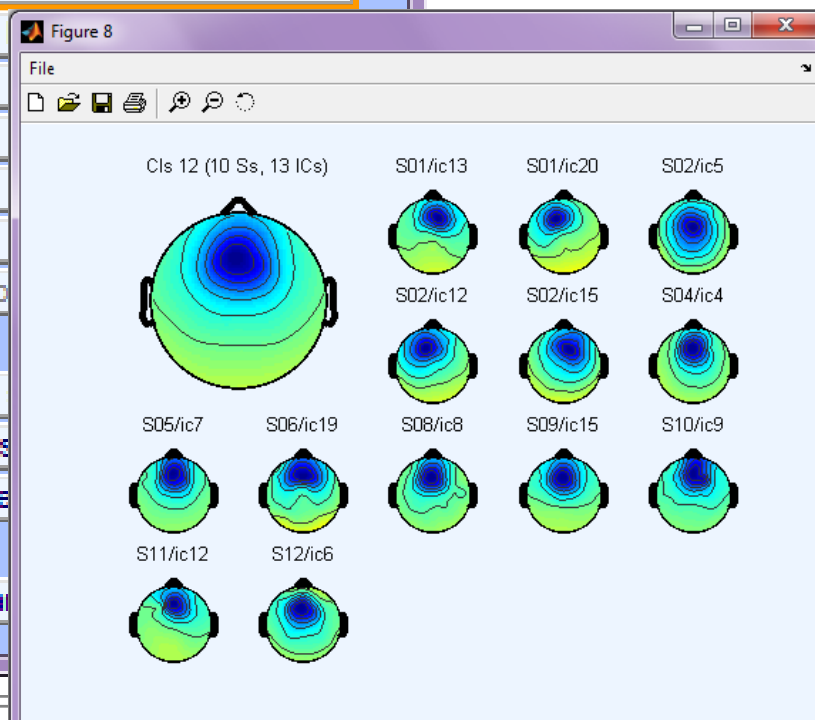
# Plot cluster data



Choose which cluster



Choose which components



# STUDY clustering overview



## STEP 1

Build a STUDY

## STEP 2

Precompute the data

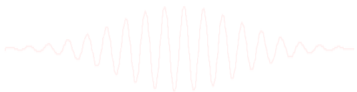
## STEP 3

Precluster the data

## STEP 4

Cluster the data

**Exercise...**



# Exercise



- **Novice**
  - Use the GUI to build a STUDY (for practice, try just a few subjects).
  - Choose 'precompute' options (but do not recalculate for the sake of time).
  - Choose 'precluster' options and cluster.
- **Intermediate**
  - Script a loop to build a STUDY from the commandline
  - Precluster (pre-computation already done) and cluster components using measures of your choice.
- **Advanced**
  - Load raw data measures and run PCA to determine the relative size of PCA dimensions for each data measure.
  - Try preclustering/clustering based on your observations

\*\* All scripts for exercises can be found in:  
.../EEGLAB\_Workshop/Scripts/Tutorial\_7\_BuildSTUDY.m

(initial variables missing a 'conds' variable:

```
conds = {'memorize', 'ignore', 'probe'};
```