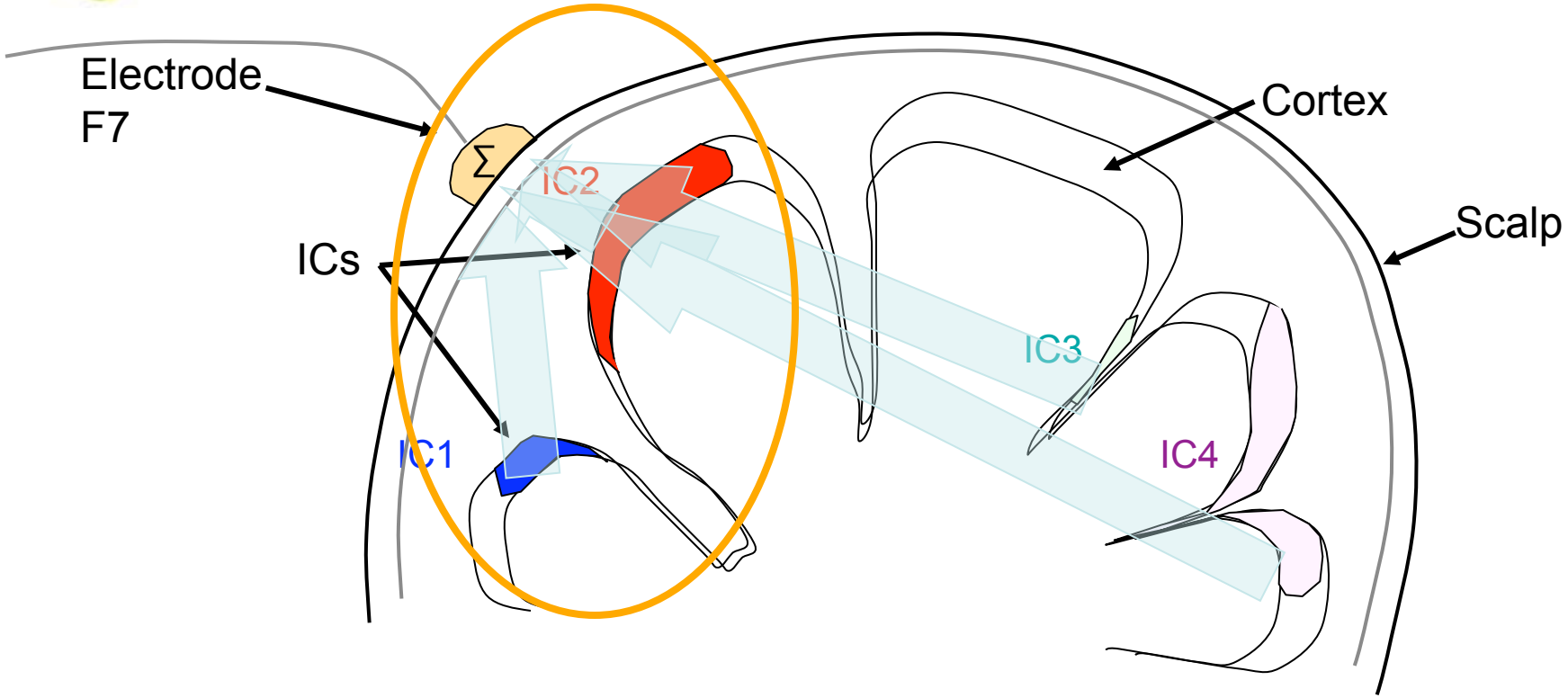
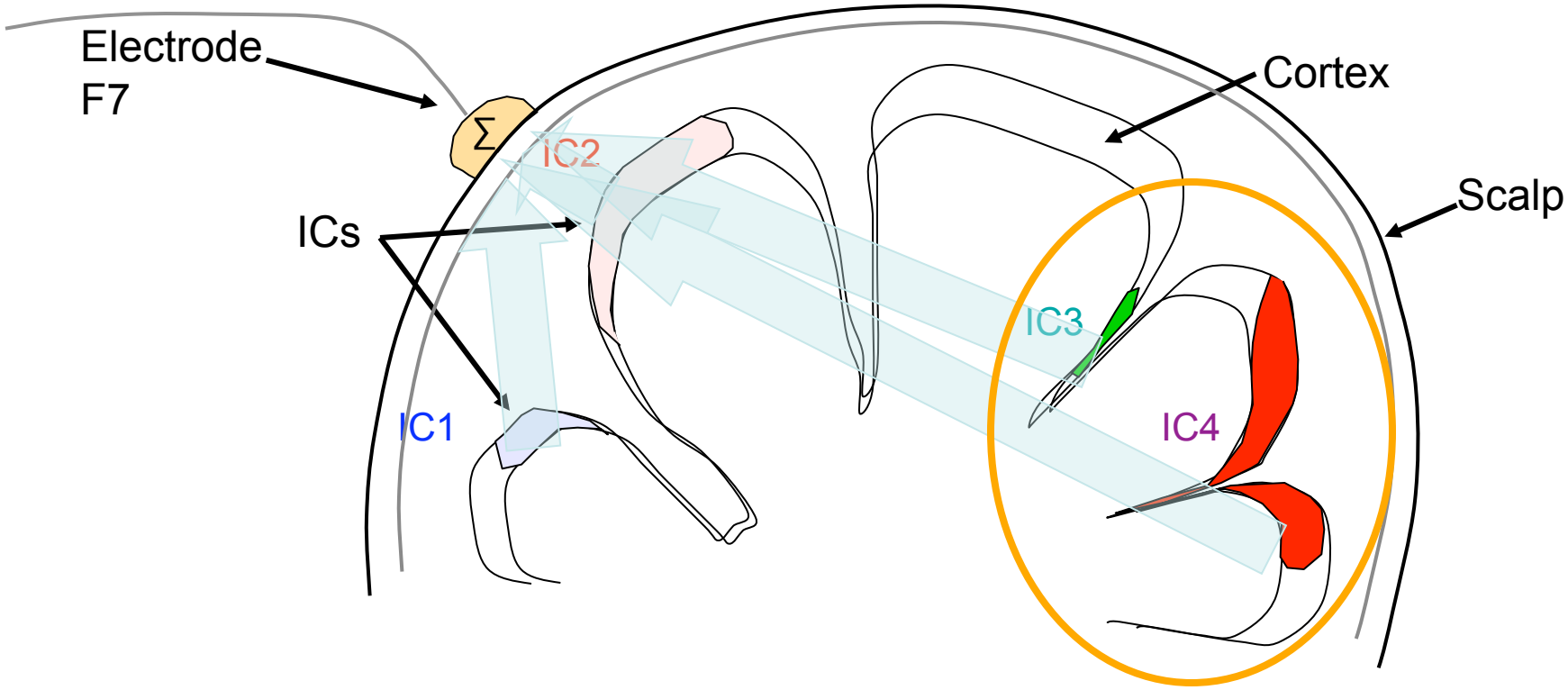
Why cluster components?

- ICA transforms the data from a channel basis (activity recorded at each channel)
  - to a component basis (activity computed at each independent spatially-filtered cortical or non-cortical component process).
- Normally, EEG researchers assume that electrode, say F7 == F7 == F7 ... in each subject – and then ‘cluster‘ their data by channel ...
- But this is only *roughly* correct!

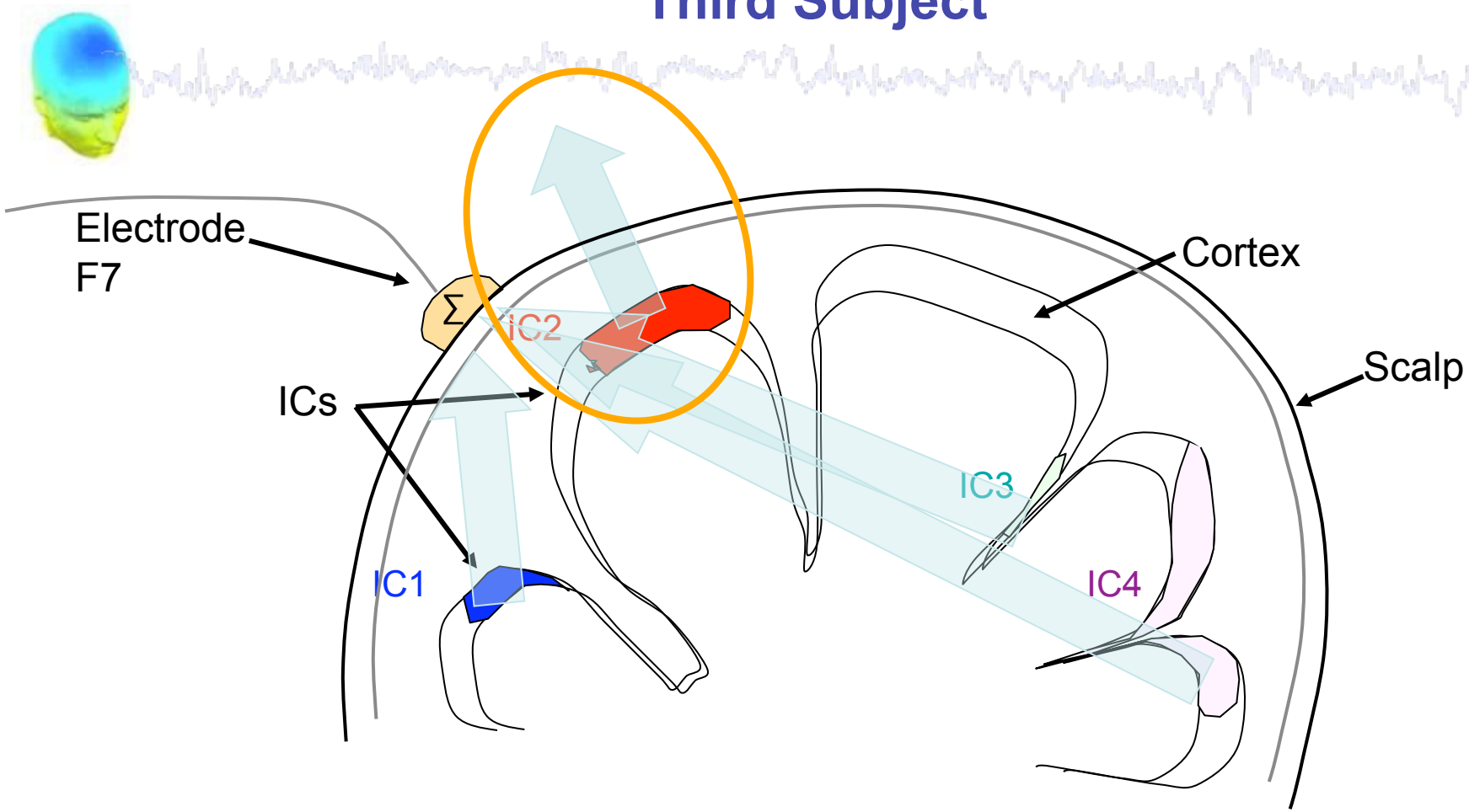
Example: First Subject



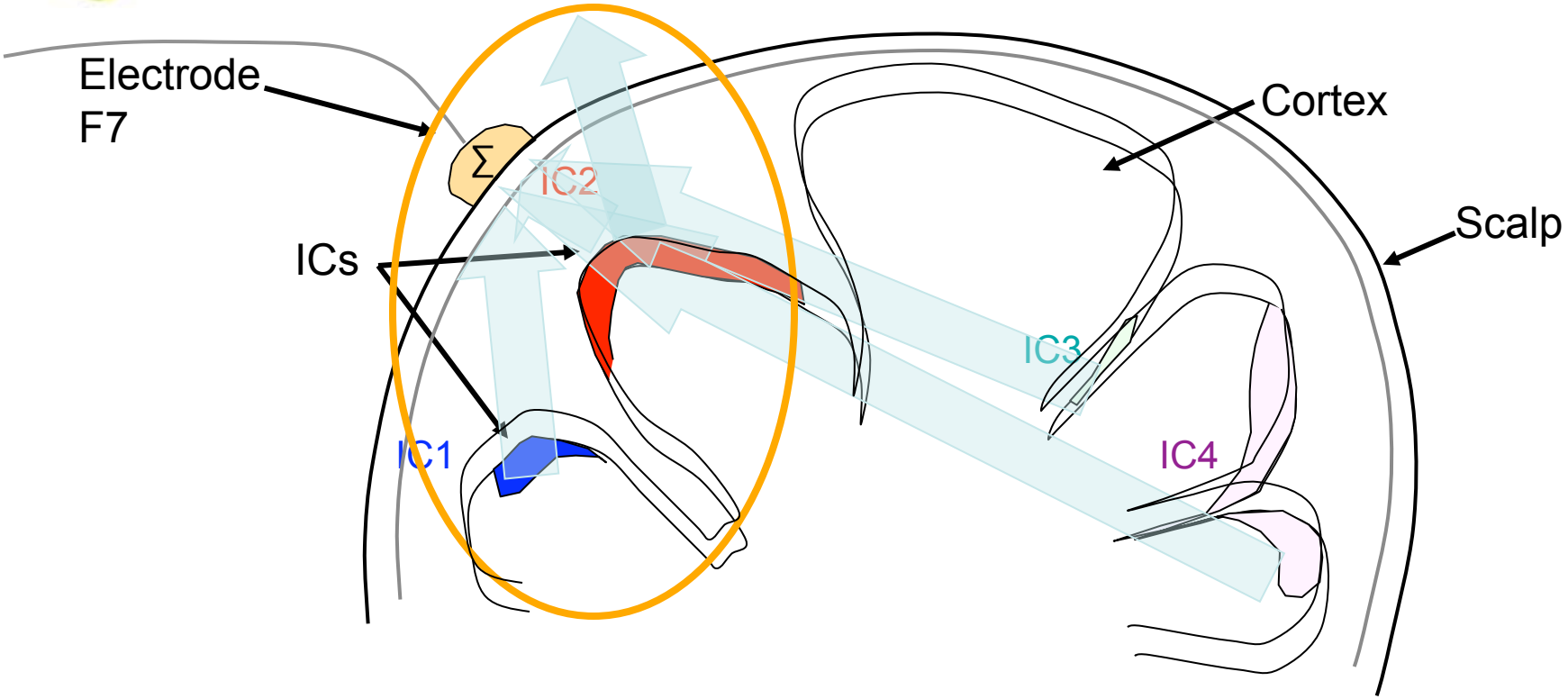
Second Subject



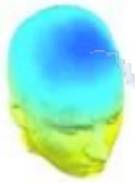
Third Subject

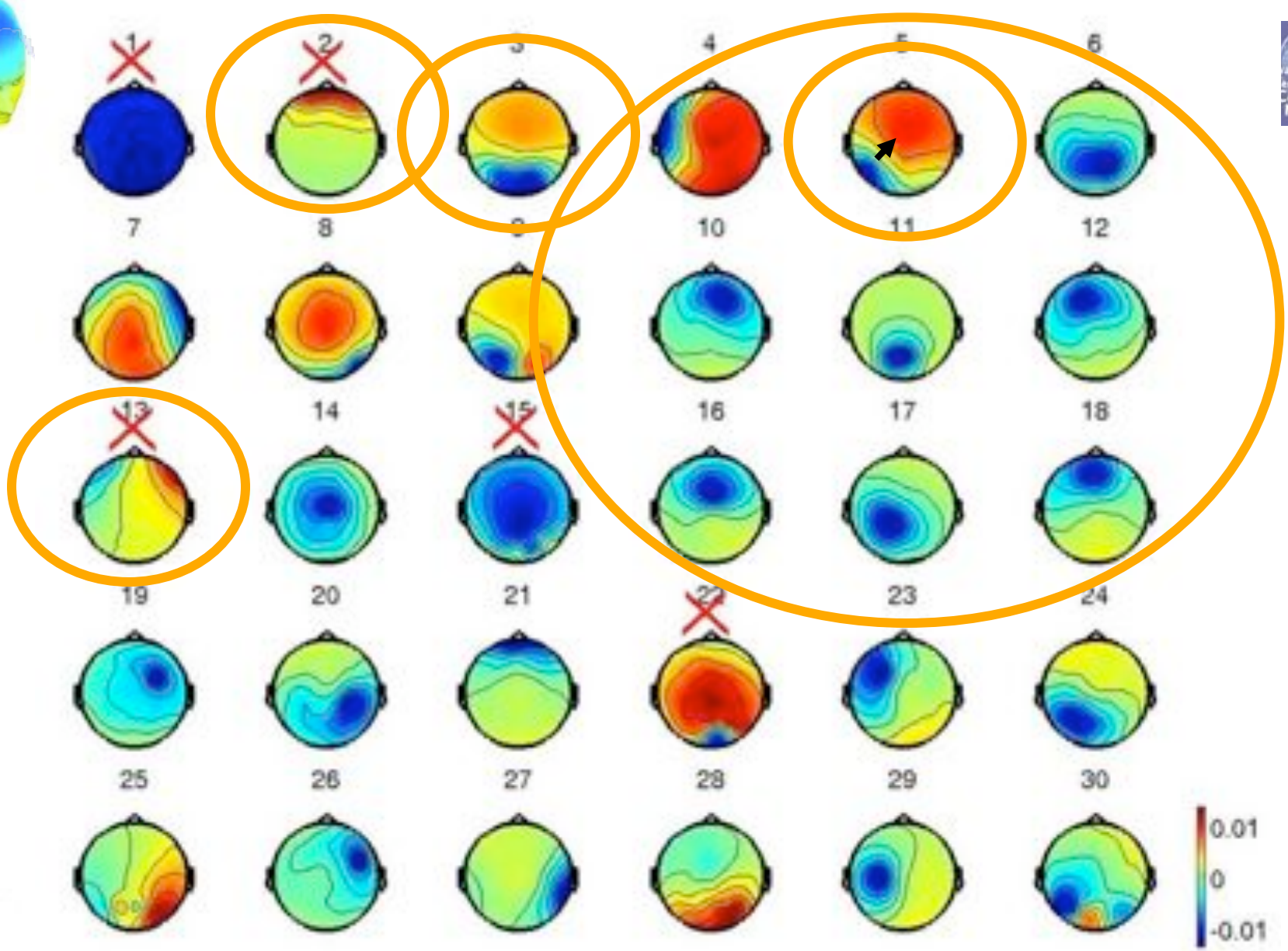


Fourth Subject



Largest 30 independent components (single subject)





So how to cluster components?

The same problems hold for clustering independent components

Across Ss, components don't even have "the same" scalp maps!

→ Are "the same" components found across subjects?

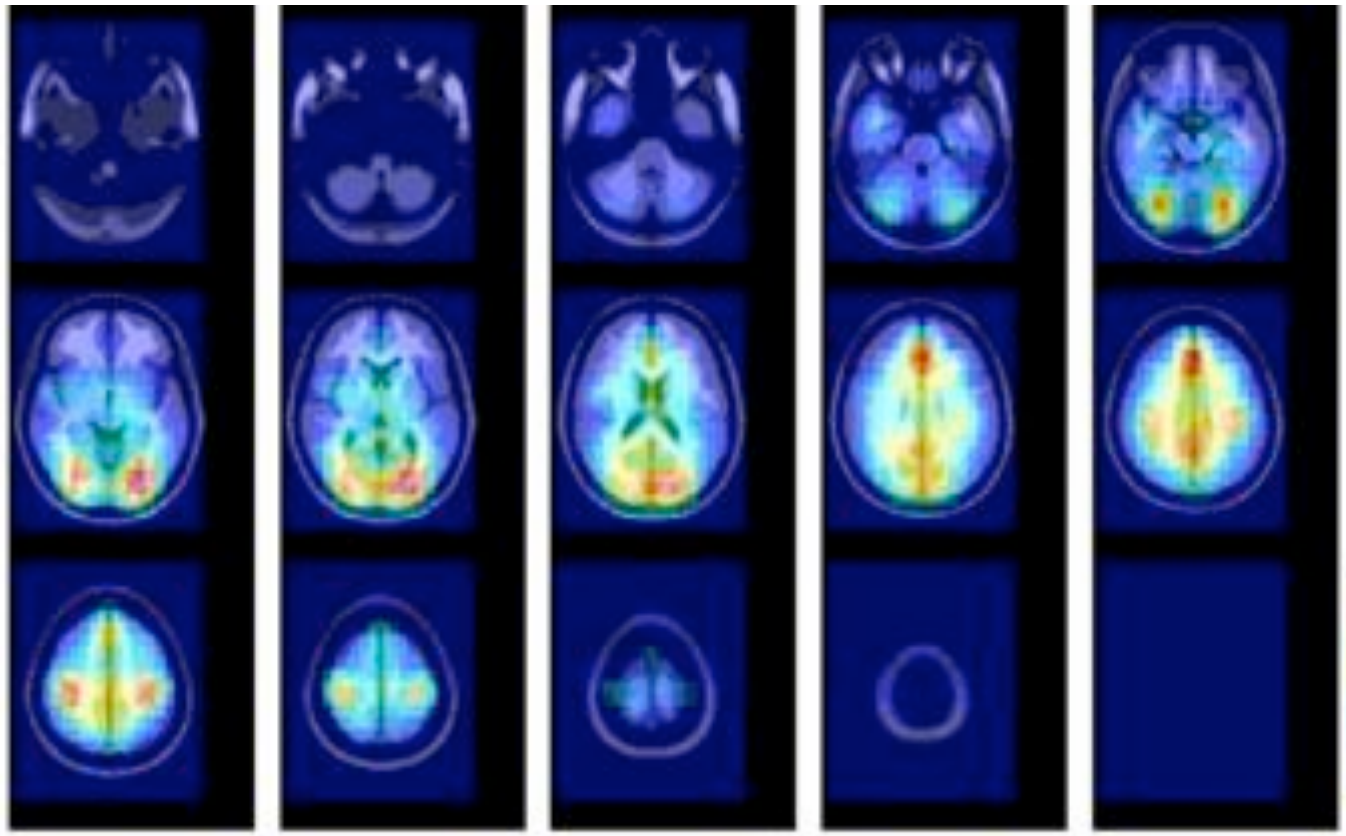
- What should define "the same" (i.e., "component equivalence")?
  - Similar scalp maps?
  - Similar cortical or 3-D equivalent dipole locations?
  - Similar activity power spectra?
  - Similar ERPs?
  - Similar ERSPs?
  - Similar ITCs?
  - OR …, Similar *combinations* of the above? …

Does the spatial distribution of independent components depend on the task the subject performs?

i.e.

Do “the same” components (and clusters) appear for every task?

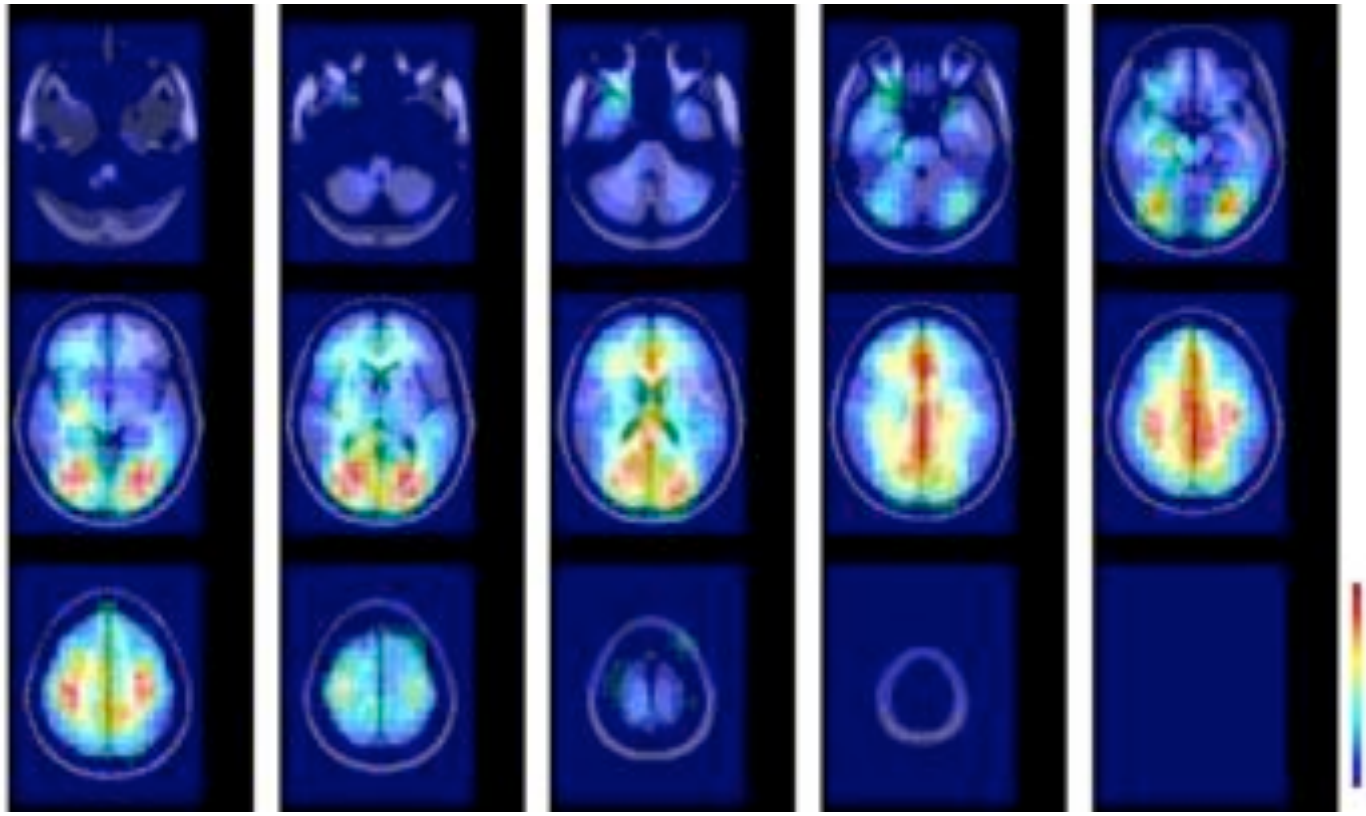
Equivalent dipole density



Sternberg letter memory task

>> dipoledensity()

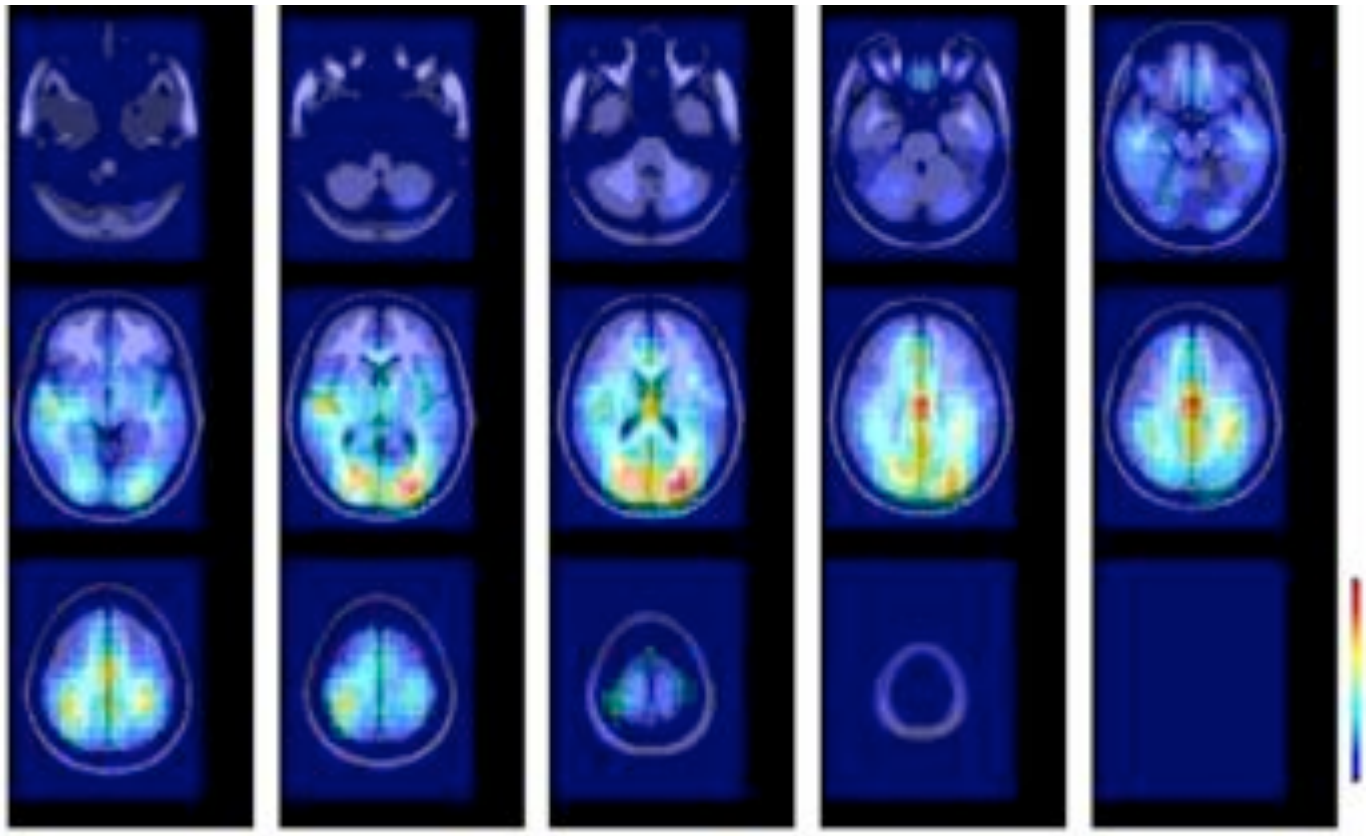
Equivalent dipole density



Letter twoback with feedback

>> dipoledensity()

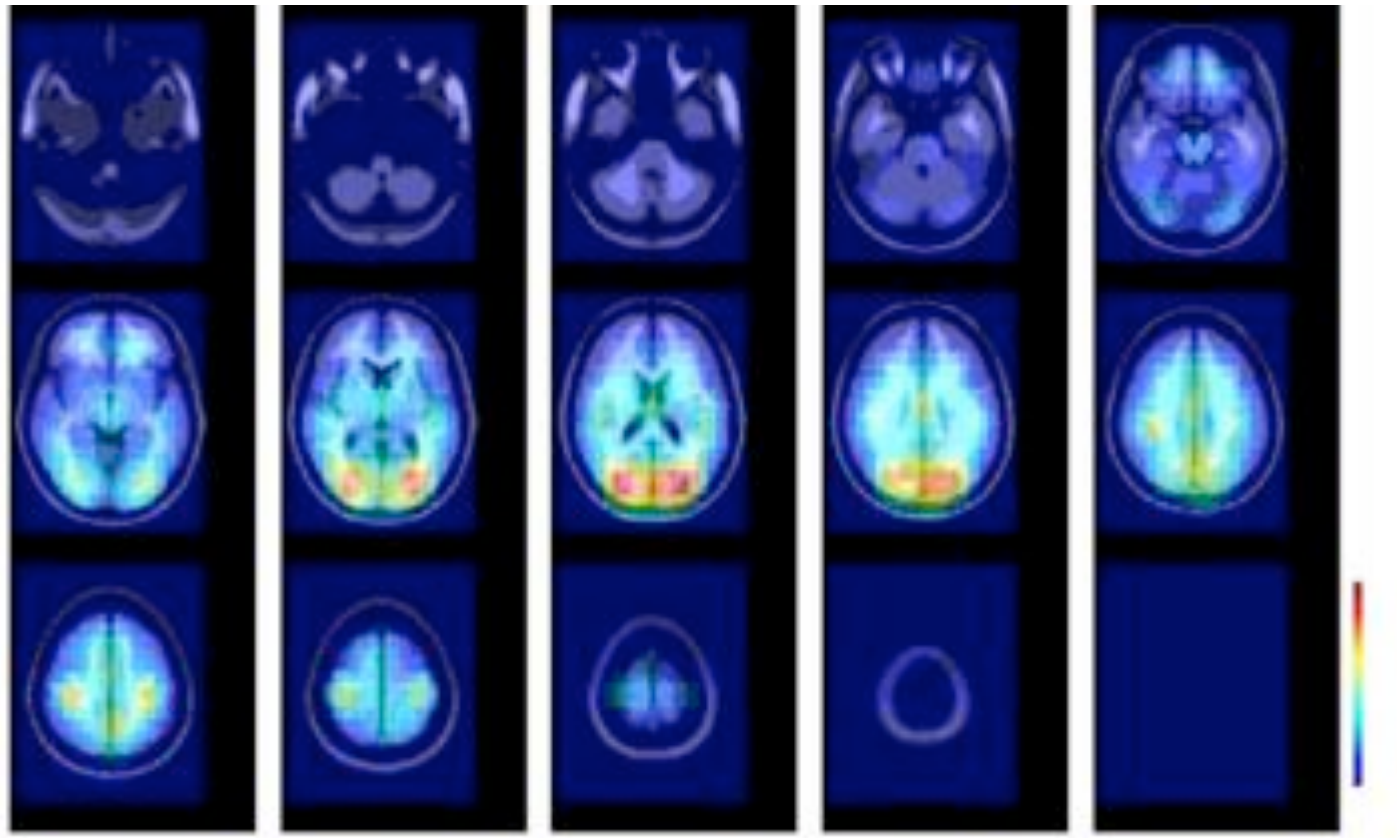
Equivalent dipole density



Auditory oddball plus novel sounds

>> dipoledensity()

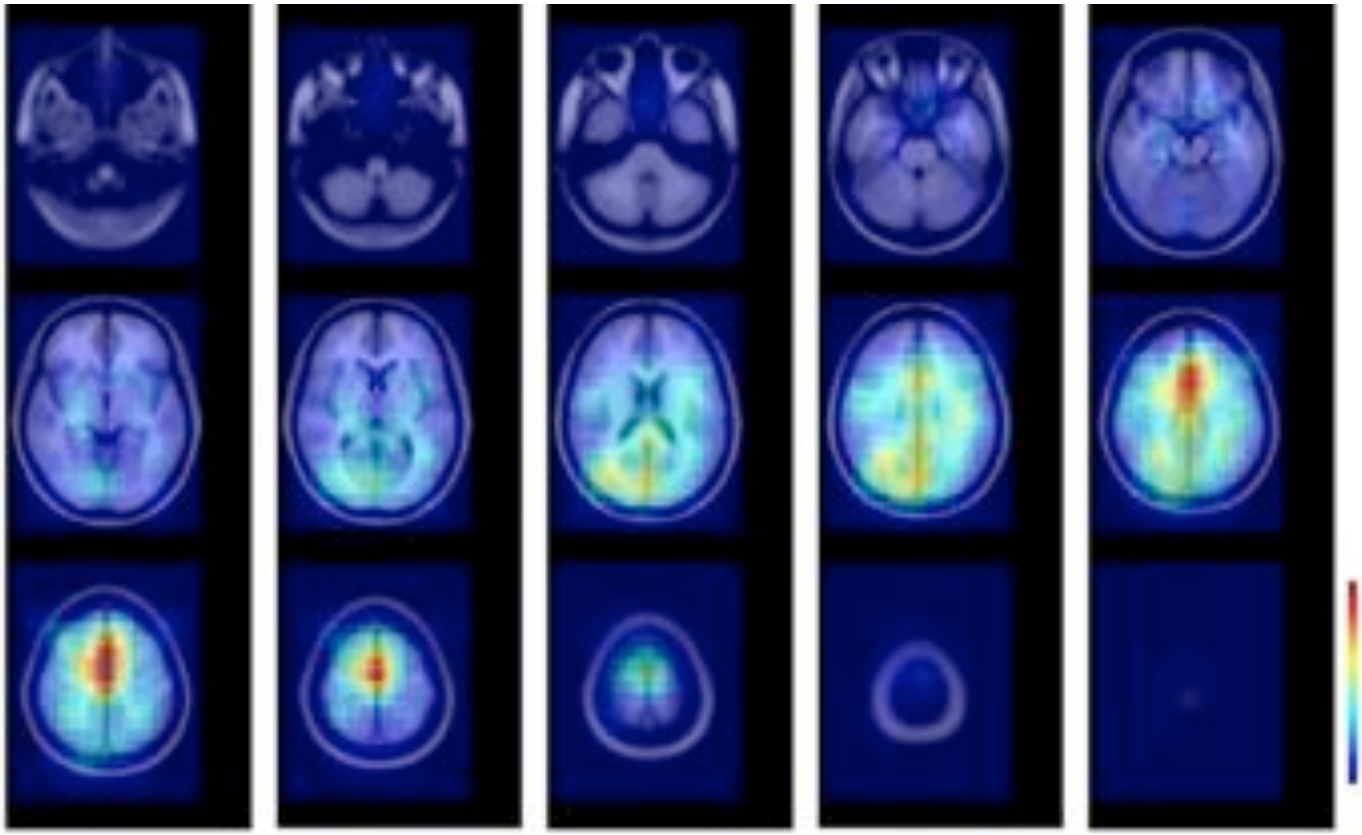
Equivalent dipole density



Emotion imagery task

>> dipoledensity()

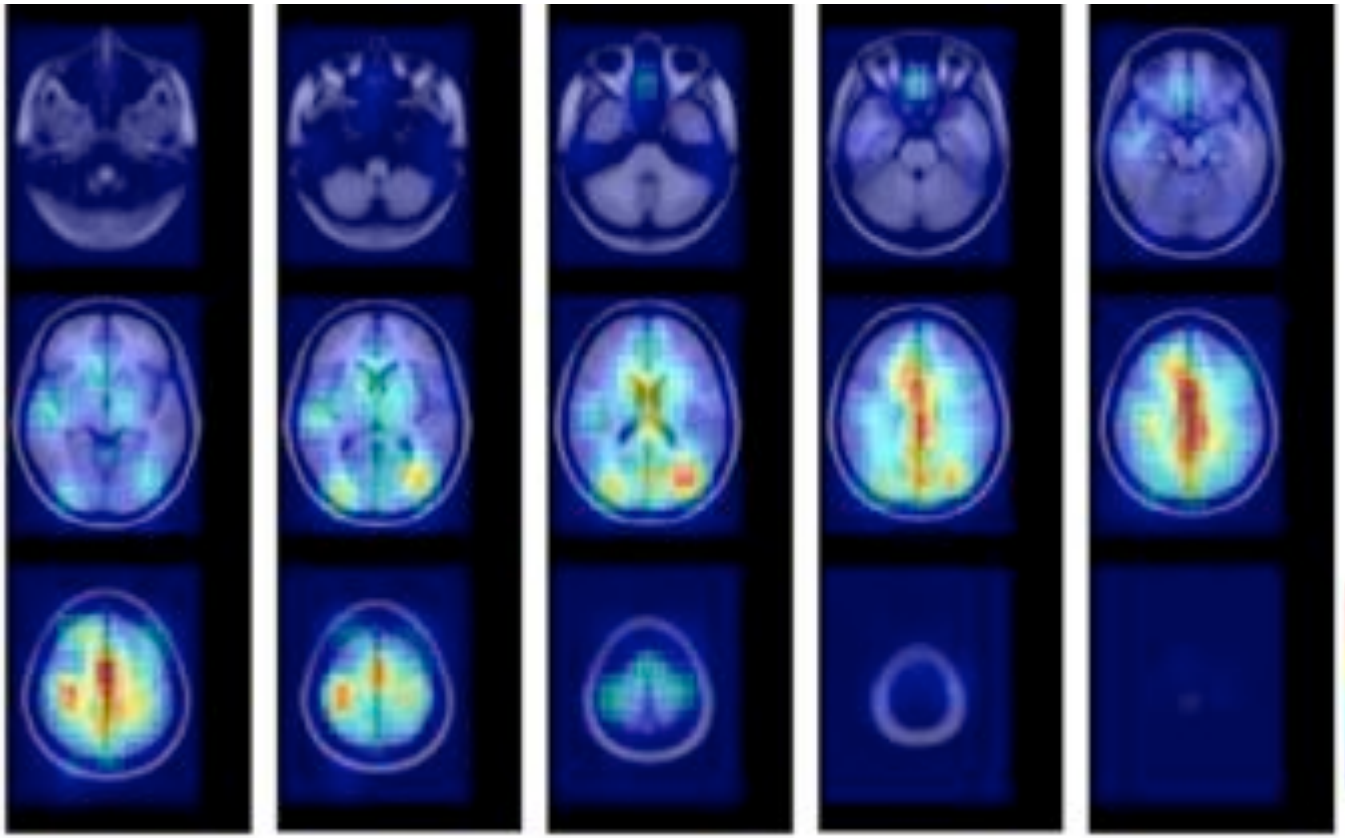
Equivalent dipole density Exp I



Word memory (old/new) task

>> dipoledensity()

Equivalent dipole density Exp II



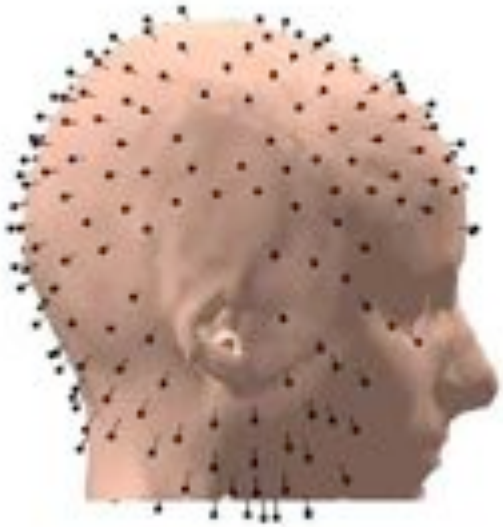
Visually cued button press task

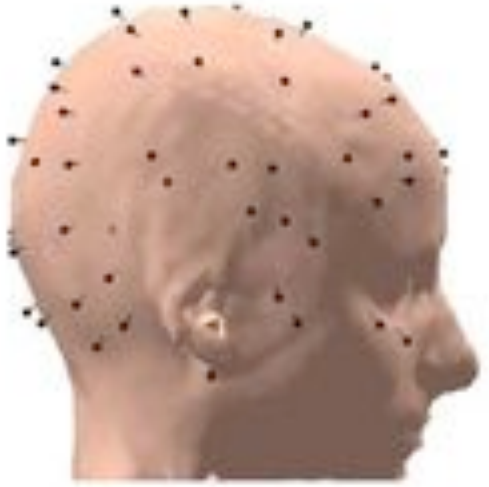
>> dipoledensity()

… Some caveats

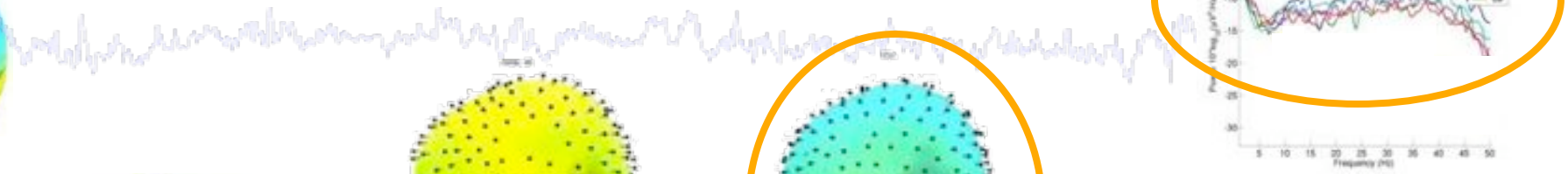
In this preliminary study …

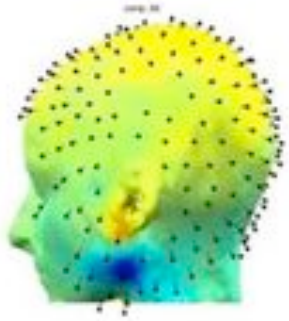
- The electrode locations were not individualized.
- MR images were not available → co-registration crude.
- Single versus dual-dipole model selection was subjective.
- Different electrode montages → possible location effects

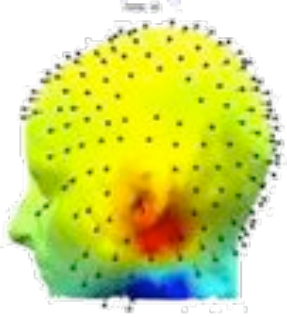


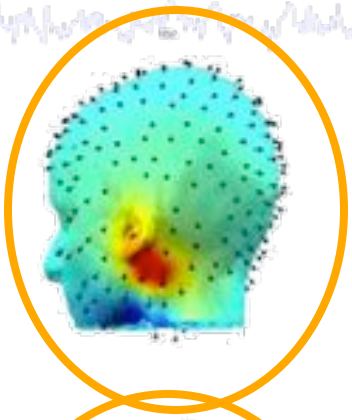


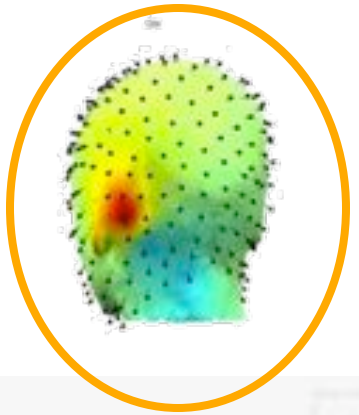
Clustering by spectra (1 subject)

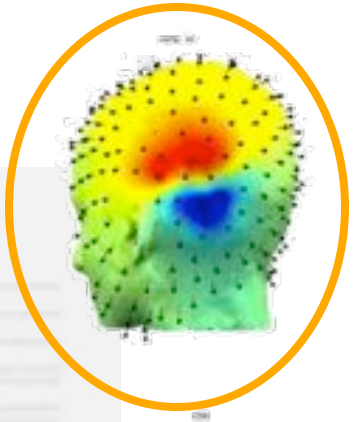


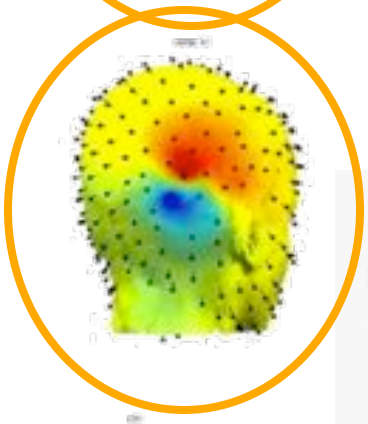


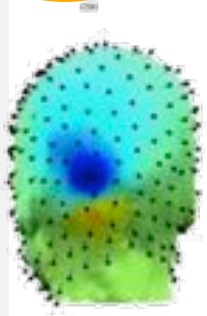


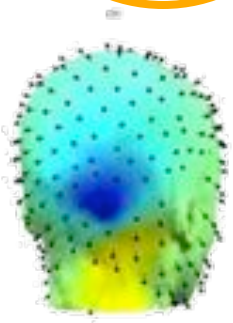


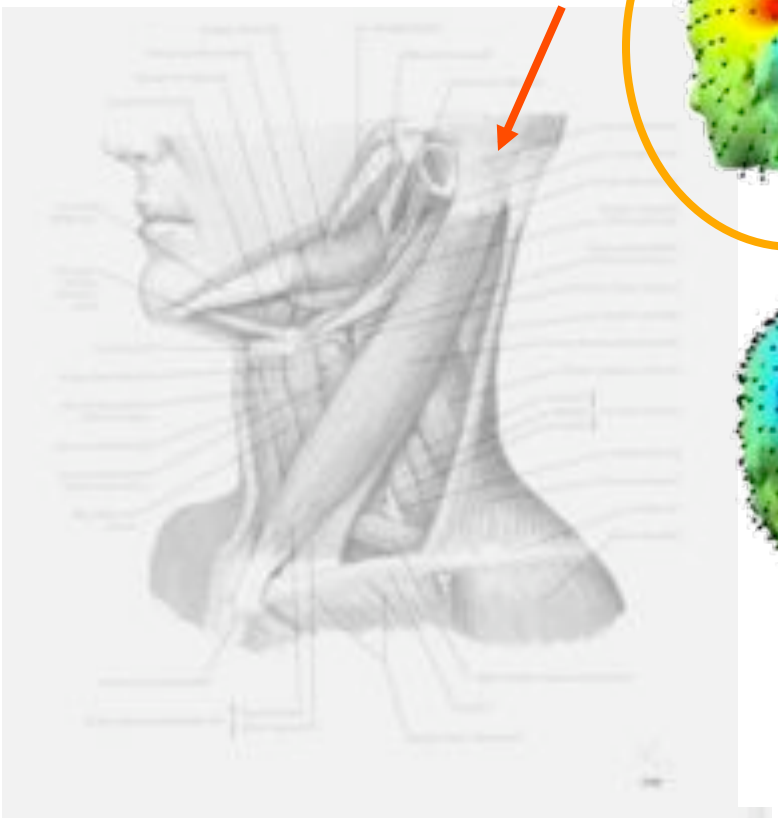


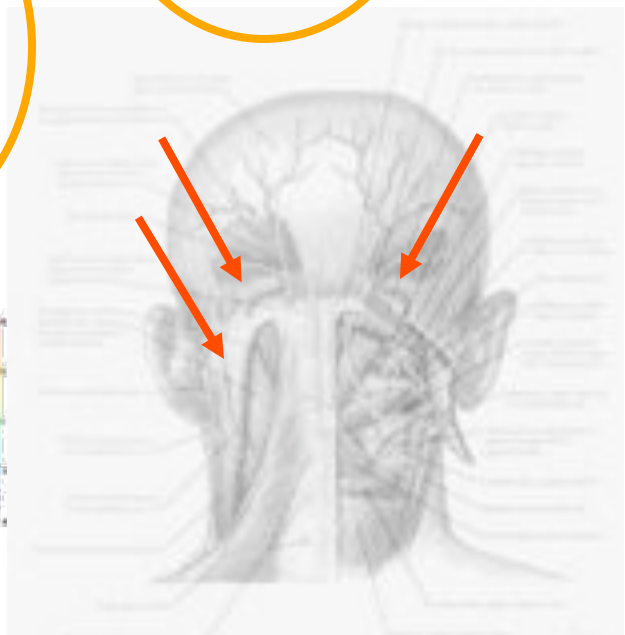












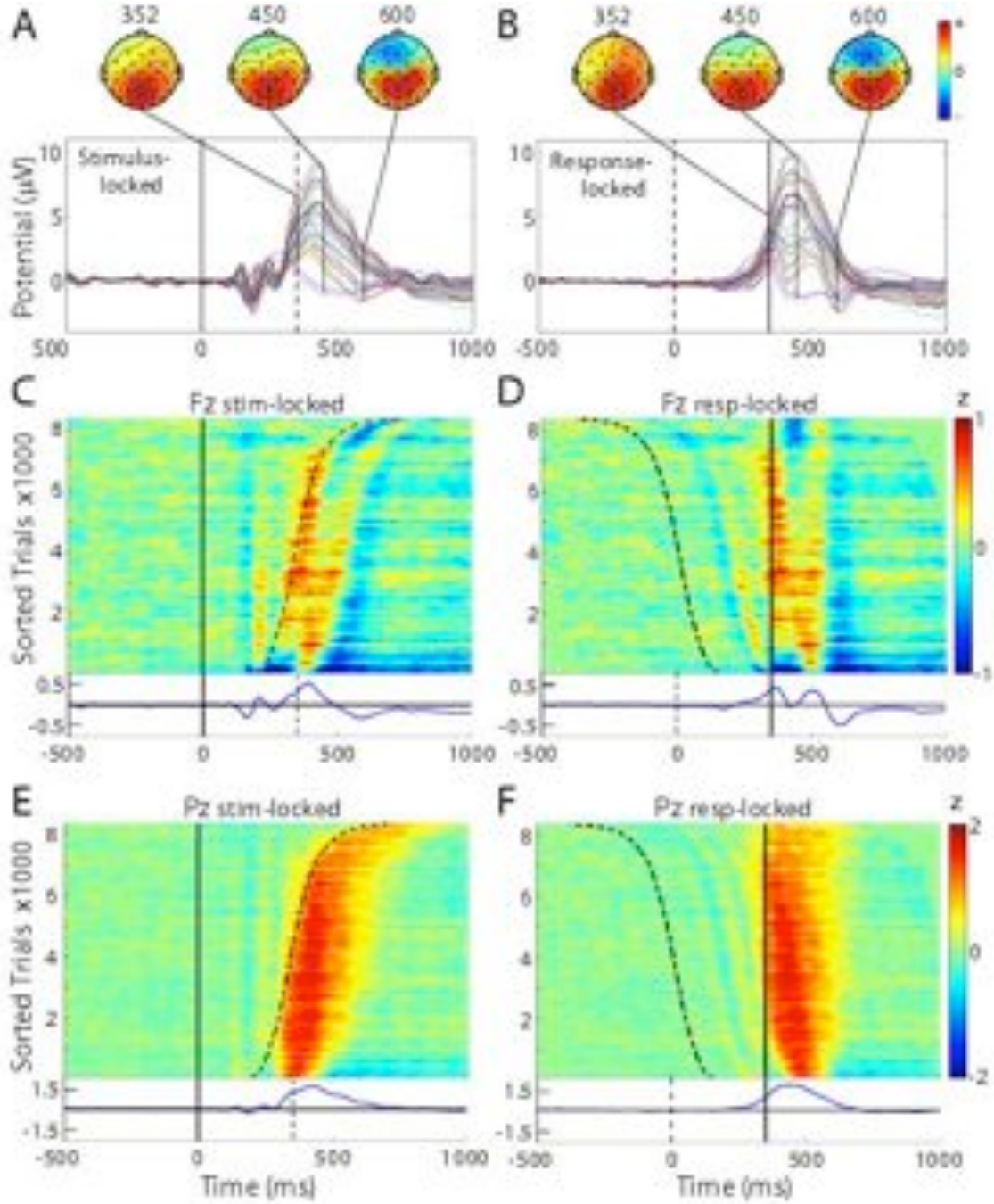
ıber, 2007, Santiago, Chile: Scott Makeig – Comp

Makeig et al., *unpublished*

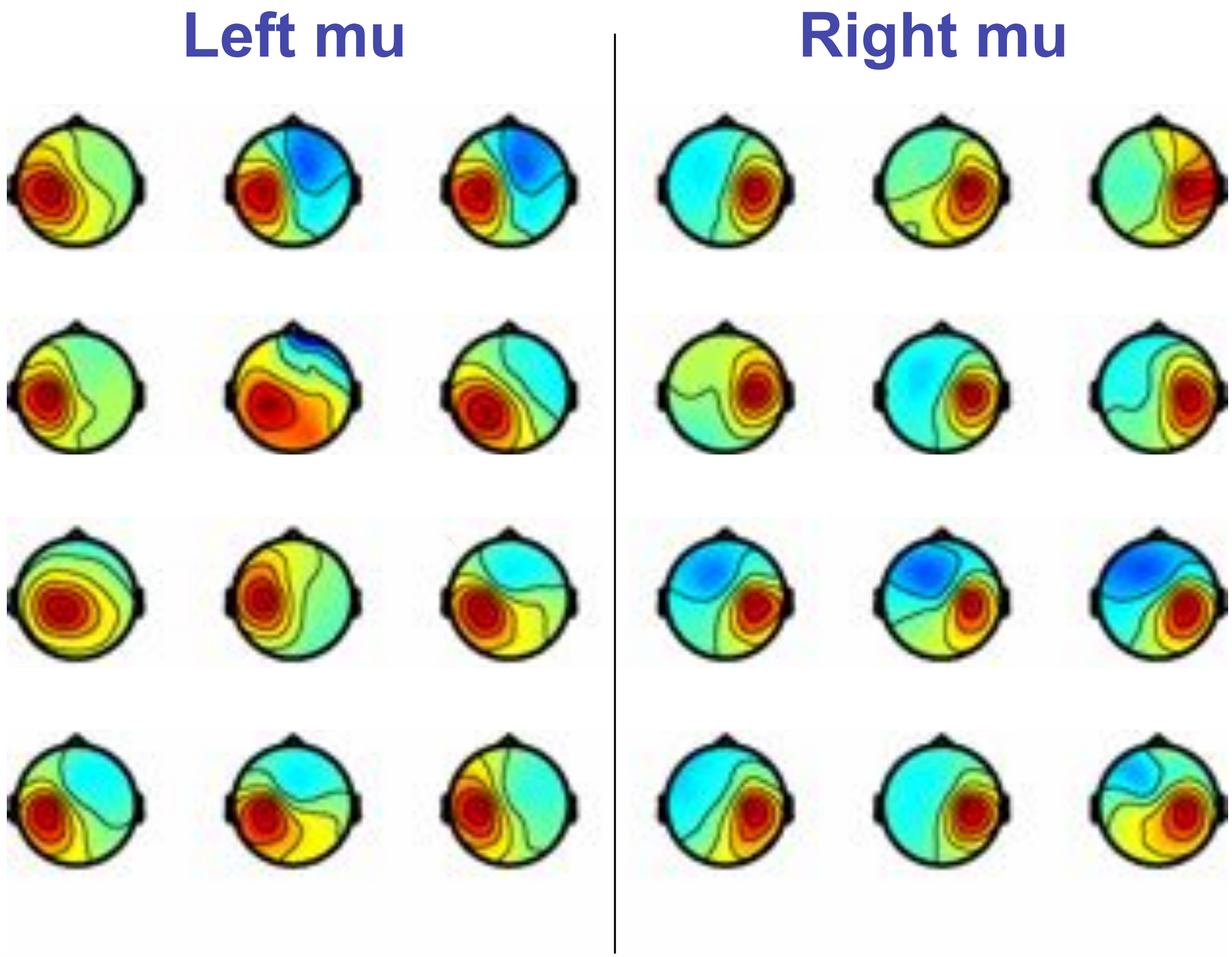
Visual Selective Attention Task

+

15 subjects


A
352
450
600
Stimulus-locked
Potential (μV)
B
Response-locked
C
Fz stim-locked
Sorted Trials x1000
D
Fz resp-locked
E
Pz stim-locked
F
Pz resp-locked
Time (ms)

Clustering ICA components by eye



Makeig et al., *submitted*

Study IC Clustering: Assumptions

- Assumes there are *functionally equivalent* ICs across most subjects.
- Assumes these ICs have *similar responses* to experimental conditions across **~all** measures (ERP, ERSP, ITC…)
- Creates *non-overlapping partitions* so that each IC belongs only to one cluster.

EEGLAB Clustering strategy

1. Cluster on **multiple measures** (dipole locations, scalp maps, spectra, ERPs, ITCs, ERSPs) in **one or more conditions**.
2. **Reduce the dimension** of each measure to a principal component subspace.
3. Compose a PCA-reduced **position vector** for each component.
4. **Cluster** the composed component vectors using k-means or other.
5. Use the computed component measures (not PCA-reduced) to **visualize the activities and spatial properties** of the clustered components.
6. Compute and visualize the **cluster-mean measures**.
7. Use the **clustered study set data** as input into **std_** functions.

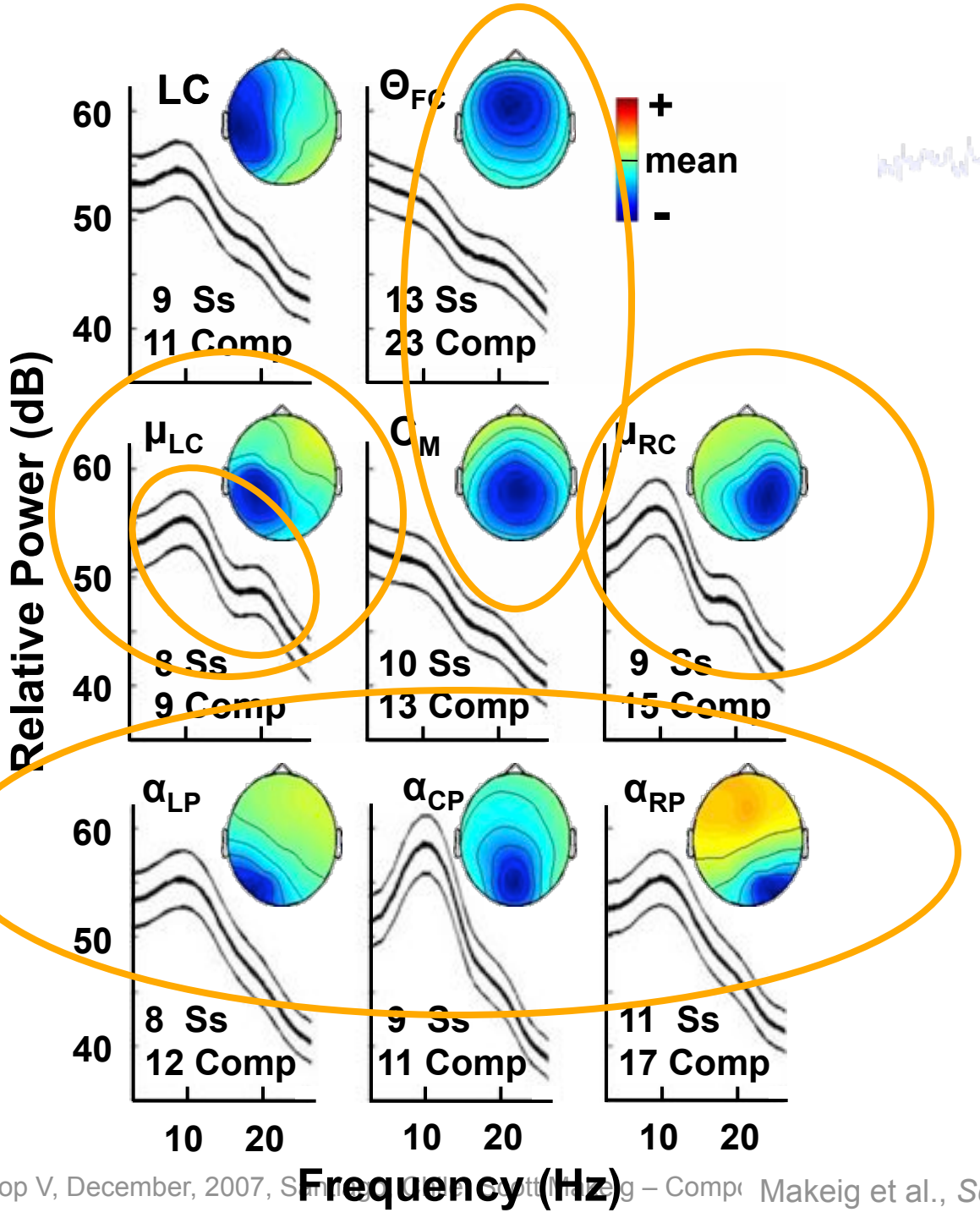
EEGLAB clustering procedure

1. Identify a set of datasets as an EEGLAB **study** or '**studyset**'.
2. Specify the subject **group**, **subject** code, **condition** and **session** of each dataset in the study.
3. Identify **components to cluster** in each study dataset.
4. Decide on **component measures** to use in clustering the study and/or to evaluate the obtained component clusters.
5. Compute the component measures for each study dataset.
6. **Cluster the components** on these component measures.
7. Review the obtained **clusters** (e.g., their scalp maps, dipoles, and activity measures).
8. **Edit the clusters** (manually remove/shift components, make sub-clusters, merge clusters, re-cluster).
9. Perform **signal processing** within or between selected clusters.

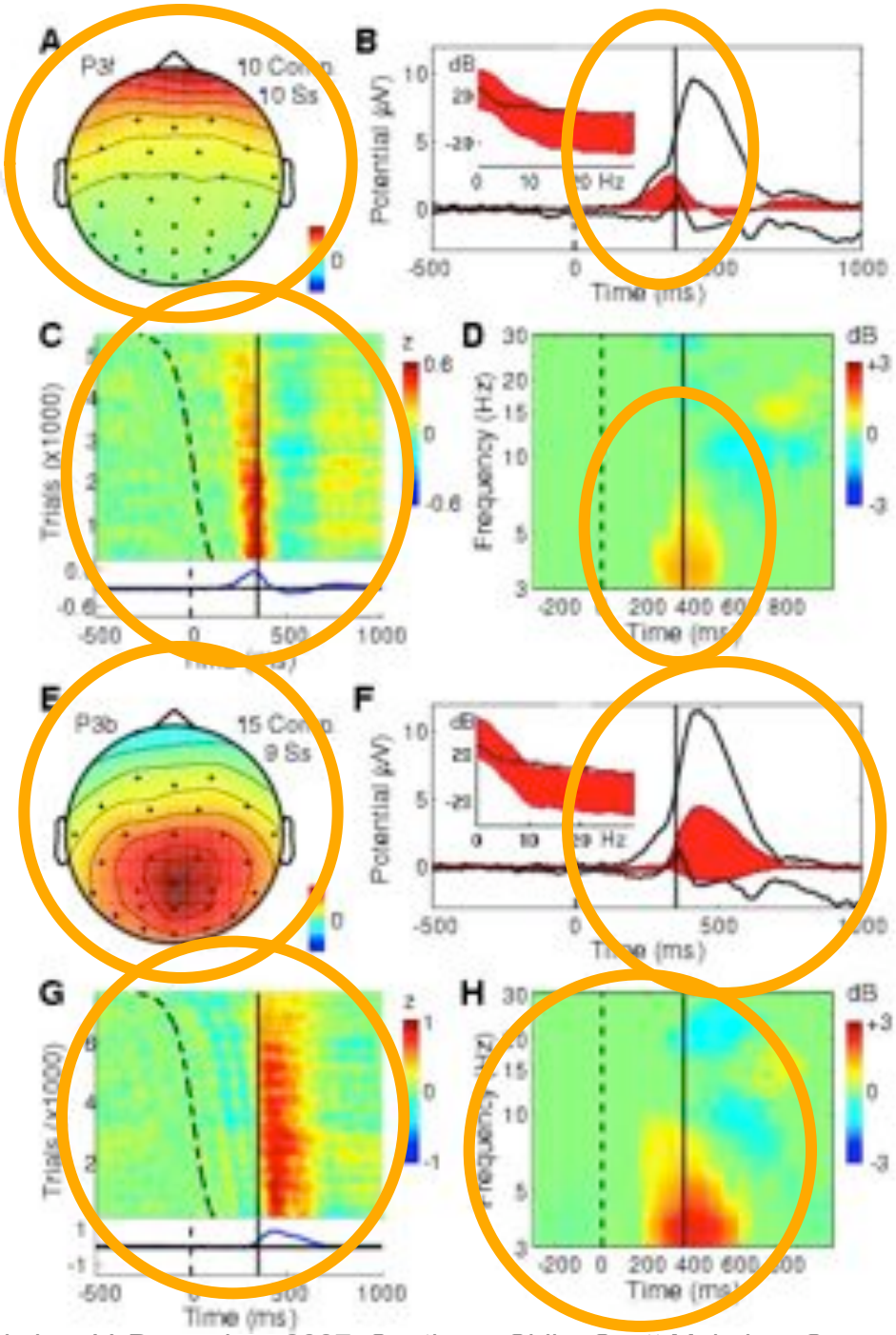
P300 -- Semi-automated clustering

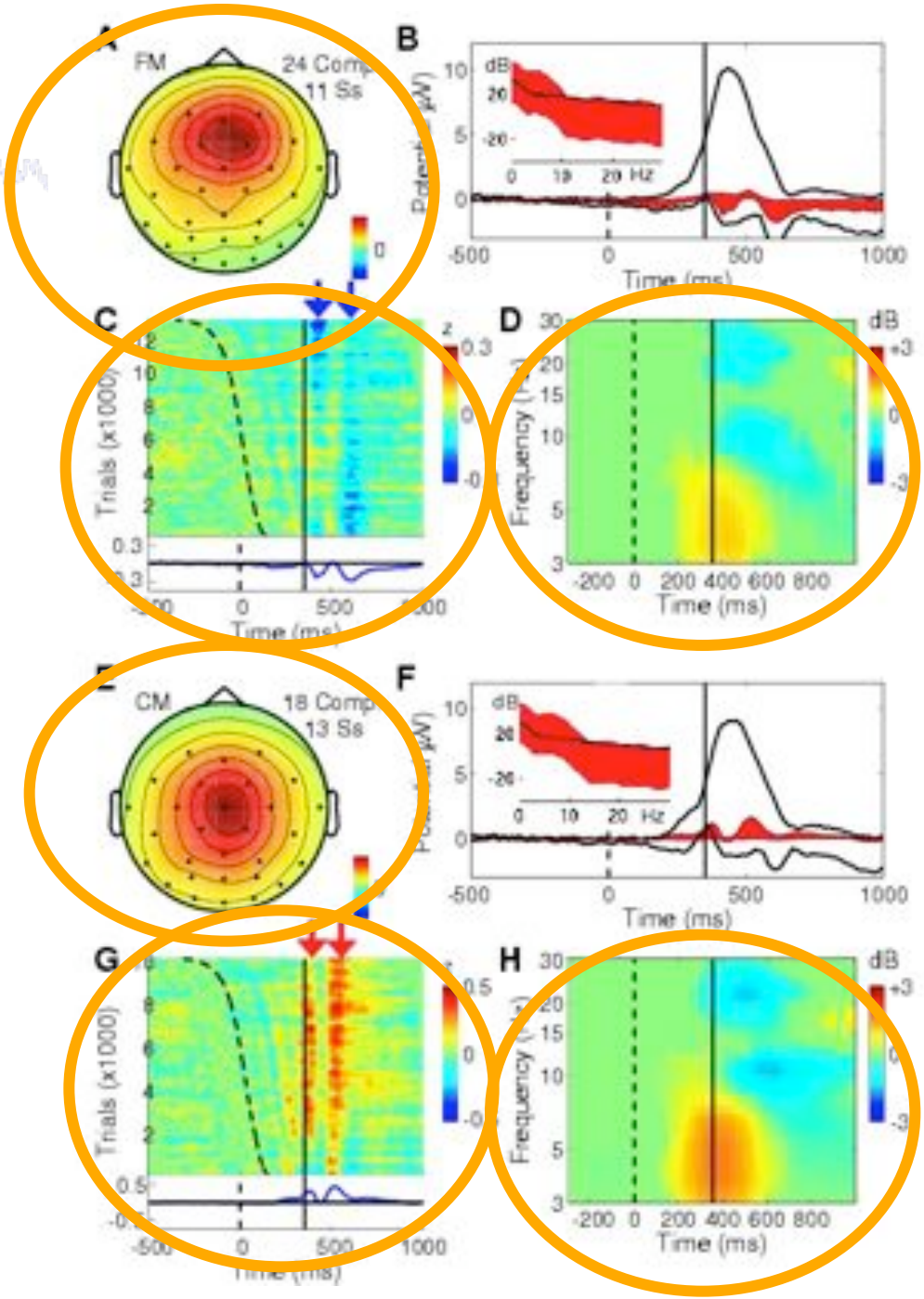
- Clustered components from 15 Ss using a ‘component distance metric’ incorporating differences between their (weighted) scalp maps, dipole locations, spectra, ERP, ERSP, and ITC patterns.
- Hand-adjusted clusters to remove outliers.
- Determined time/frequency regions of significant ERSP and ITC for each component using permutation-based statistics.
- Used binomial statistics to highlight time/ frequency regions significantly active within clusters.

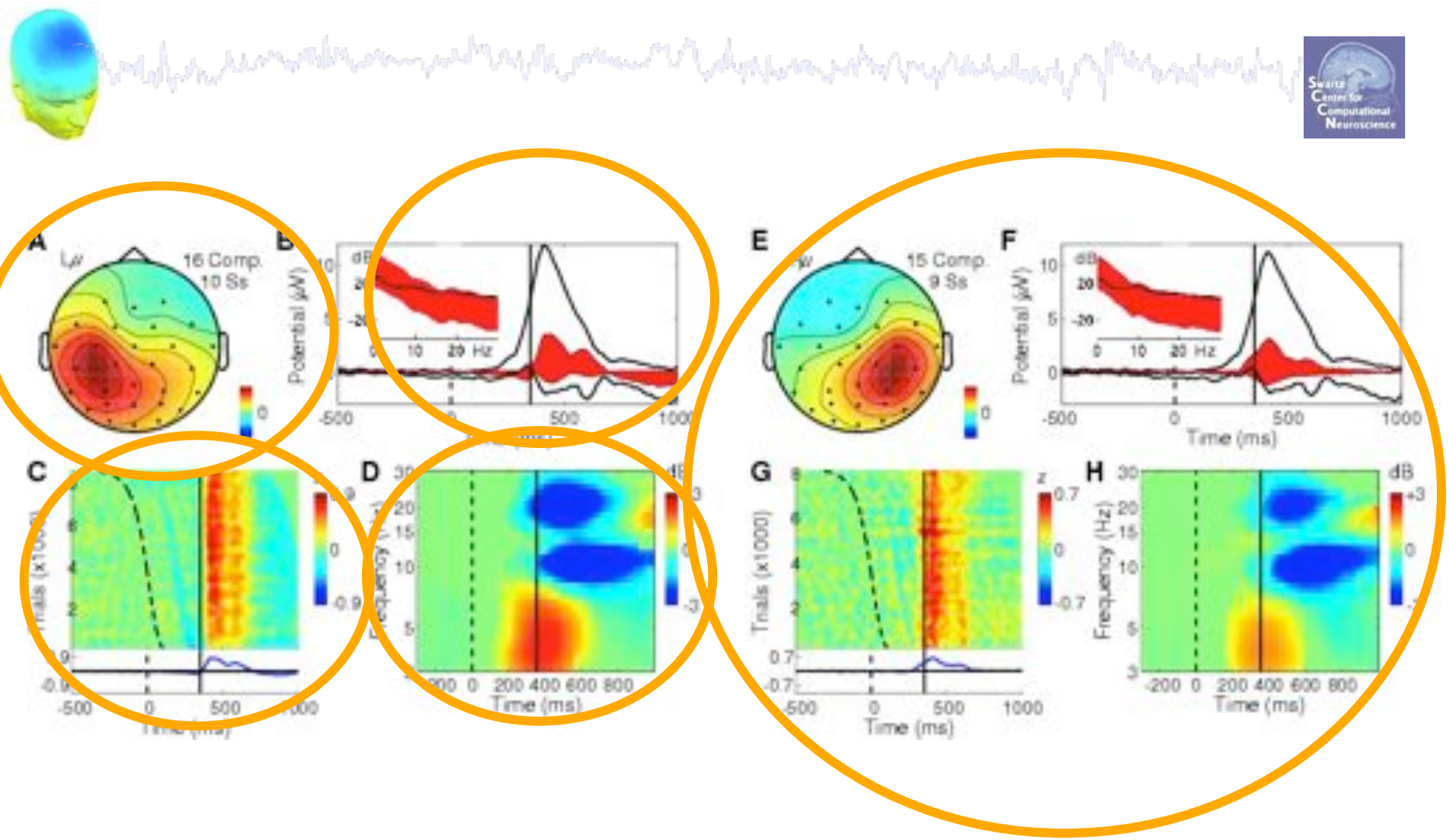
N1 Component Clusters

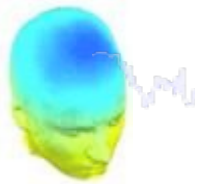


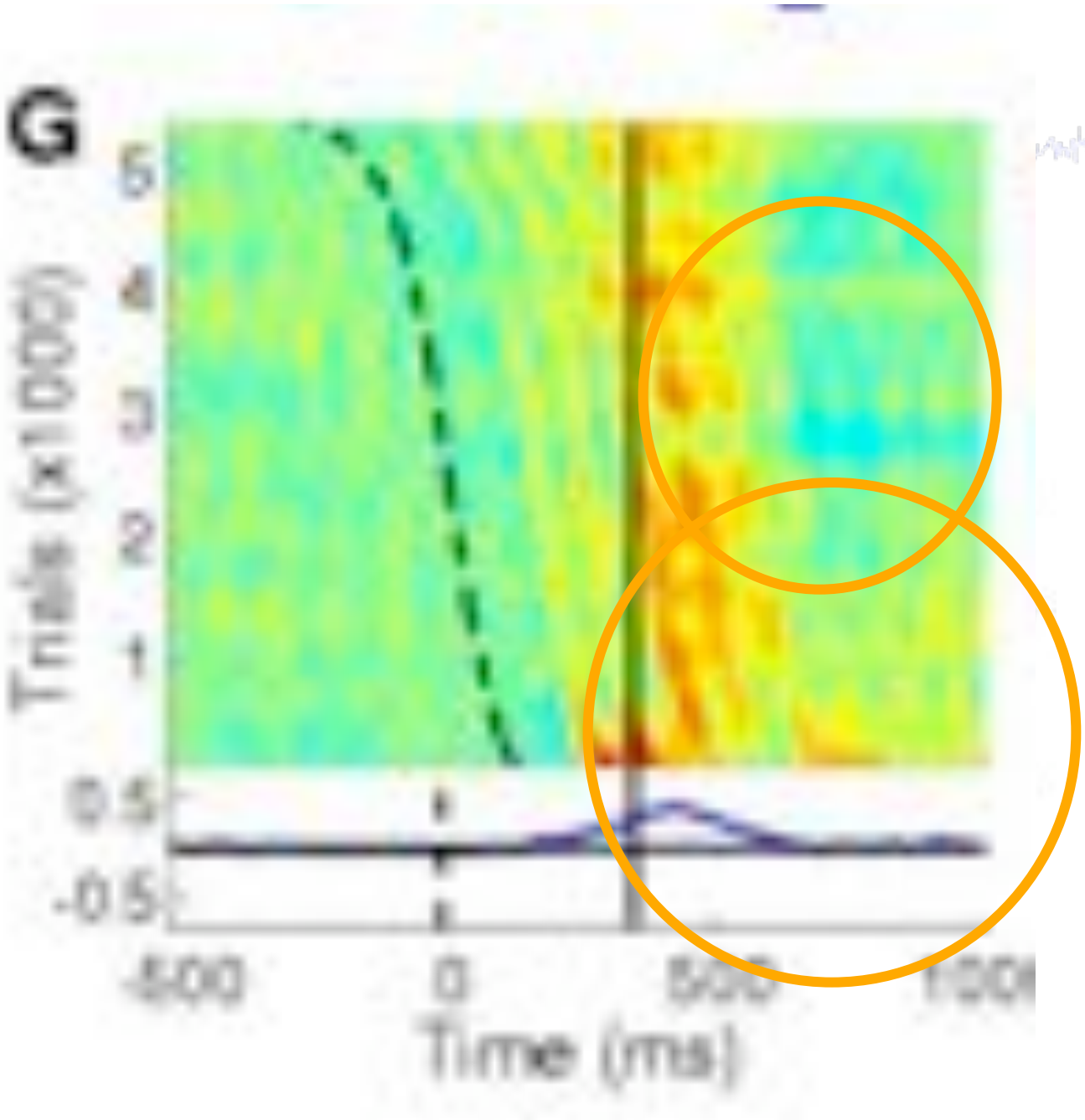
Makeig et al., *Science* 2002

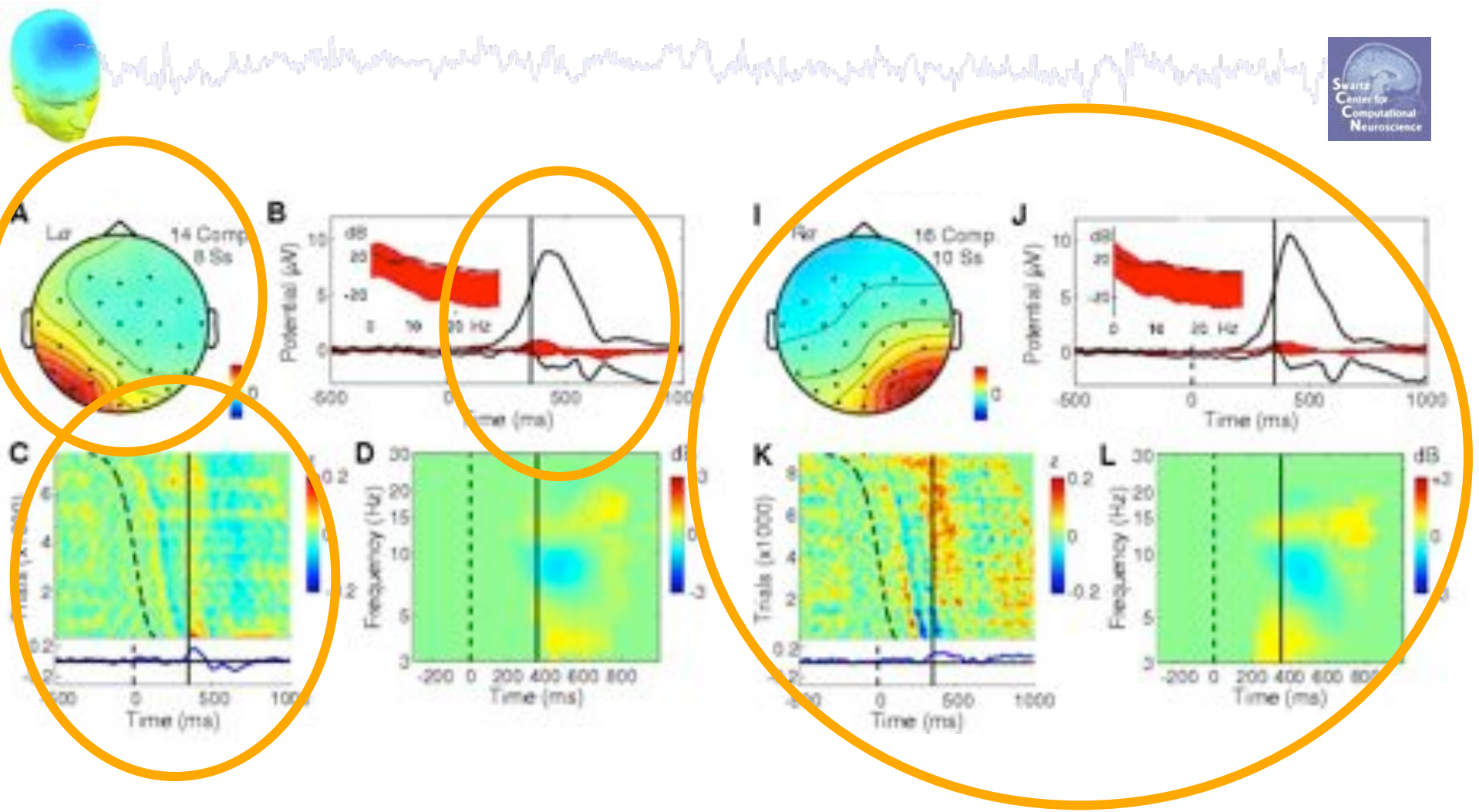

A
P3f
10 Co
10 Ss
B
dB
Potential (μV)
Time (ms.)
C
Trials (x1000)
Time (ms.)
D
dB
Frequency (Hz)
Time (ms.)
E
P3b
15 Co
9 Ss
F
dB
Potential (μV)
Time (ms.)
G
Trials (x1000)
Time (ms.)
H
dB
Frequency (Hz)
Time (ms.)


A
FM
24 Comp
11 Ss
B
dB
20
-20
0 10 20 Hz
10
5
0
-500 0 500 1000
Time (ms)
C
Trials (x1000)
Time (ms)
D
dB
+3
0
-3
Frequency (Hz)
30
20
15
10
5
3
-200 0 200 400 600 800
Time (ms)
E
CM
18 Comp
13 Ss
F
dB
20
-20
0 10 20 Hz
-500 0 500 1000
Time (ms)
G
Trials (x1000)
Time (ms)
H
dB
+3
0
-3
Frequency (Hz)
30
20
15
10
5
3
-200 0 200 400 600 800
Time (ms)

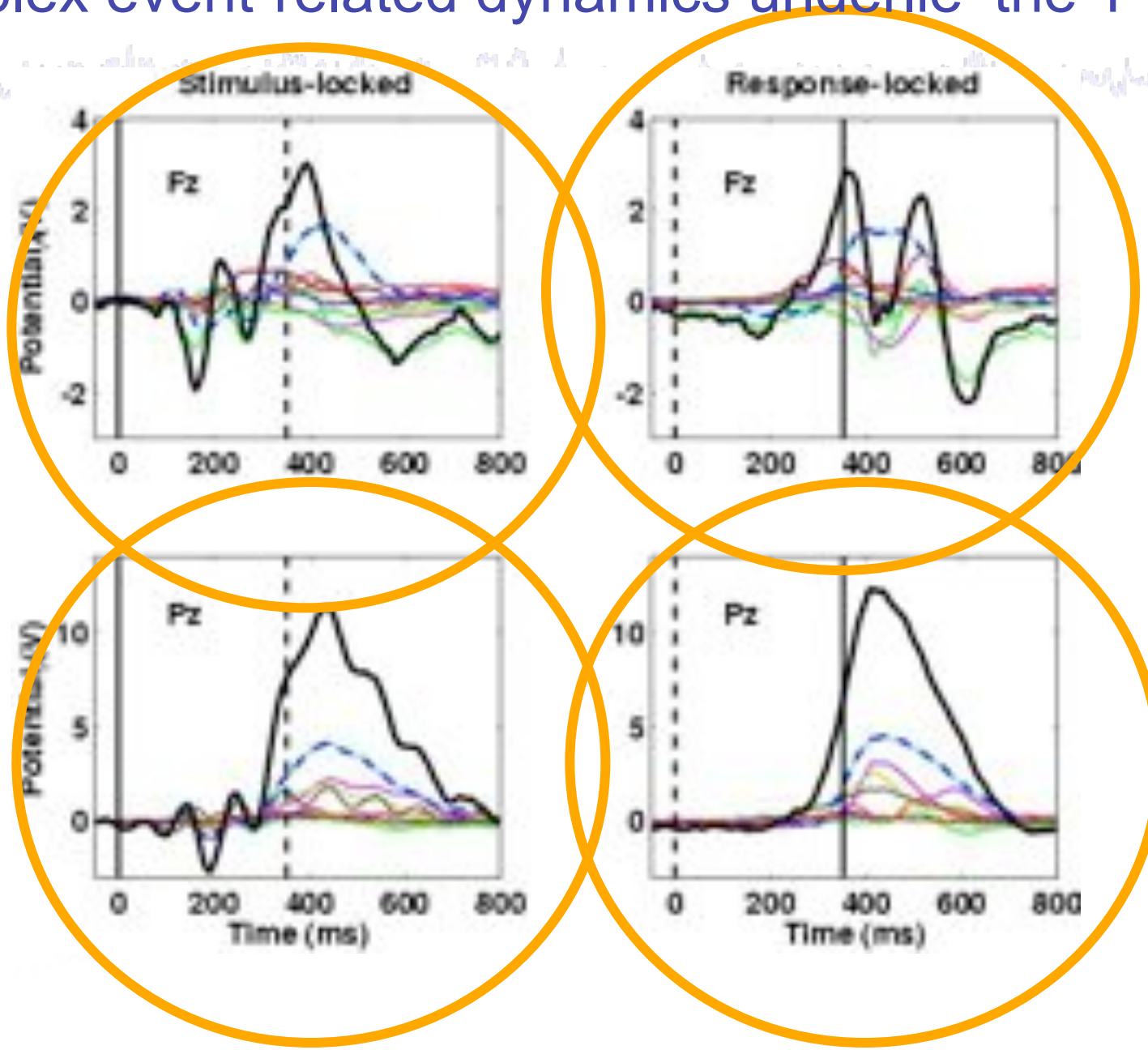

A
16 Comp.
10 Ss
B
Potential (μV)
10 20 Hz
-500
500
1000
C
Trials (x1000)
-500
0
500
D
Frequency (Hz)
200 400 600 800
Time (ms)
E
15 Comp.
9 Ss
F
Potential (μV)
10 20 Hz
-500
0
500
1000
Time (ms)
G
Trials (x1000)
-500
0
500
1000
Time (ms)
H
Frequency (Hz)
-200 0 200 400 600 800
Time (ms)
dB



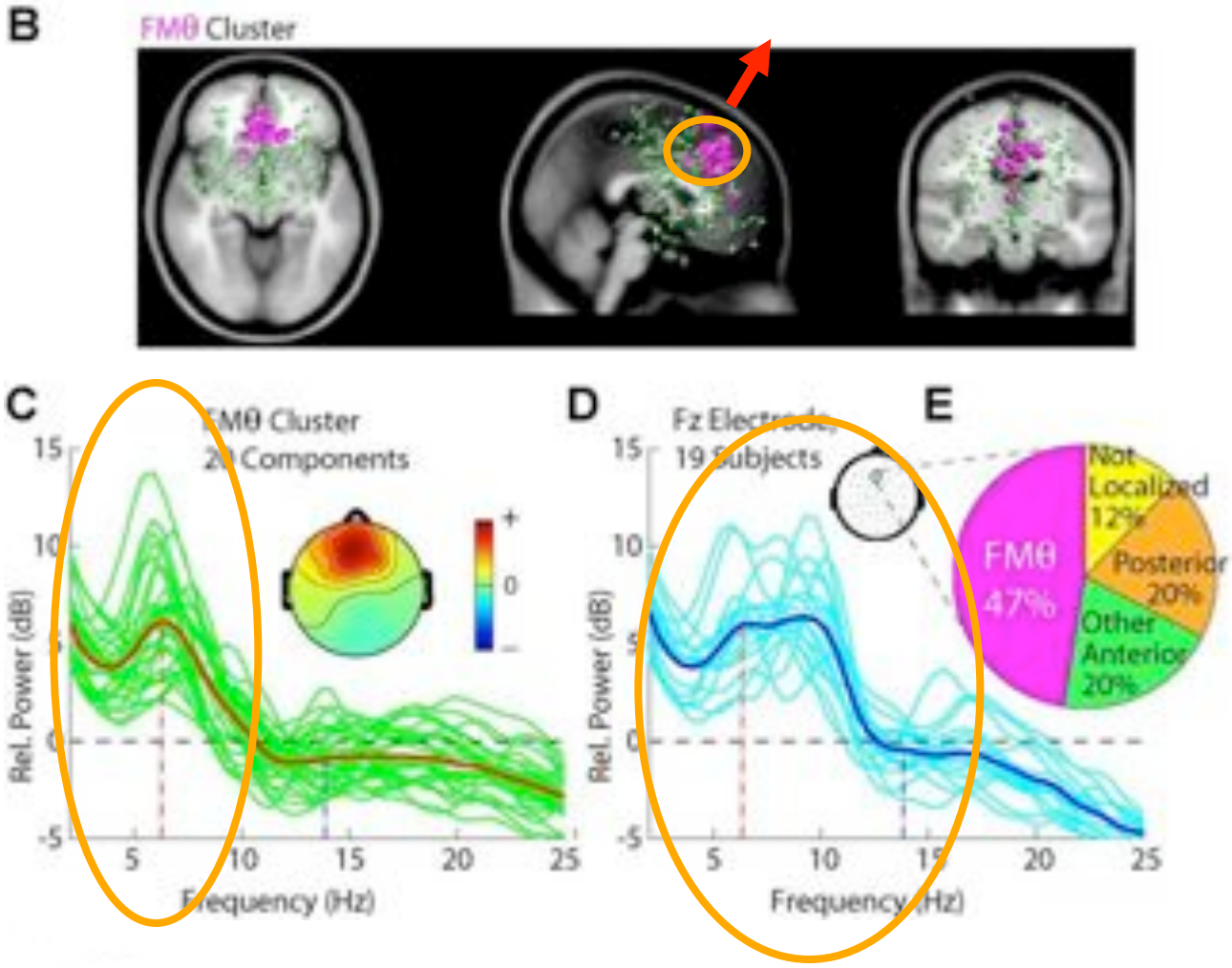

G
Trials (x1000)
6
4
3
2
1
0.5
-0.5
-500
0
Time (ms)


A
Lar
14 Comp
8 Ss
B
Potential (μV)
Time (ms)
C
Trials (x1000)
Time (ms)
D
Frequency (Hz)
Time (ms)
I
Rar
16 Comp
10 Ss
J
Potential (μV)
Time (ms)
K
Trials (x1000)
Time (ms)
L
Frequency (Hz)
Time (ms)

Complex event-related dynamics underlie ‘the’ P300

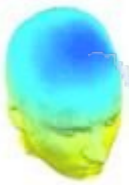


A FMΘ cluster during working memory



Onton et al., *NeuroImage* 2005

Are obtained component clusters “real“?



- **Naïve realism** (a.k.a. “expertise”)
  - “Yes! … because I know one when I see one!”
    - “If it appears where Mu components appear,
      and acts like Mu components act,
      then it IS a Mu component!”
- **Convergent evidence** (a.k.a., “doublechecking”)
  - Two possible approaches:
    - Cluster on PLACE → Check ACTIVITY consistency (re task)
    - Cluster on ACTIVITY → Check PLACE consistency
- **Absolute truth:**
    - More ideal forward and inverse models
    - Invasive multiscale recordings + modeling

Should all subjects be included in each cluster?

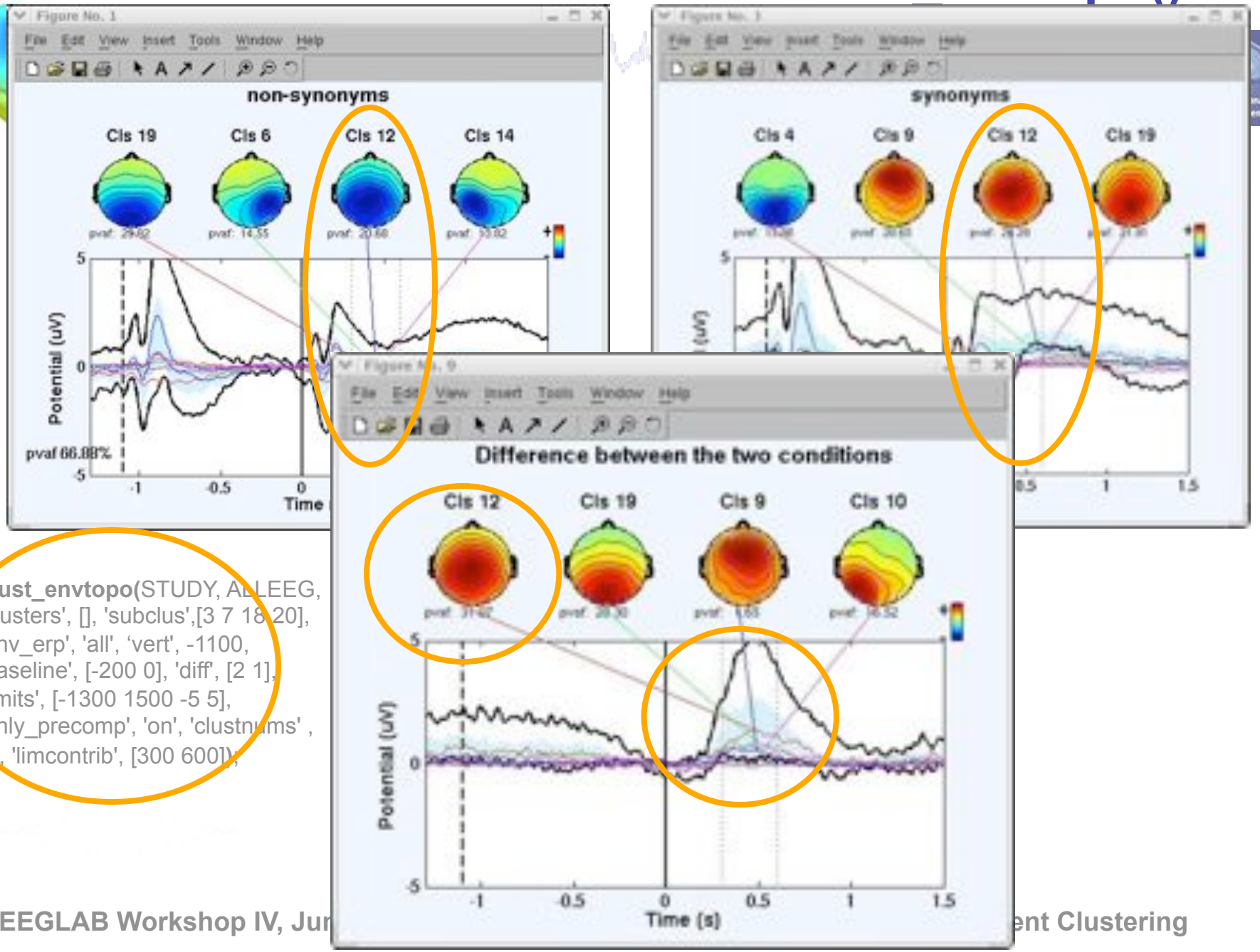
Not all subjects contribute components to each cluster.

Why not?

- Different numbers of artifact components (~INR)
- Subject differences!?
- Is my subject group a Gaussian cloud??

  → subject space

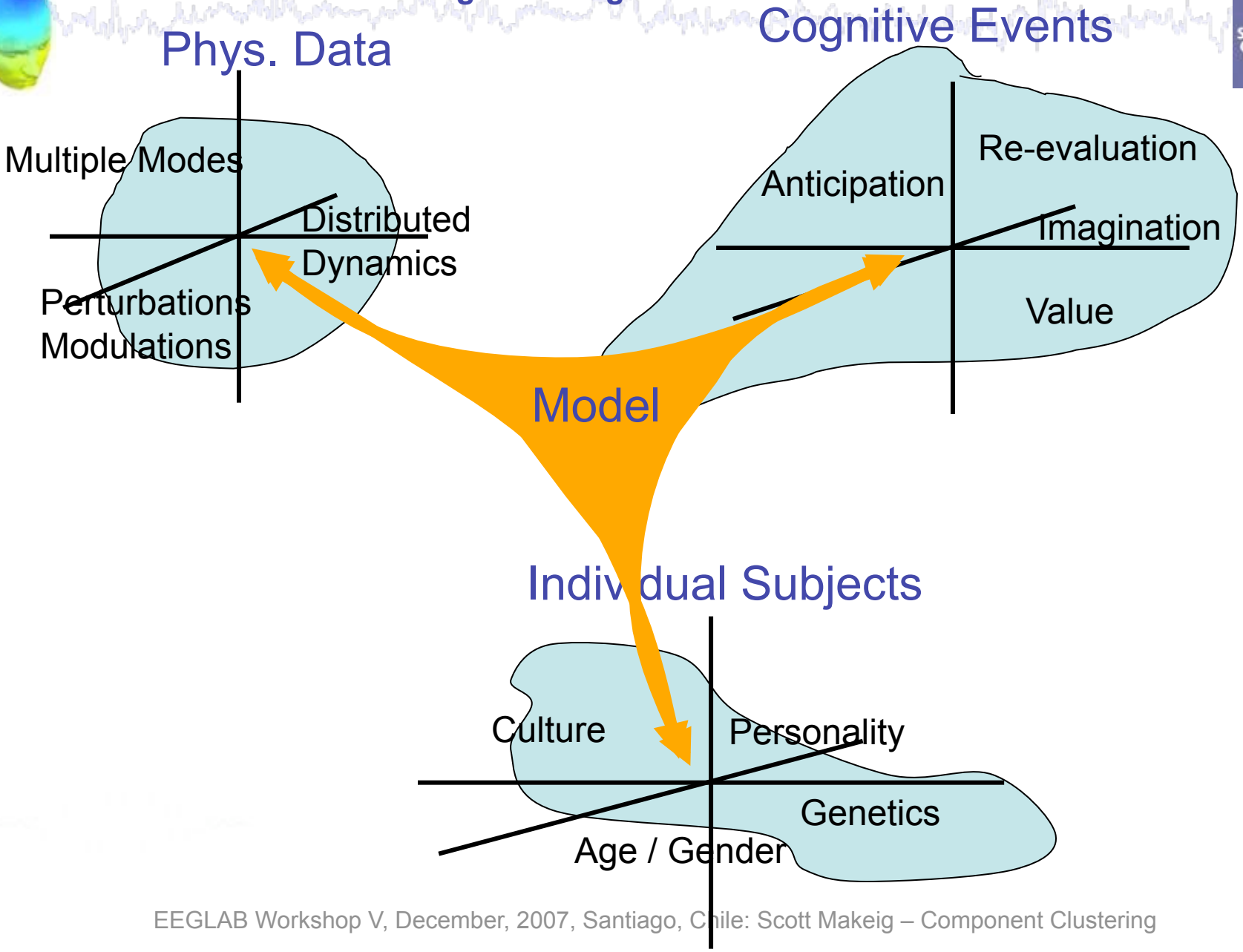
Cluster ERP contributions - clust_envtopo()



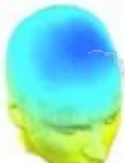
clust_envtopo(STUDY, ALLEEG,
'clusters', [], 'subclus',[3 7 18 20],
'env_erp', 'all', 'vert', -1100,
'baseline', [-200 0], 'diff', [2 1],
'limits', [-1300 1500 -5 5],
'only_precomp', 'on', 'clustnums' ,
4, 'limcontrib', [300 600]);

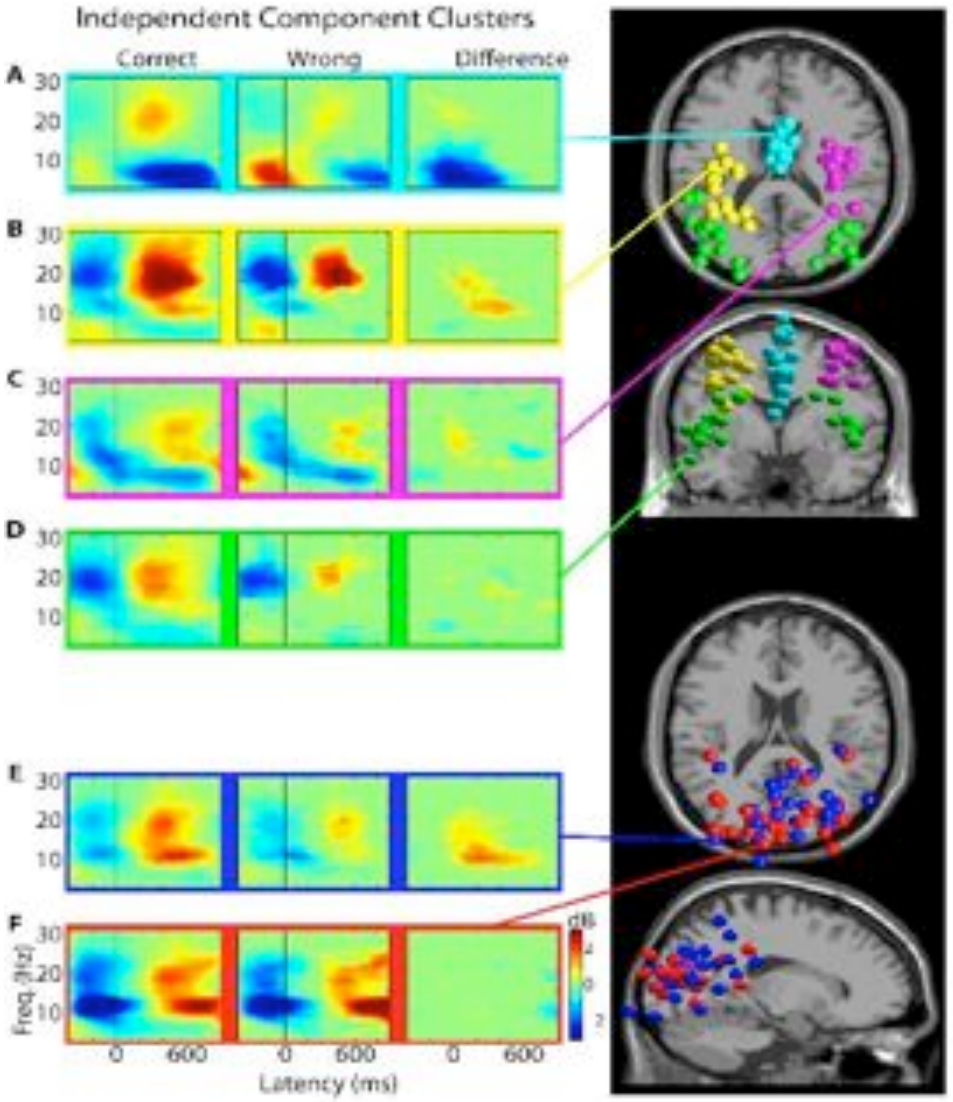
Beyond Clustering

The goal of cognitive neuroscience



Study IC Clustering





Sometime clusters are spatially separate AND have distinct responses.

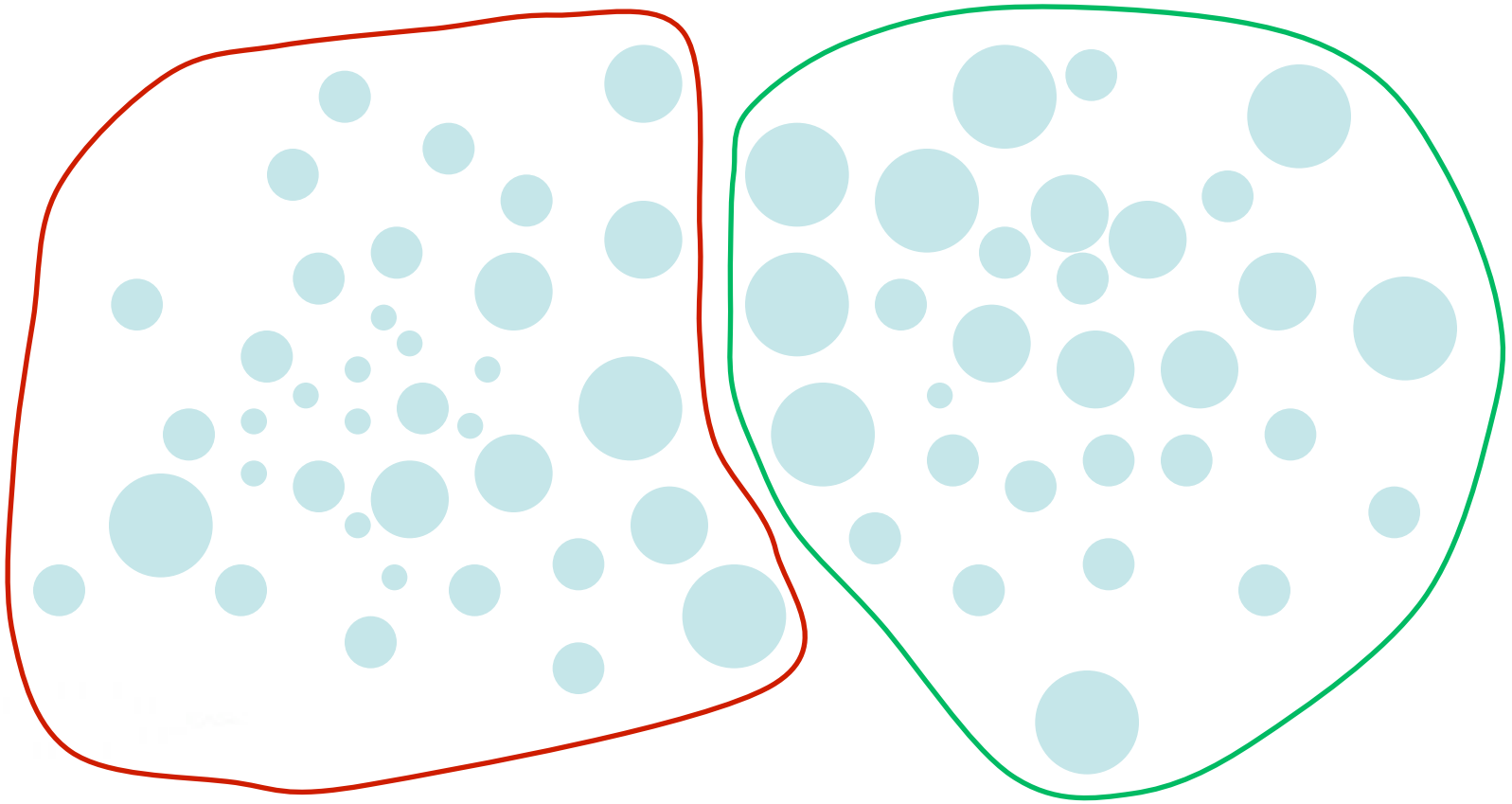
In other cases, they have similar responses or they overlap spatially.

Problems with multi-measure clustering

What are the clusters according to location?

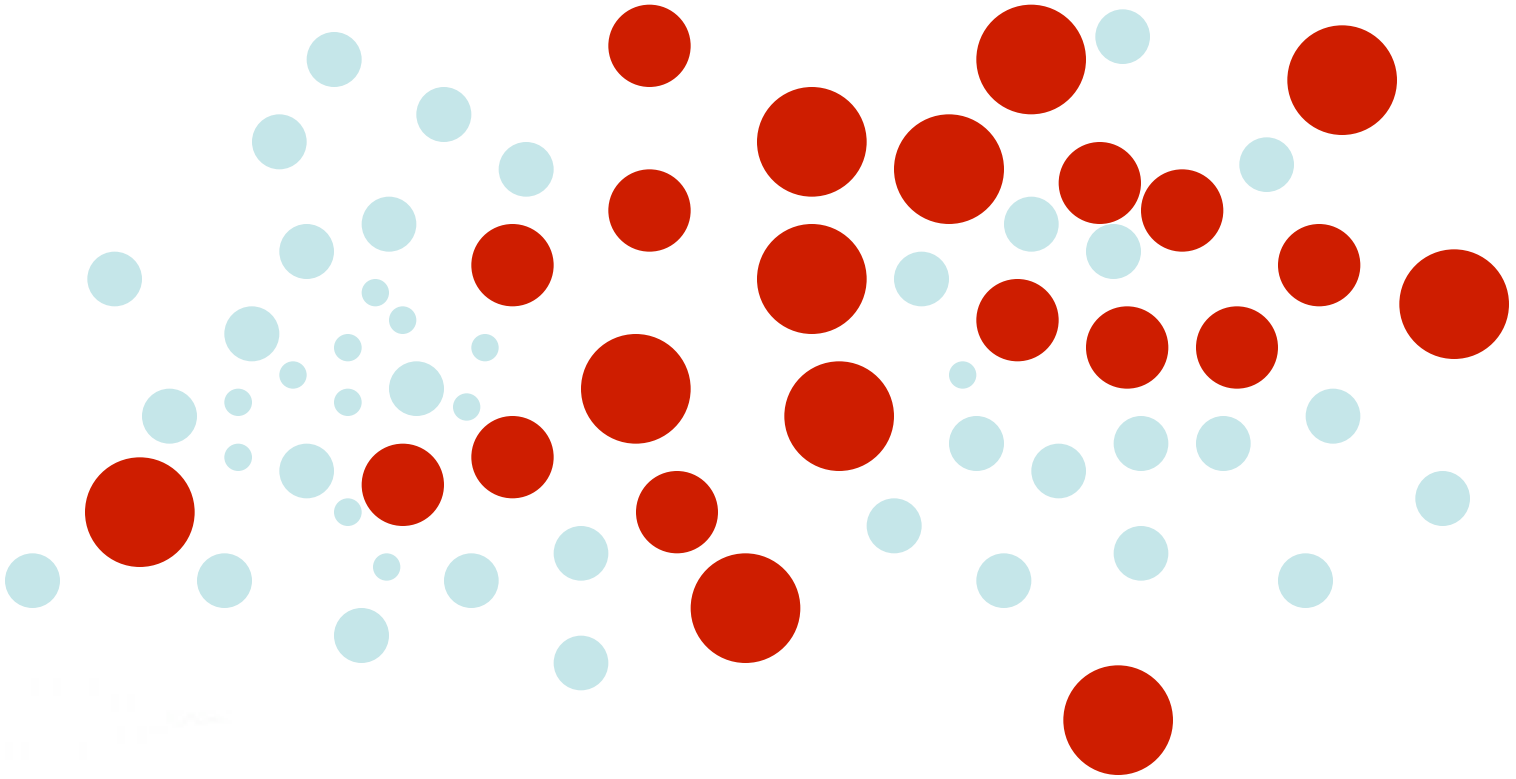
Problems with multi-measure clustering

What are the clusters according to location?



Problems with multi-measure clustering

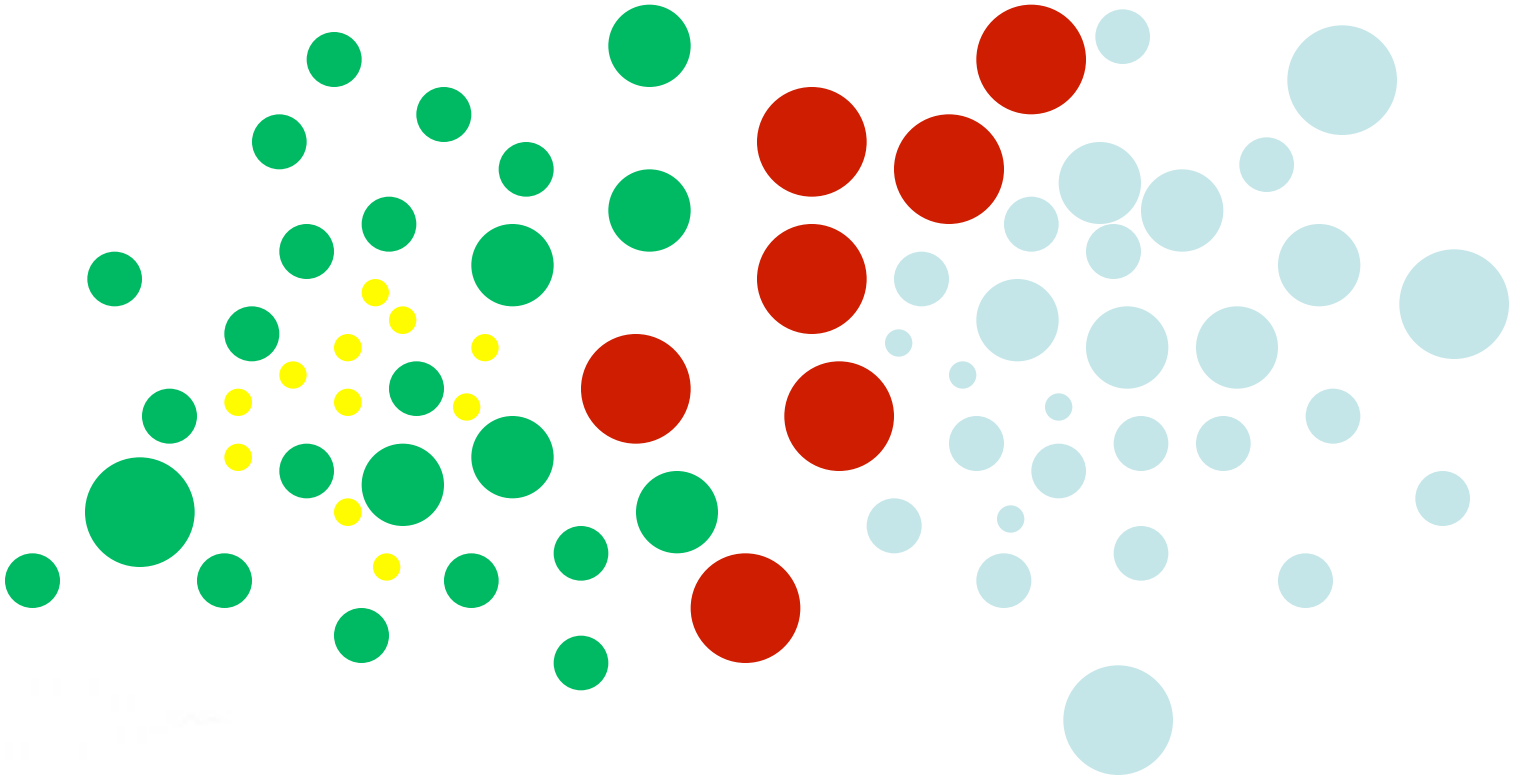
What are the clusters according to size ?



Problems with multi-measure clustering

What are the clusters according to location and size?

Depends on how much weight we give each measure...



Study IC Clustering: Conceptual Problems

1. Condition responses in the same brain area may significantly differ across subject groups. (Often the goal of the study is finding these differences)
2. Components may have similar responses for one measure (e.g., ERSP) but not for another (e.g., ERP). This is one of the most serious issues.
3. *Boosts* evidence by rejecting ICs that are in the same brain area but show different responses. This makes calculating unbiased significance values difficult.

Measure Projection

- Instead of clustering, we assume each location in the brain has a unique EEG response.
- The response at each location is calculated as the weighted sum of IC responses in its dipole neighborhood.
- Weights are assigned by passing the distance between the location and IC dipole through a Gaussian function.
- The std. dev. of this function should represent expected error in dipole localization plus inter-subject variability.

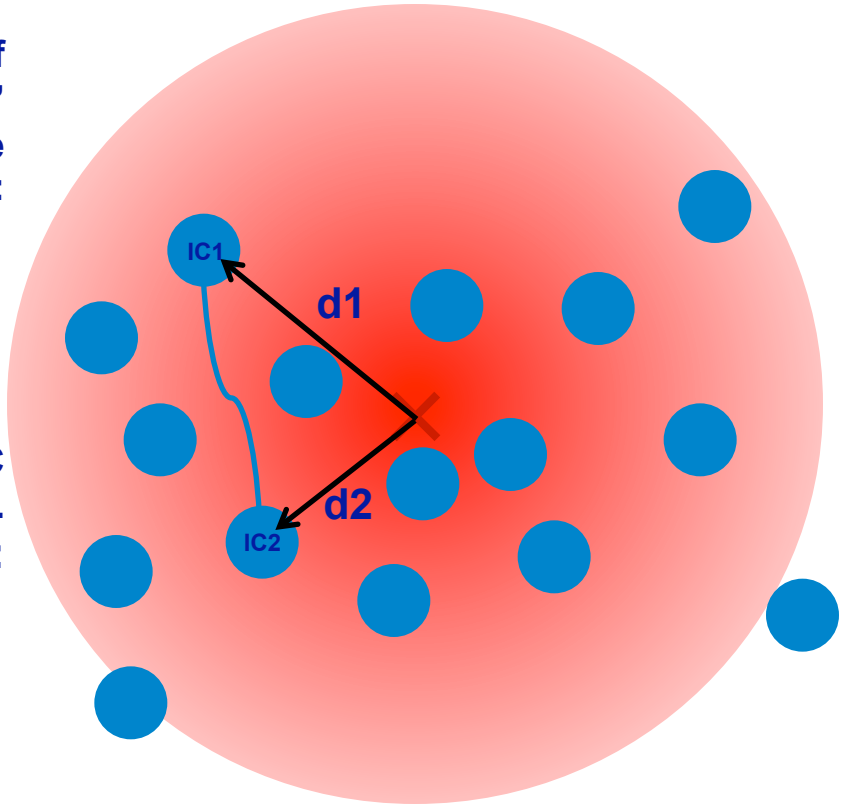
Measure Projection: Definition

Now we can extend this concept of convergence to neighborhoods with 'soft' Gaussian boundaries, for each IC pair we modify the membership function:

$$MG(IC_i, IC2_j) = MG(IC_i)MG(IC_j)$$

Where $MG(IC) = e^{-\frac{d^2}{2\sigma^2}}$ (d is distance from IC equiv dipole to neighborhood center).

Convergence can now be defined as:



$$convergence = \frac{\sum_{i=1}^{n}\sum_{j=1,j<>i}^{n} S(i,j)e^{-\frac{d_1^2+d_2^2}{2\sigma^2}}}{\sum_{i=1}^{n}\sum_{j=1,j<>i}^{n} e^{-\frac{d_1^2+d_2^2}{2\sigma^2}}}$$

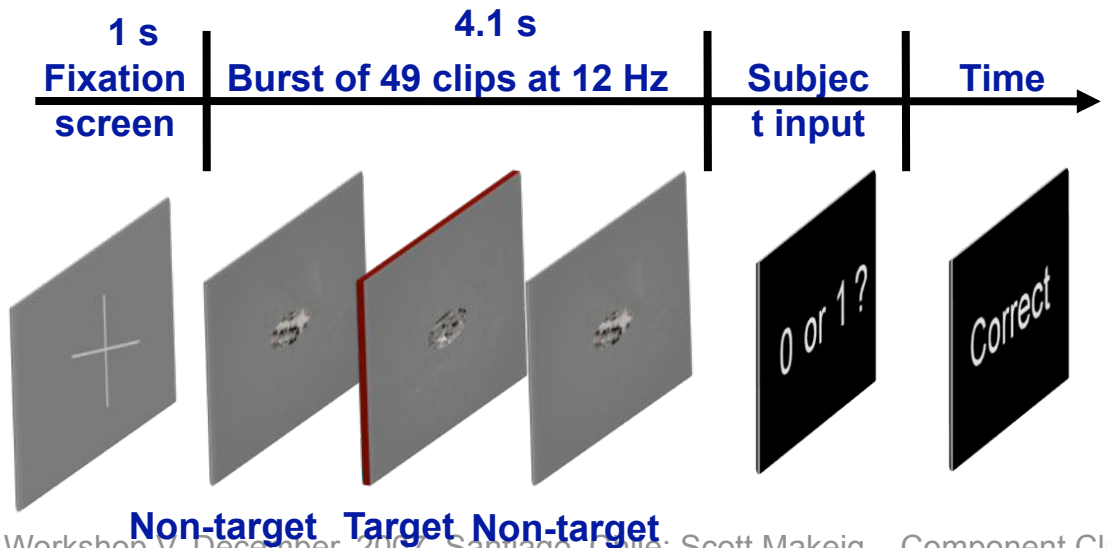
Where S is the pairwise similarity matrix.

This is basically the weighted mean of IC similarities around a location in the brain.

Measure Projection: RSVP Example

Rapid Serial Visual Presentation Experiment

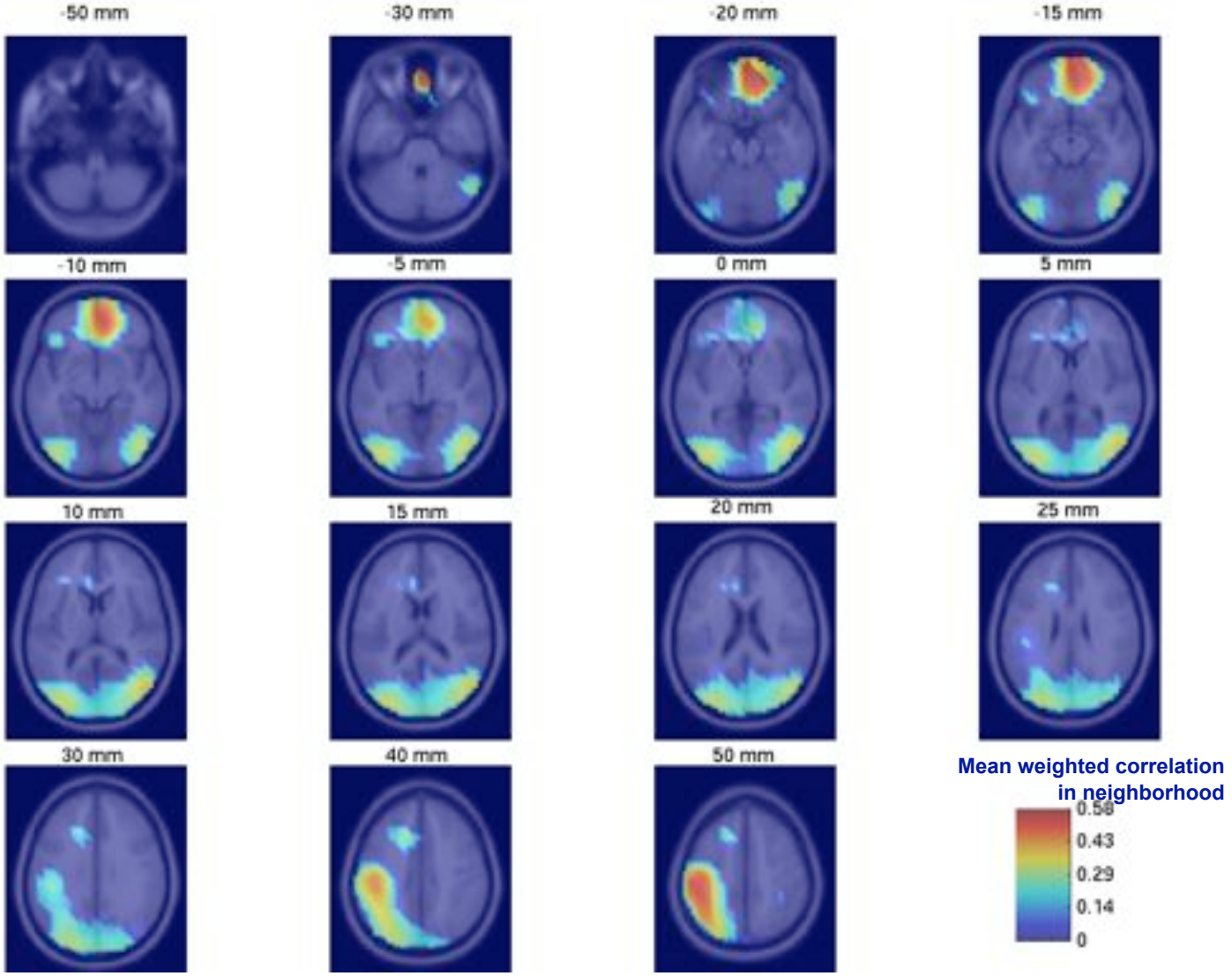
- **8 subjects**
- **15 Sessions**
- **Visual target detection**
- **257 components with equiv. dipoles inside the brain**



Measure Projection: RSVP Example

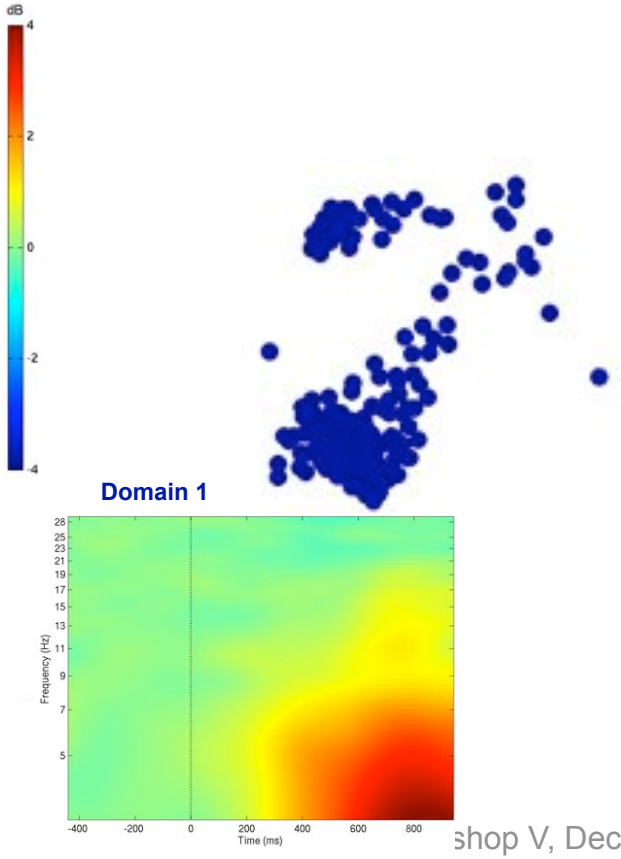
Areas in which convergence is significant (p<0.01).

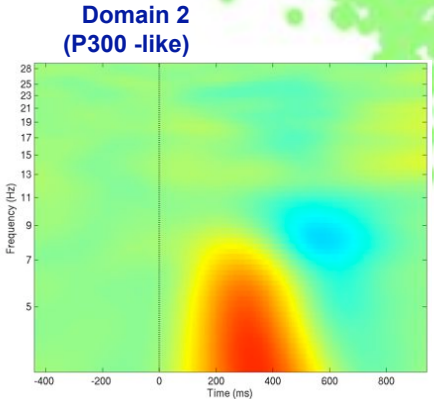
(STD = 12 mm)

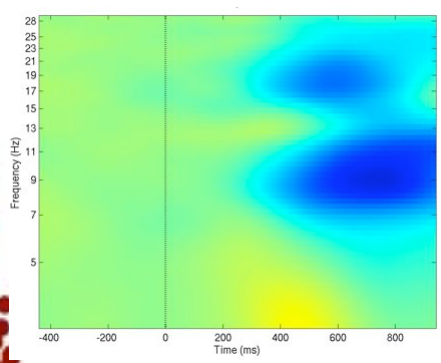


Measure Projection: RSVP Example

To better visualize measure responses in areas with significant convergence, they can be summarized into different *domains*. The exact number of these domains depends on how similar their exemplars are allowed to be. Below you can see ERSP responses in the RSVP experiment form three (3) domains (with the correlation threshold of 0.8).





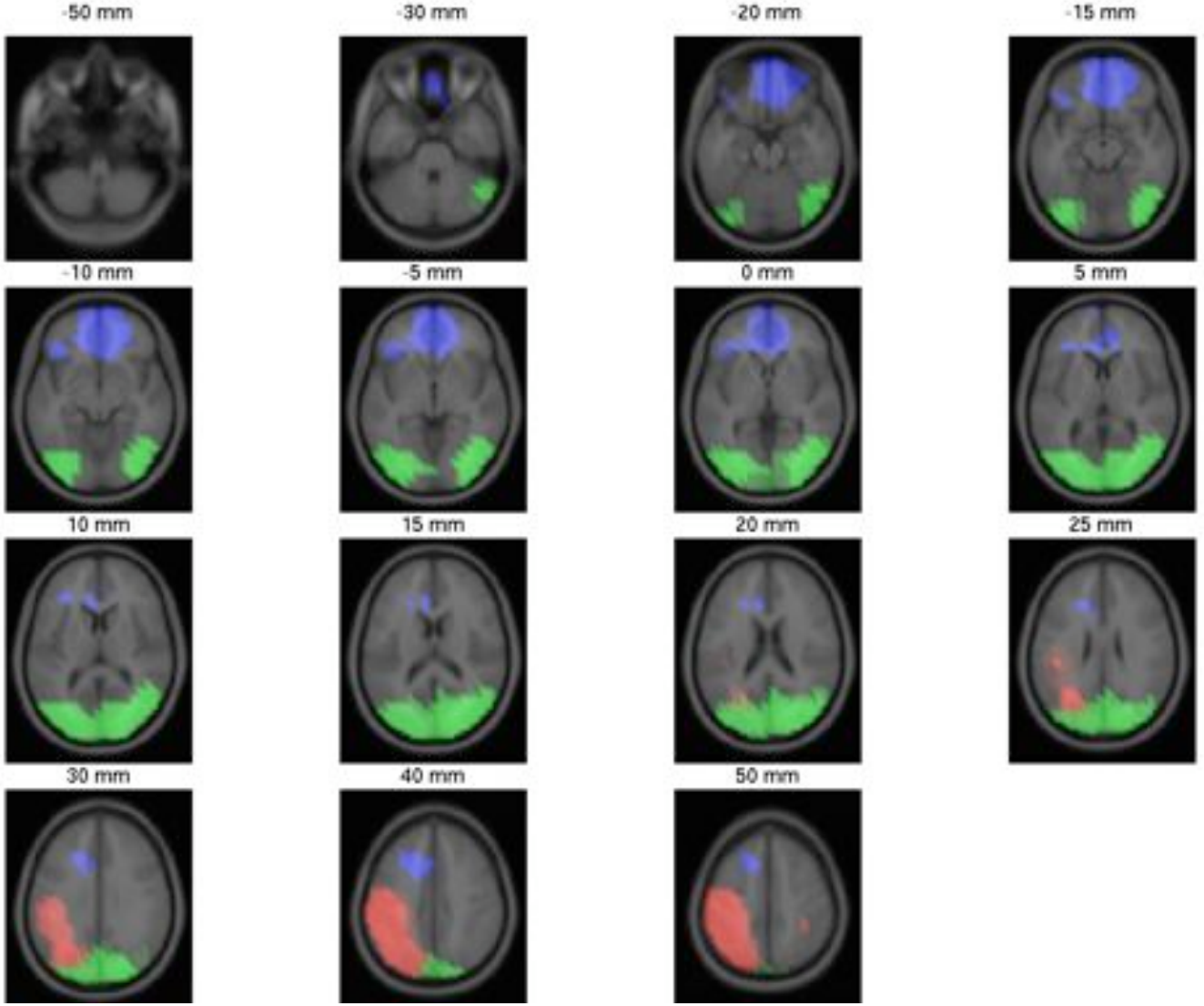


Multi-dimensional scaling visualization of ERSP projections for convergent locations.

Measure Projection: RSVP Example

ERSP domains (exemplar similarity <0.8)

- **Domain 1**
- **Domain 2**
- **Domain 3**

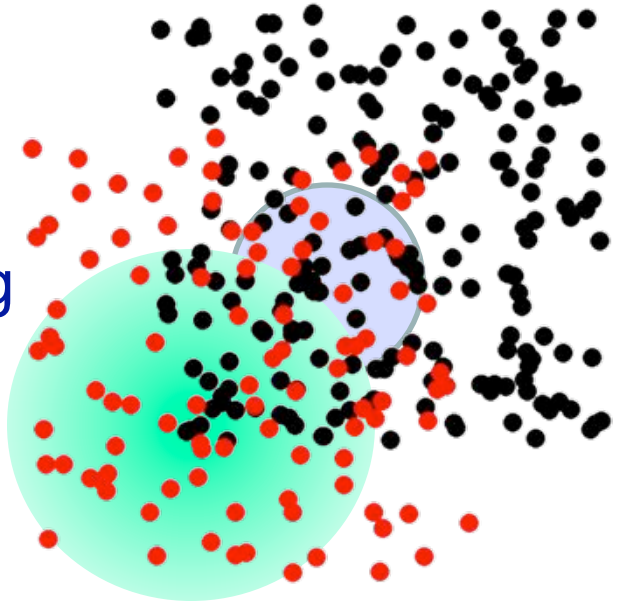


Measure Projection: ADHD Twin Study

- One of the largest studies we have analyzed: 132 subjects.
- Some are Monozygotic or Dizygotic twin pairs.
- Includes 42 ADHD subjects
- Four-choice RT task.
- Includes 1604 IC equiv. brain dipoles from the Control group and 996 IC equiv. dipoles from the ADHD group.

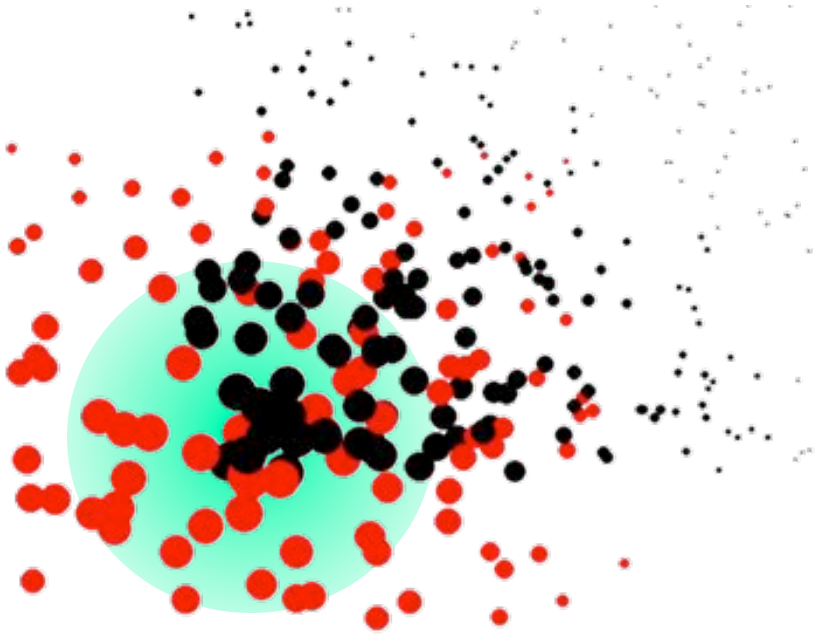
Measure Projection: ADHD Twin Study

- Group Separation = Mean cross-group distances – Mean within-group distances
- Group significance can be evaluated by permuting group identities and calculating the Group Separation values.
- Group significance can be defined in a neighborhood.
- Neighborhood can be a ‘soft’ Gaussian.



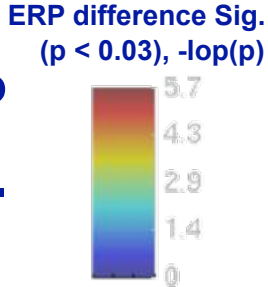
Measure Projection: ADHD Twin Study

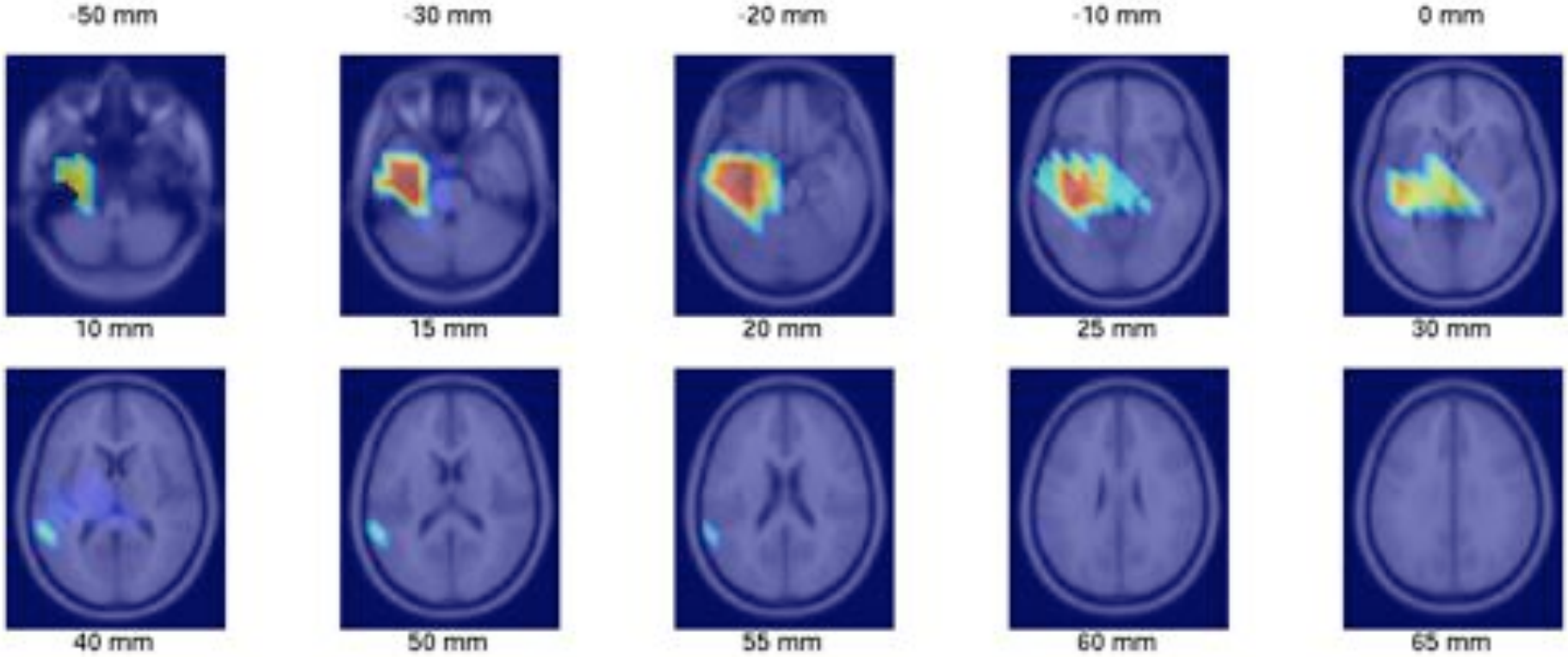
- We can use weighted means in Group Separation formula to emphasize group differences in a Gaussian neighborhood.
- This allows us to investigate EEG measure differences between two groups at different brain locations, by only looking at their pair-wise IC measure similarities.



Measure Projection: ADHD Twin Study

ADHD and control subjects have significant ERP differences in the area highlighted below.





Measure Projection: ADHD Twin Study

ADHD and control subjects have significant IC density differences in the area highlighted below.

Note: By fMRI, right insula may regulate the salience of selective attention vs. bottom-up salience.