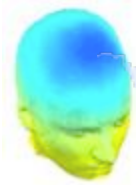
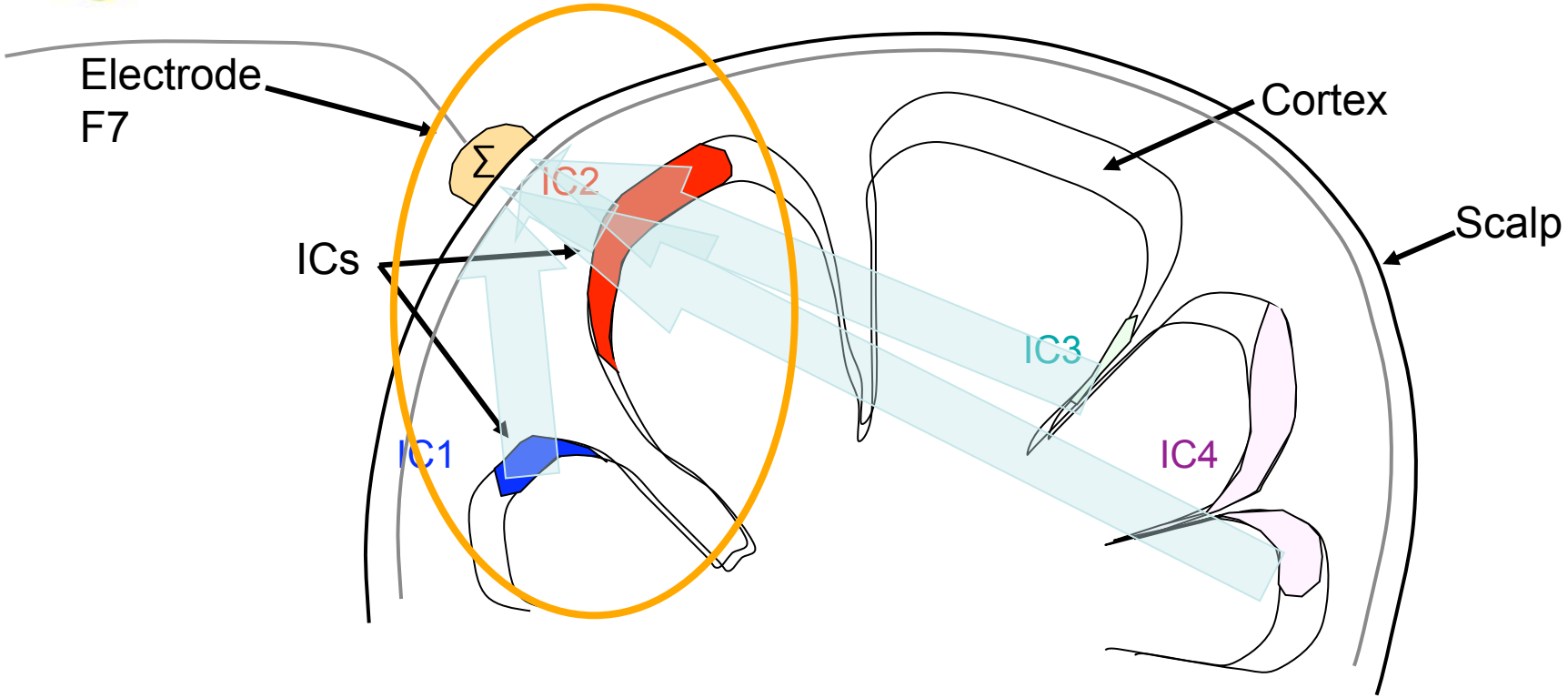
Why cluster independent components across subjects or sessions?



- 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!

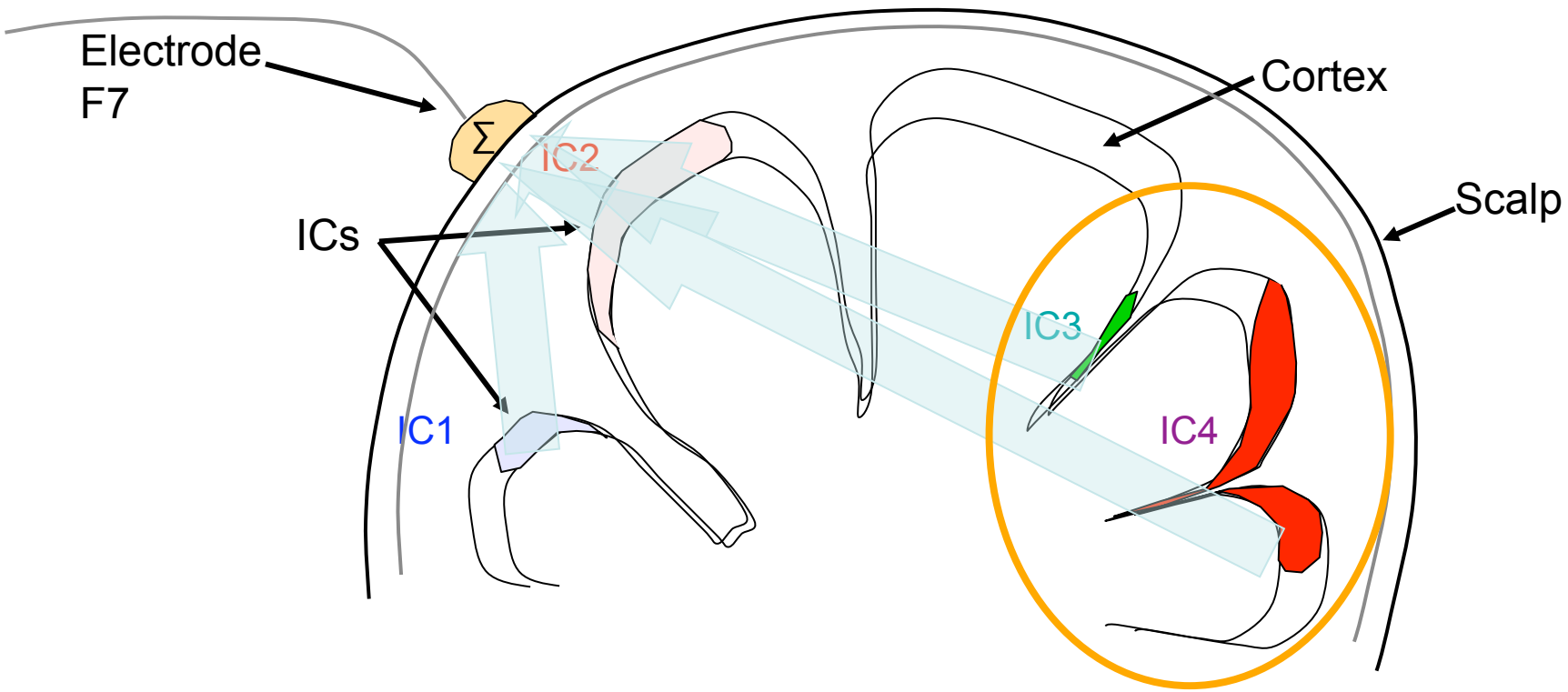
Makeig, 2007

Example: First Subject



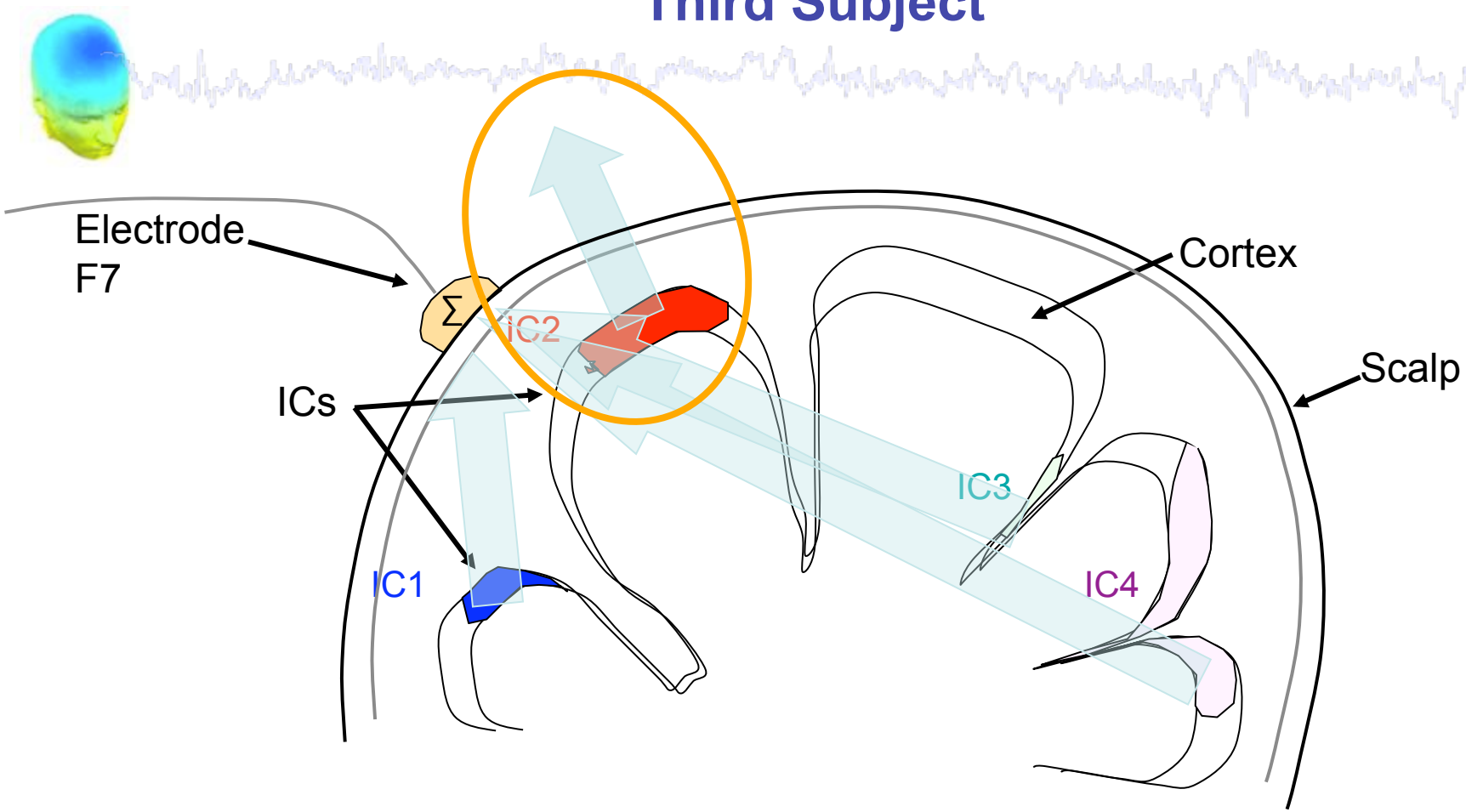
Makeig, 2005

Second Subject



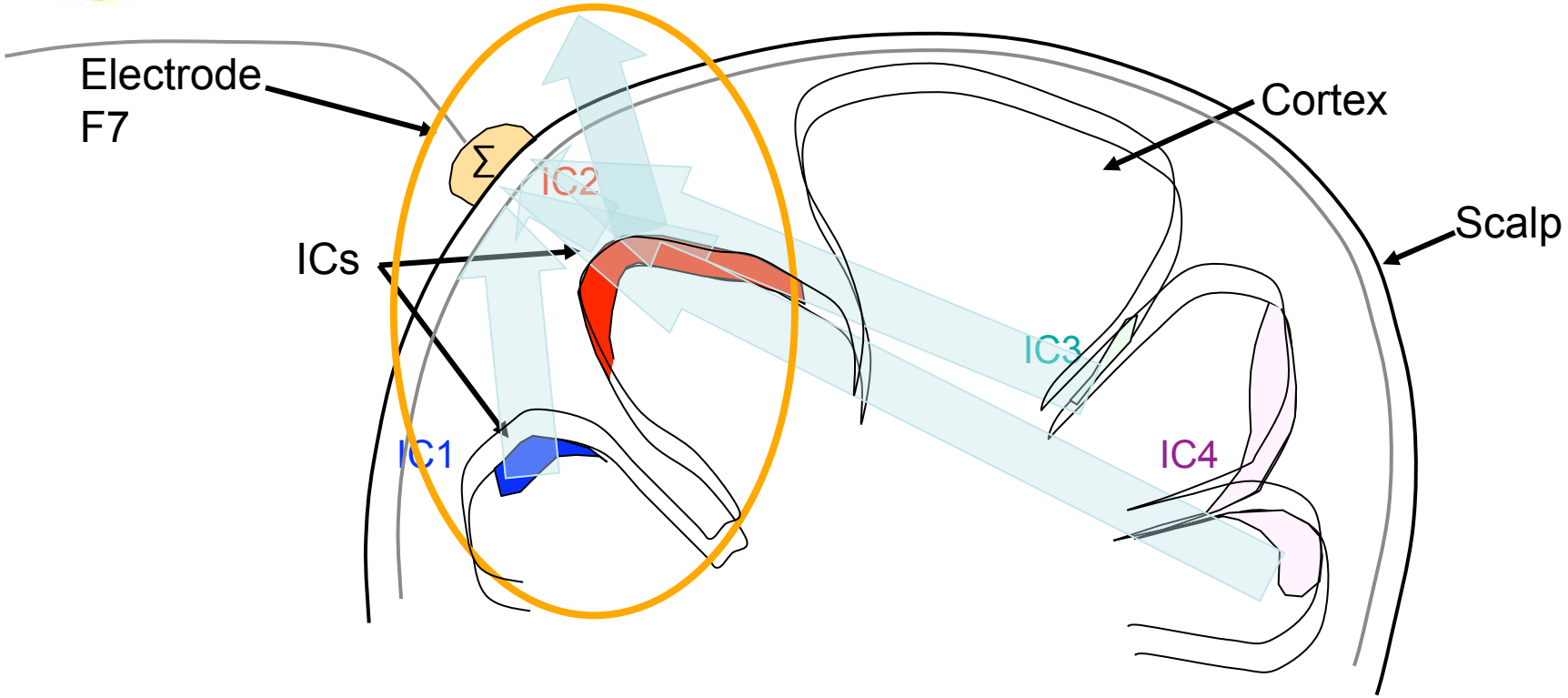
Makeig, 2005

Third Subject



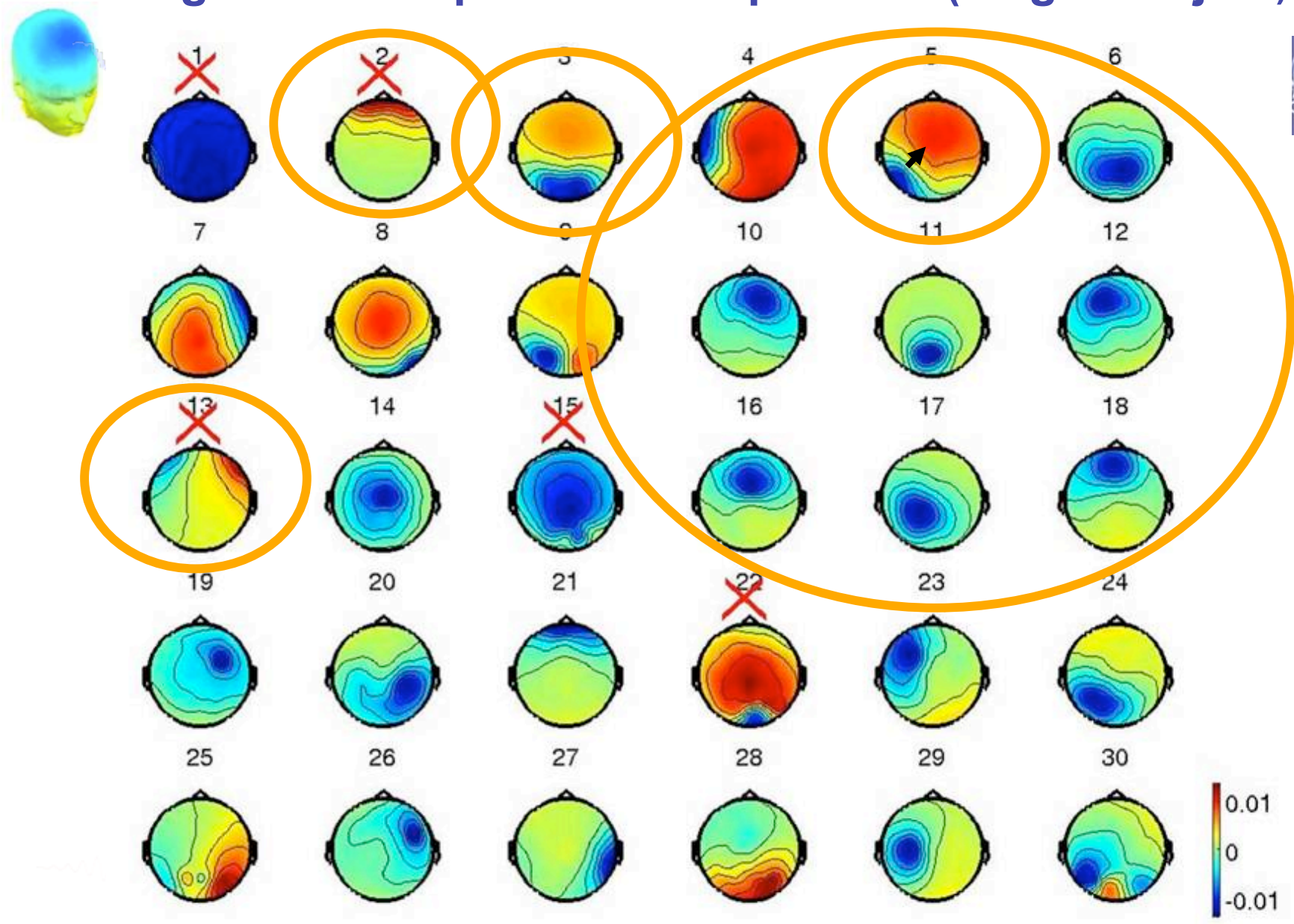
Makeig, 2005

Fourth Subject



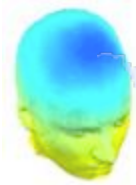
Makeig, 2005

Largest 30 independent components (single subject)



Makeig, 2007

So how to cluster components?



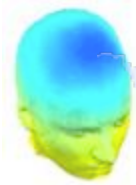
The same problems hold for clustering independent components

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? …

Makeig, 2007



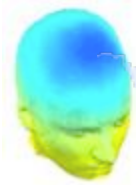
Does the spatial distribution of independent components depend on the task the subject performs?

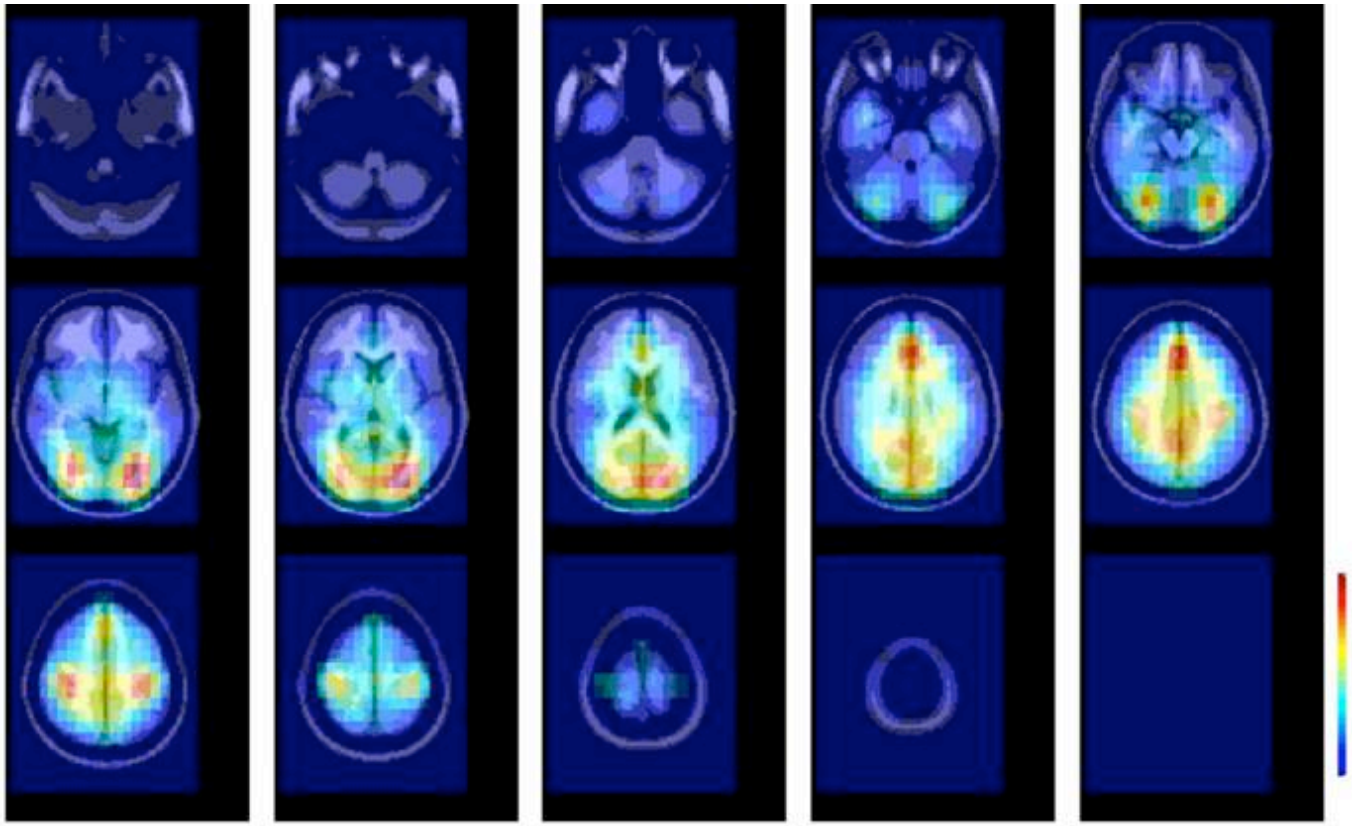
i.e.

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

Makeig, 2007

Equivalent dipole density



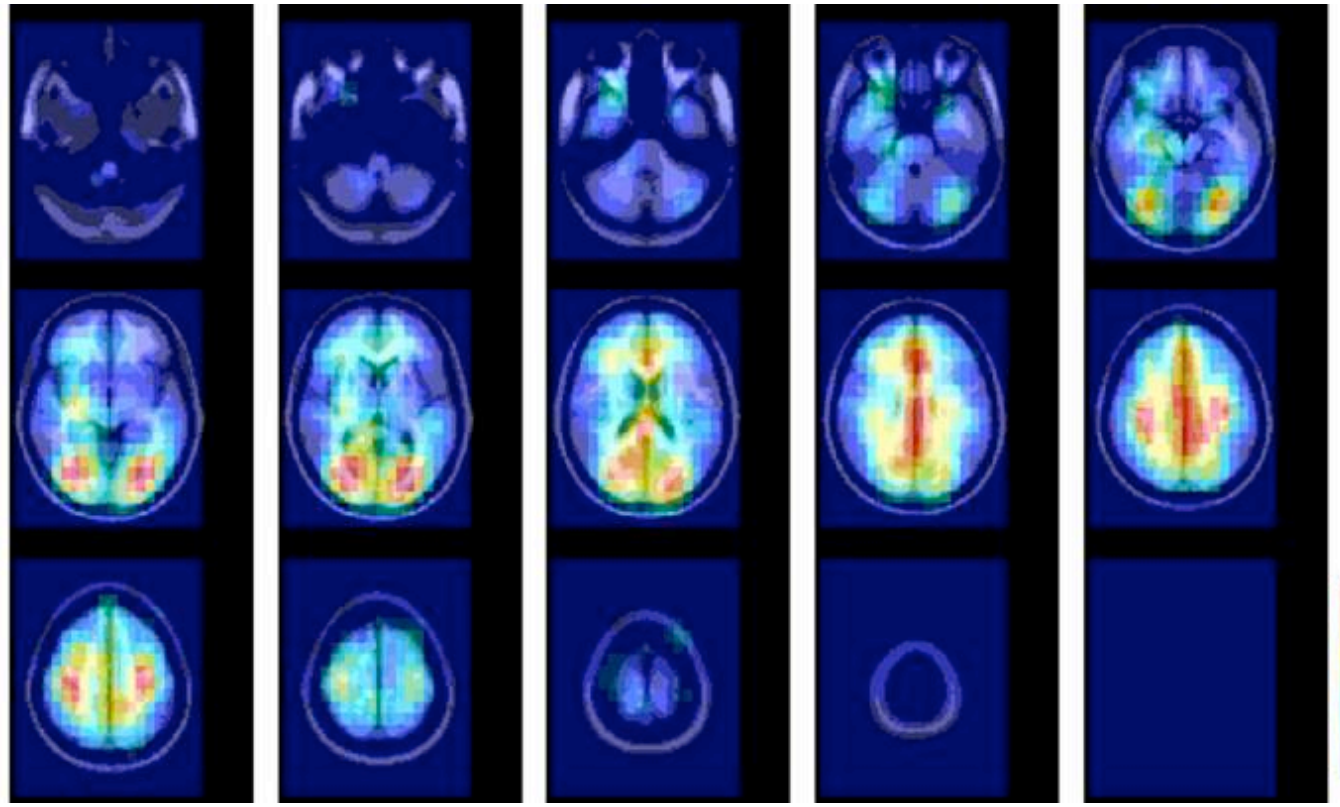


Sternberg letter memory task

>> dipoledensity()

Onton et al., ‘05

Equivalent dipole density

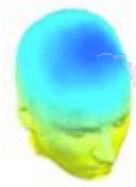


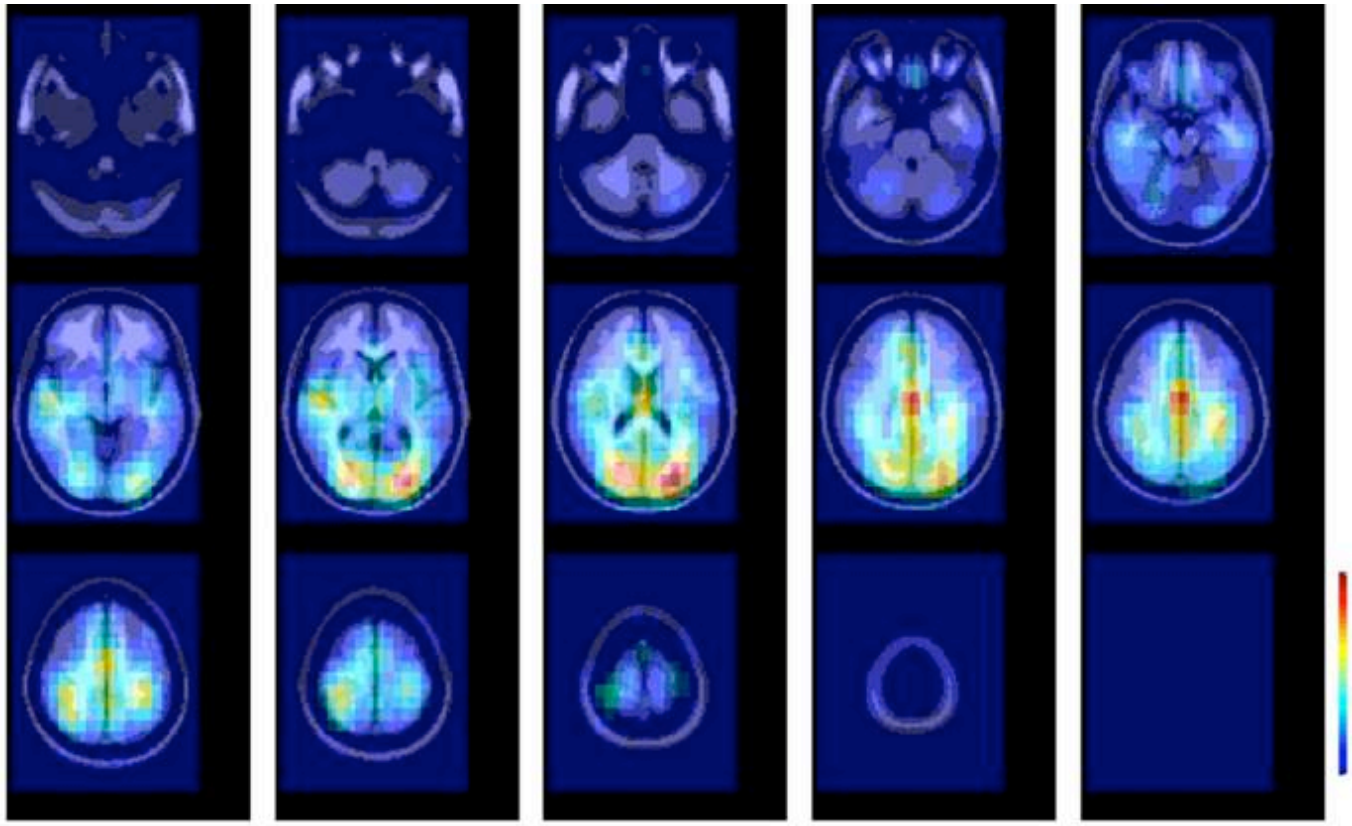
Letter twoback with feedback

>> dipoledensity()

Onton et al., ‘05

Equivalent dipole density

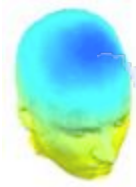


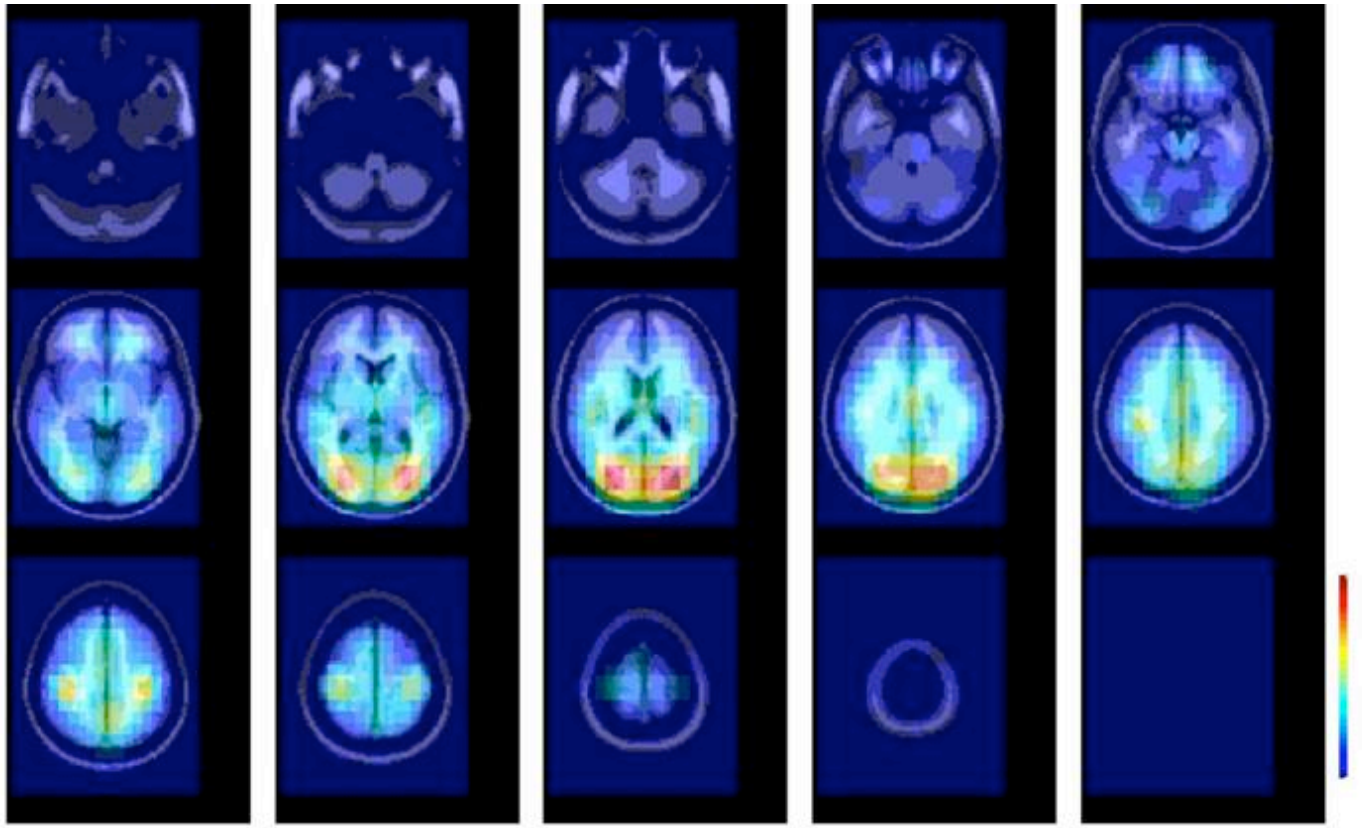


Auditory oddball plus novel sounds

>> dipoledensity()

Equivalent dipole density



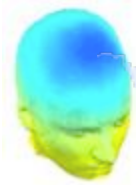


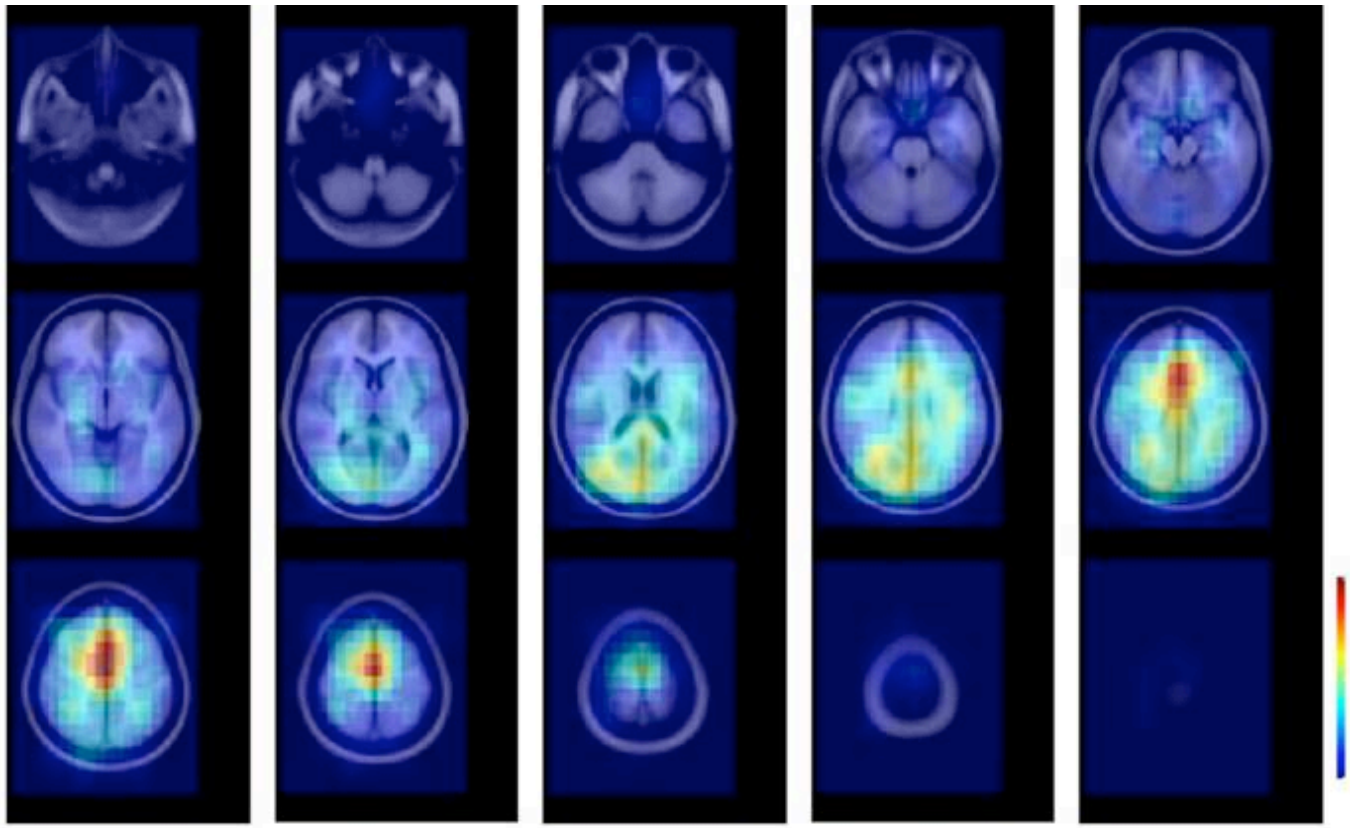
Emotion imagery task

>> dipoledensity()

Onton et al., ‘05

Equivalent dipole density Exp I



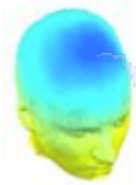


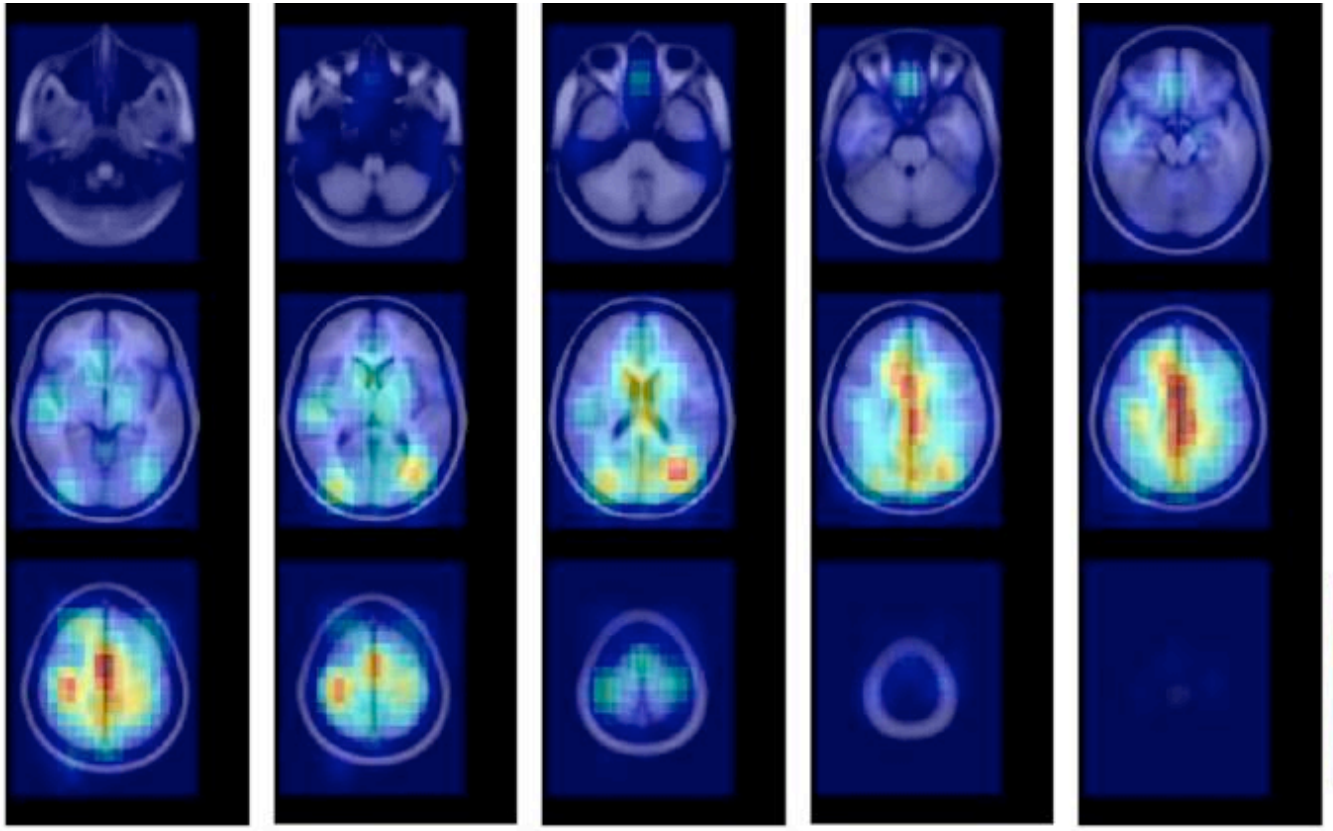
Word memory (old/new) task

>> dipoledensity()

Onton et al., ‘05

Equivalent dipole density Exp II



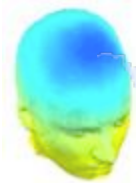


Visually cued button press task

>> dipoledensity()

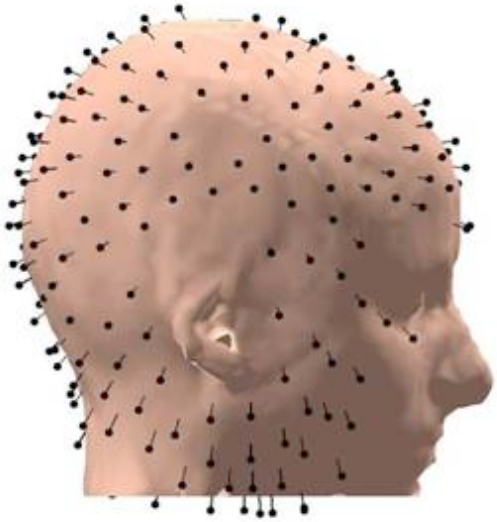
Onton et al., ‘05

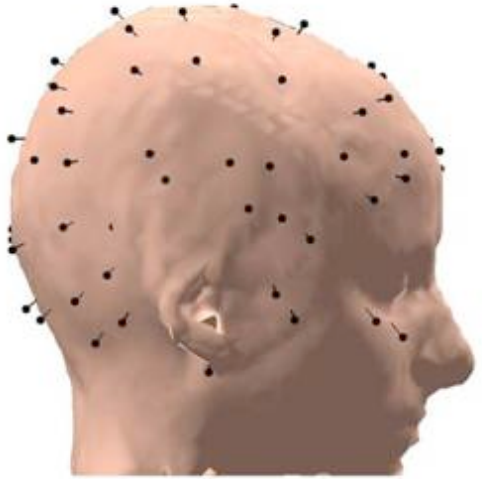
… Some caveats



In this preliminary study …

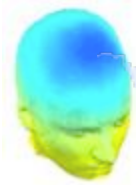
- The electrode locations were not individualized.
- MR images were not available → co-registration crude.
- Single versus dual-dipole model selection was subjective.
- Different electrode montages → possible location effects



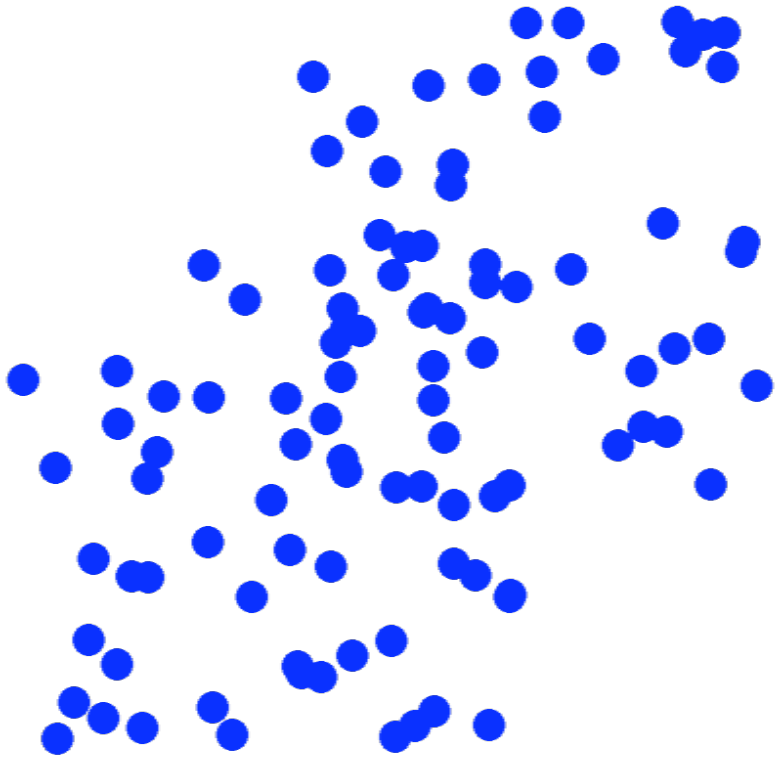


Onton, 2005

Problems with multi-measure clustering

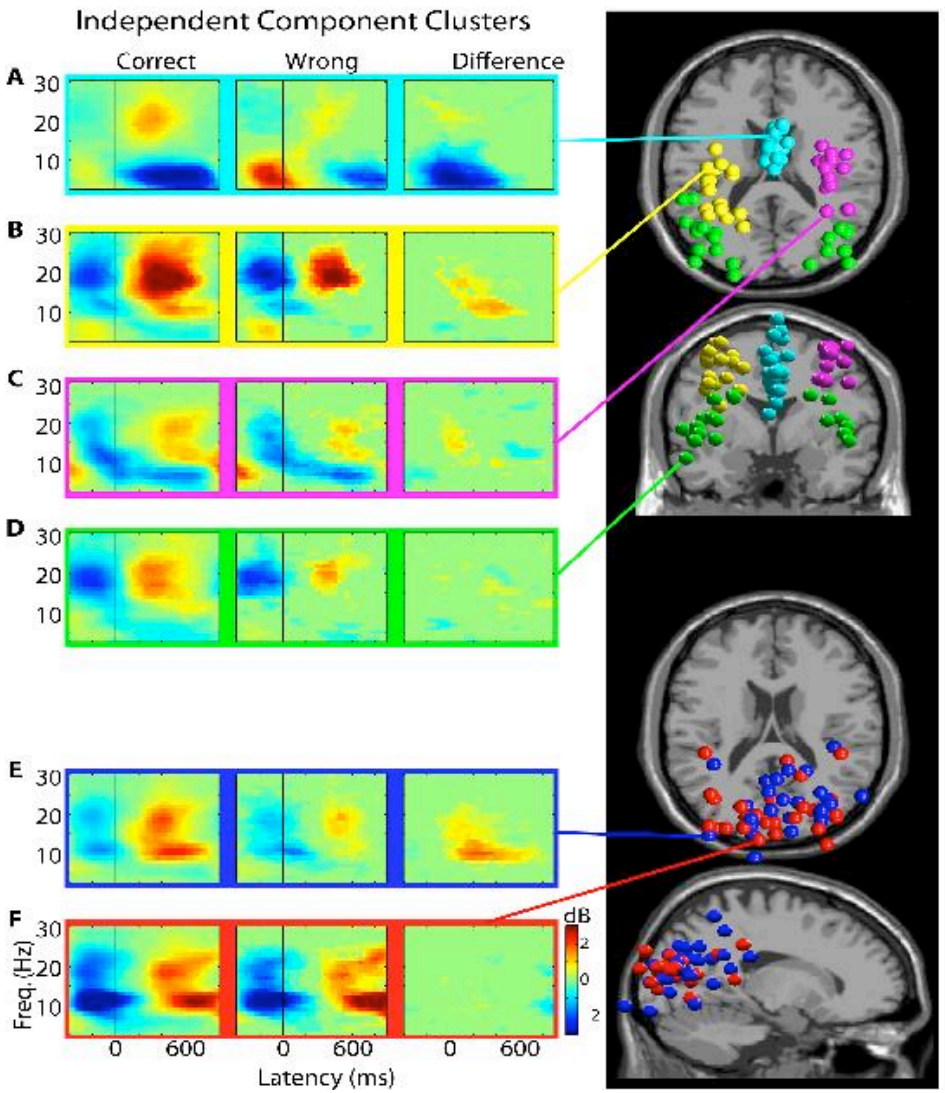


What are the clusters according to location?



N. Bigdely-Shamlo, 2010

Study IC Clustering



Sometime clusters are spatially separate AND have distinct responses.

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

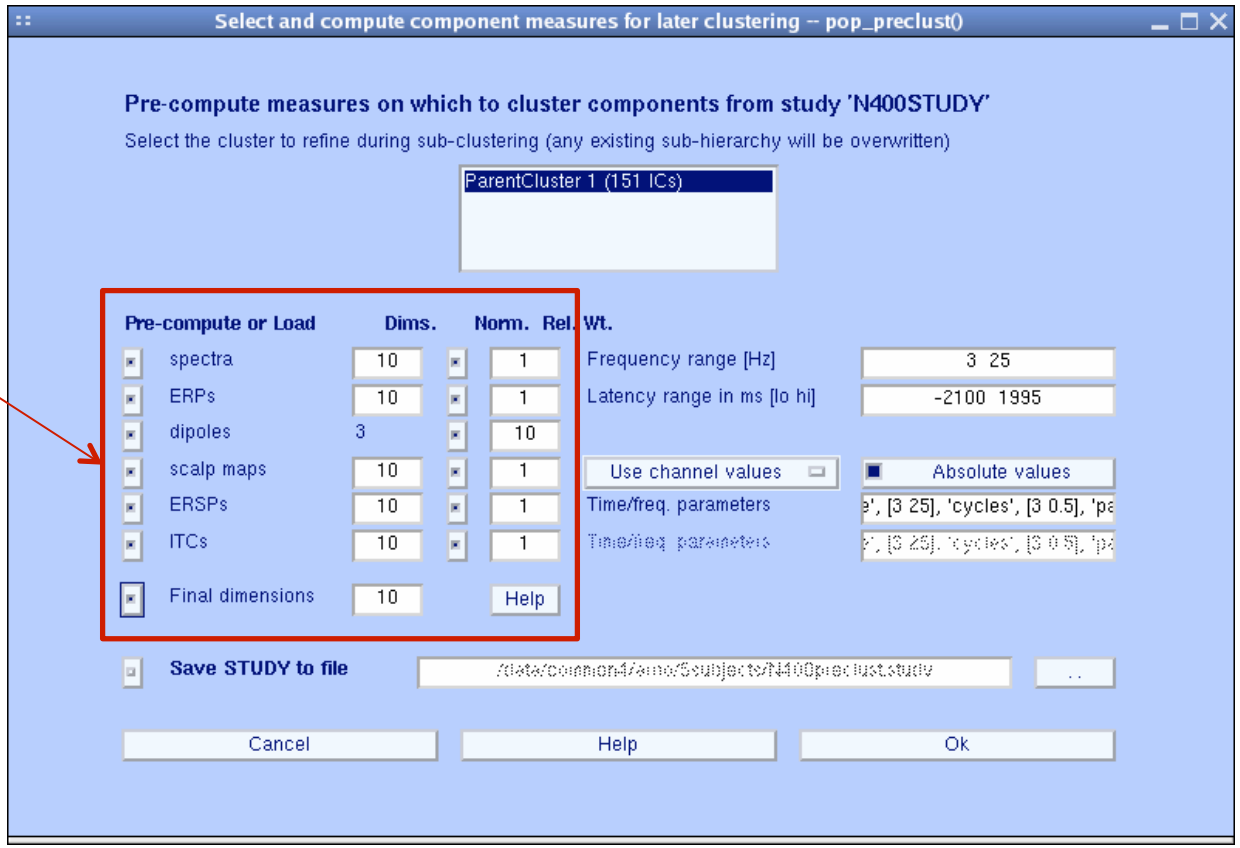
Onton & Makeig, 2007

Study IC Clustering: Practical Problems

Large parameter space problem: many different clustering solutions can be produced by changing parameters and measure subsets. Which one should we choose?

EEGLAB original clustering has ~12 parameters

N. Bigdely-Shamlo, 2010

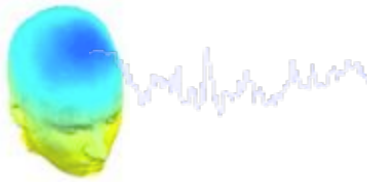


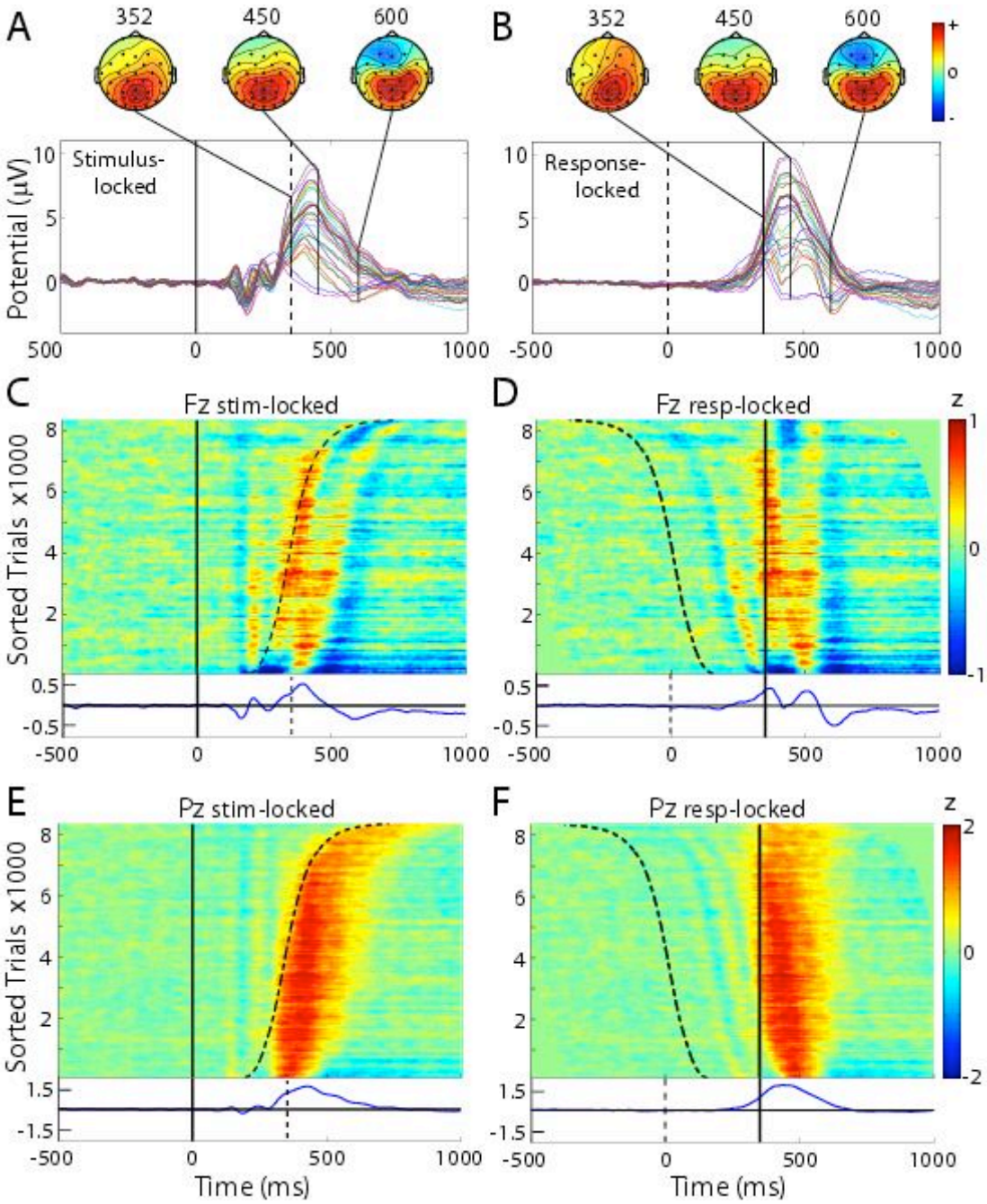
Visual Selective Attention Task

+

15 subjects

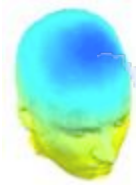
Makeig et al., PLPS 2004





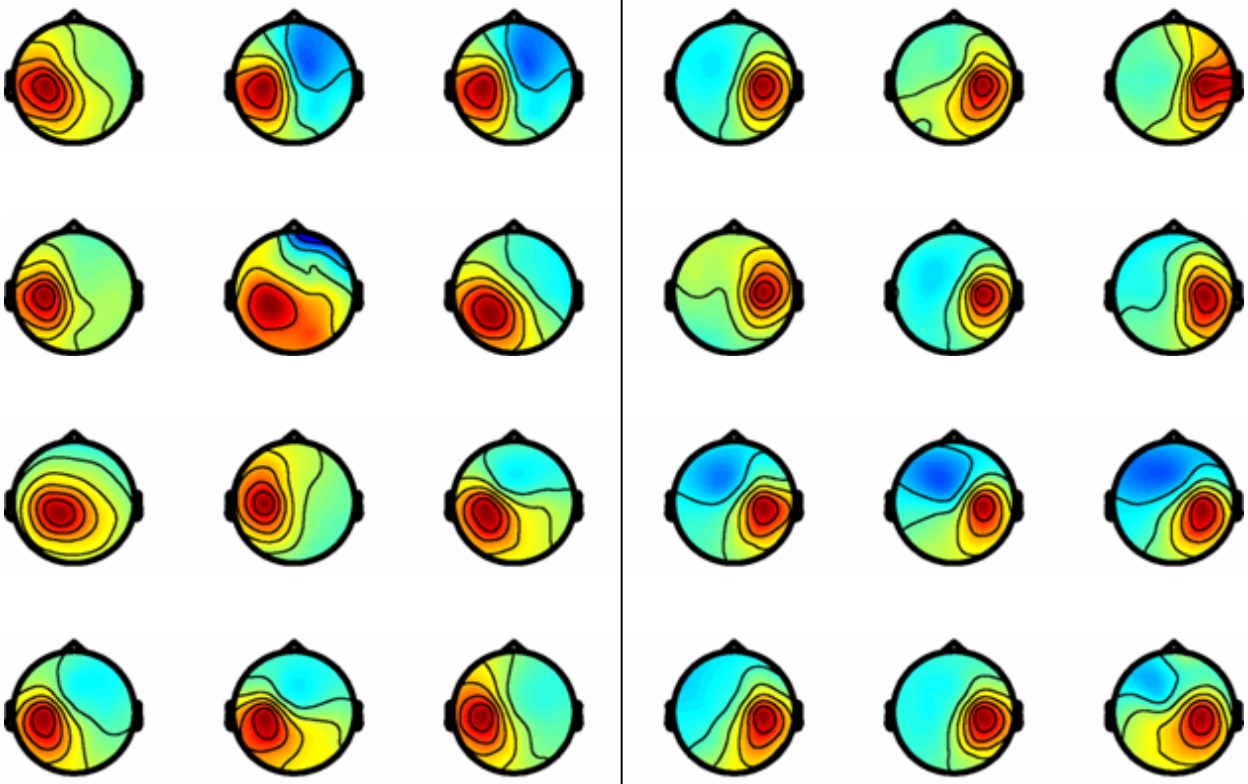
Makeig et al., PLOS 2004

Clustering ICA components by eye



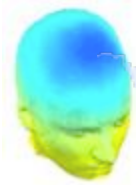
Left mu

Right mu



Makeig et al., *unpublished*

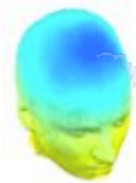
Study IC Clustering: Assumptions



- 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.

Makeig, 2007

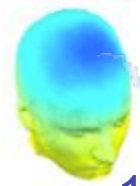
EEGLAB Study Clustering strategy



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.

Makeig, 2007

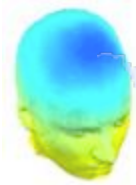
EEGLAB Study Clustering procedure



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.

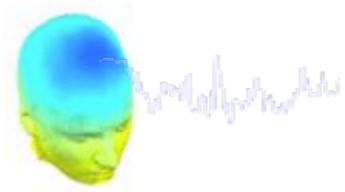
Makeig, 2007

P300 ← Semi-automated IC clustering

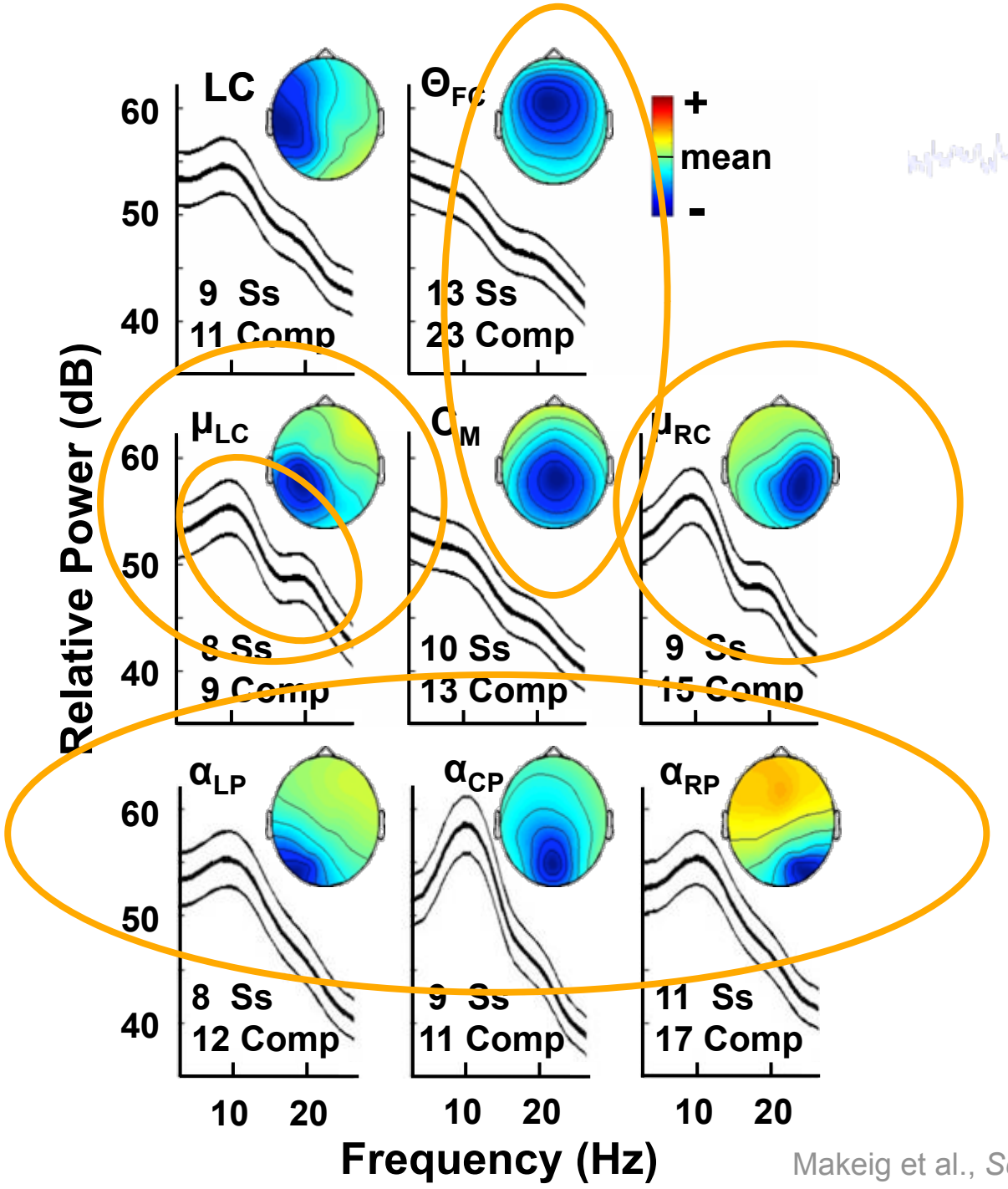


- Clustered components from 15 Ss using a 'component distance metric' incorporating differences between their (weighted) scalp maps, dipole locations, spectra, ERP, ERSP, and ITC patterns.
- Hand-adjusted clusters to remove outliers.
- Determined time/frequency regions of significant ERSP and ITC for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

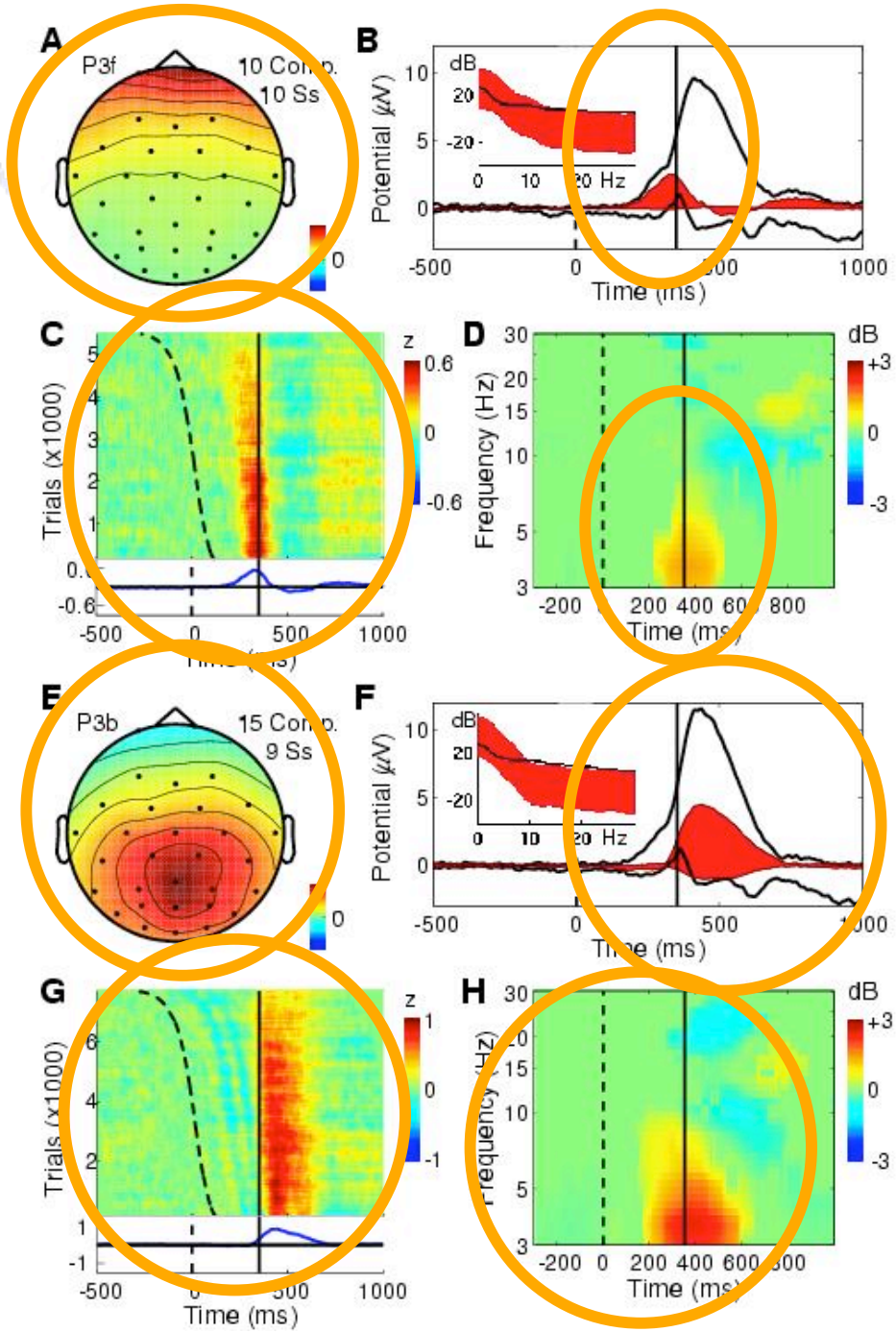
Makeig, 2007



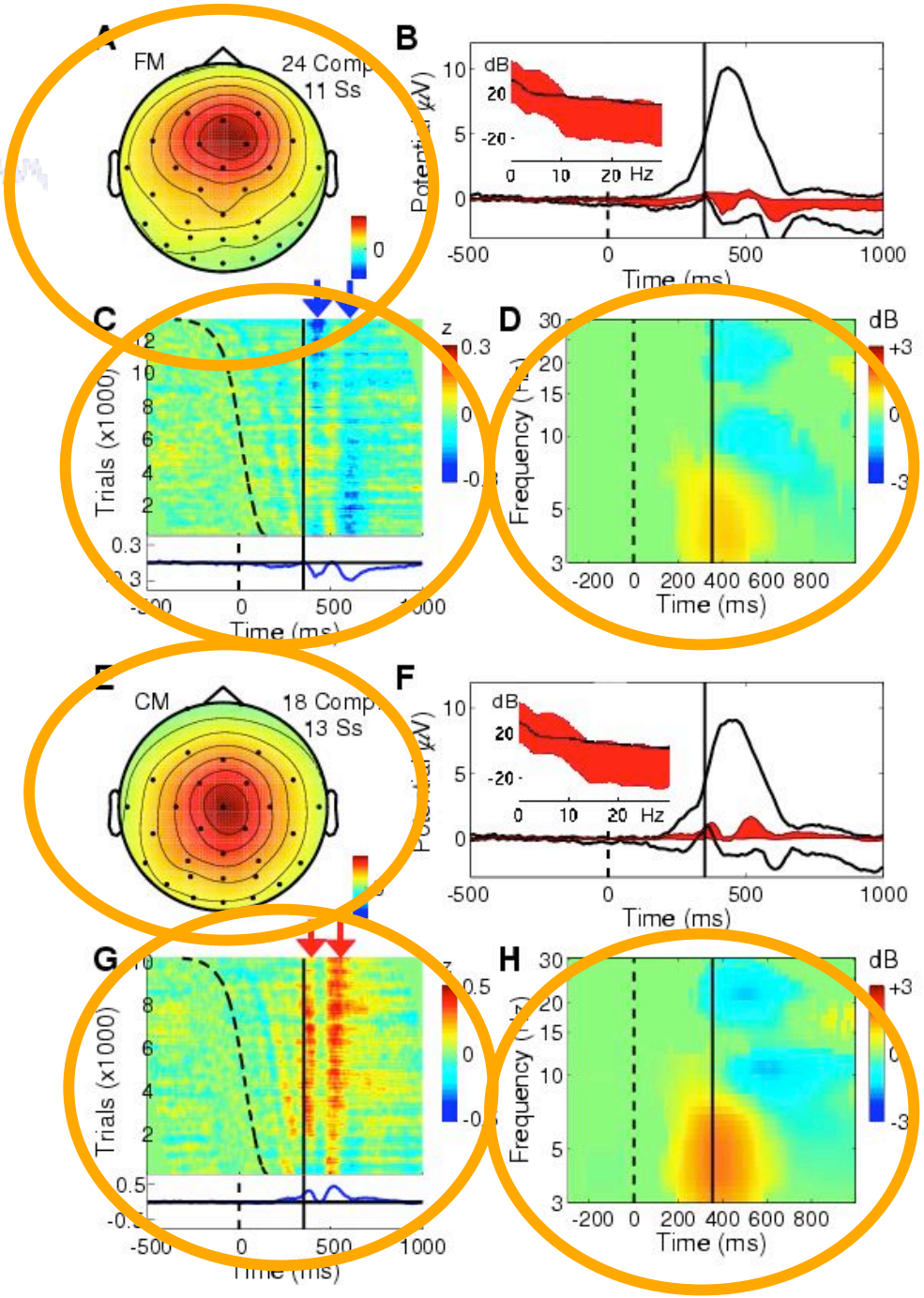
N1 Component Clusters



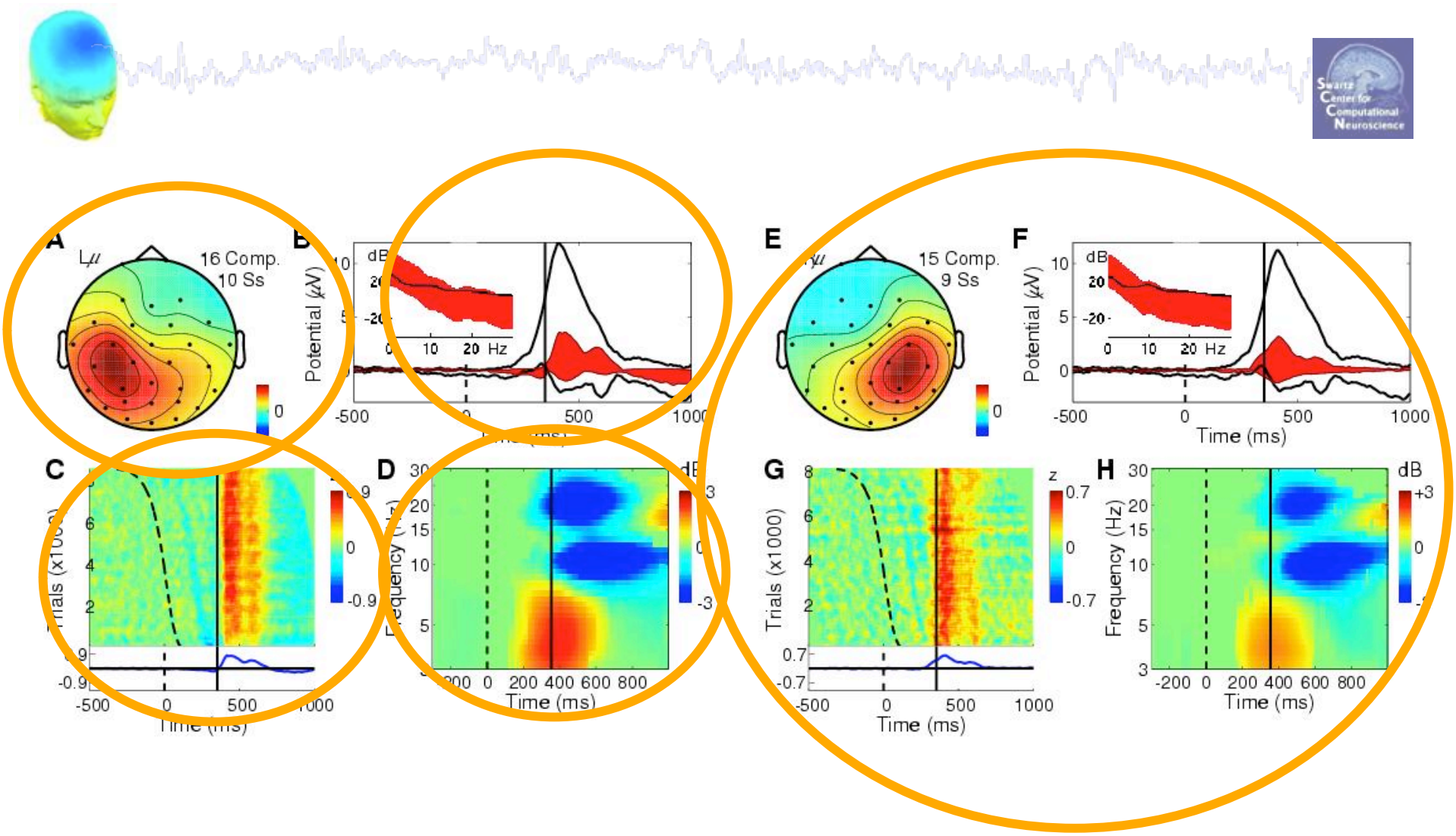
Makeig et al., *Science* 2002



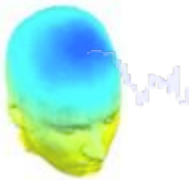
Makeig et al. PLOS 2004

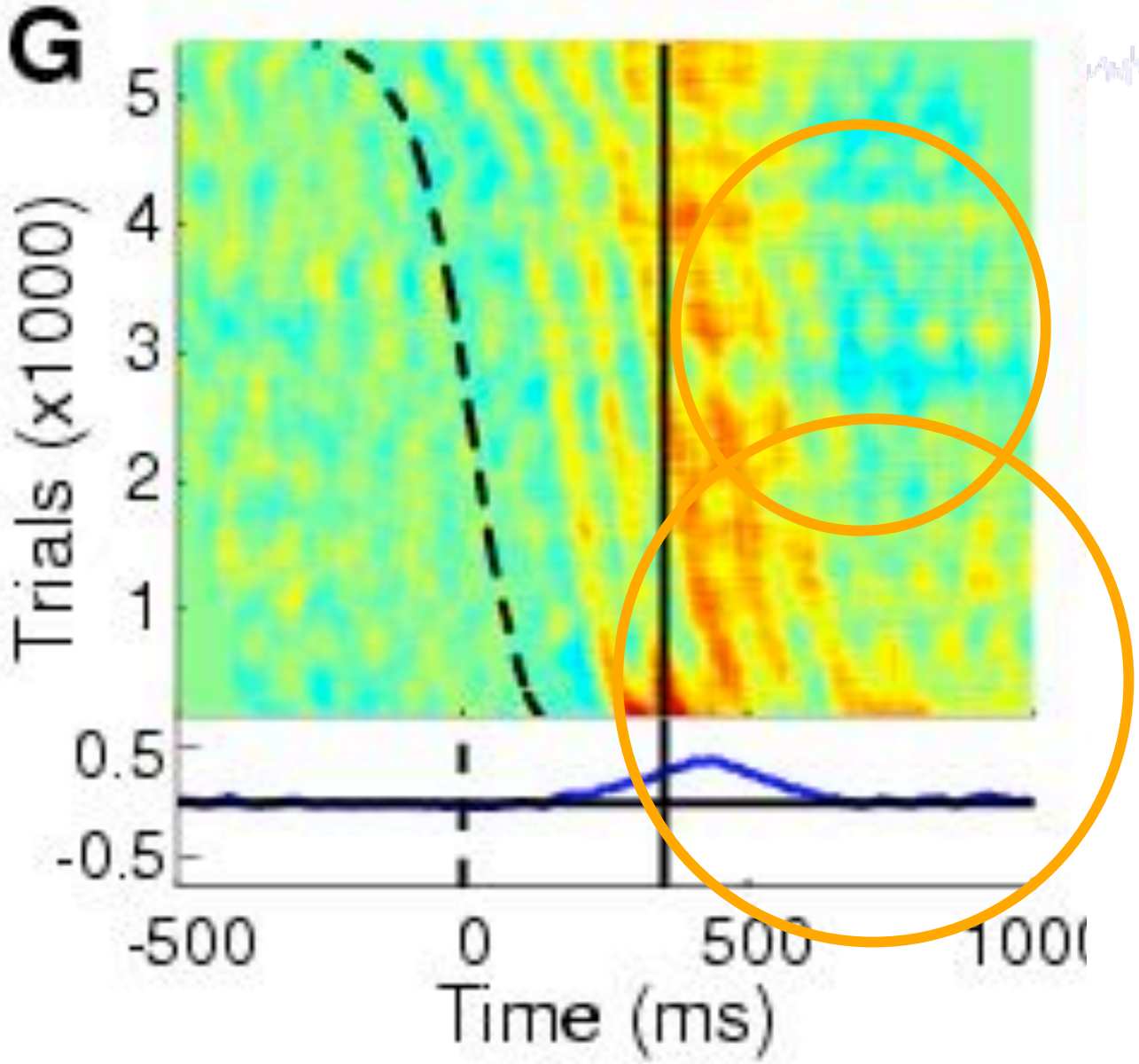


Makeig et al. PLOS 2004

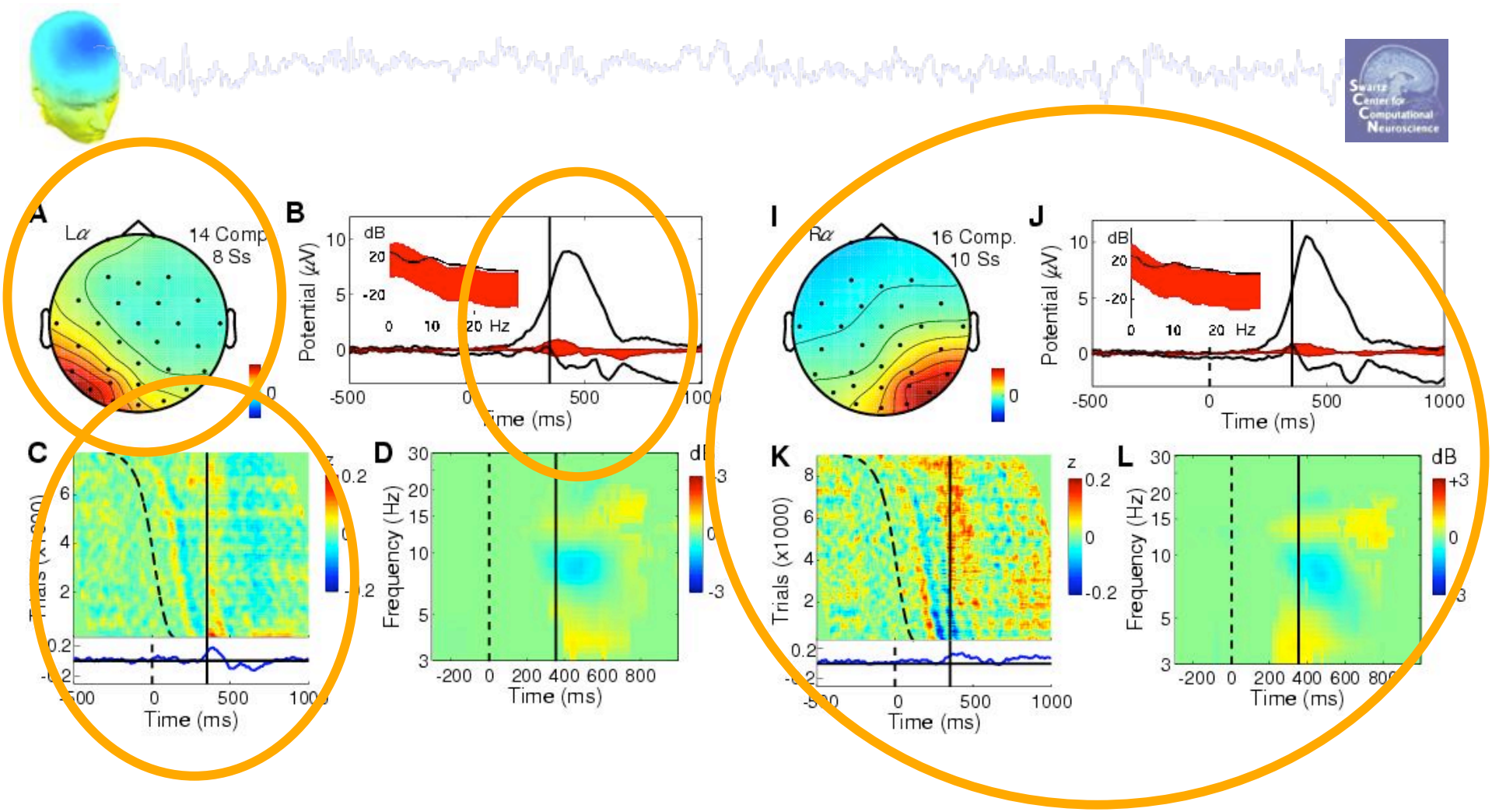


Makeig et al. PLOS 2004



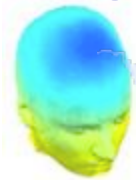


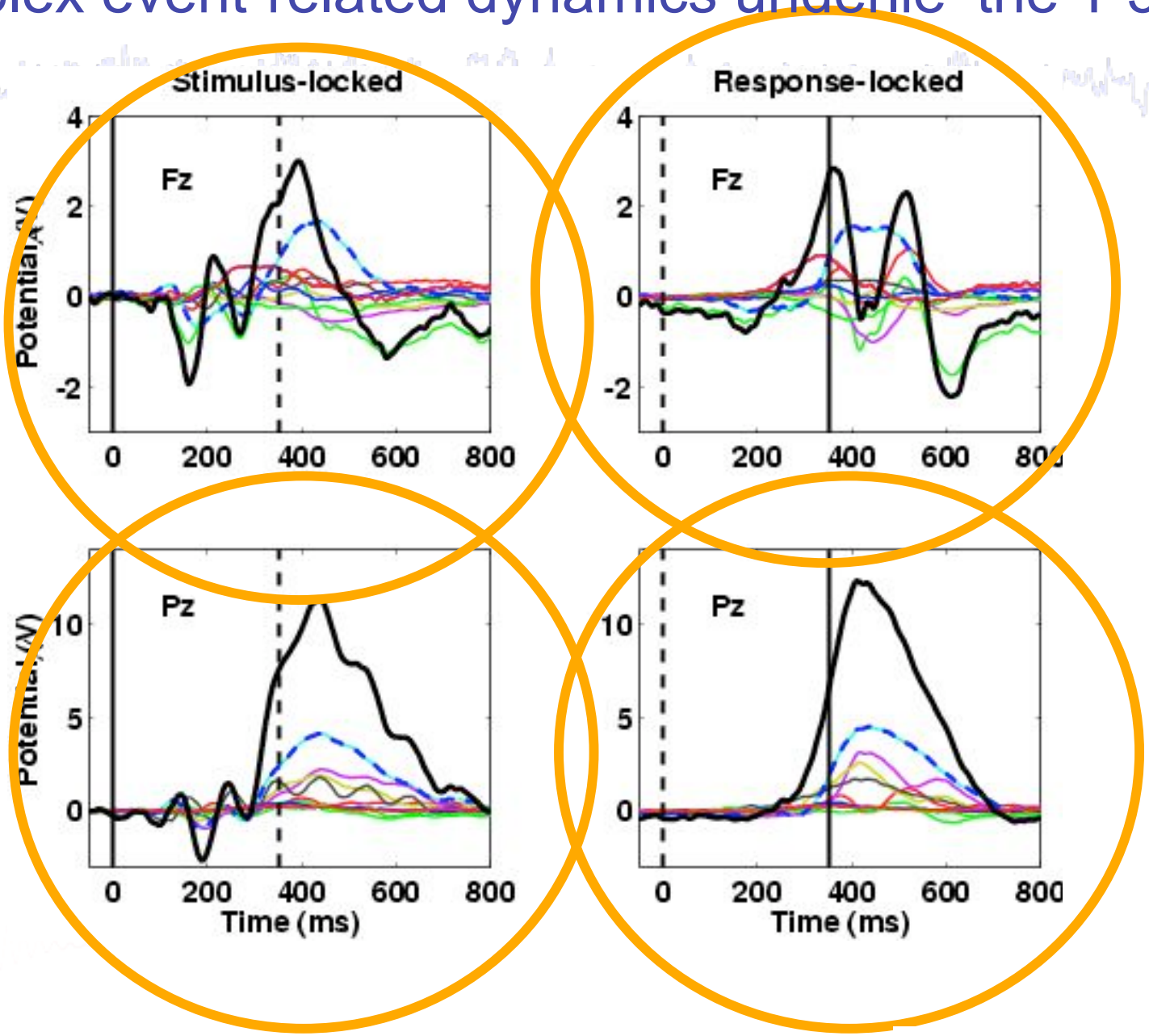
Makeig et al. PLOS 2004



Makeig et al. PLOS 2004

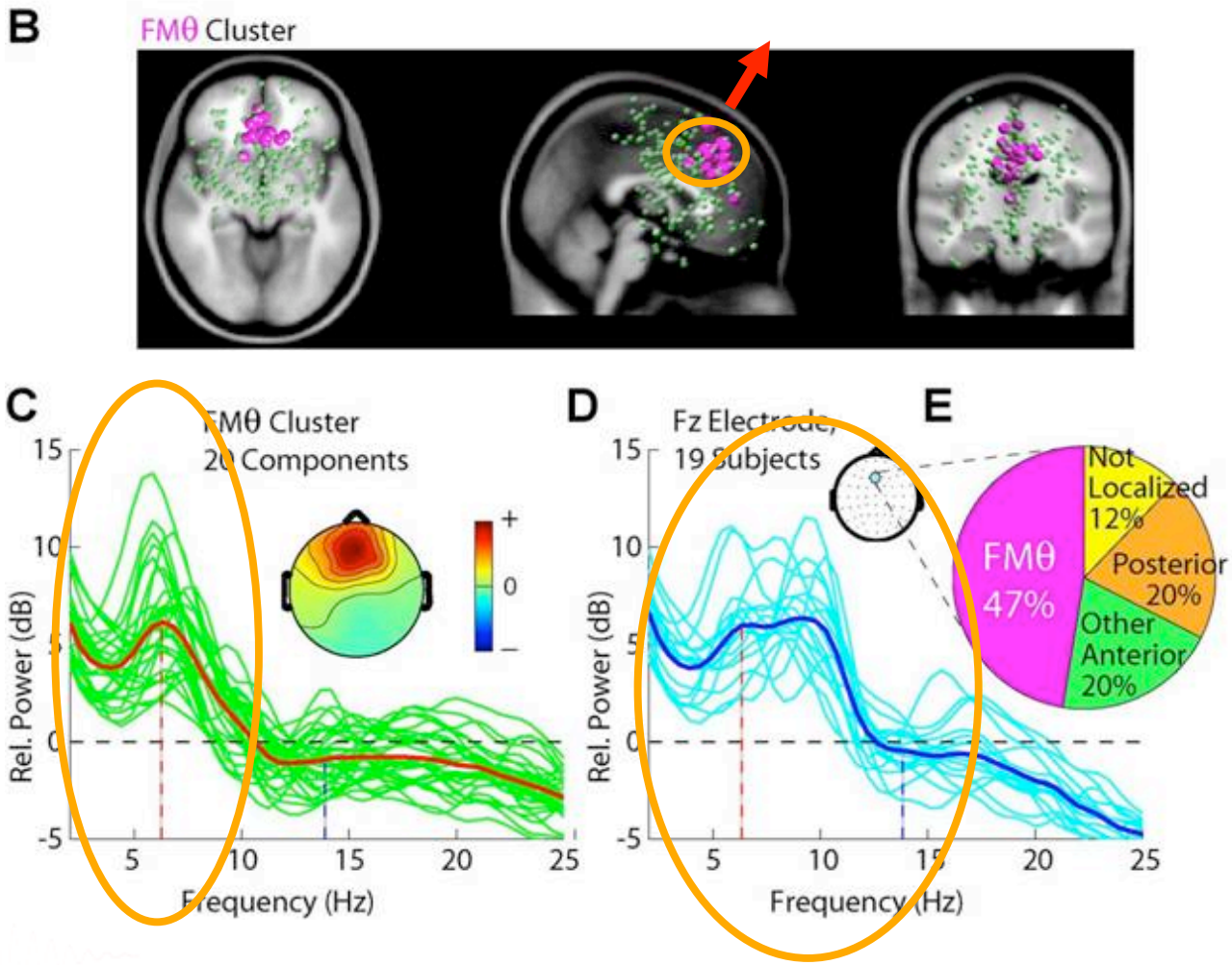
Complex event-related dynamics underlie ‘the’ P300





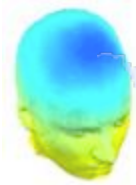
Makeig et al. PLOS 2004

A FMΘ cluster during working memory



Onton et al., *NeuroImage* 2005

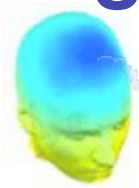
Are the obtained IC clusters “real“?



- **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

Makeig, 2007

Should all subjects be included in each cluster?



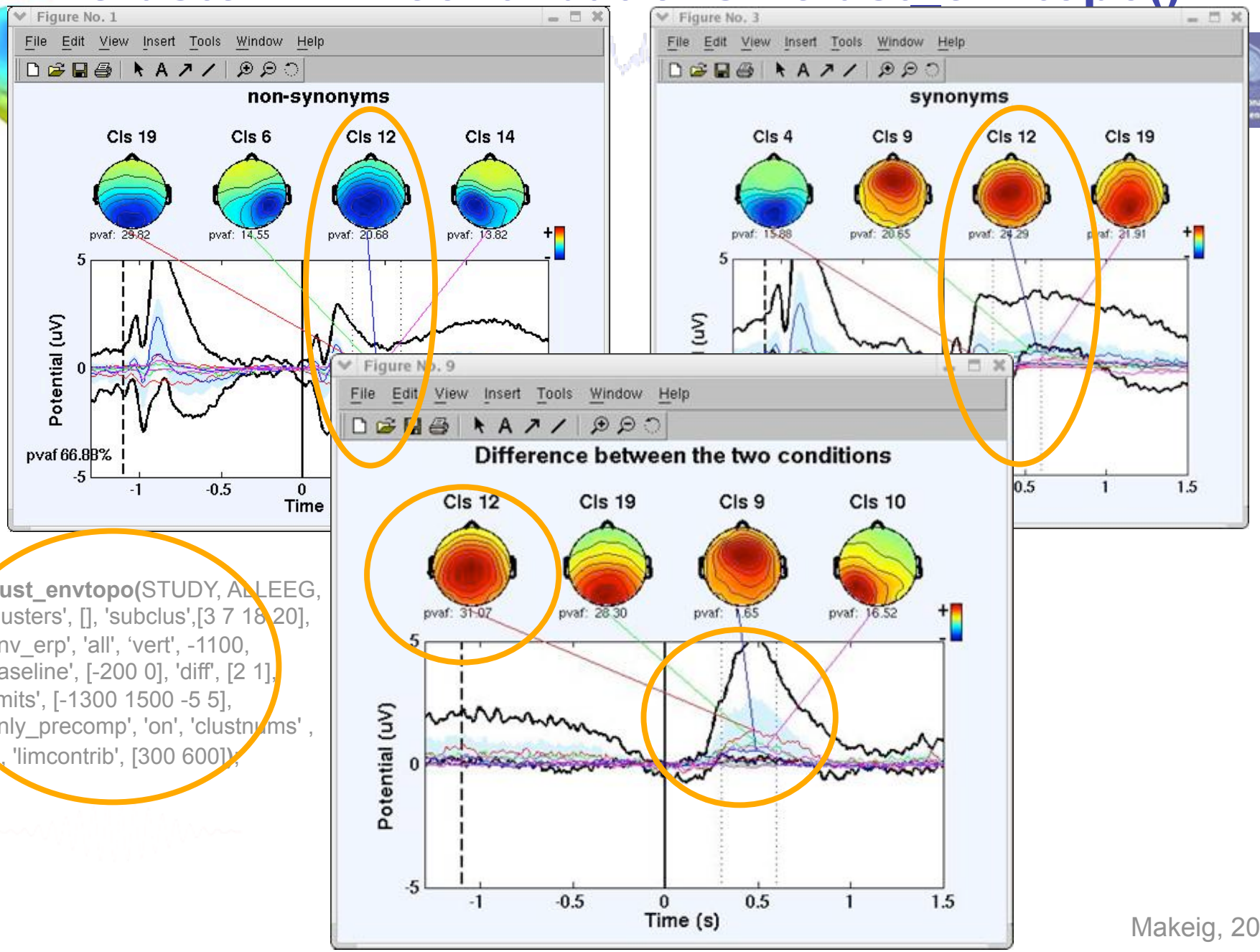
Not all subjects contribute components to each cluster.

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]);

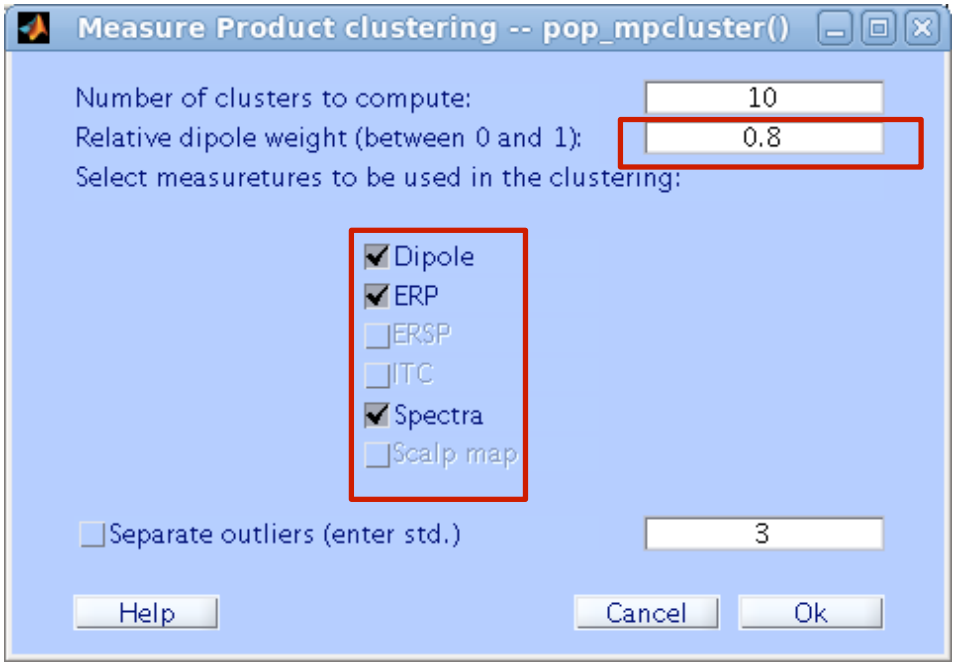
Makeig, 2007

Study IC Clustering: New Developments

The Affinity Clustering method

(EEGLAB plug-in by Nima Bigdely Shamlo)

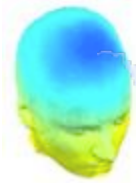
only has one pre-clustering parameter.



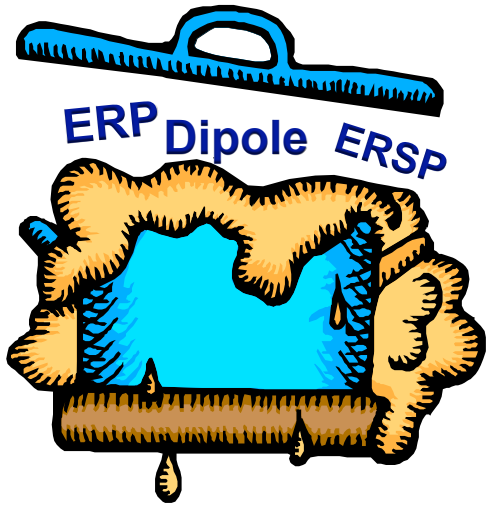
Of course, one still has to select a subset of measures and the number of clusters….

N. Bigdely-Shamlo, 2010

Study IC Clustering: New Developments

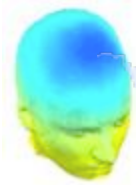


- We still have to select the number of clusters.
- With both these clustering methods, we basically mix (either add or multiply) distances for a subset of EEG measures (ERP, ERSP, ITC, mean spectrum, dipole location) together.
- This makes clustering parameters less meaningful.



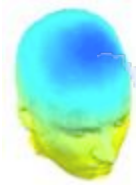
N. Bigdely-Shamlo, 2010

Study IC Clustering: New Developments



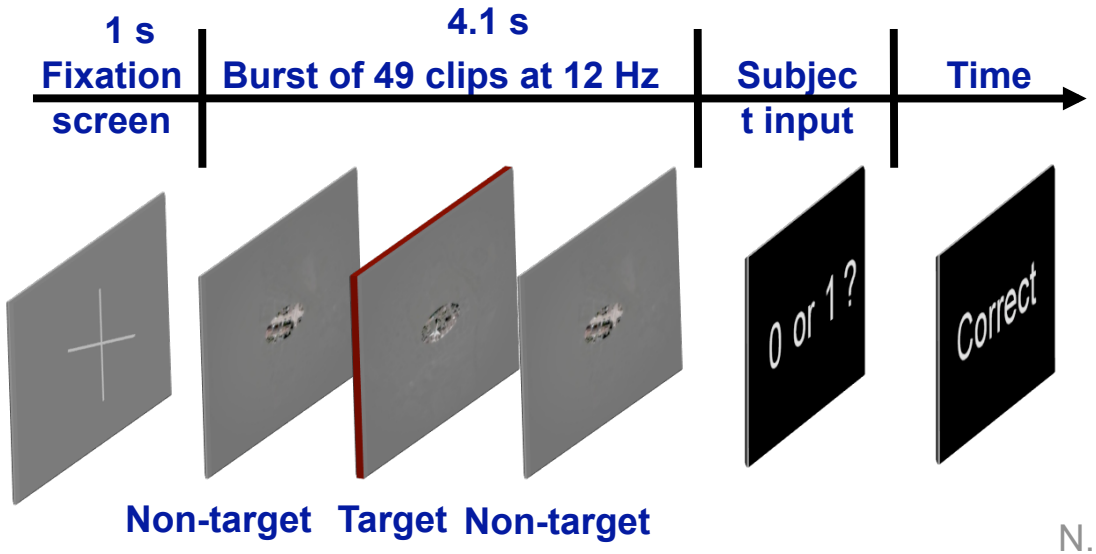
- **Instead**, we can directly work on pair-wise similarity matrices and prevent ICs with similarities less than certain threshold (e.g., ERSP corr. < 0.5) to be clustered together.
- The most important measure is **equivalent dipole location**.
- Assuming a certain variability estimate for dipole location (due to error in localization and subject variability), one can also estimate an optimum number of clusters.

Measure Projection: RSVP Example



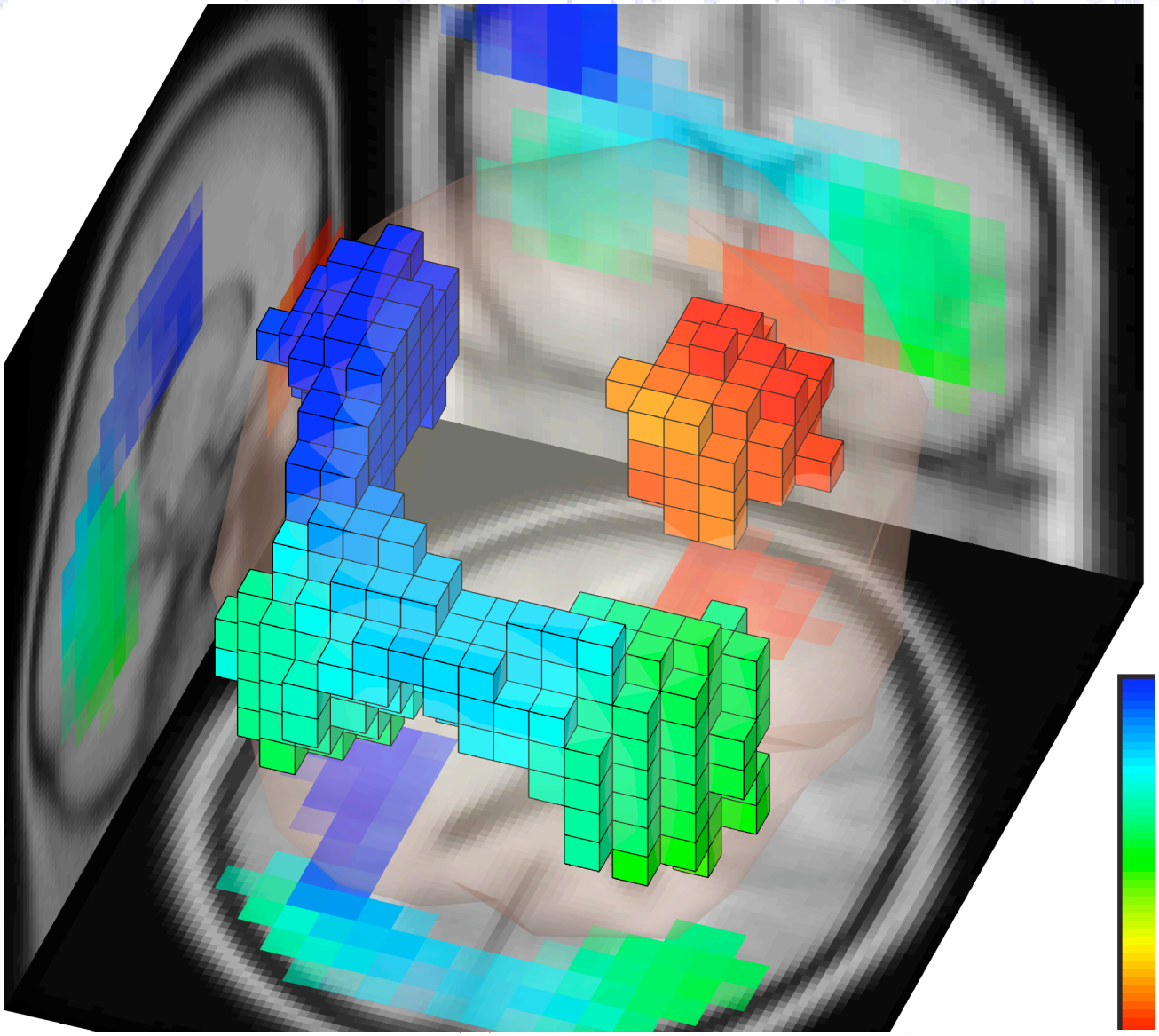
Rapid Serial Visual Presentation Experiment

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



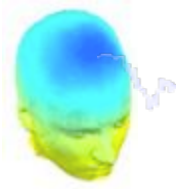
N. Bigdely-Shamlo, 2010

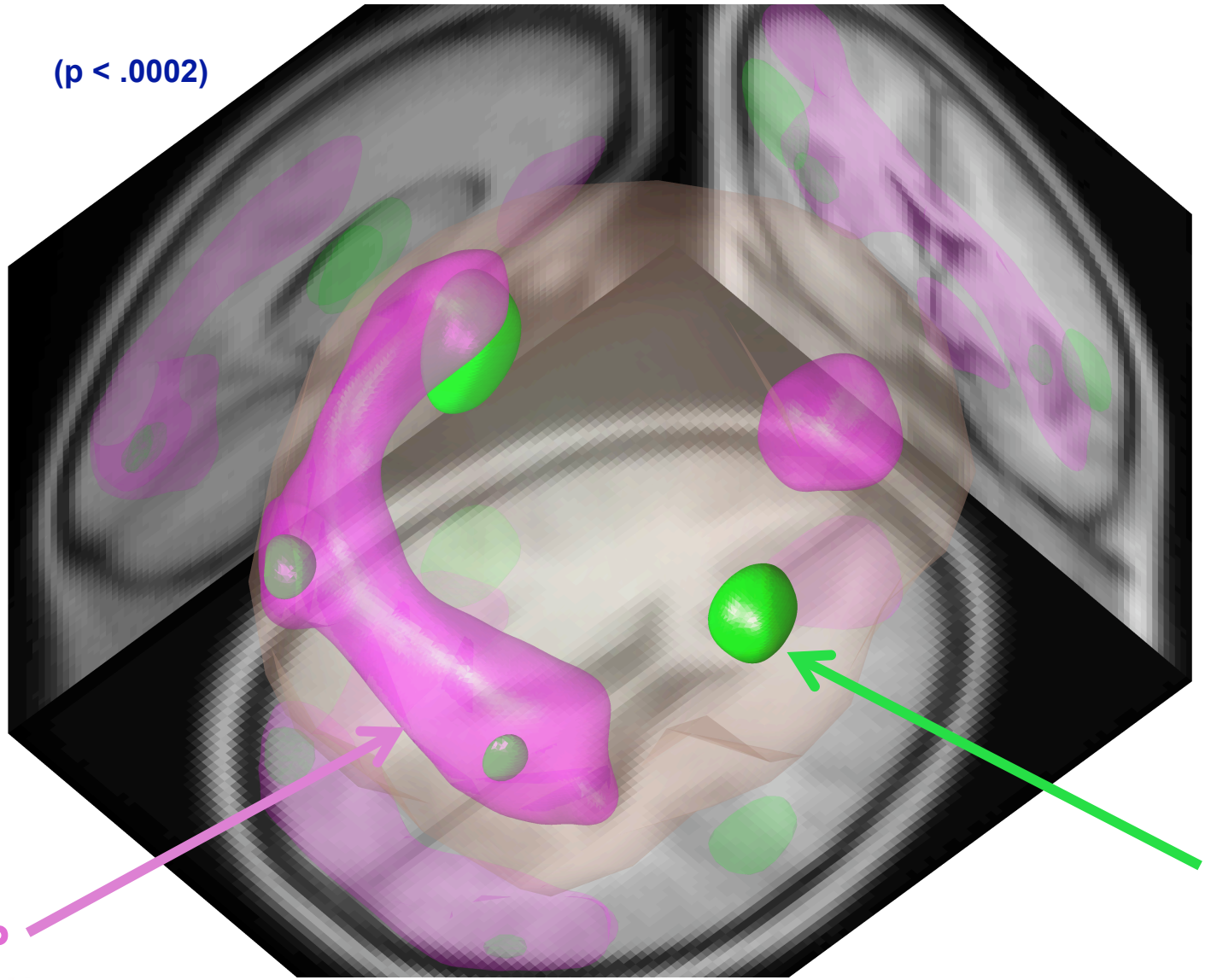
Measure Projection: RSVP Example



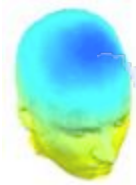
N. Bigdely-Shamlo, 2011

Measure Projection: RSVP Example





N. Bigdely-Shamlo, 2011



Questions?