Independent Component Analysis of Electrophysiological Data

Swartz Center for Computational Neuroscience

Scott Makeig
Institute for Neural Computation
University of California San Diego

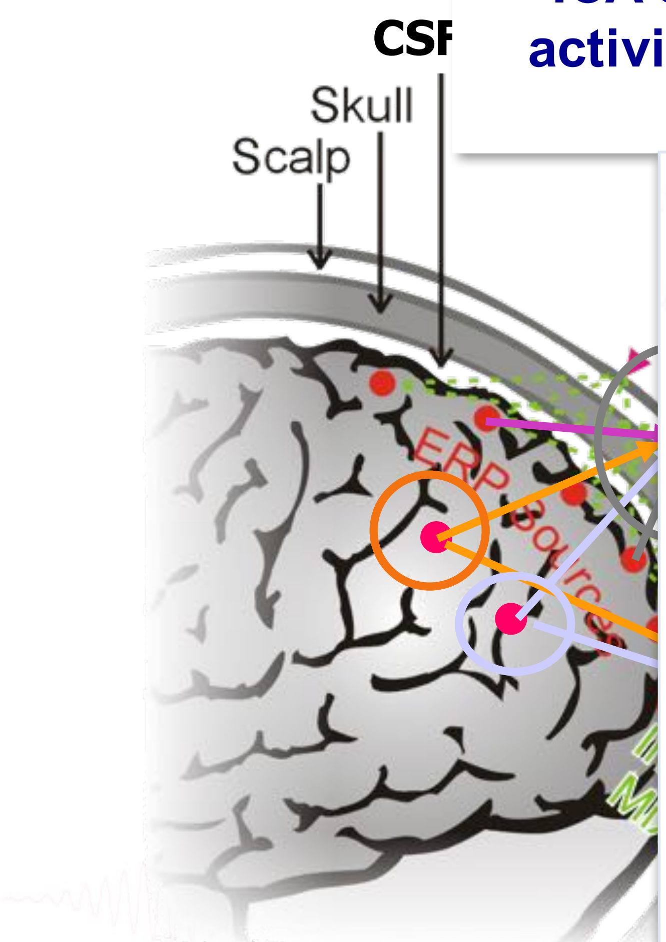
23rd EEGLAB Workshop
Mysuru, India

January, 2017

Blind EEG Source Separation by Independent Component Analysis

Tony Bell, developer of Infomax ICA

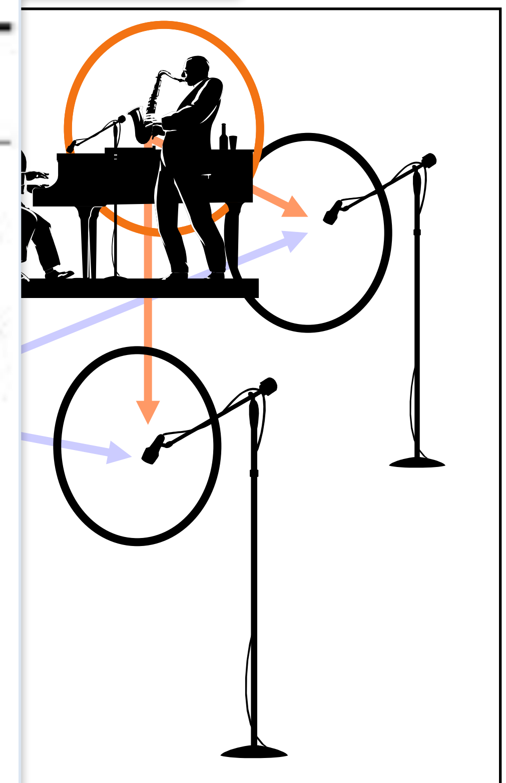
ICA can find distinct EEG source activities -- and their 'simple' scalp maps!



Independent Component Analysis of Electroencephalographic Data

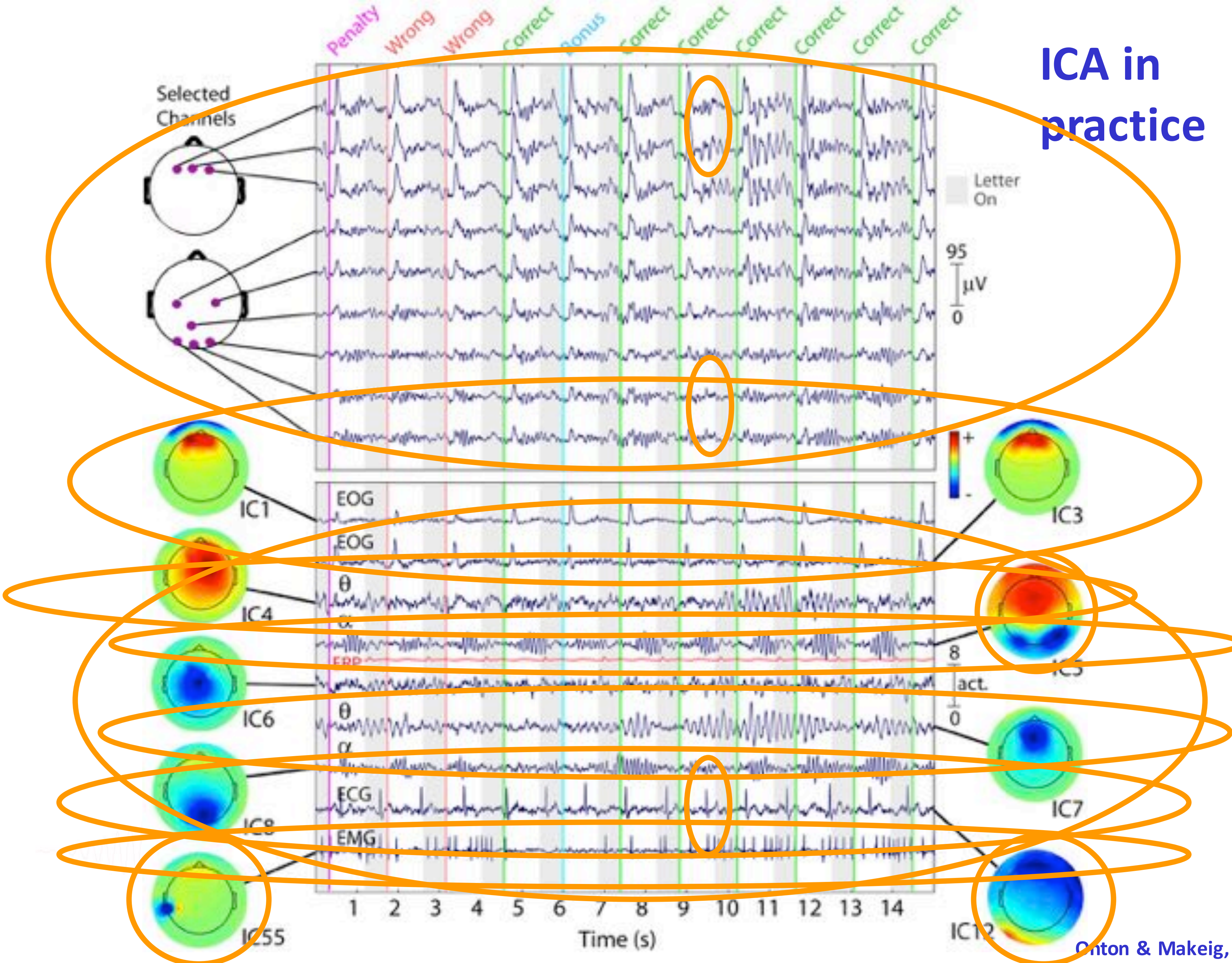
Scott Makeig
Anthony J. Bell
Tzyy-Ping Jung
Terrence J. Sejnowski

Abstract

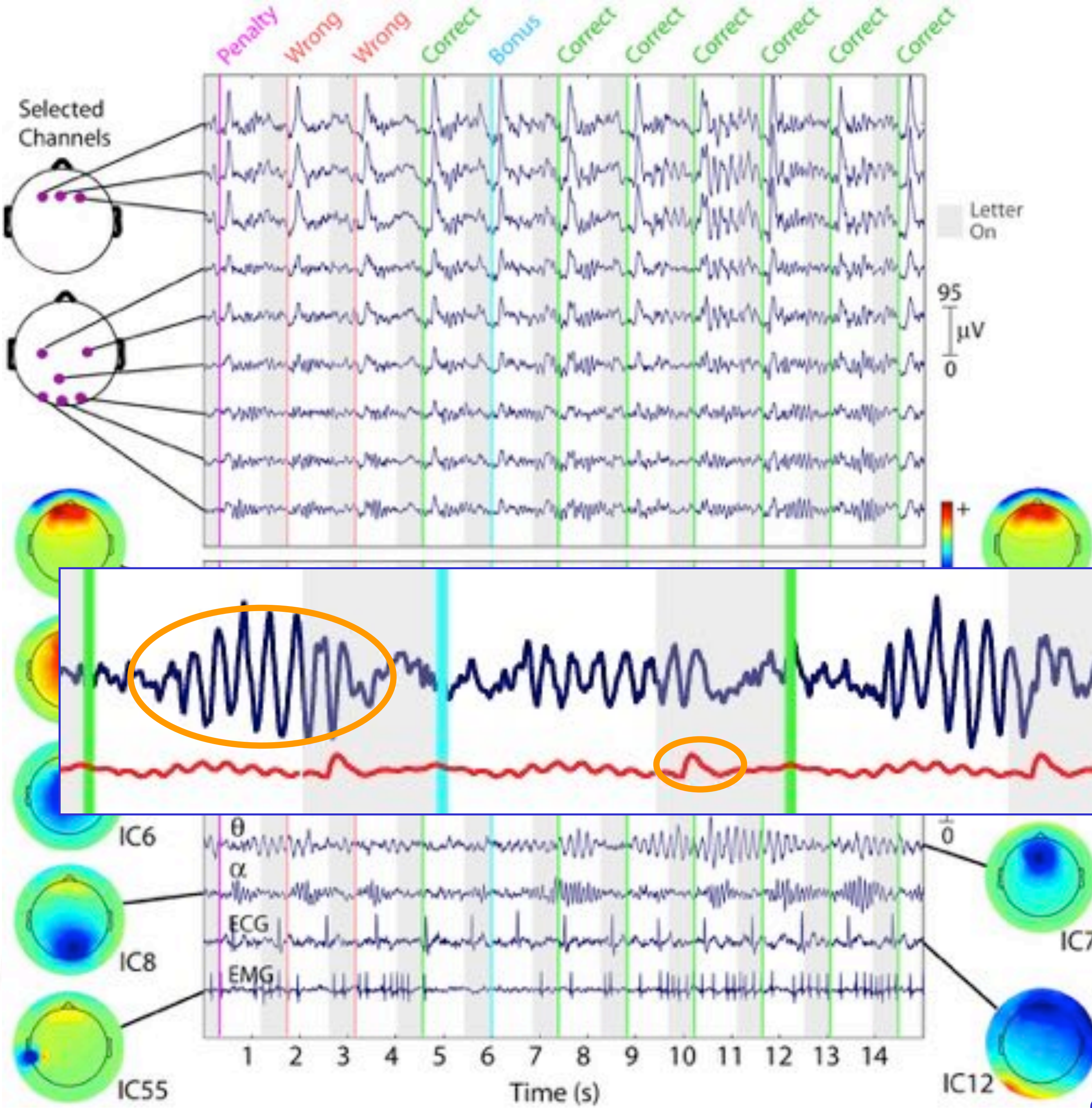


S. Makeig, S. Enghoff (2000)

ICA in practice



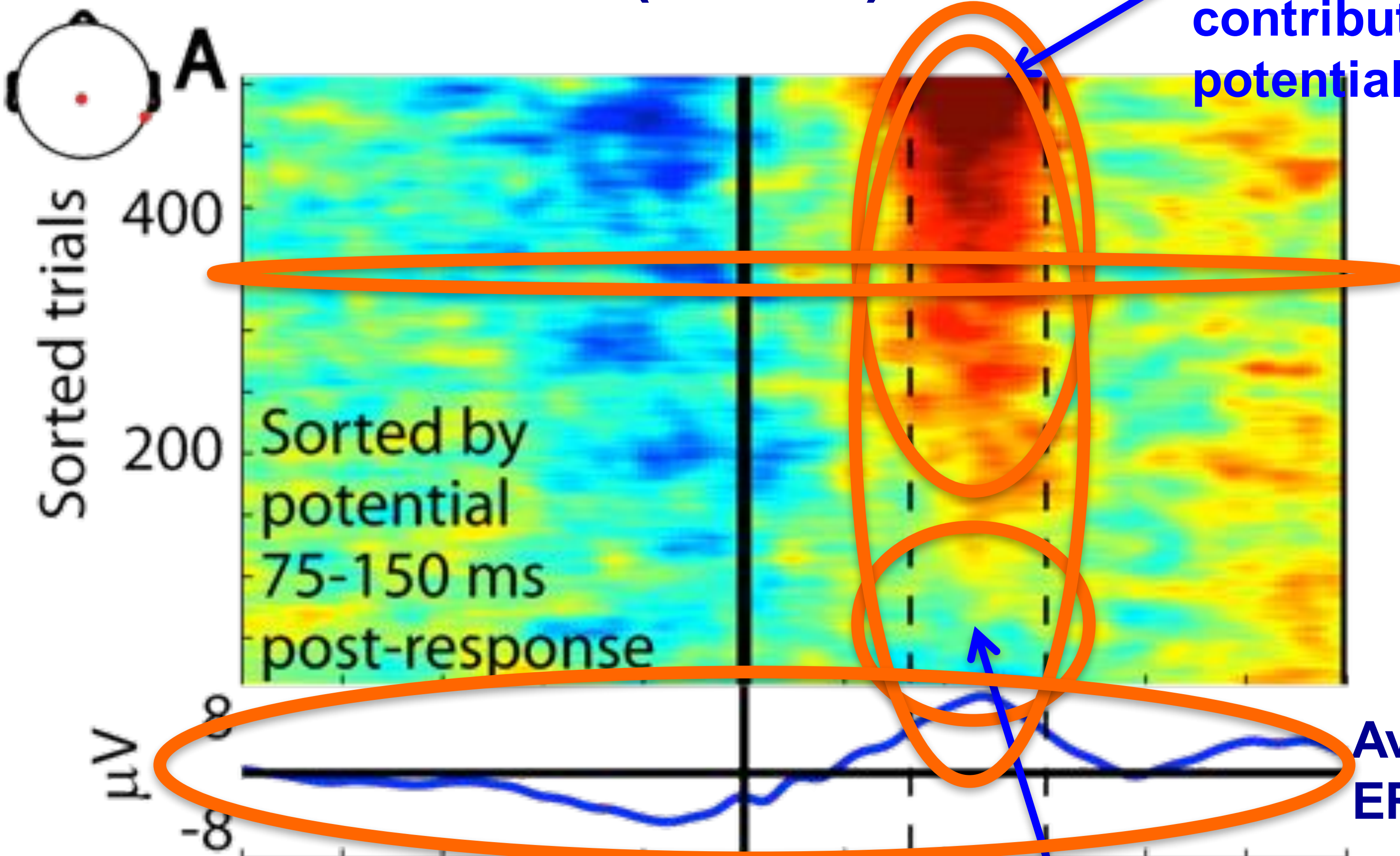
Onton & Makeig, 2006



Onton, Makeig (2006)

A P300' visual target response at electrode Cz (vertex)

What sources contribute to this potential?

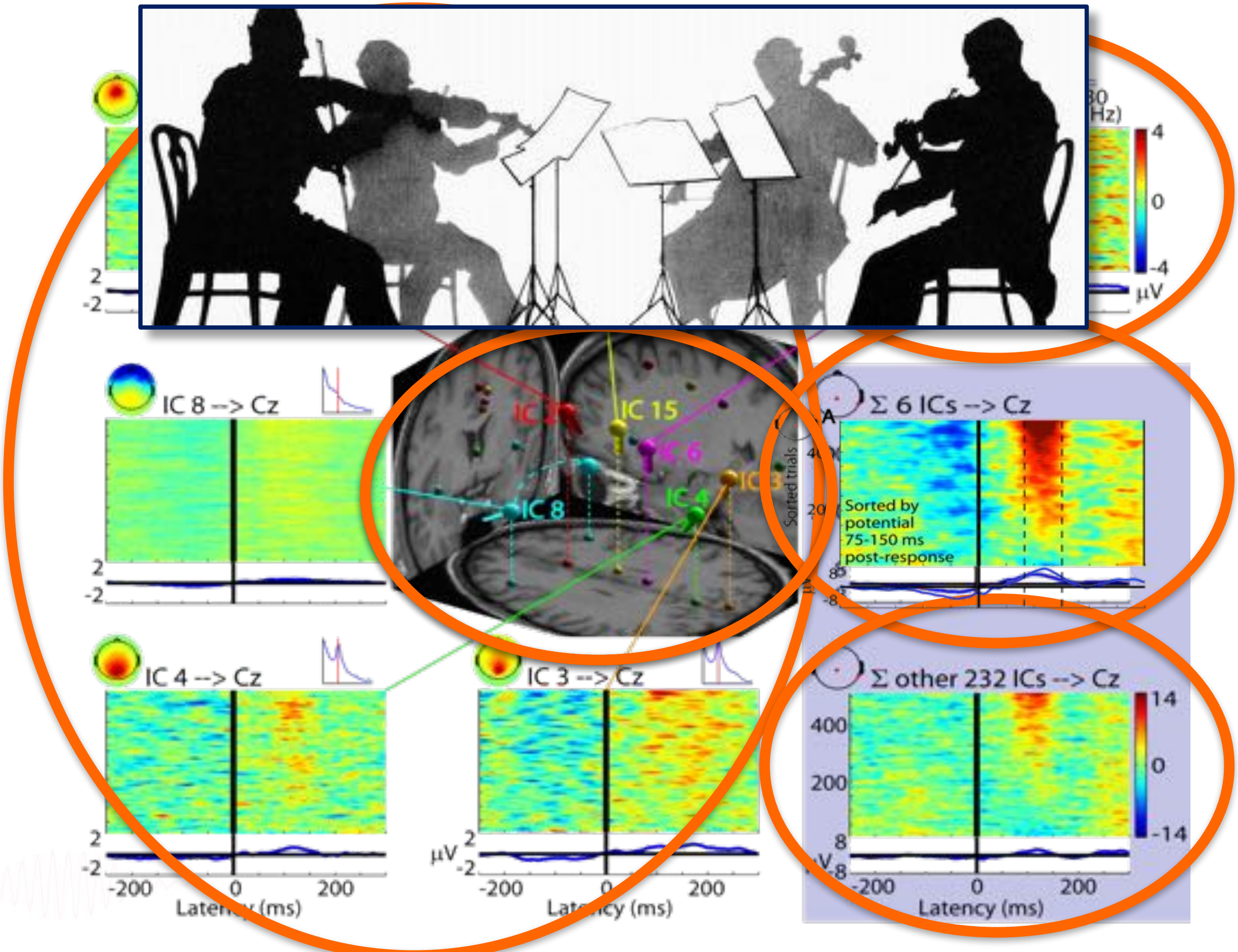


Average ERP

No scalp response in these trials … Why not?

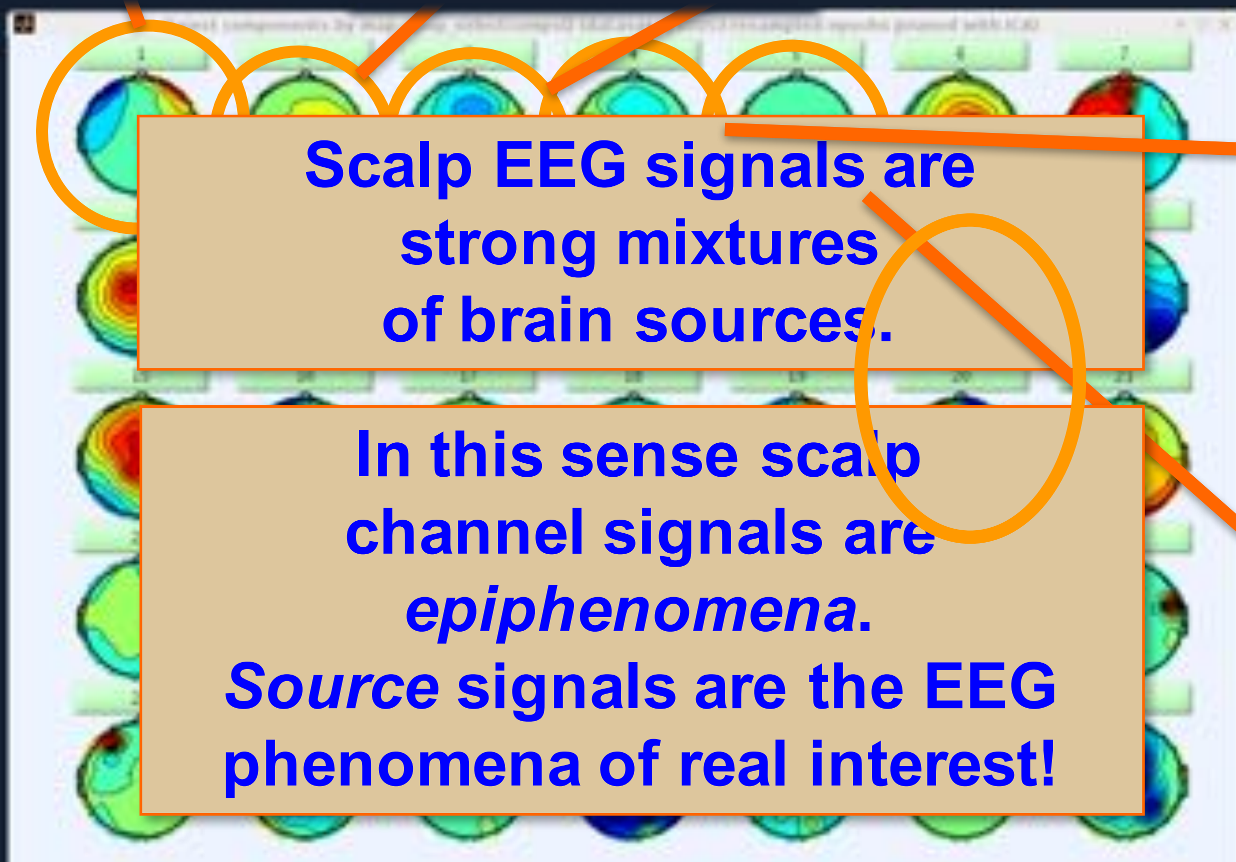
J. Onton & S. Makeig (2009)

The response (at Cz) sums 238 independent sources



J. Onton & S. Makeig (2009)

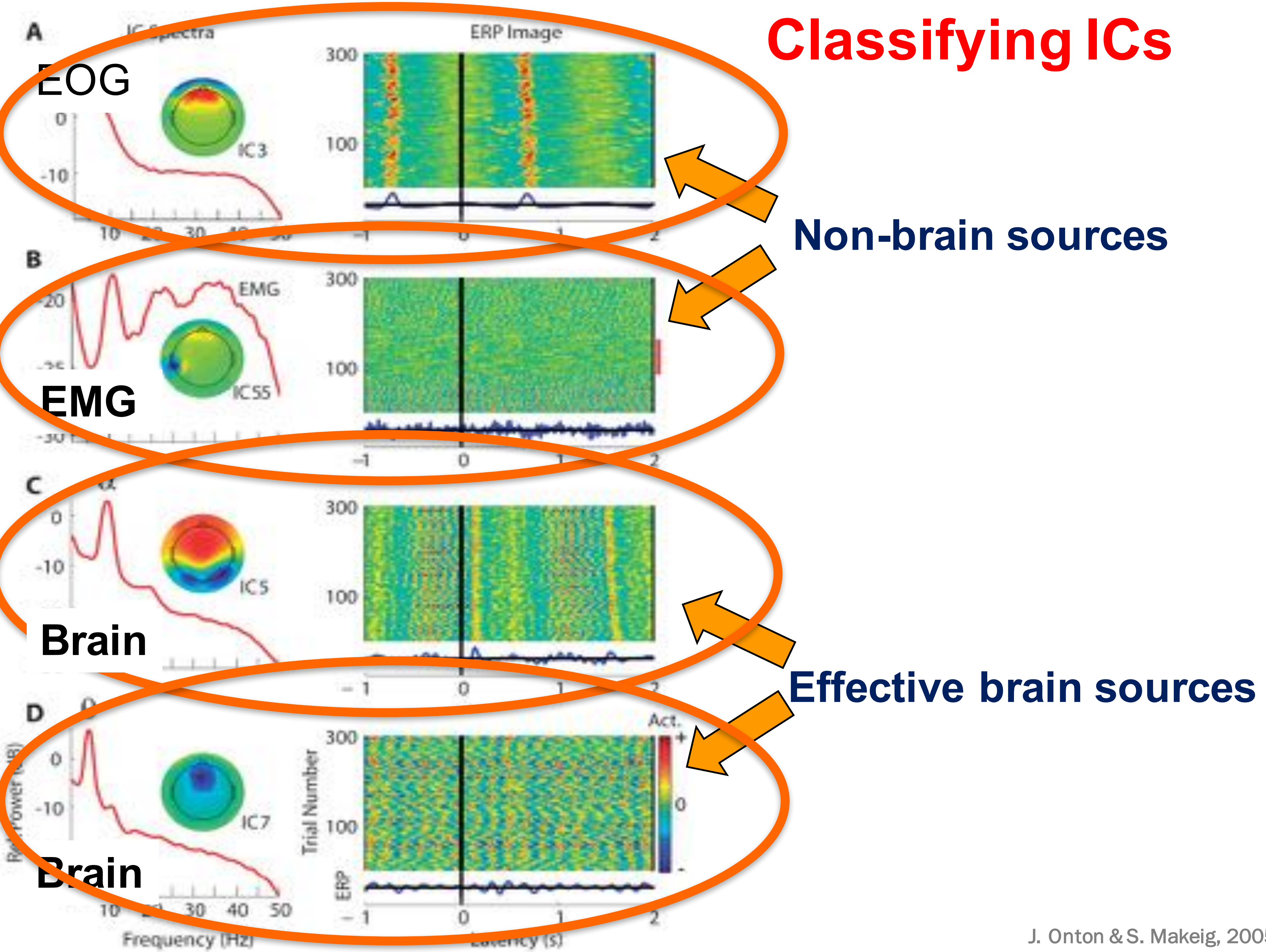
No more than ~30% of any scalp channel variance is produced by any one brain source!



Scalp EEG signals are strong mixtures of brain sources.

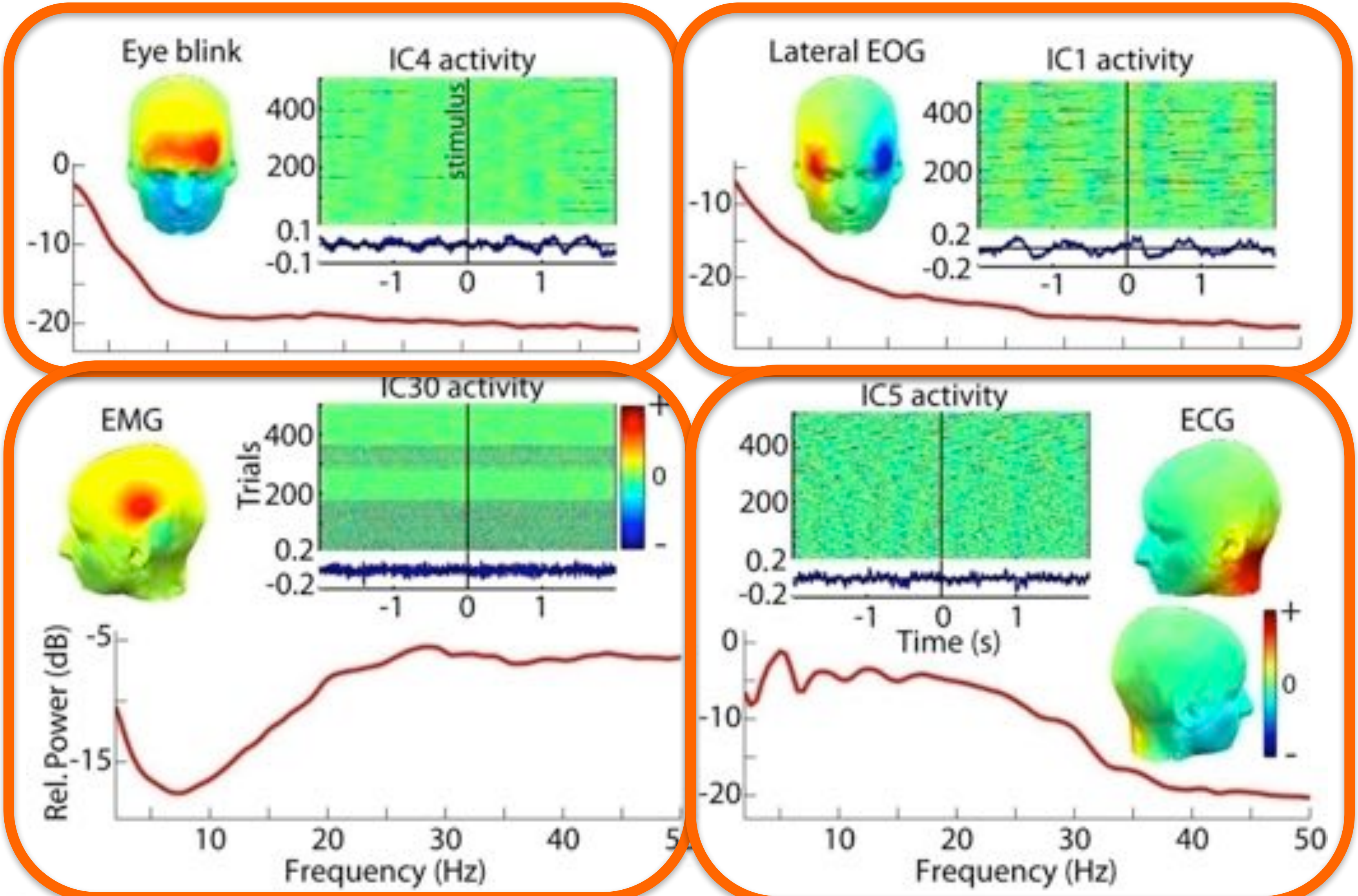
In this sense scalp channel signals are *epiphenomena*.
***Source* signals are the EEG phenomena of real interest!**

Classifying ICs



J. Onton & S. Makeig, 2005

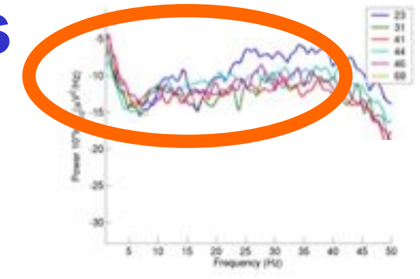
ICA finds Non-Brain Independent Component (IC) Processes ...

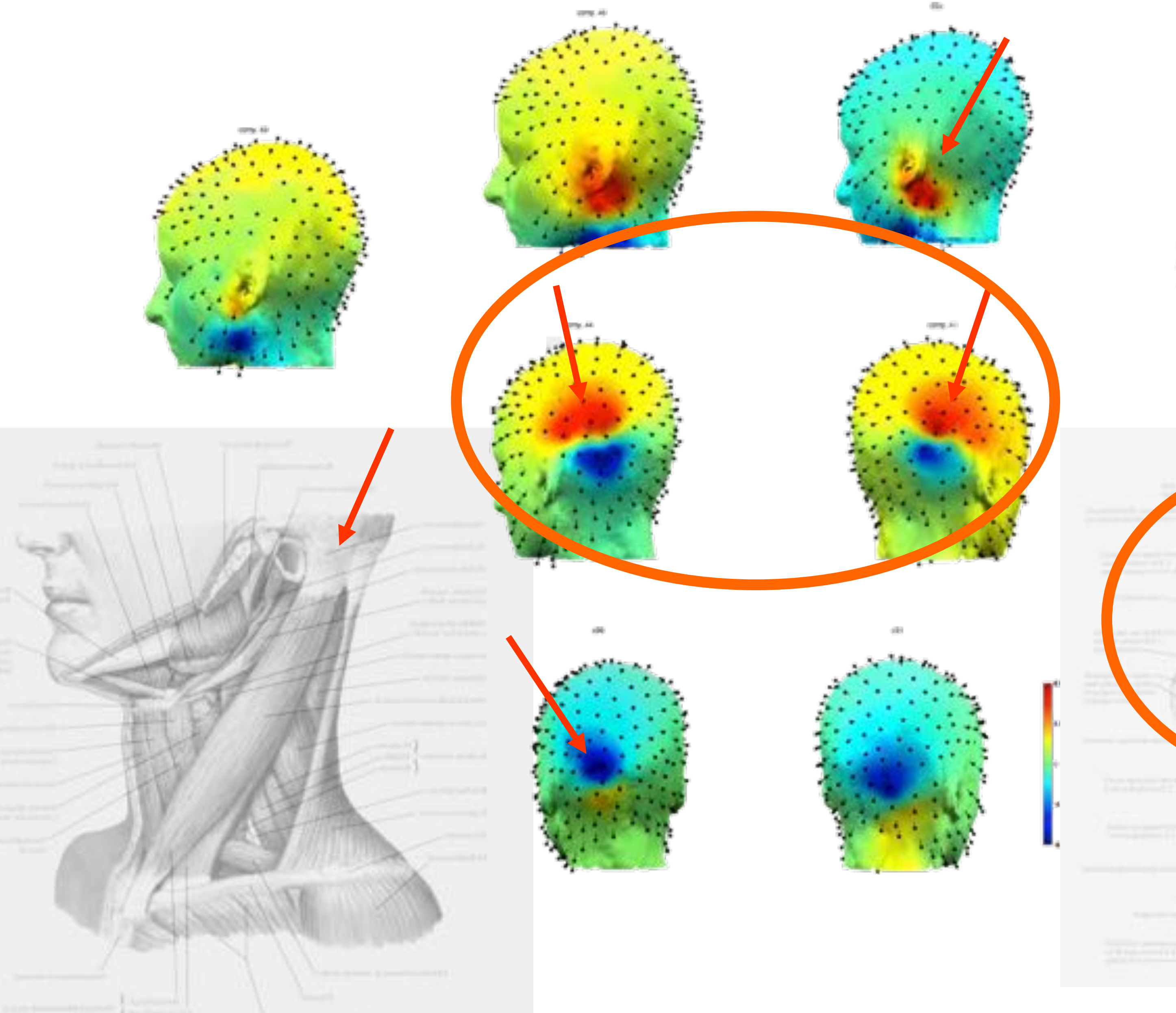


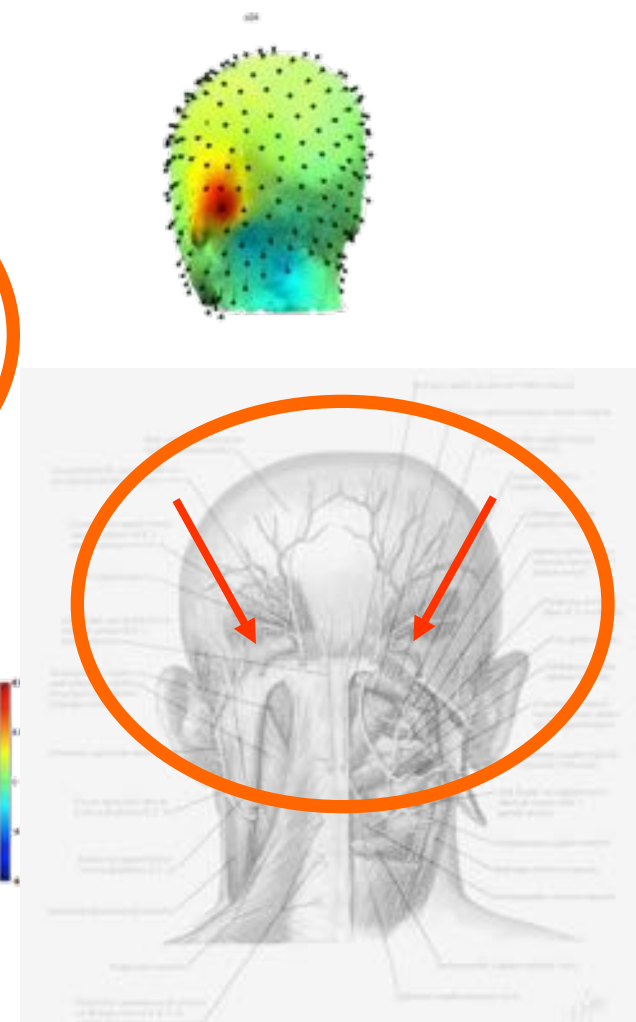
... separates them from the remainder of the data ...

J. Onton & S. Makeig

Independent muscle signals

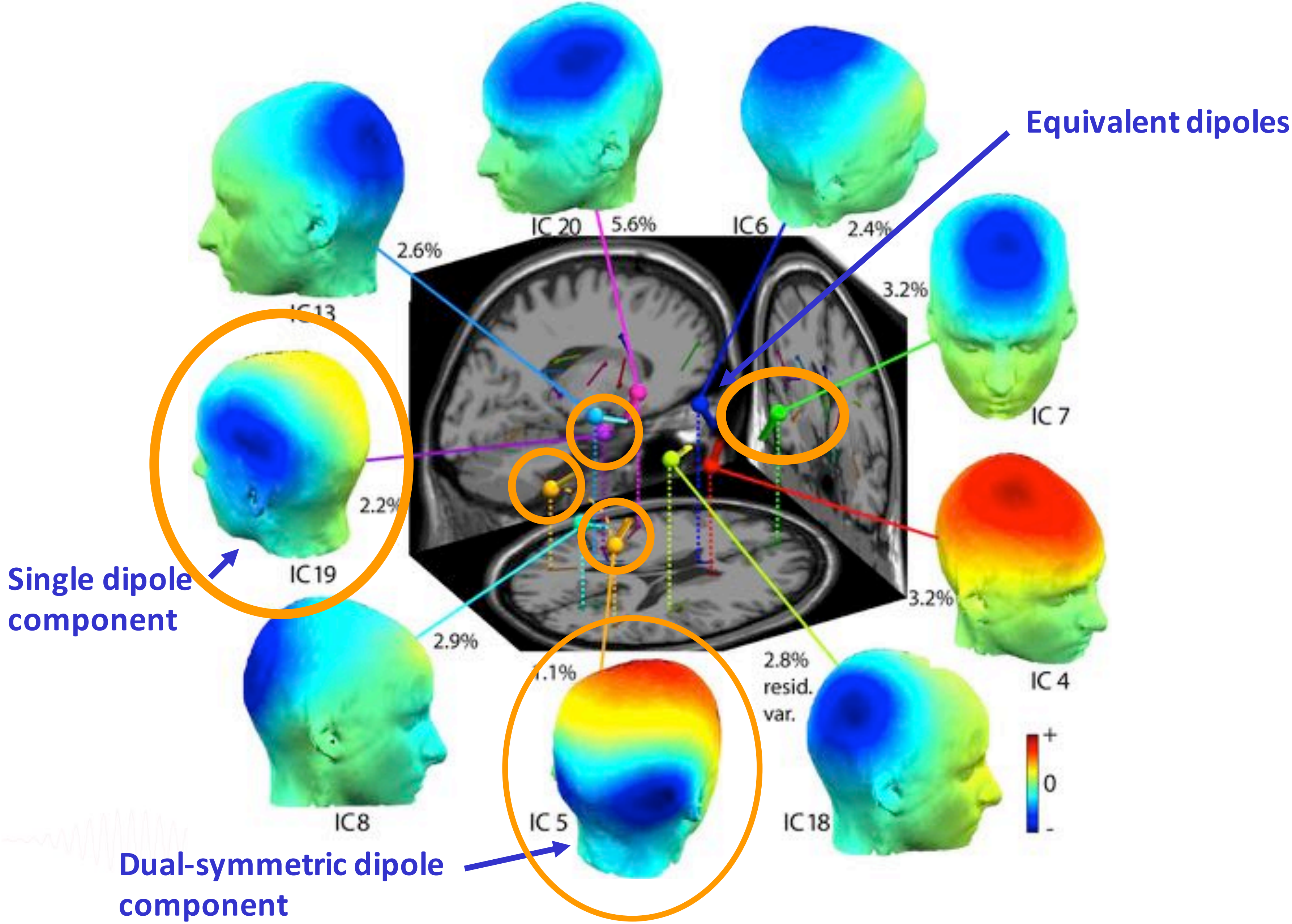






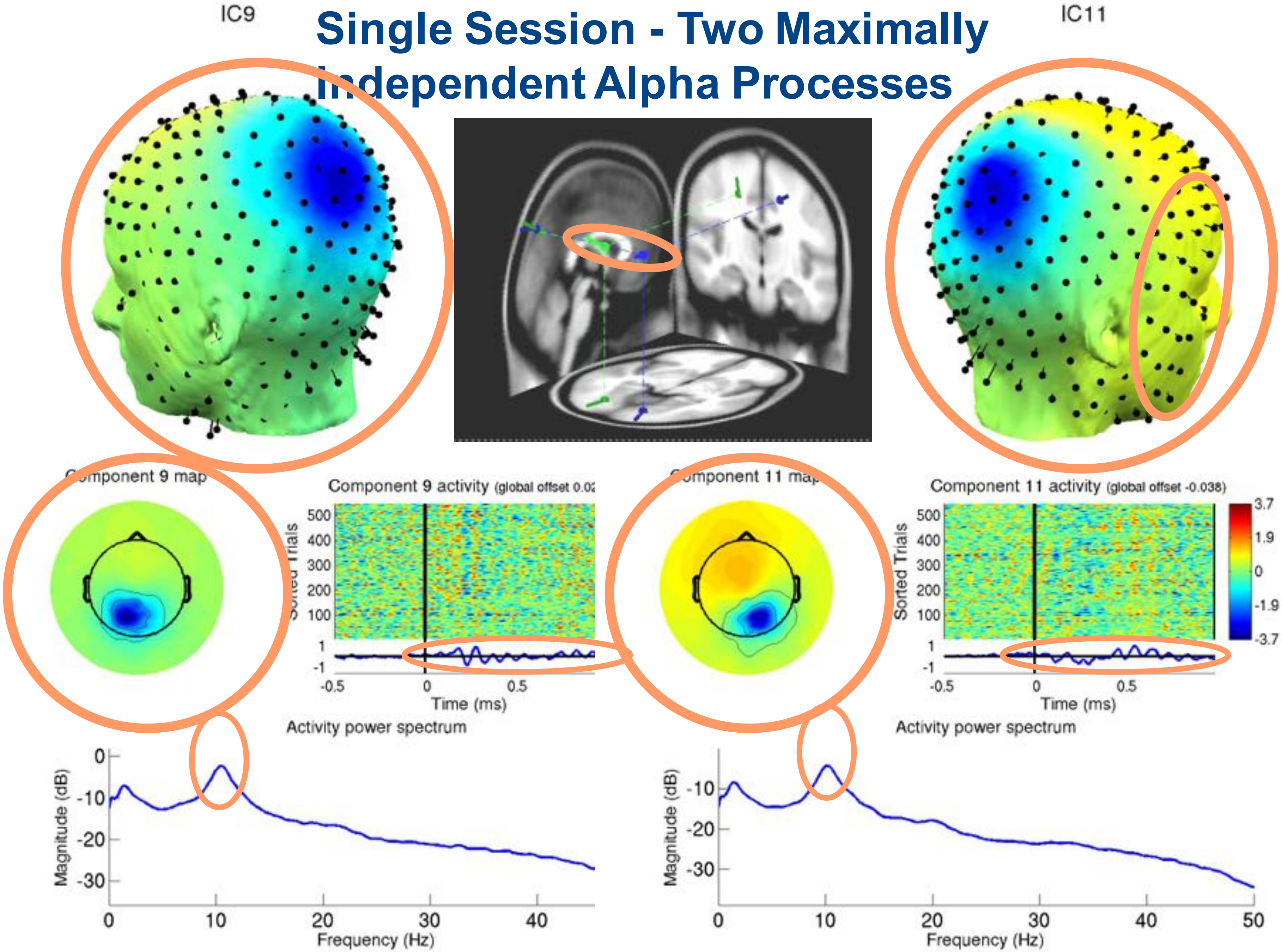
S. Makeig, J. Onton 2005

… and also separates cortical brain IC processes

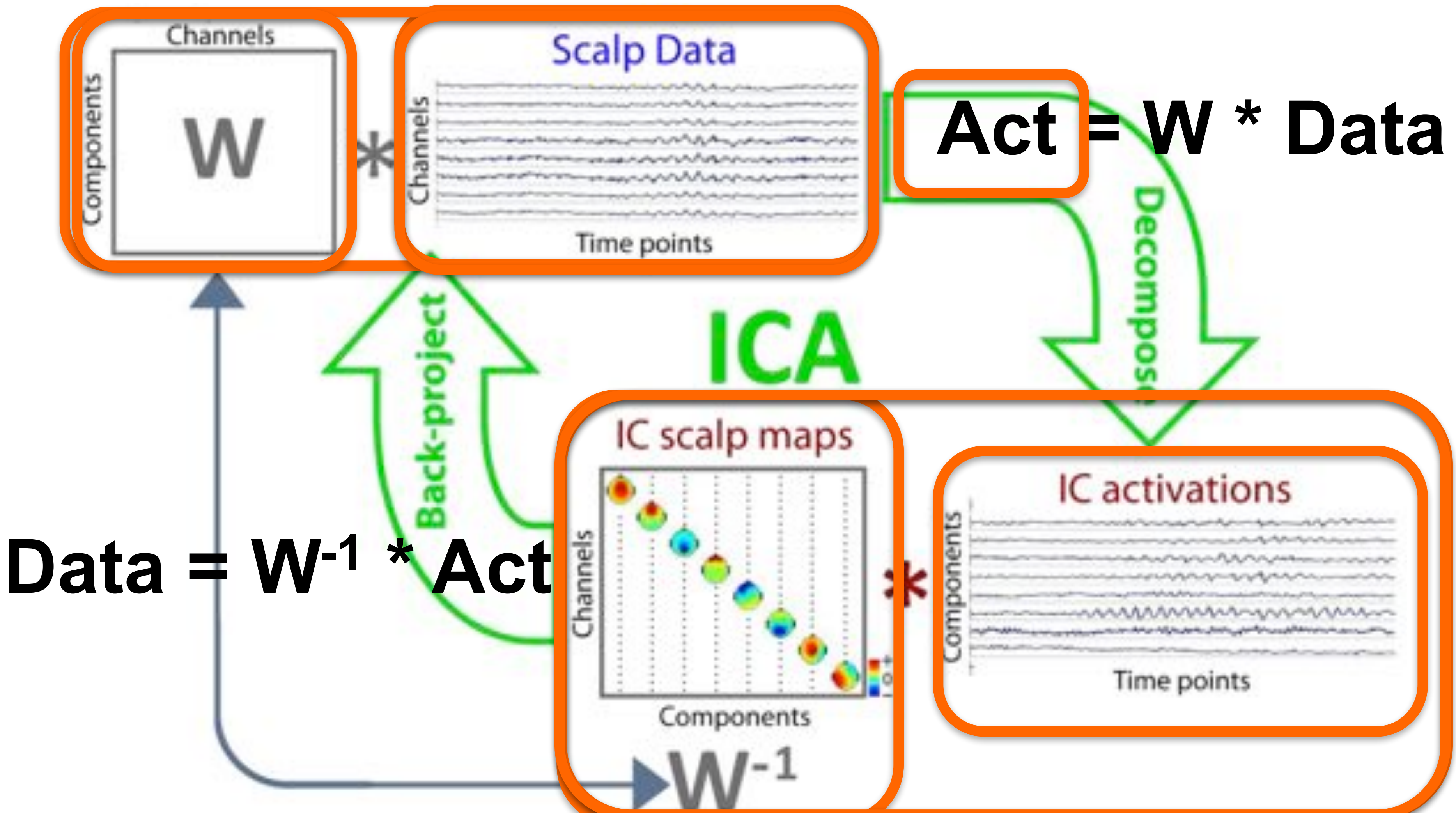


Julie Onton & S. Makeig (2006)

Single Session - Two Maximally Independent Alpha Processes



ICA is a linear data decomposition method



Data = W-1 * (W * Data)

Makeig & Onton, 2011

Infomax ICA learning approach

How to make the outputs statistical independent?

Minimize their redundancy or mutual information.

Consider the joint entropy of two components,

$$\boxed{H(y_1, y_2) = H(y_1) + H(y_2) - I(y_1, y_2).}$$

$$\boxed{\text{Maximizing } H(y_1, y_2)} \Longrightarrow \boxed{\text{minimizing } I(y_1, y_2).}$$

Infomax

Is 0 if the two variables are independent

The learning rule:

$$\boxed{\Delta \mathbf{W}} \propto \boxed{\frac{\partial H(\mathbf{y})}{\partial \mathbf{W}}} \underbrace{\mathbf{W}^T \mathbf{W}}_{\text{Natural gradient normalization (Amari)}}$$

Historical Remarks

- Herault & Jutten ("Space or time adaptive signal processing by neural network models“, *Neural Nets for Computing Meeting*, Snowbird, Utah, 1986): Seminal paper
- Bell & Sejnowski (1995): Information maximization (Infomax)
- Makeig, Bell, Jung, Sejnowski (1996); ICA decomposition of EEG
- Amari et al. (1996): Natural gradient learning
- Cardoso (1996): Joint approximate diagonalization (JADE)
- Hyvarinen (1999): (fastICA)
- Lee/Girolami (1999): Mixture model ICA (Extended Infomax)

Applications of ICA to biomedical signals

- EEG/ERP analysis (Makeig, Bell, Jung & Sejnowski, *NIPS* 1996)
- fMRI analysis (McKeown et al., 1998)
- Fetal/mother ECG separation (Cardoso, 1998)
- Electrocorticography (ECoG) (Whitmer, 2010)

Important Recent Result (2012)

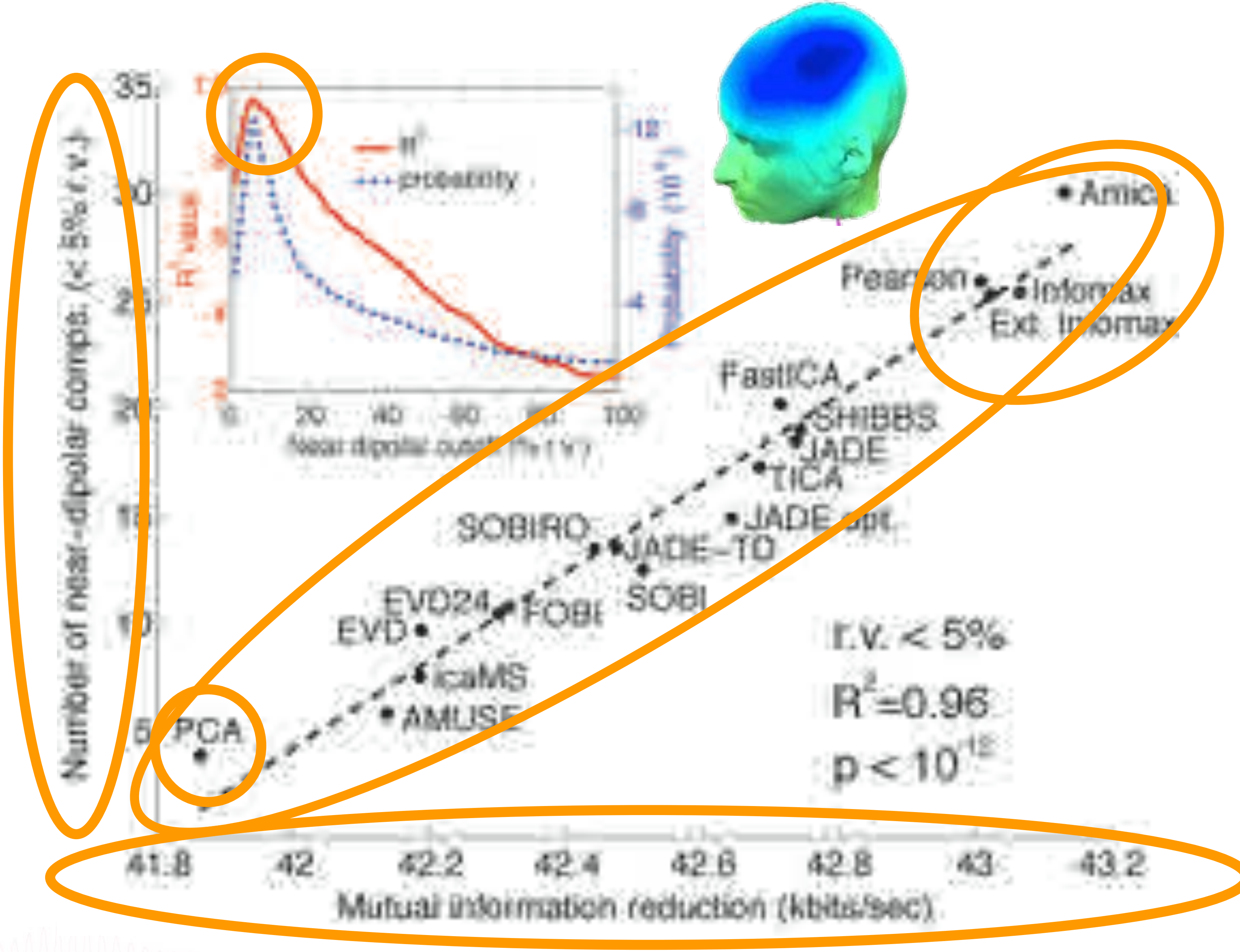
Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent …**

Also find more ICs whose scalp maps are highly '**dipolar**' – i.e., ICs compatible with the spatial projection of a single local cortical (or non-brain, artifactual) source process – whose location can be accurately estimated.

More independent time courses ←→ Larger number of dipolar ICs

Hypothesis: Dipolar ICs = Localized cortical source processes

Delorme et al., *PLOS One*, 2012



Delorme et al., *PLOS One*, 2012

S. Makeig, 2011

Important Recent Result

Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent …**

Also find more ICs whose scalp maps are highly '**dipolar**' – i.e., ICs compatible with the spatial projection of a single local cortical (or non-brain, artifactual) source process – whose location can be accurately estimated.

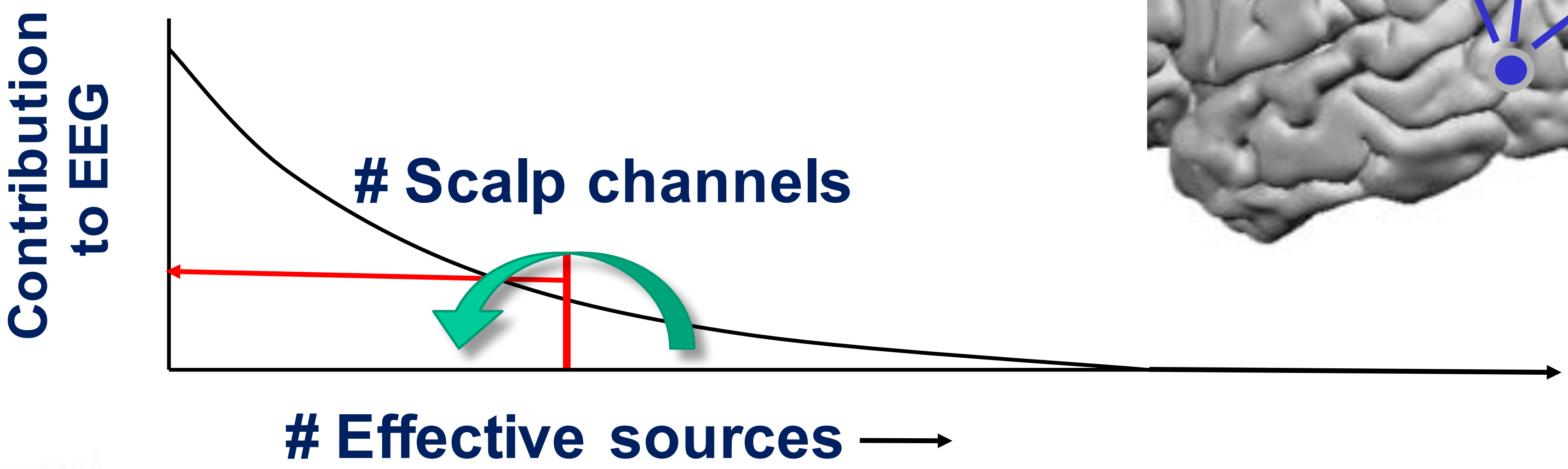
More independent time courses ←→ Larger number of dipolar ICs

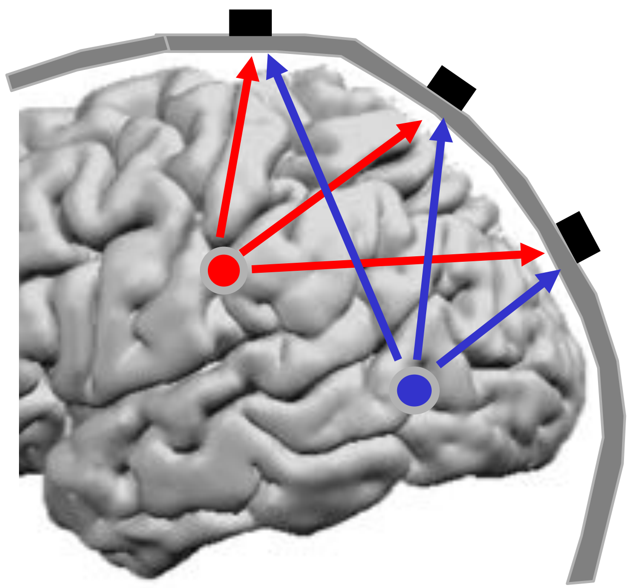
Dipolar ICs = Localized cortical source processes

Delorme et al., *PLOS One*, 2012

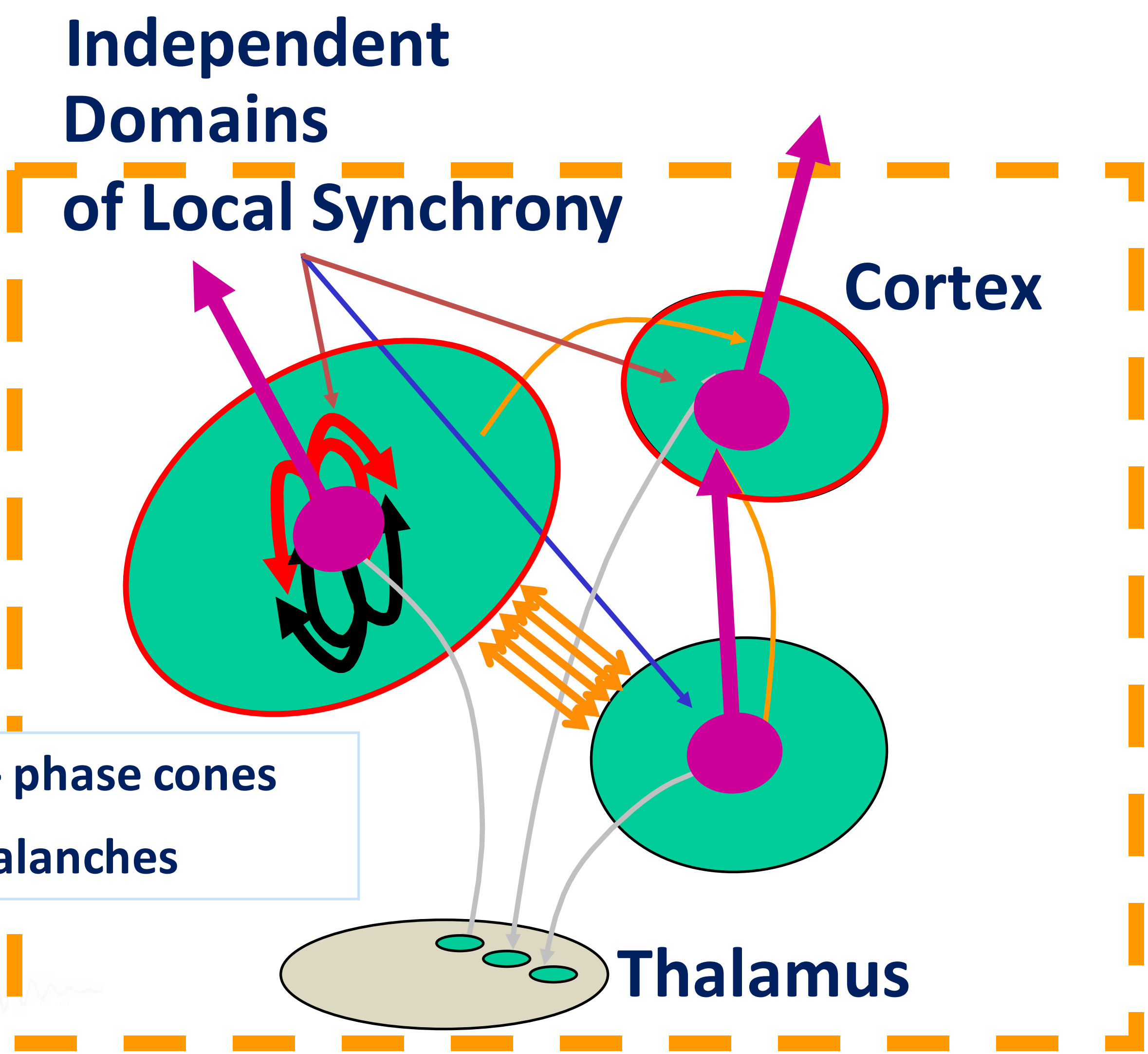
ICA Assumptions

- Mixing is linear at electrodes ✔
- Propagation delays are negligible ✔
- Component locations are fixed ?
- Component time courses are independent ?
- # components <= # scalp channels ?





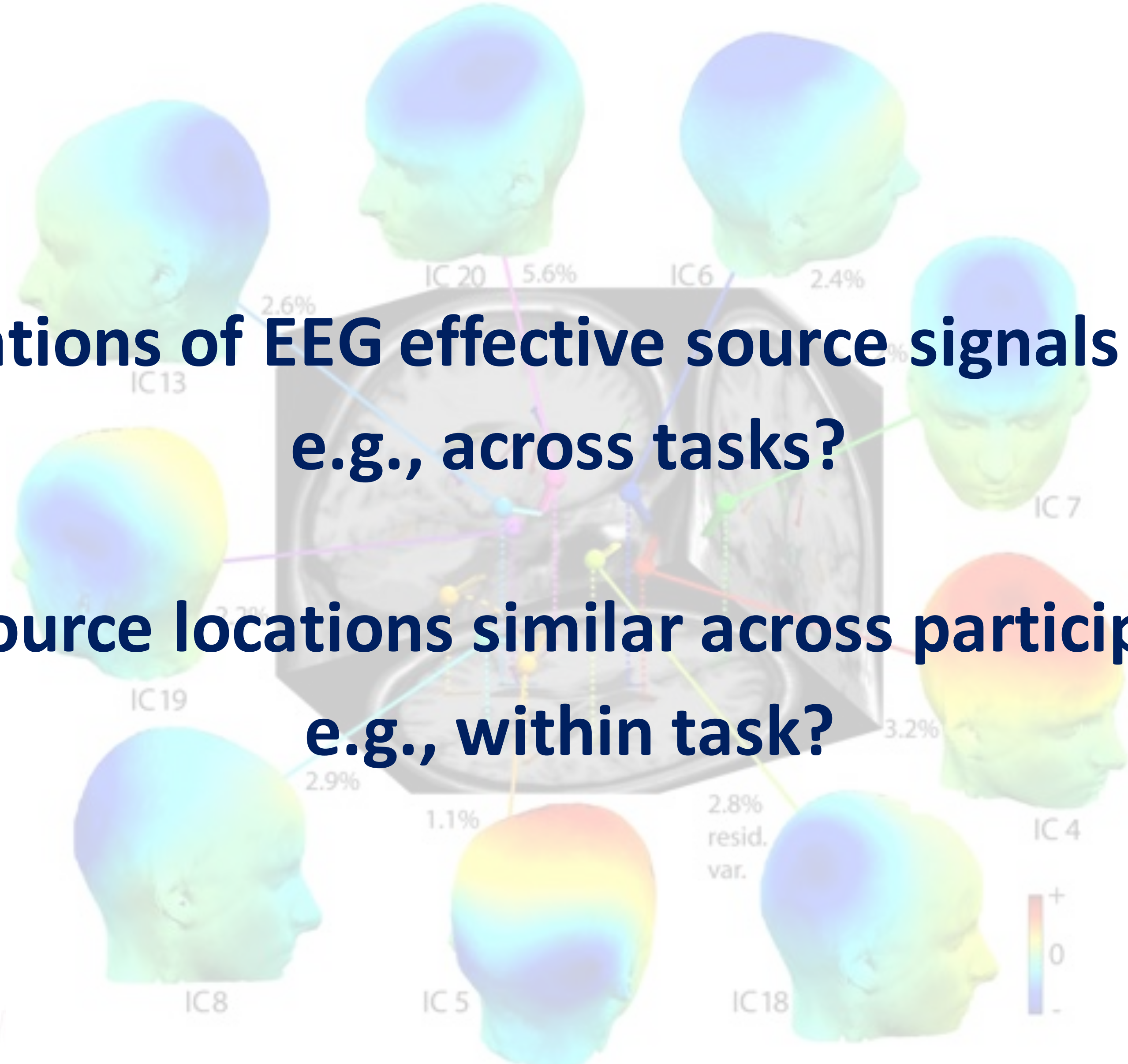
Are EEG effective source signals independent?



Freeman - phase cones

Plenz - avalanches

S. Makeig (2007)



Are locations of EEG effective source signals similar?

e.g., across tasks?

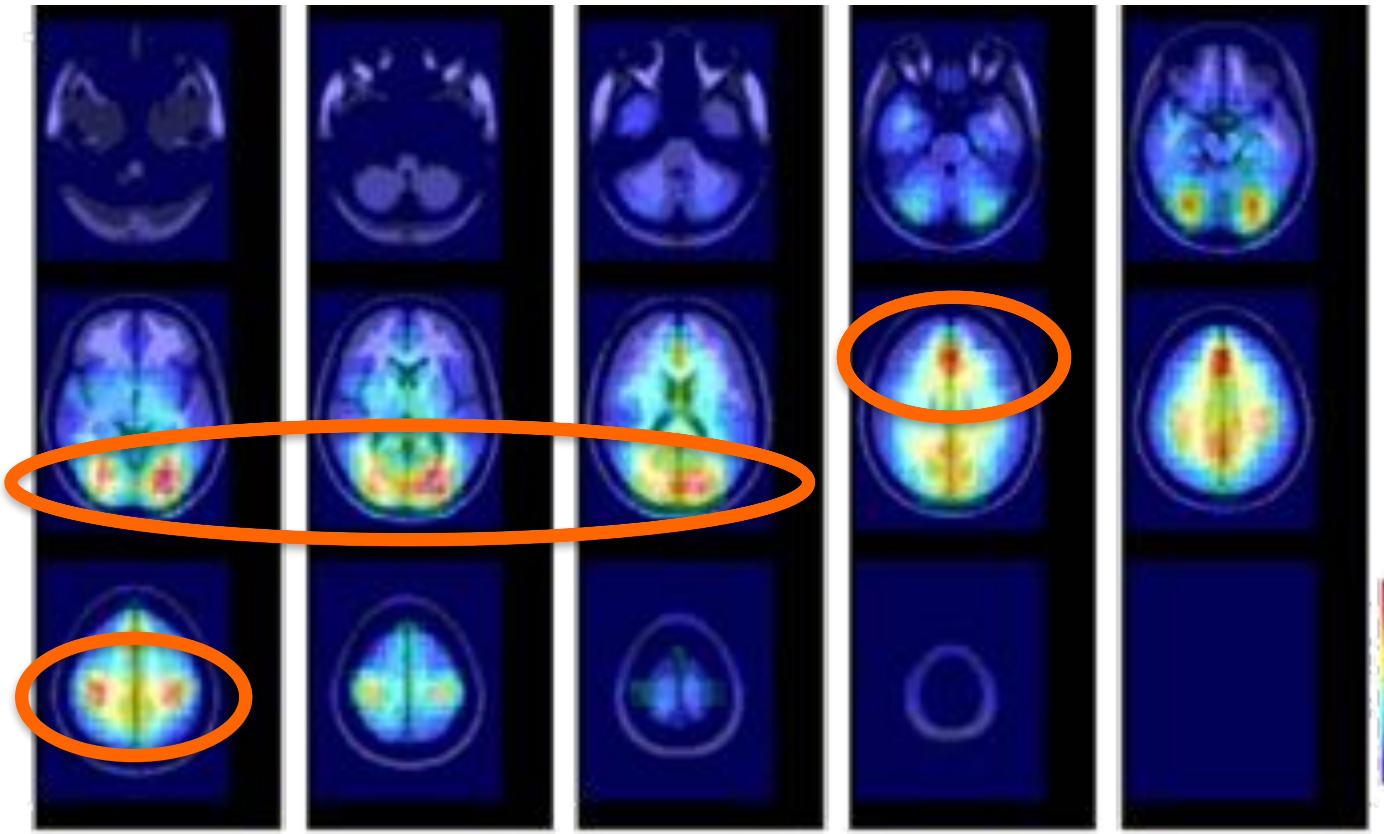
Are source locations similar across participants?

e.g., within task?

S. Makeig (2007)

Effective Source Density

Visual Working Memory



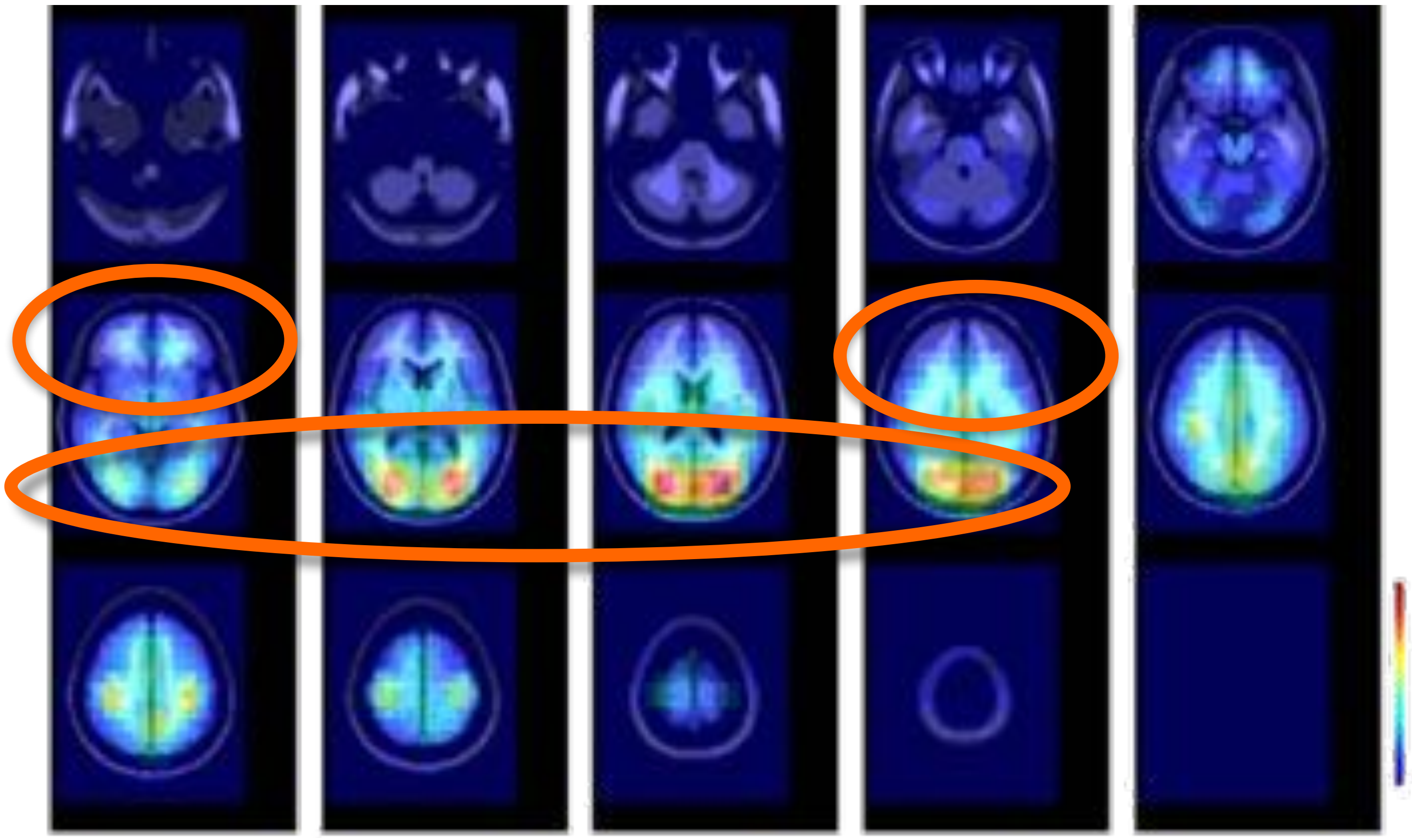
Sternberg letter memory task

dipoledensity()

Onton & Makeig, 2005

Effective Source Density

Eyes-closed emotion imagination

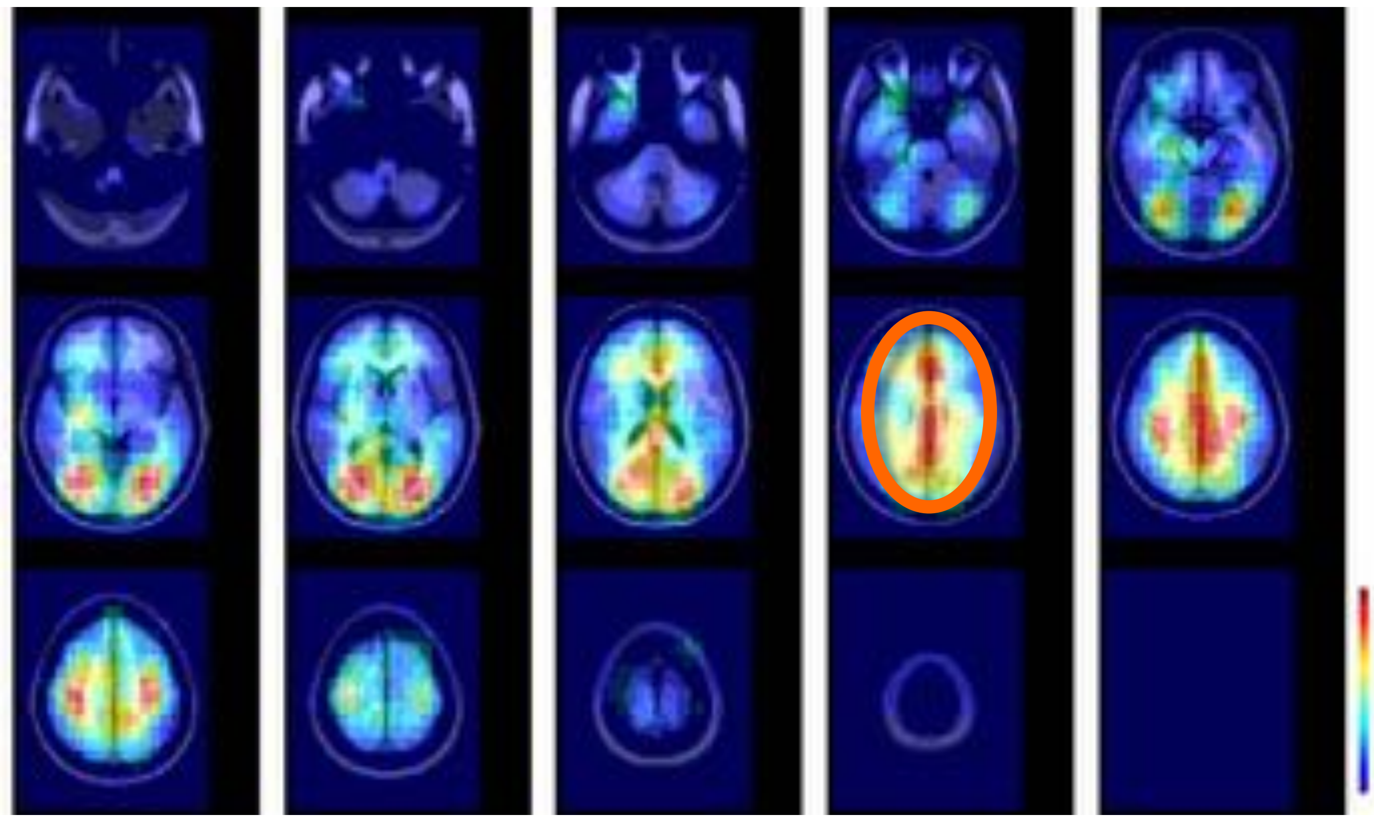


>> dipoledensity()

Onton & Makeig, 2005

Effective Source Density

Letter twoback with feedback

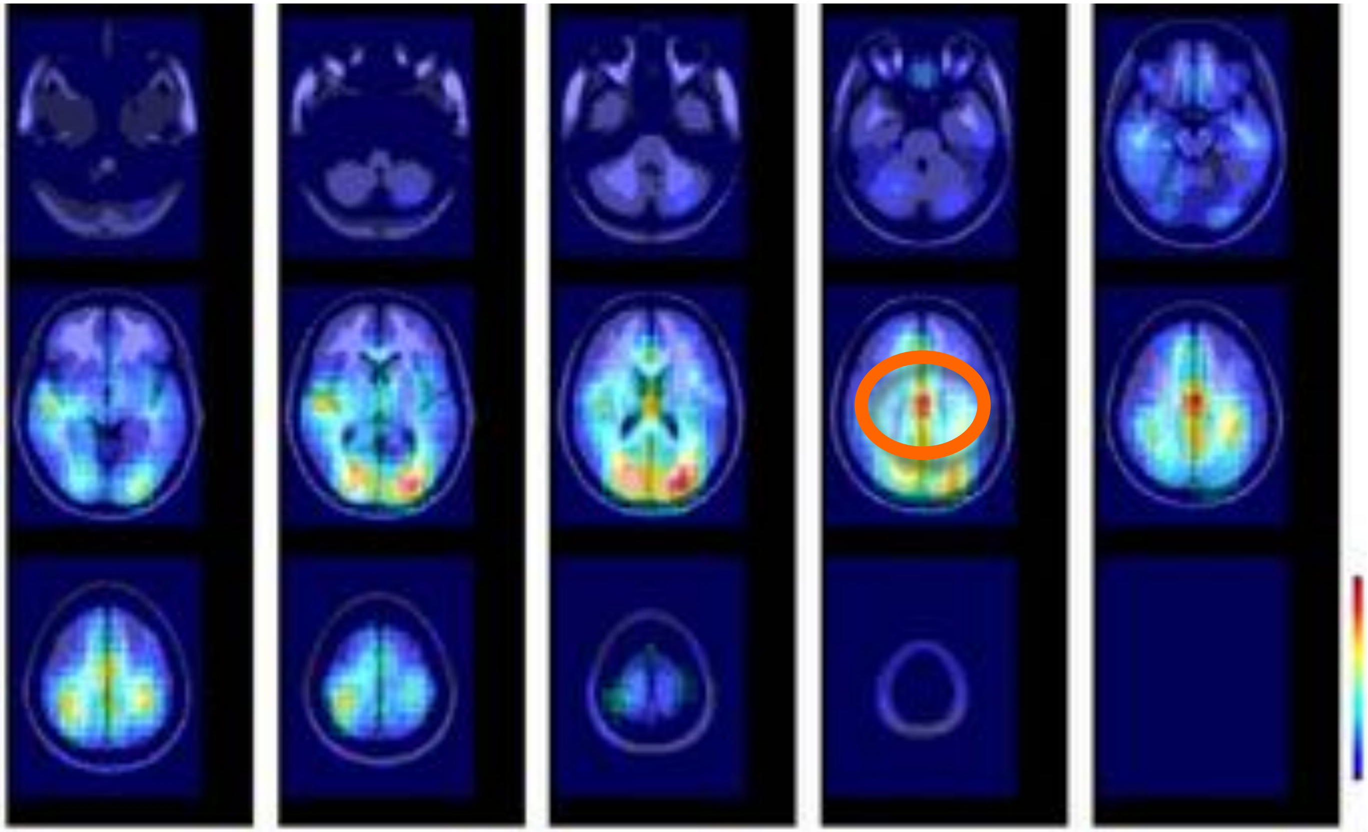


Letter twoback with feedback

Onton & Makeig, 2005

Effective Source Density

Auditory novelty oddball

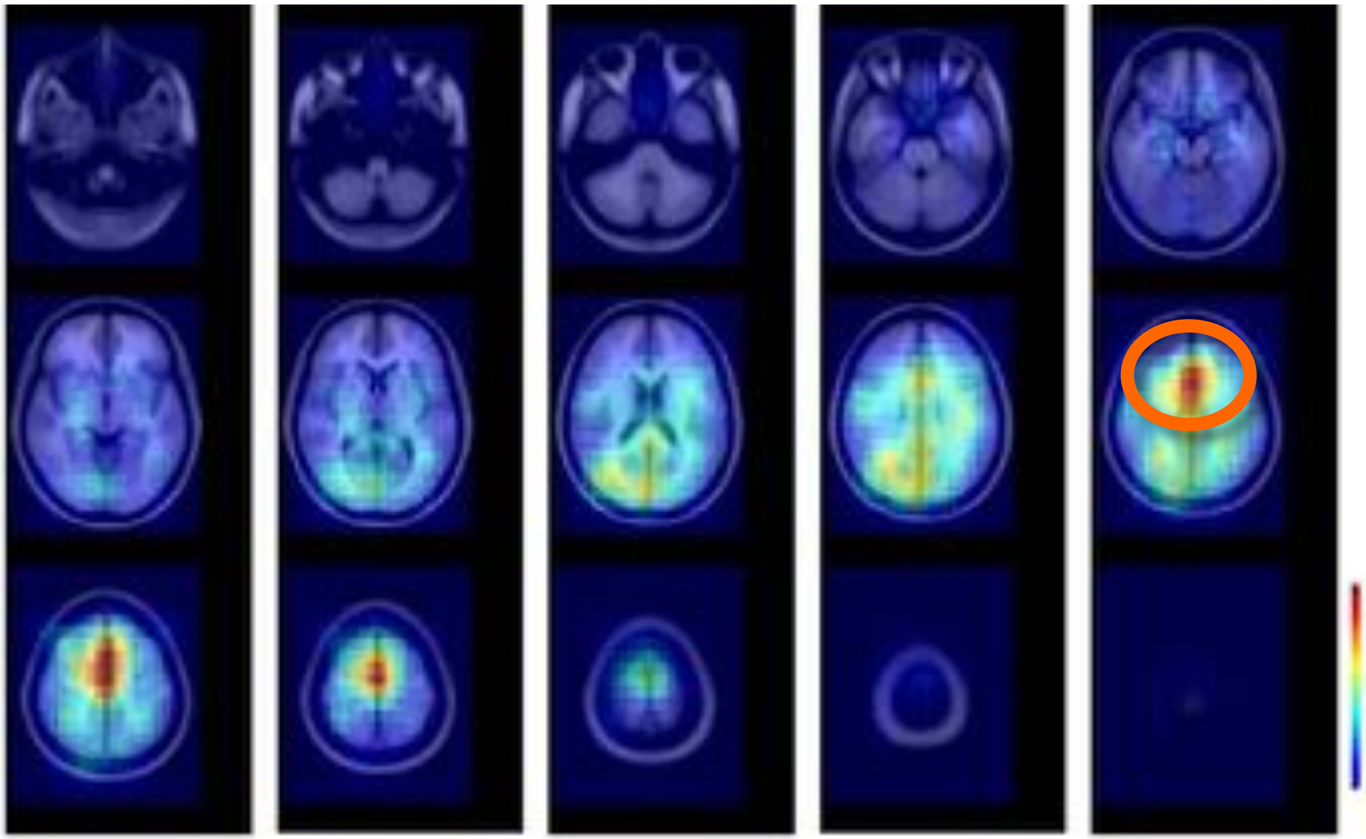


Auditory oddball plus novel sounds

Onton & Makeig, 2005

Effective Source Density

A. Old/new word memory

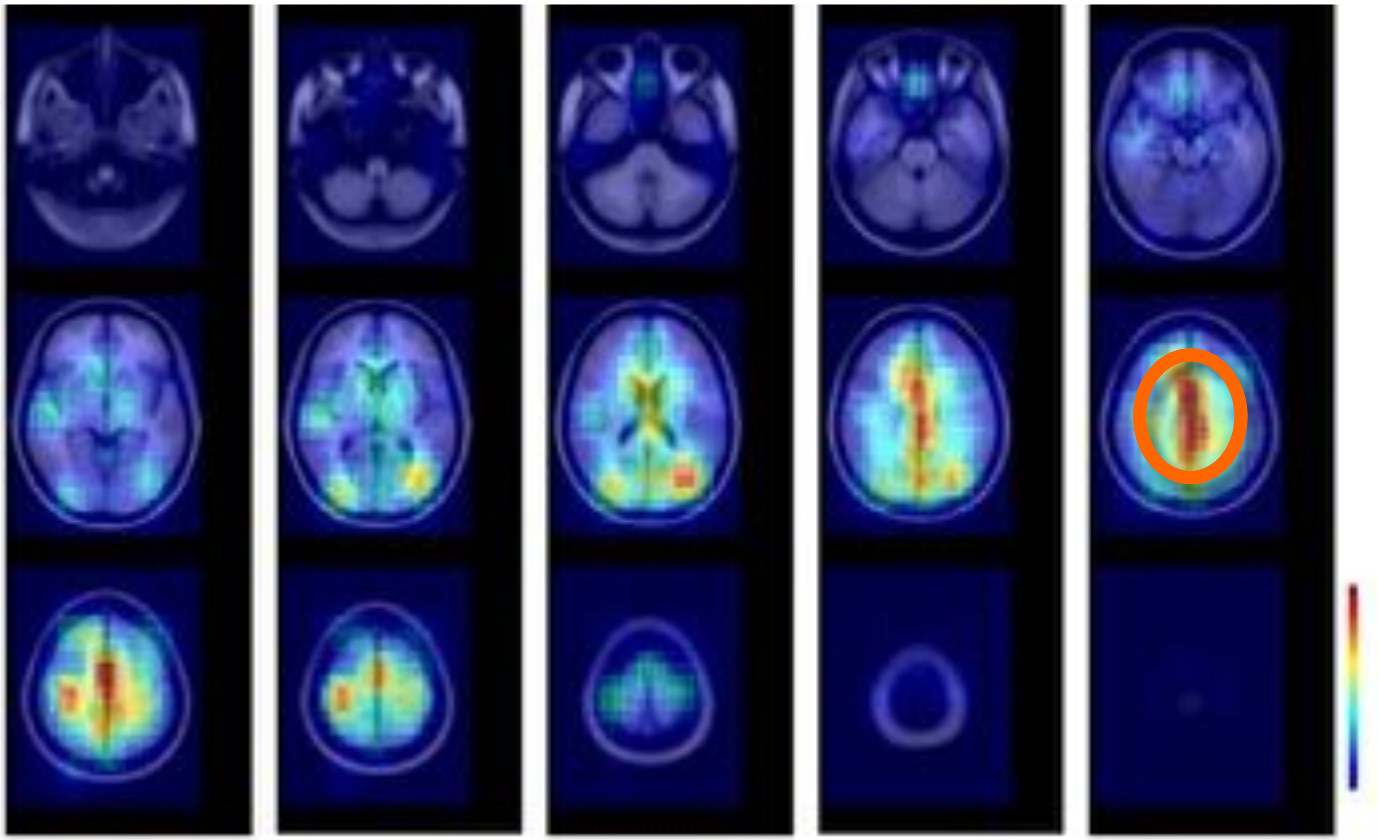


Word memory (old/new) task

Onton & Makeig, 2005

Effective Source Density

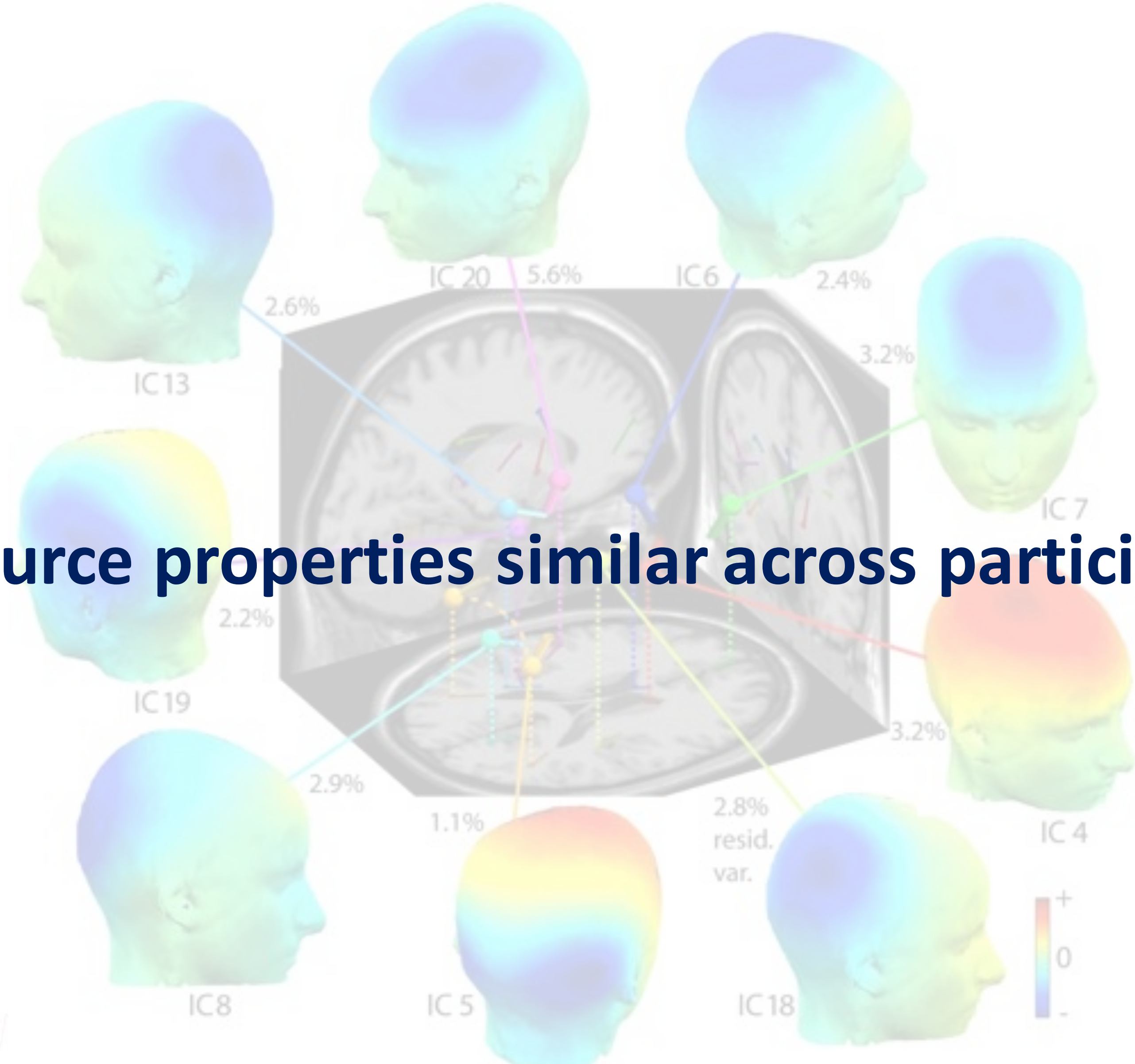
B. Visually cued selective response



Visually cued button press task

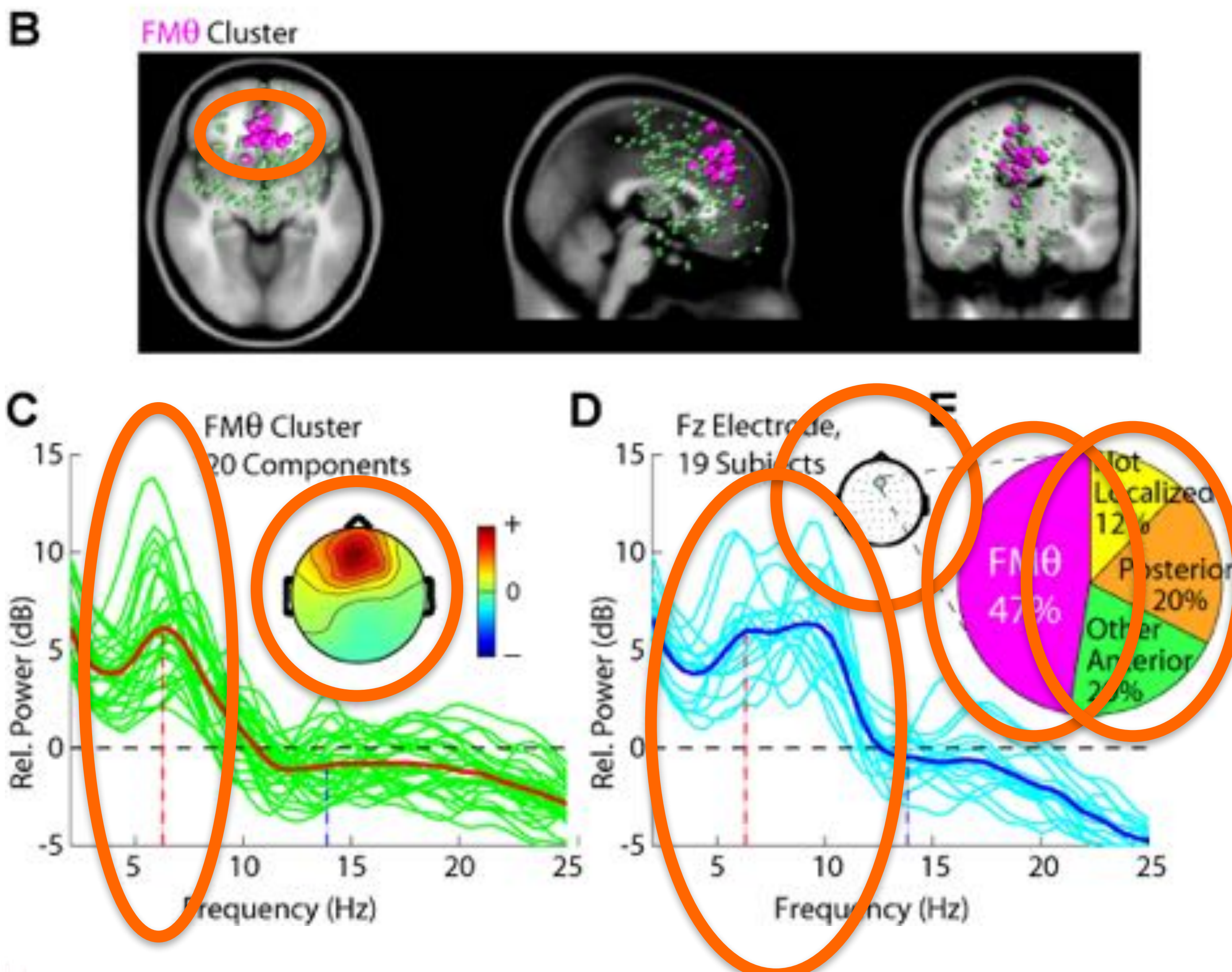
Onton & Makeig, 2005

Are source properties similar across participants?



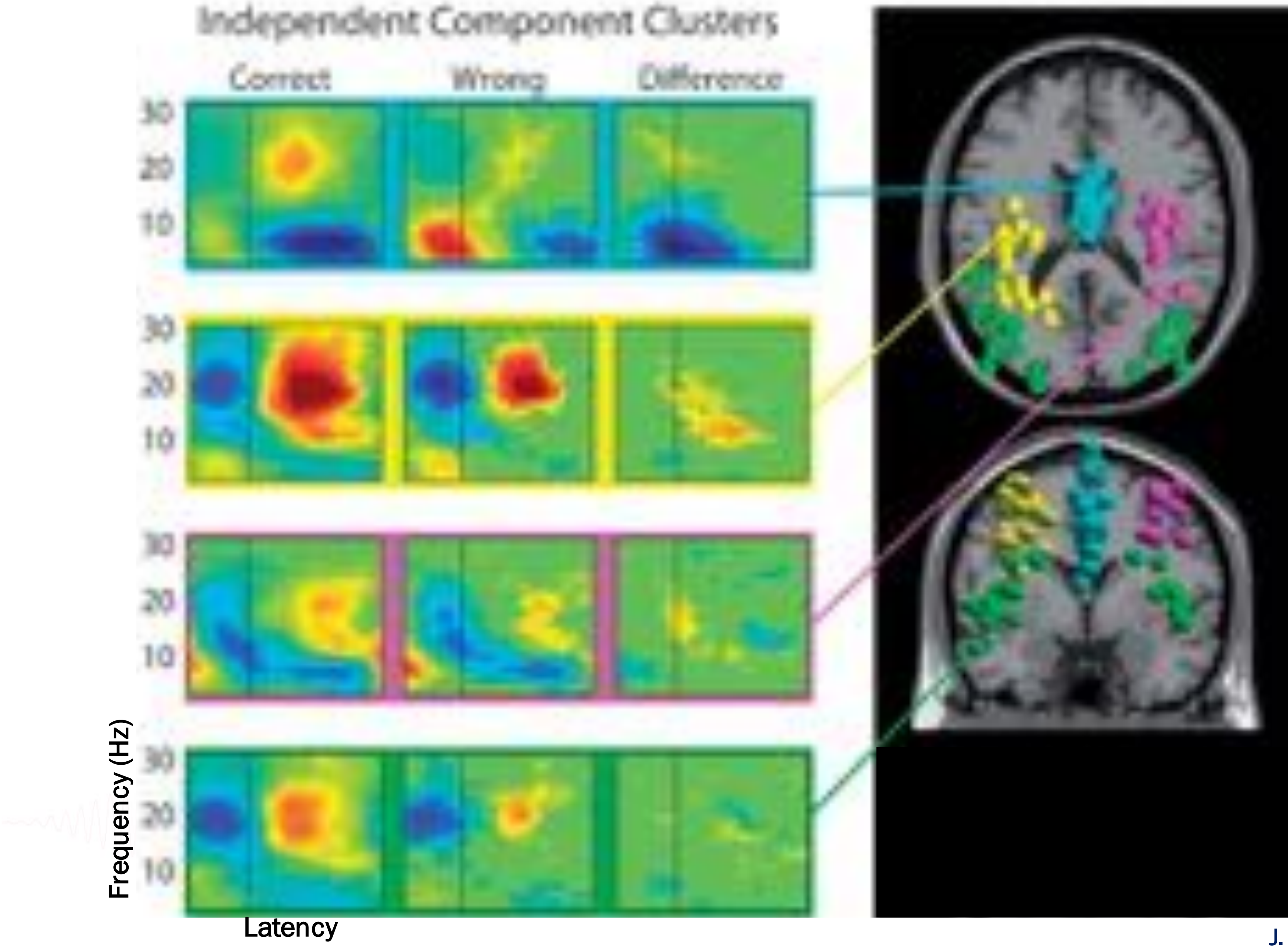
S. Makeig (2007)

Example: frontal midline theta cluster



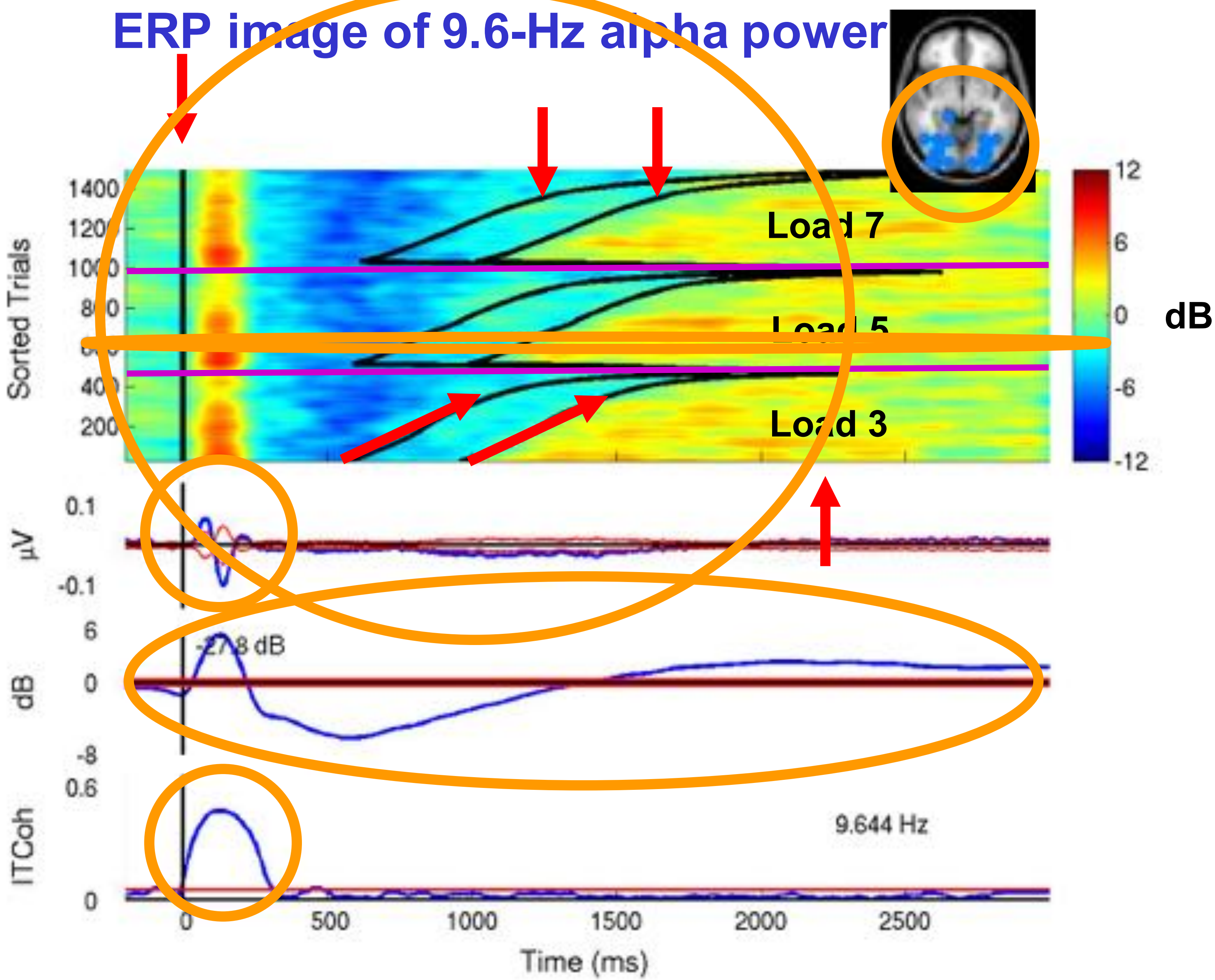
Onton, Delorme & Makeig, *NeuroImage* (2005)

Goal: To cluster equivalent ICs across subjects



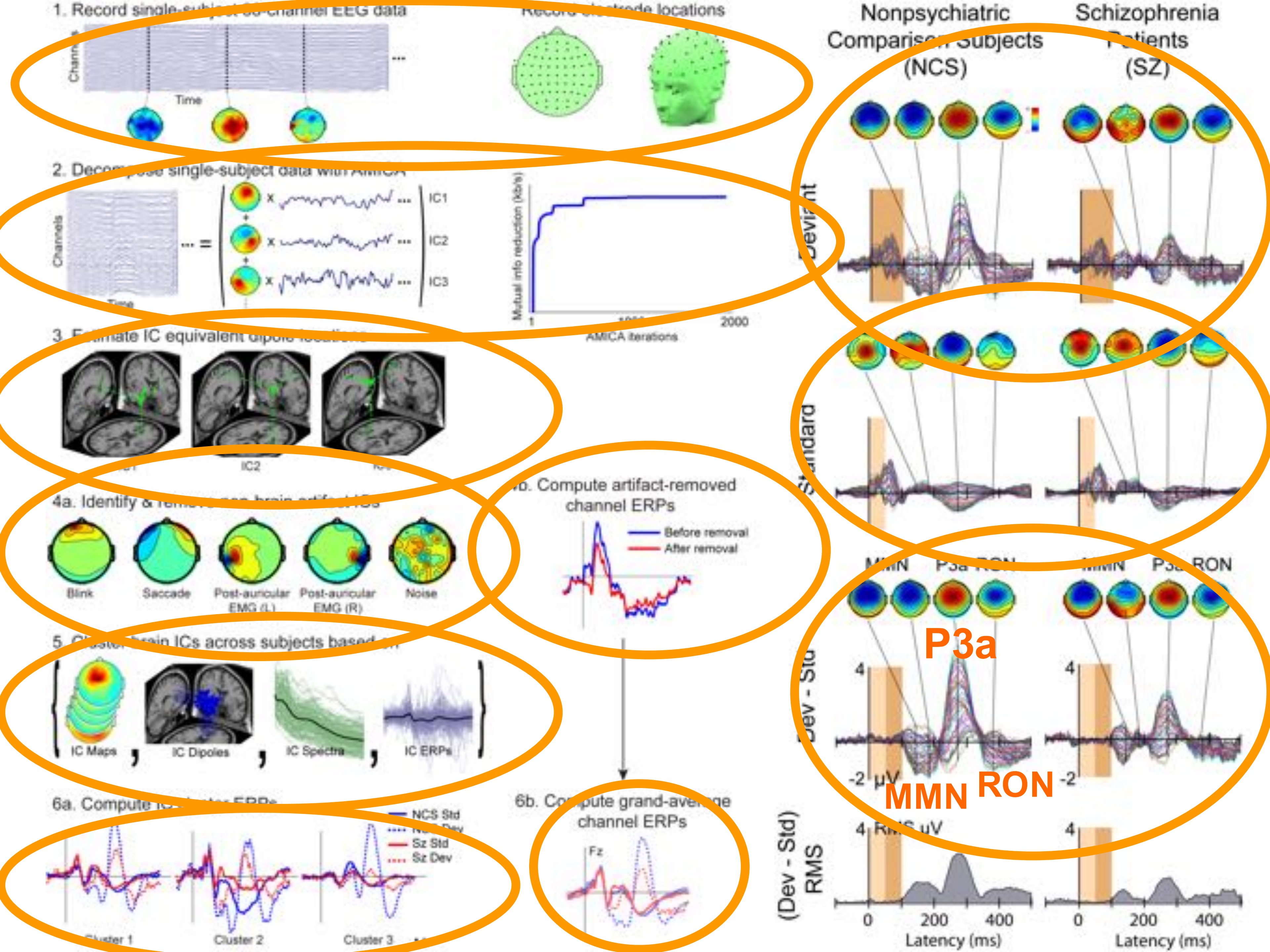
J. Onton, 2005

ERP image of 9.6-Hz alpha power

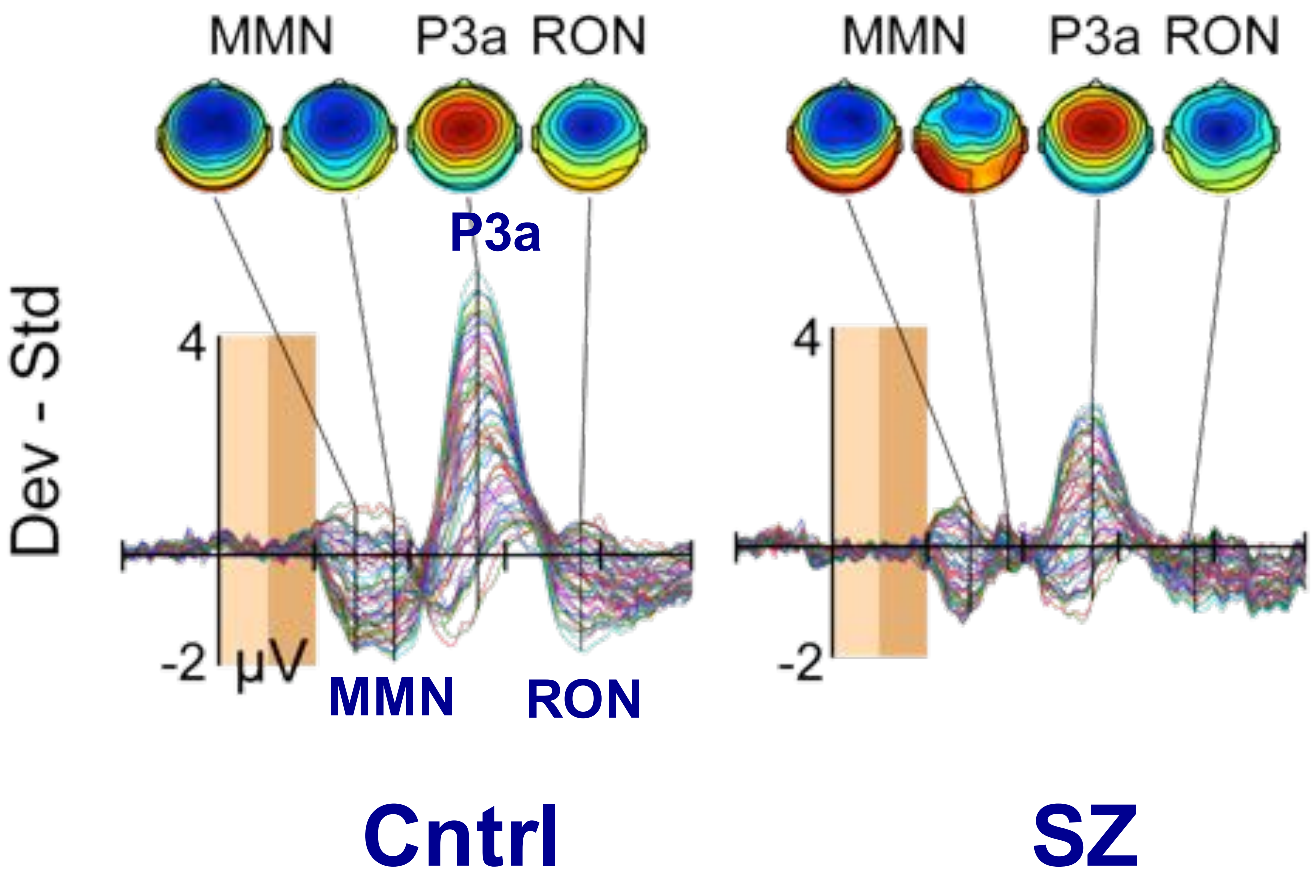


erpimage()

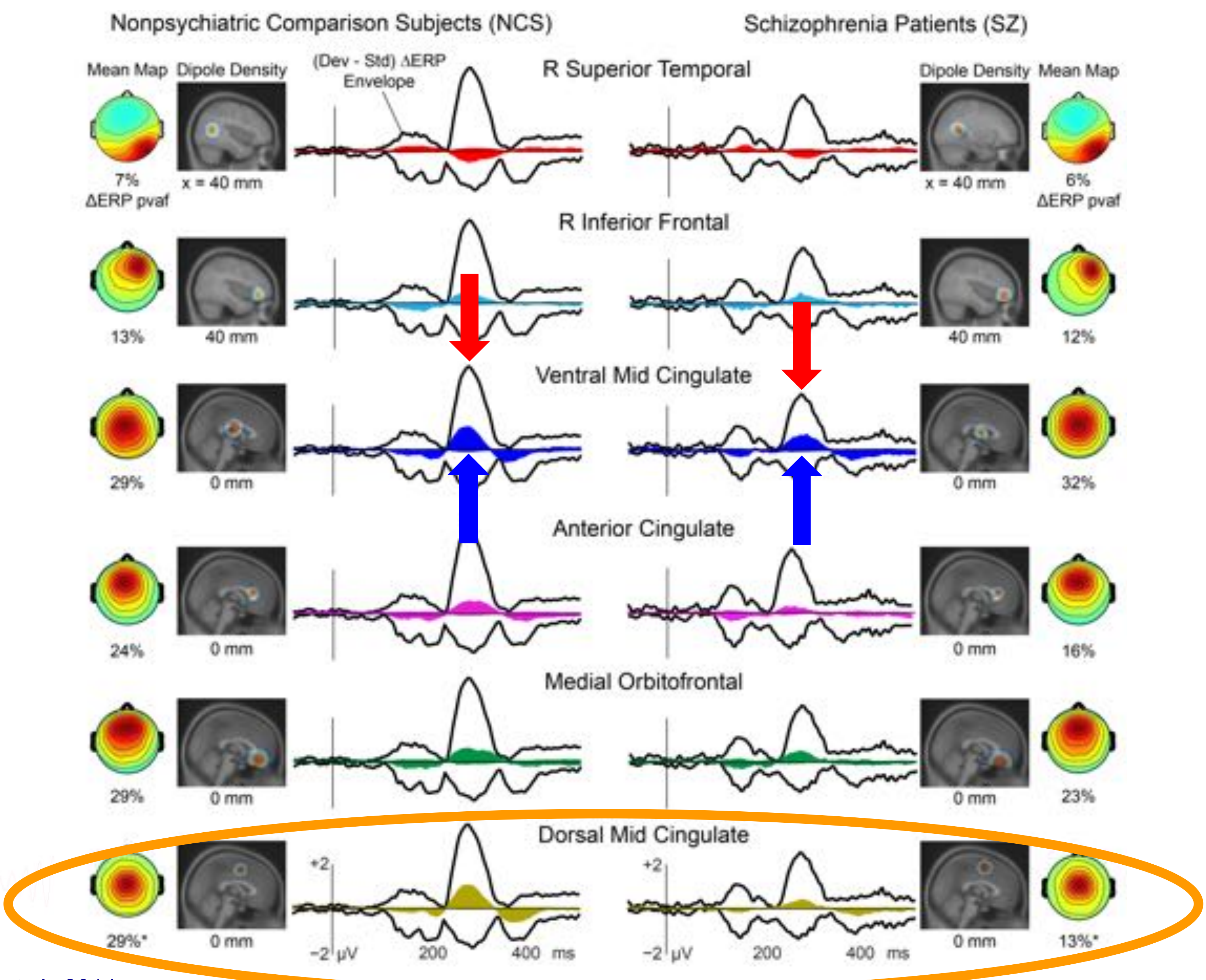
Onton, Delorme & Makeig, 2005.


1. Record single-subject 80-channel EEG data
Record electrode locations
Time
2. Decompose single-subject data with AMICA
Channels
IC1
IC2
IC3
Mutual info reduction (kb/s)
AMICA iterations
1
2000
3. Estimate IC equivalent dipole locations
IC2
4a. Identify & remove non-brain artifact ICs
Blink
Saccade
Post-auricular EMG (L)
Post-auricular EMG (R)
Noise
4b. Compute artifact-removed channel ERPs
Before removal
After removal
5. Cluster brain ICs across subjects based on
IC Maps
IC Dipoles
IC Spectra
IC ERPs
6a. Compute IC cluster ERPs
NCS Std
NCS Dev
Sz Std
Sz Dev
Cluster 1
Cluster 2
Cluster 3
6b. Compute grand-average channel ERPs
Fz
Nonpsychiatric Comparison Subjects (NCS)
Schizophrenia Patients (SZ)
Deviant
Standard
MMN
P3a
RON
Dev - Std
4
-2 μV
-2
P3a
MMN
RON
(Dev - Std) RMS
4 RMS μV
0
200
400
Latency (ms)

Auditory Deviance Response



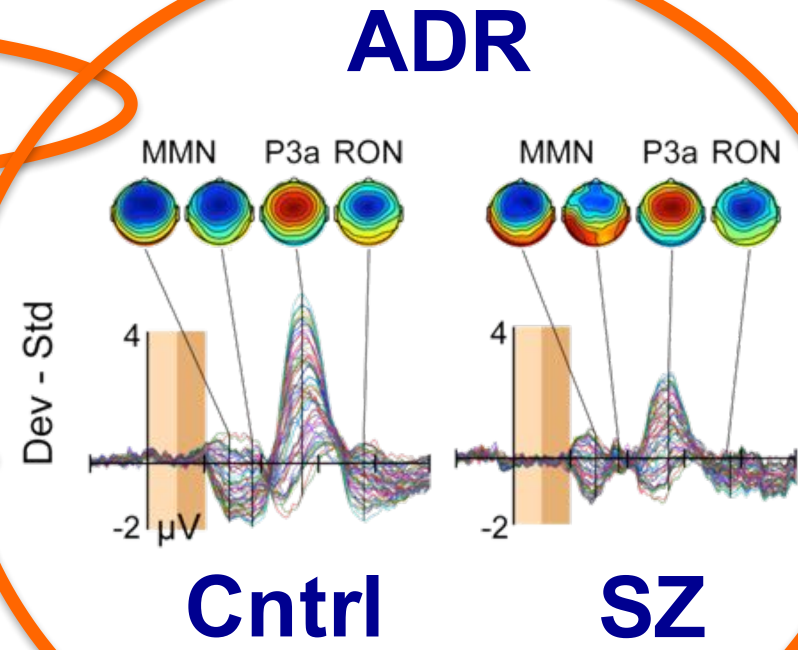
The deepest mental trap in electrophysiology lurks in the word "THE" !!!



Rissling et al., 2014

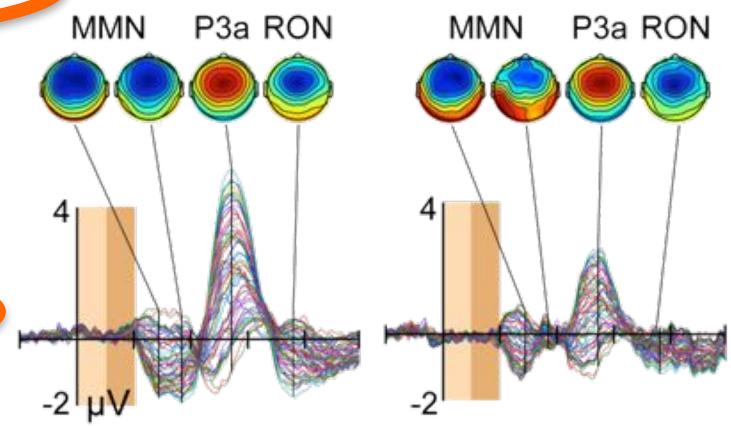
PEAK AMPLITUDES	ERP	r^2
Scalp Electrode (Fz)		
Verbal IQ (WRAT)	P3a	0.11
Functional Capacity (UPSA)	RON	0.12
R Superior Temporal		
Working Memory (LNS Reorder)	RON	0.15
Verbal IQ (WRAT)	RON	0.15
Immediate Verbal Memory (CVLT)	**RON**	**0.28**
Delayed Verbal Memory (CVLT)	RON	0.26
Functional Capacity (UPSA)	**MMN**	**0.48**
Functional Capacity (UPSA)	RON	0.26
R Inferior Frontal		
Negative Symptoms (SANS)	**RON**	**0.36**
Psychosocial Functioning (SOF)	RON	0.24
Auditory Attention (LNS Forward)	**MMN**	**0.38**
Working Memory (LNS Reorder)	**MMN**	**0.30**
Verbal IQ (WRAT)	**MMN**	**0.46**
Ventral Mid Cingulate		
Positive Symptoms (SAPS)	**RON**	**0.29**
Negative Symptoms (SANS)	**P3a**	**0.36**
Immediate Verbal Memory (CVLT)	**RON**	**0.41**
Delayed Verbal Memory (CVLT)	RON	0.24
Verbal IQ (WRAT)	**RON**	**0.29**
Executive Functioning (WCST)	RON	0.24
Anterior Cingulate		
Functional Status (GAF)	MMN	0.18
Functional Status (GAF)	RON	0.17
Immediate Verbal Memory (CVLT)	RON	0.25
Delayed Verbal Memory (CVLT)	RON	0.17
Medial Orbitofrontal		
Positive Symptoms (SAPS)	**P3a**	**0.40**
Negative Symptoms (SANS)	**P3a**	**0.54**
Psychosocial Functioning (SOF)	**P3a**	**0.37**
Functional Capacity (UPSA)	**P3a**	**0.32**
Dorsal Mid Cingulate		
Verbal IQ (WRAT)	P3a	0.15
Executive Functioning (WCST)	MMN	0.18

X



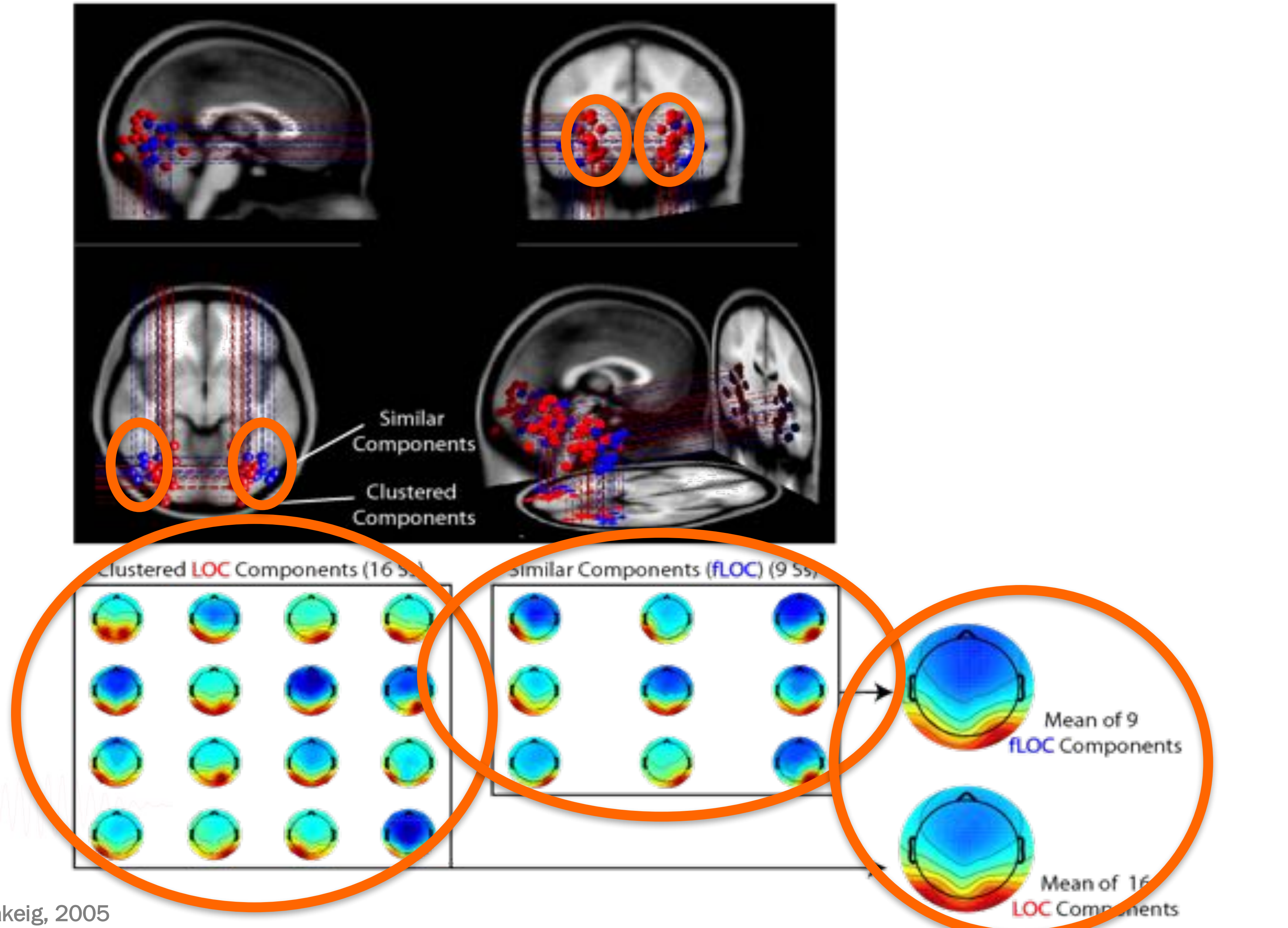
PEAK LATENCIES	ERP	r^2
Scalp Electrode (Fz)		
---n/a---	---	---
R Superior Temporal		
Functional capacity (UPSA)	MMN	0.25
Delayed Verbal Memory (CVLT)	MMN	0.17
R Inferior Frontal		
Negative Symptoms (SANS)	**RON**	**0.51**
Psychosocial Functioning (SOF)	RON	0.25
Executive Functioning (WCST)	**MMN**	**0.30**
Executive Functioning (WCST)	**P3a**	**0.28**
Ventral Mid Cingulate		
Negative Symptoms (SANS)	**P3a**	**0.33**
Negative Symptoms (SANS)	**RON**	**0.33**
Psychosocial Functioning (SOF)	**P3a**	**0.31**
Verbal IQ (WRAT)	MMN	0.25
Executive Functioning (WCST)	**P3a**	**0.30**
Anterior Cingulate		
Functional Capacity (UPSA)	RON	0.17
Verbal IQ (WRAT)	MMN	0.24
Auditory Attention (LNS-Forward)	MMN	0.17
Medial Orbitofrontal		
Negative Symptoms (SANS)	**RON**	**0.41**
Positive Symptoms (SAPS)	**RON**	**0.40**
Auditory Attention (LNS-Forward)	**MMN**	**0.29**
Executive Functioning (WCST)	**P3a**	**0.32**
Dorsal Mid Cingulate		
Negative Symptoms (SANS)	MMN	0.20
Negative Symptoms (SANS)	P3a	0.17
Global Functioning (GAF)	RON	0.24
Functional Capacity (UPSA)	P3a	0.13

ADR



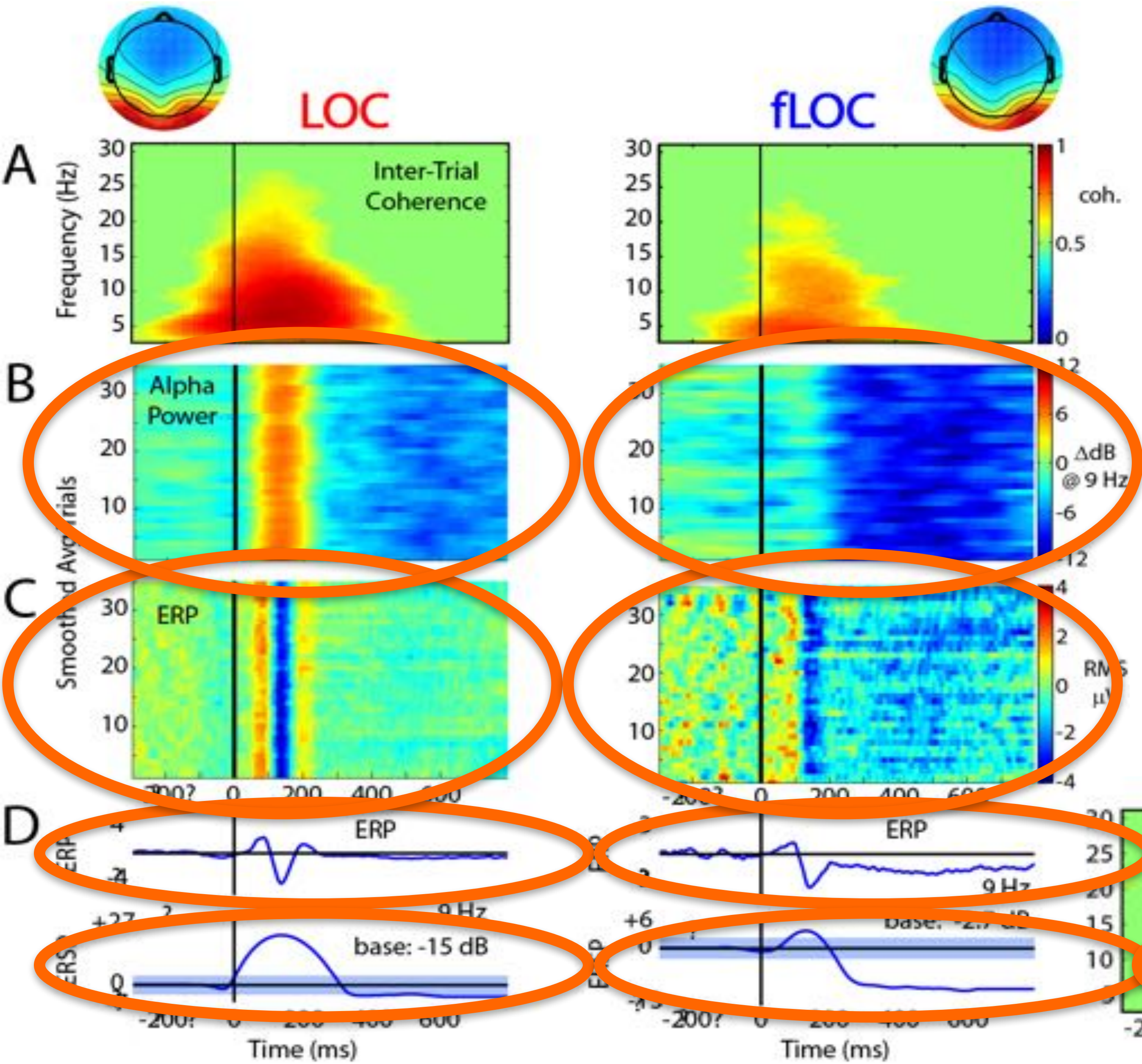
Cntrl **SZ**

Why aren't all participants in every IC cluster?

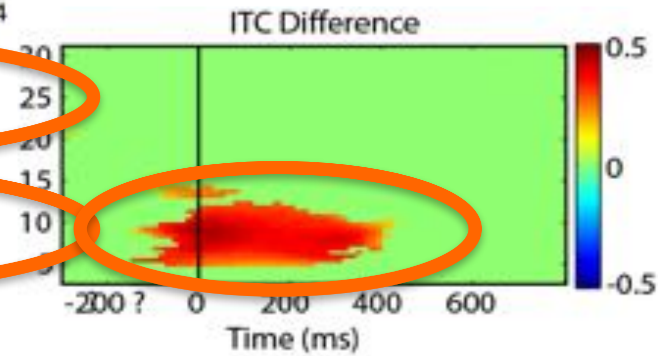


Onton & Makeig, 2005

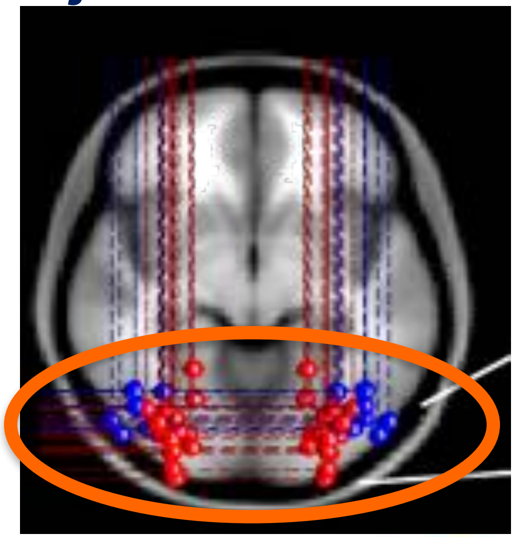
Subject differences?

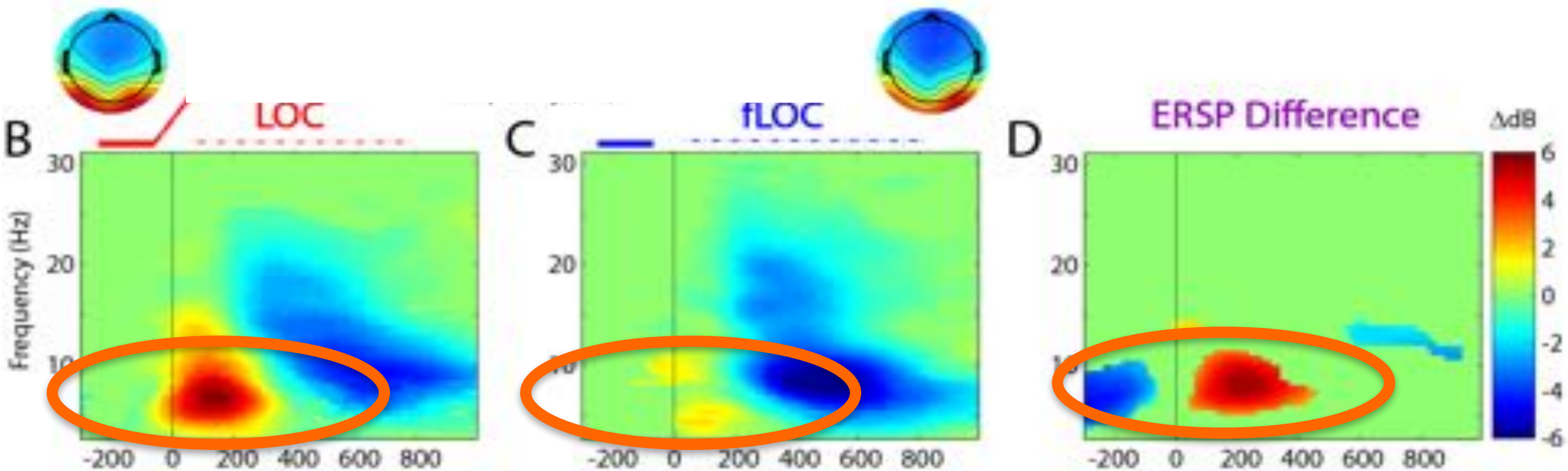


Significant ITC differences (by bootstrap) between the LOC and fLOC clusters immediately follow Probe presentation (5-11 Hz).



Subject differences?

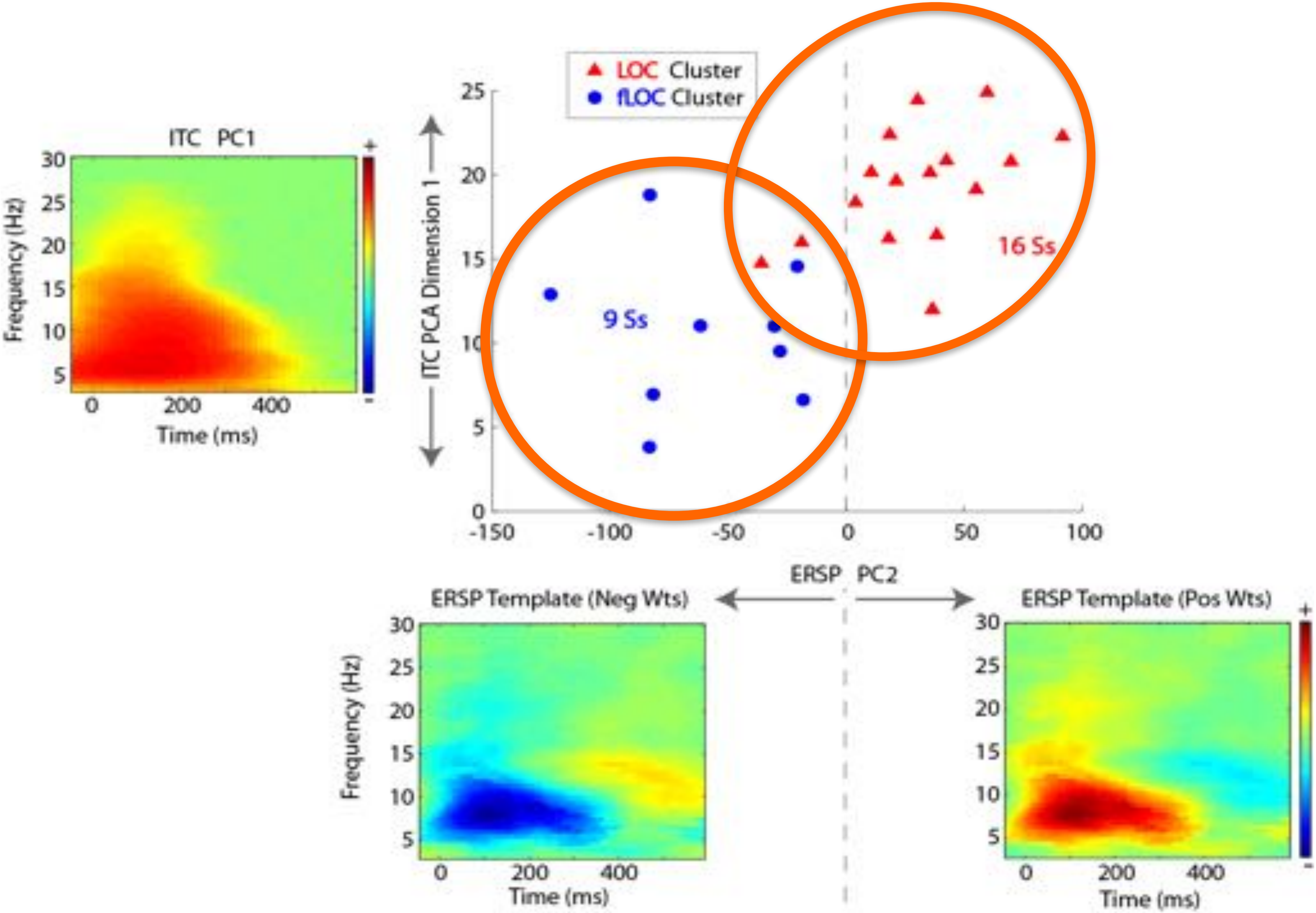




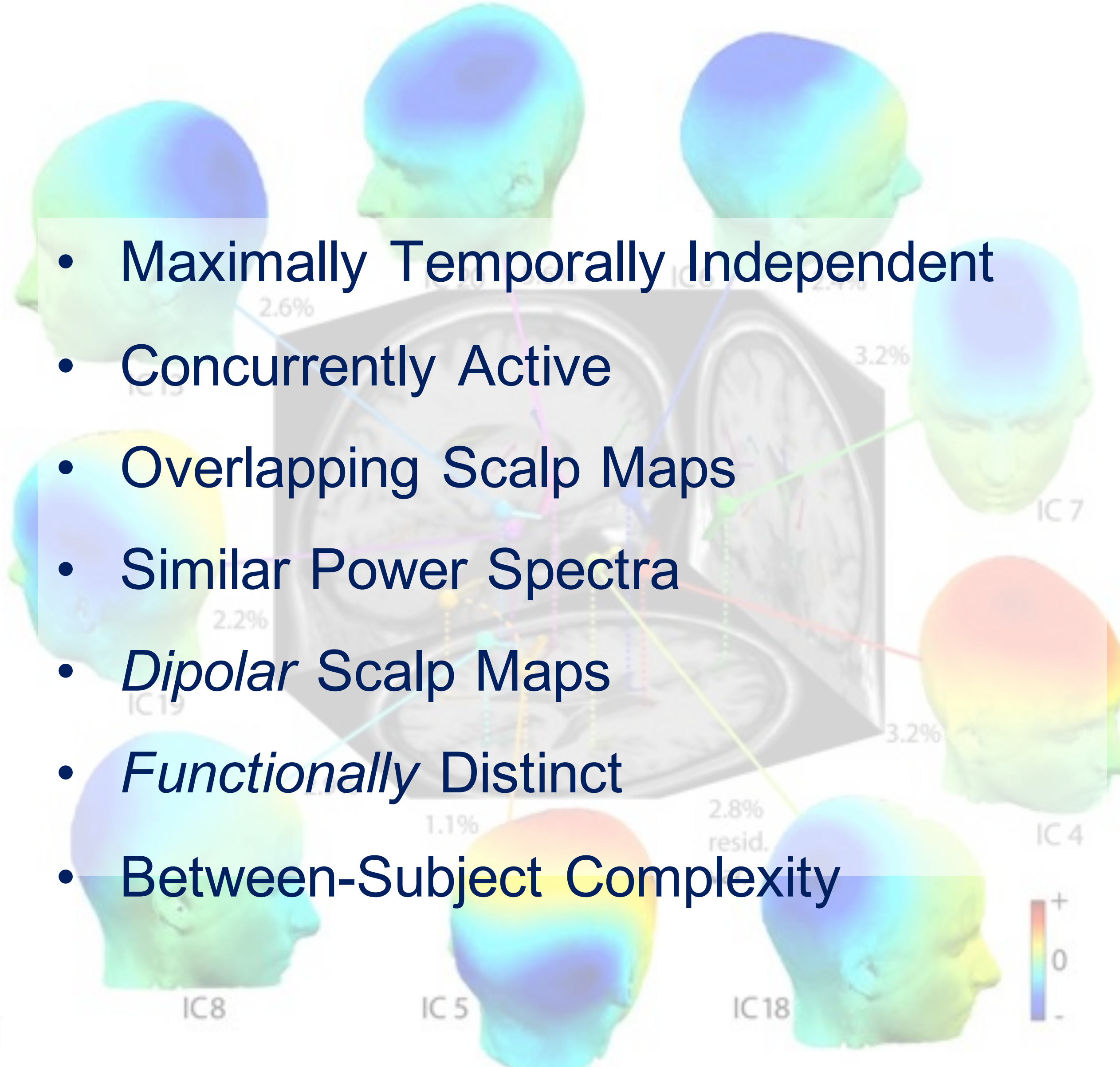
** Cluster ERSPs show significant activity determined by bootstrap statistics within subject and binomial probability between subjects ($p < 0.01$)

*** Difference ERSP shows significant differences between the two clusters by bootstrap statistics ($p < 0.001$)

Subject differences?



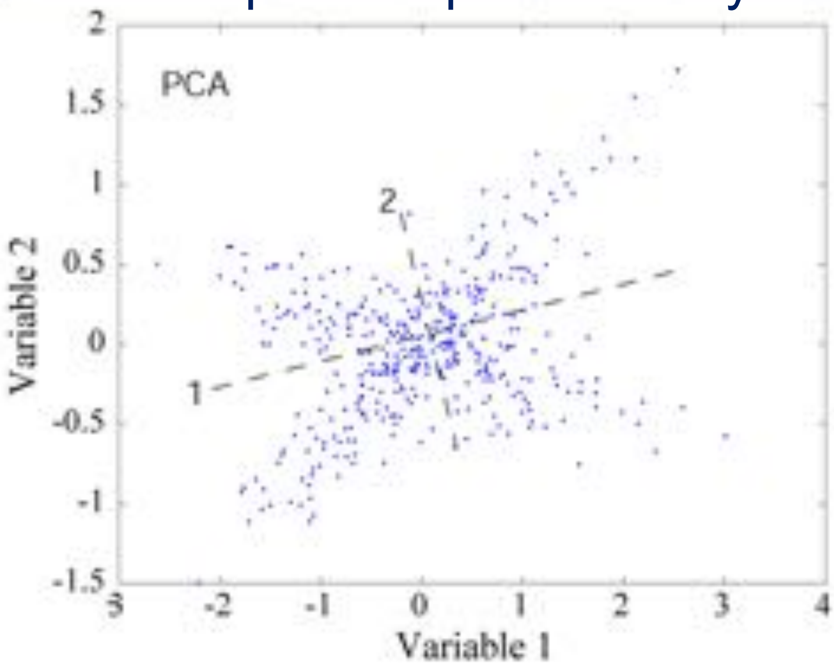
Properties of EEG Independent Components



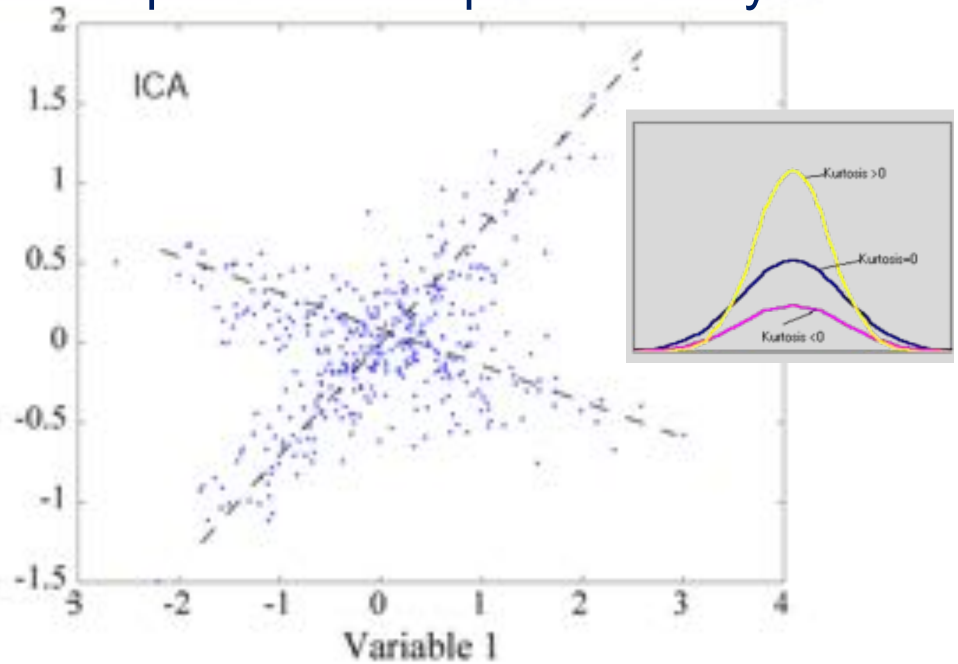
- Maximally Temporally Independent
- Concurrently Active
- Overlapping Scalp Maps
- Similar Power Spectra
- *Dipolar* Scalp Maps
- *Functionally* Distinct
- Between-Subject Complexity

ICA vs. PCA

Principal Component Analysis



Independent Component Analysis



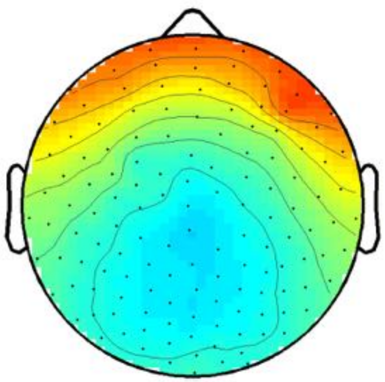
PCA simply decorrelates the outputs using an **orthogonal mixing matrix**.

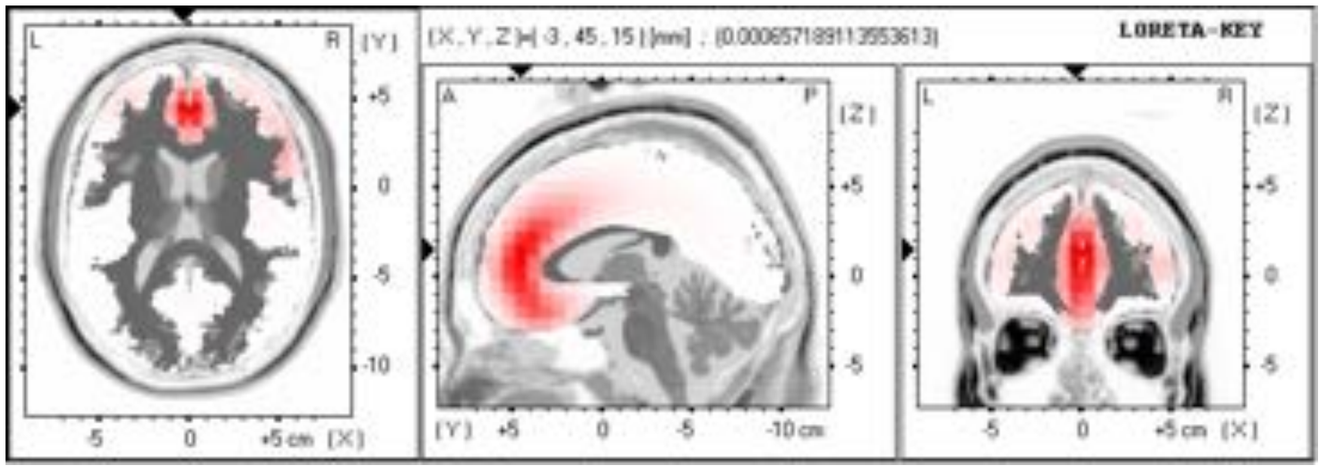
PCA makes each successive component account for as much **variance** in the data as possible.

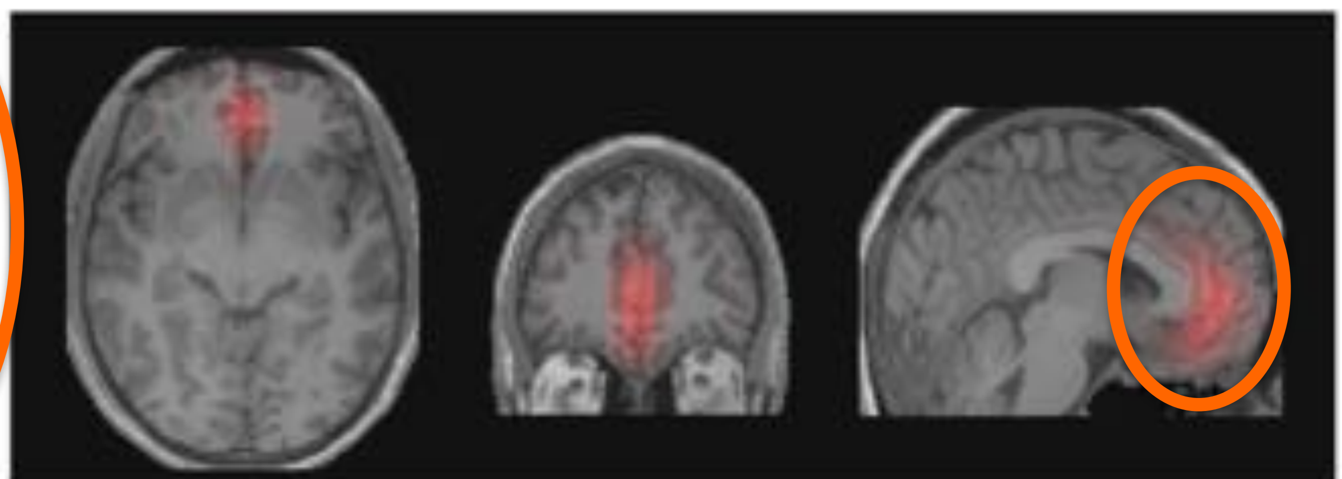
ICA makes each component account for as much **temporally independent information** in the data as possible, with no constraints on the mixing matrix.

PCA lumps – ICA splits!

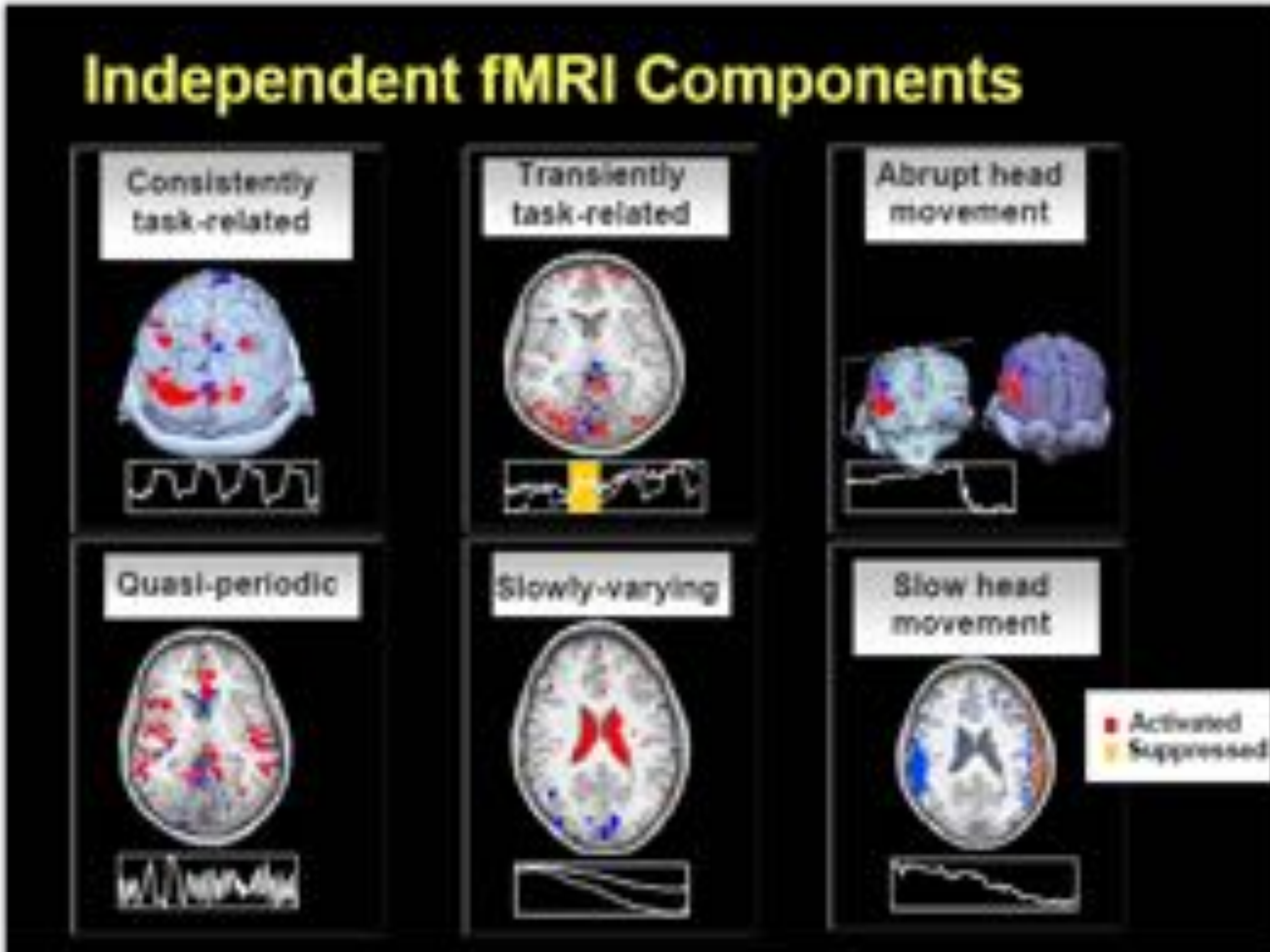
EEG Source Localization







A. Delorme


Independent fMRI Components
Consistently task-related
Transiently task-related
Abrupt head movement
Quasi-periodic
Slowly-varying
Slow head movement
Activated
Suppressed

More …