Independent Component Analysis of Electrophysiological Data

Scott Makeig

Institute for Neural Computation

University of California San Diego

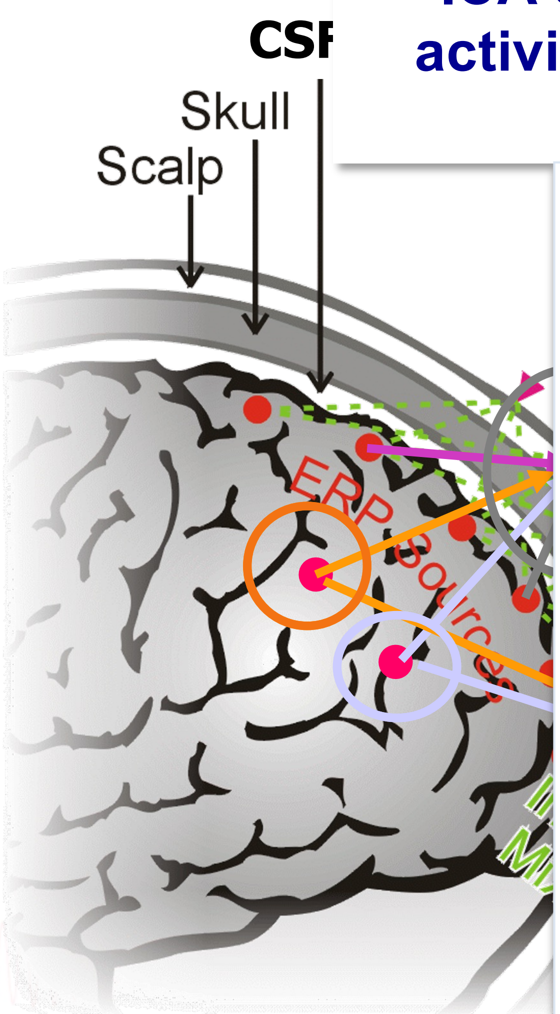
28th EEGLAB Workshop

November, 2018

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
Naval Health Research Center
P.O. Box 85122
San Diego CA 92186-5122
scott@cpl.mag.nhrc.navy.mil

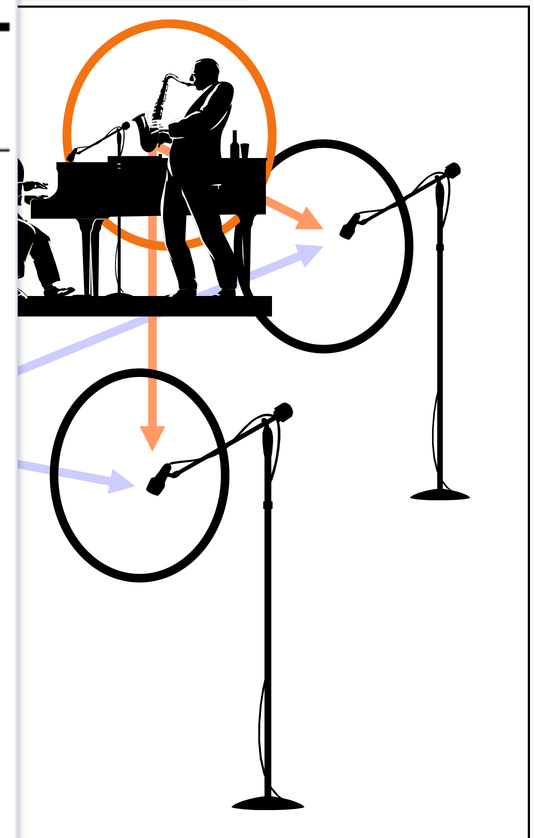
Anthony J. Bell
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
tony@salk.edu

Tzyy-Ping Jung
Naval Health Research Center and
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
jung@salk.edu

Terrence J. Sejnowski
Howard Hughes Medical Institute and
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
terry@salk.edu

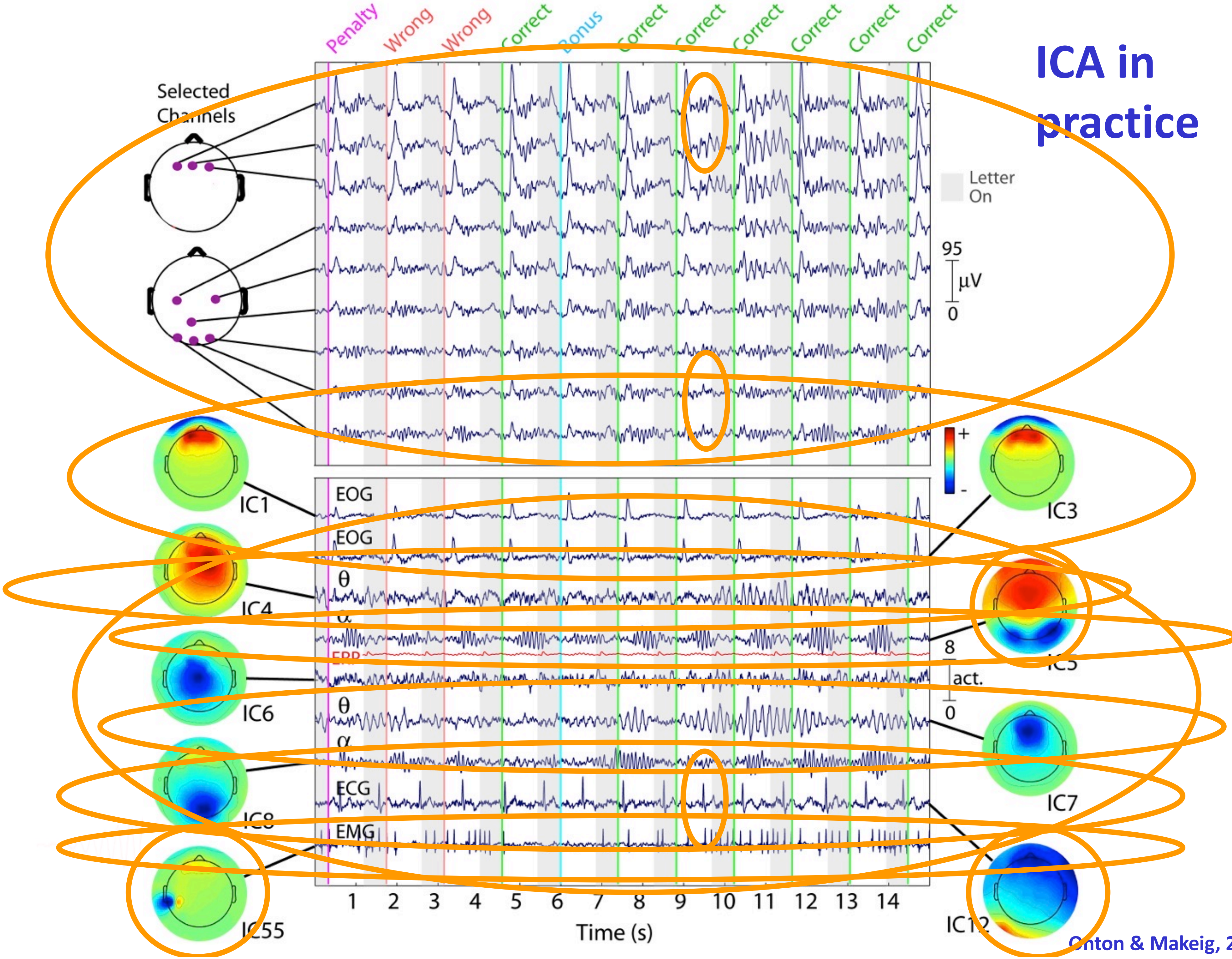
Abstract

Because of the distance between the skull and brain and their different resistivities, *electroencephalographic* (EEG) data collected from any point on the human scalp includes activity generated within a large brain area. This spatial smearing of EEG data by volume conduction does not involve significant time delays, however, suggesting that the *Independent Component Analysis* (ICA) algorithm of Bell and Sejnowski [1] is suitable for performing blind source separation on EEG data. The ICA algorithm separates the problem of source identification from that of source localization. First results of applying the ICA algorithm to EEG and *event-related potential* (ERP) data collected during a sustained auditory detection task show: (1) ICA training is insensitive to different random seeds. (2) ICA may be used to segregate obvious artifactual EEG components (line and muscle noise, eye movements) from other sources. (3) ICA is capable of isolating overlapping EEG phenomena, including alpha and theta bursts and spatially-separable ERP components, to separate ICA channels. (4) Nonstationarities in EEG and behavioral state can be tracked using ICA via changes in the amount of residual correlation between ICA-filtered output channels.

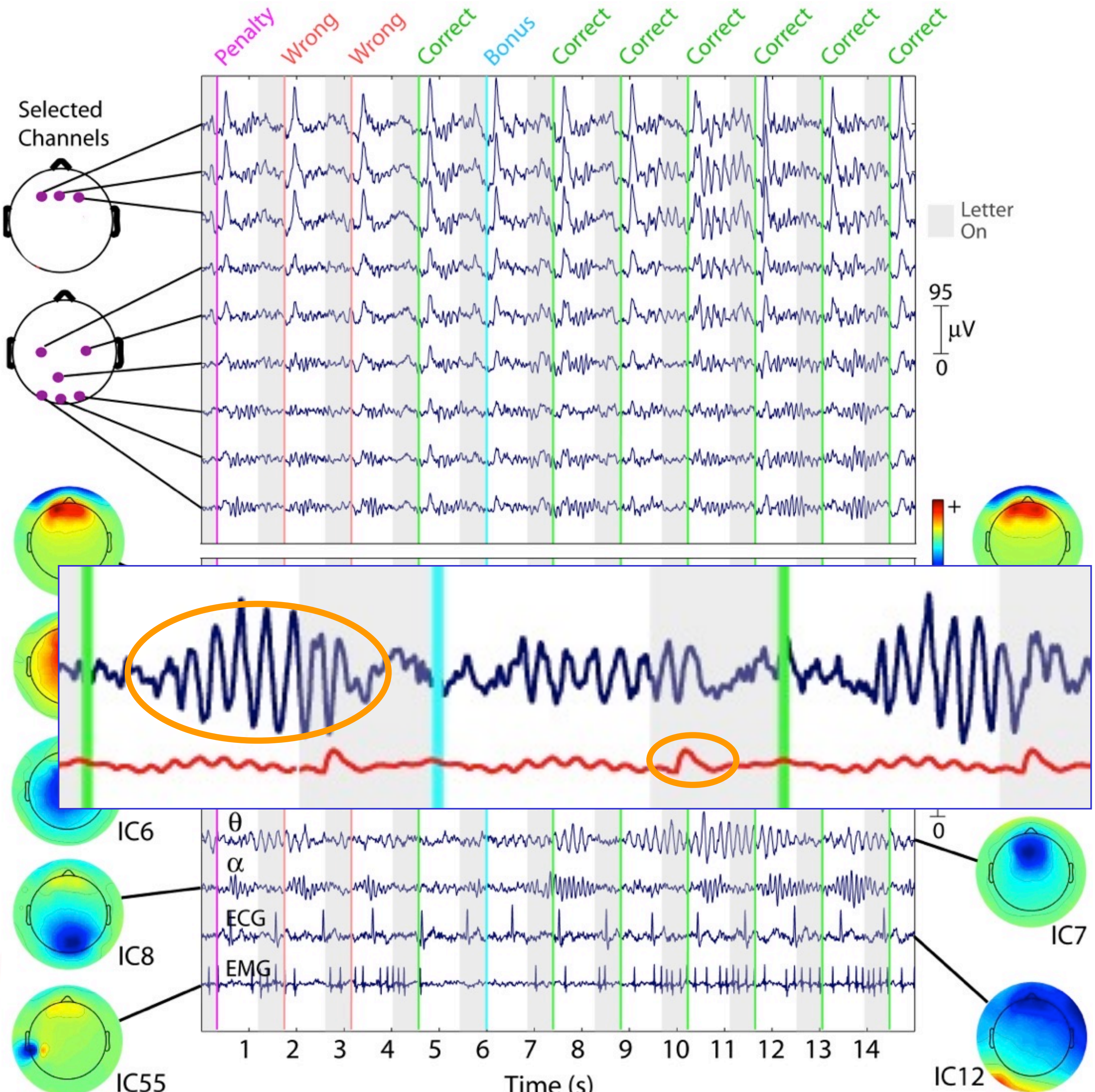


S. Makeig, S. Enghoff (2000)

ICA in practice



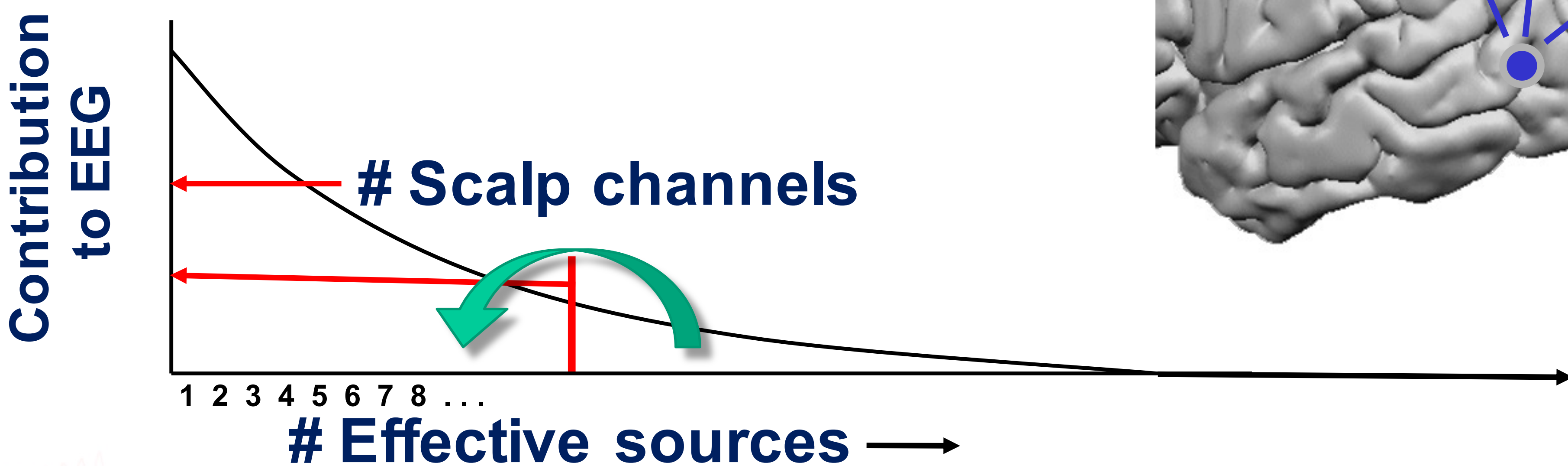
Onton & Makeig, 2006

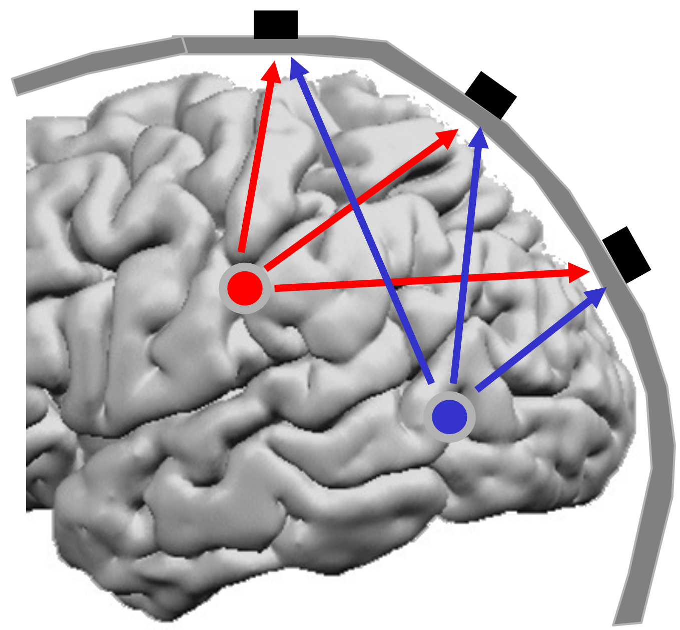


Onton, Makeig (2006)

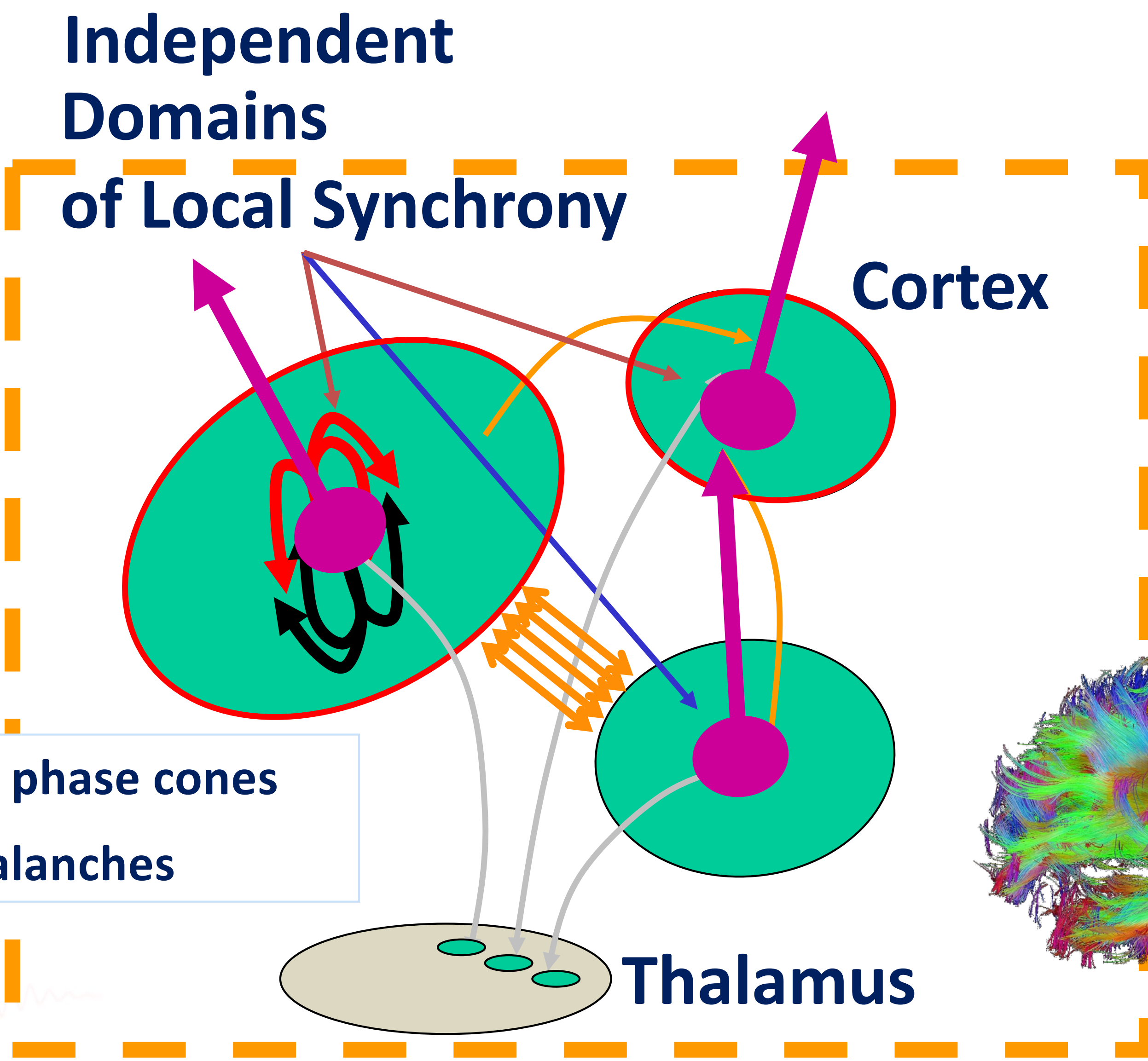
ICA Assumptions

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



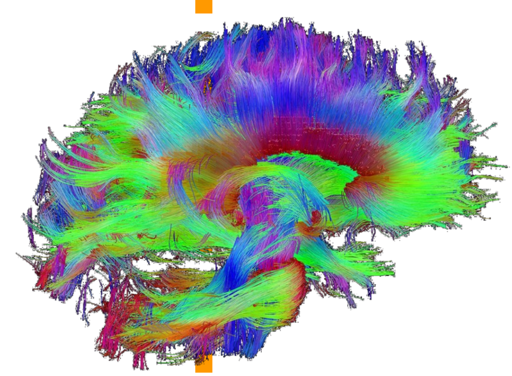


Are EEG effective source signals independent?



Freeman - phase cones

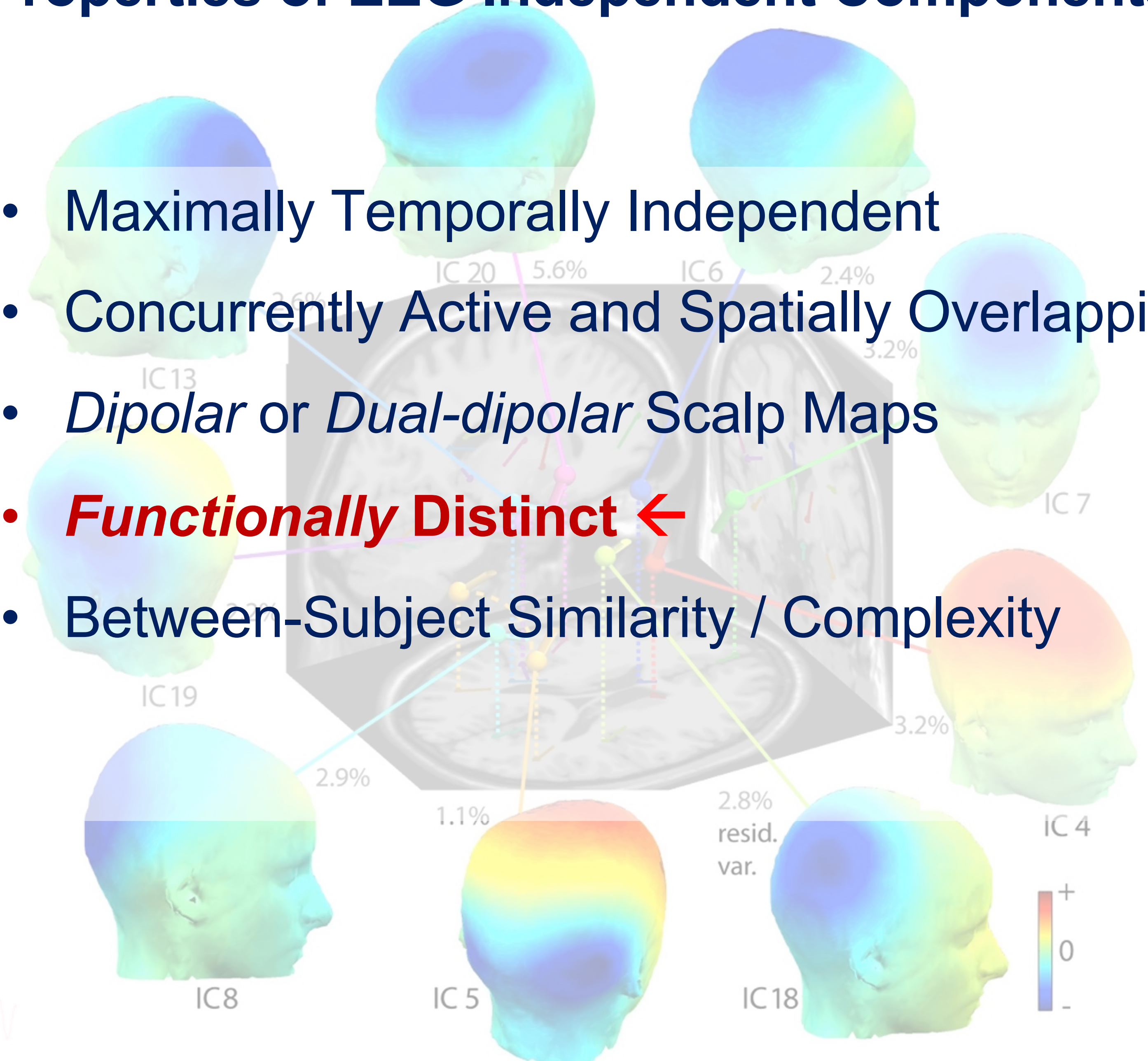
Plenz - avalanches



S. Makeig (2007)

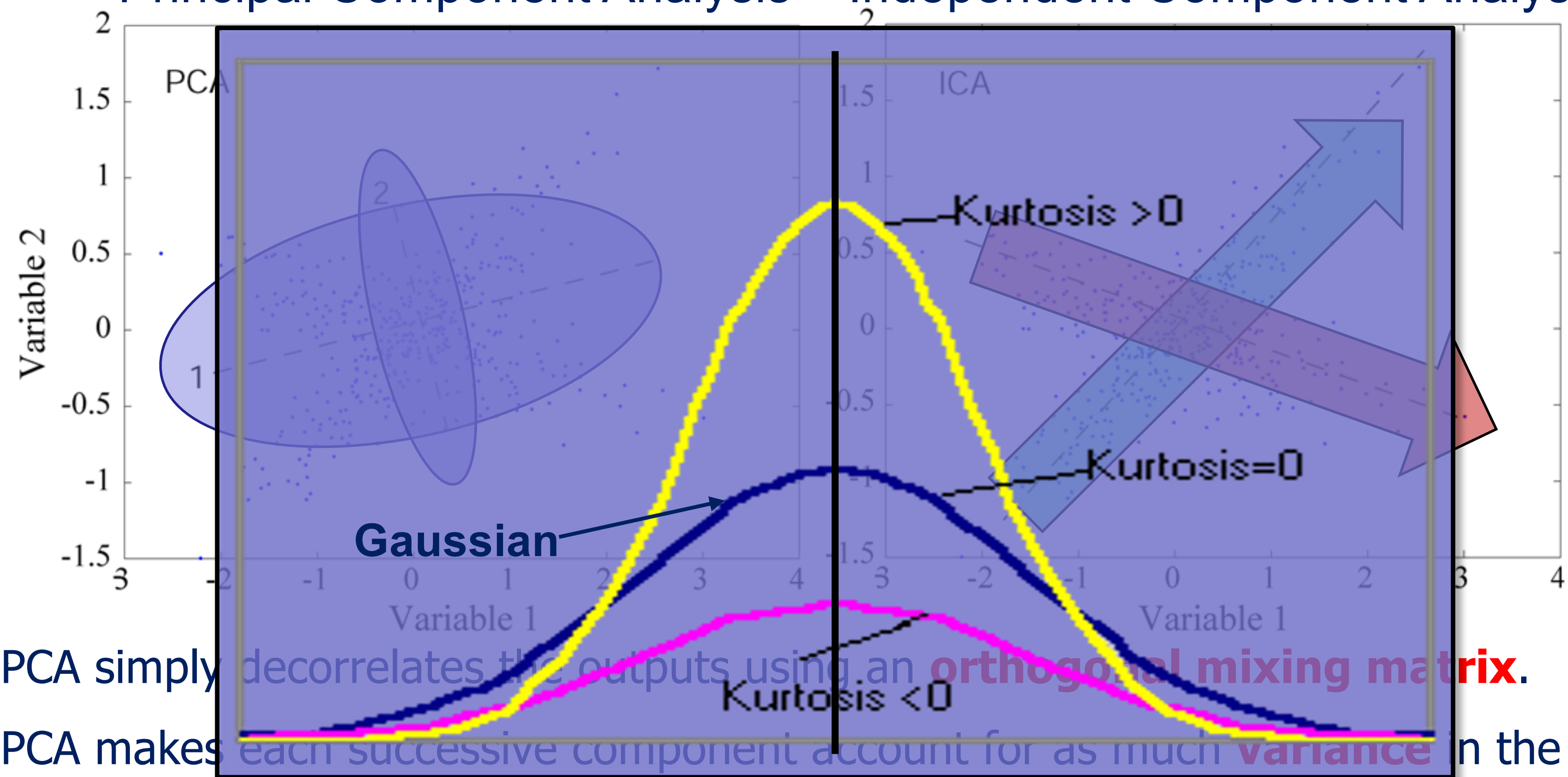
Properties of EEG Independent Components

- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar* or *Dual-dipolar* Scalp Maps
- ***Functionally* Distinct ←**
- Between-Subject Similarity / Complexity



ICA vs. PCA

Principal Component Analysis   Independent Component Analysis



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

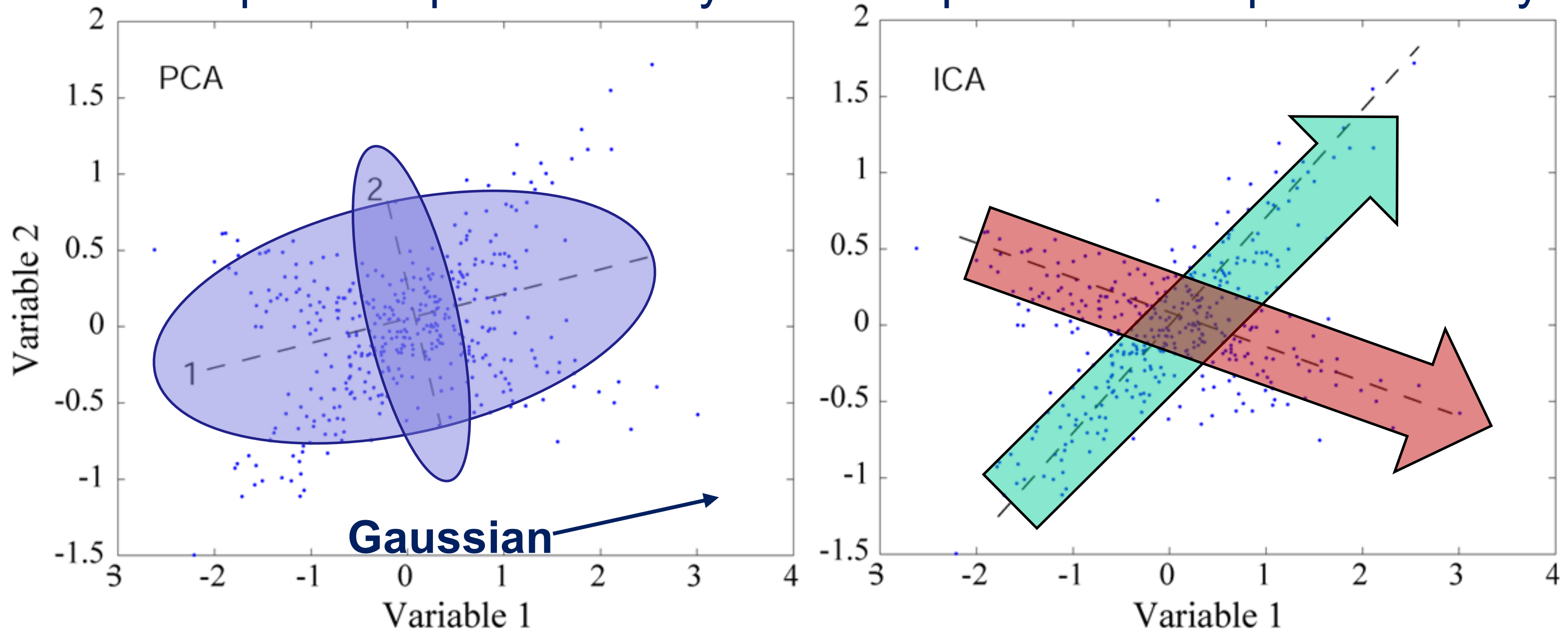
PCA makes each successive component account for as much **variance** in the data as possible.

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!

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.

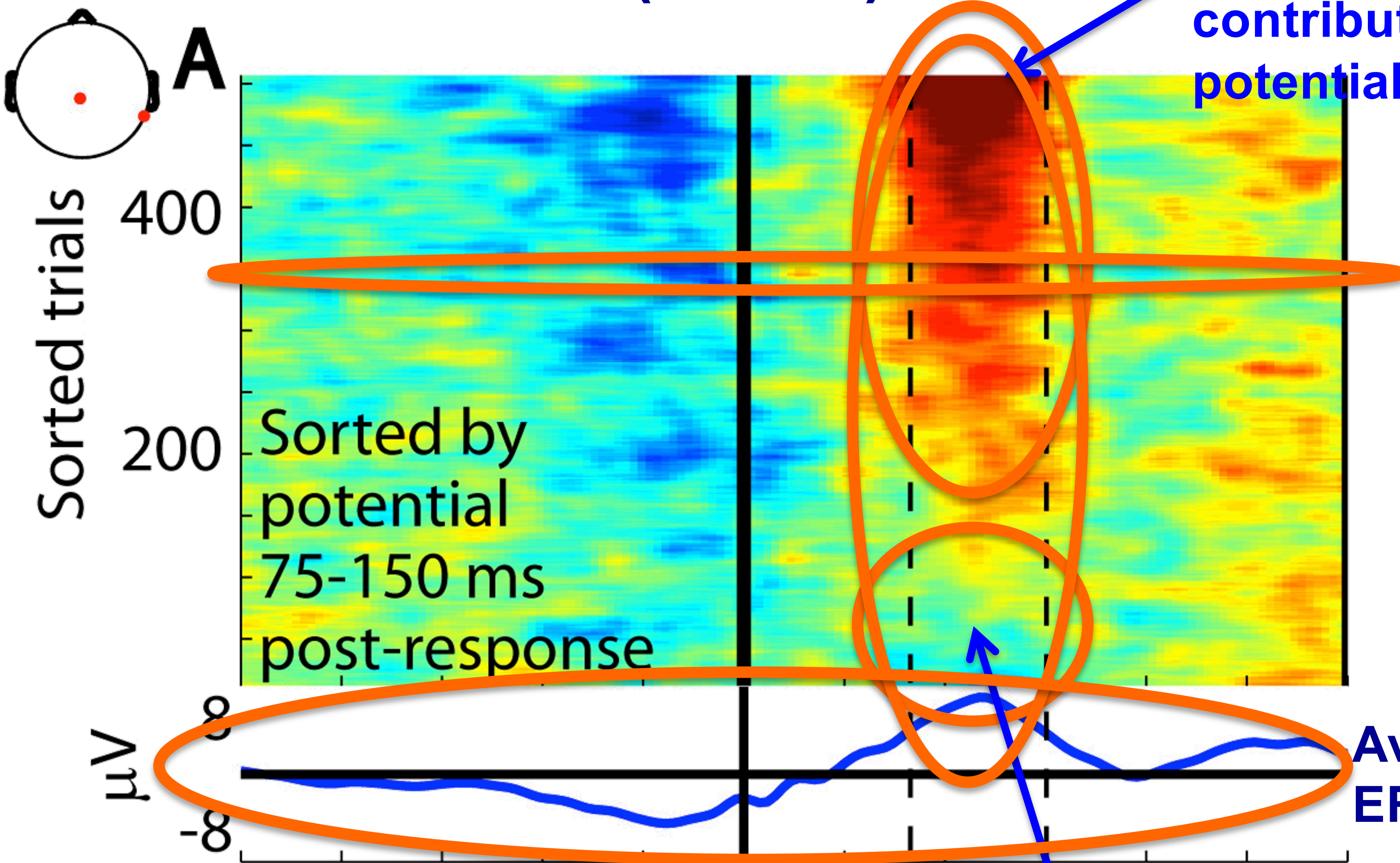
ICA makes each component account for as much **temporally independent information** in the data as possible, with no constraints on the mixing matrix.

PCA lumps – ICA splits!

Arnaud Delorme & S. Makeig, 2016

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

What sources contribute to this potential?

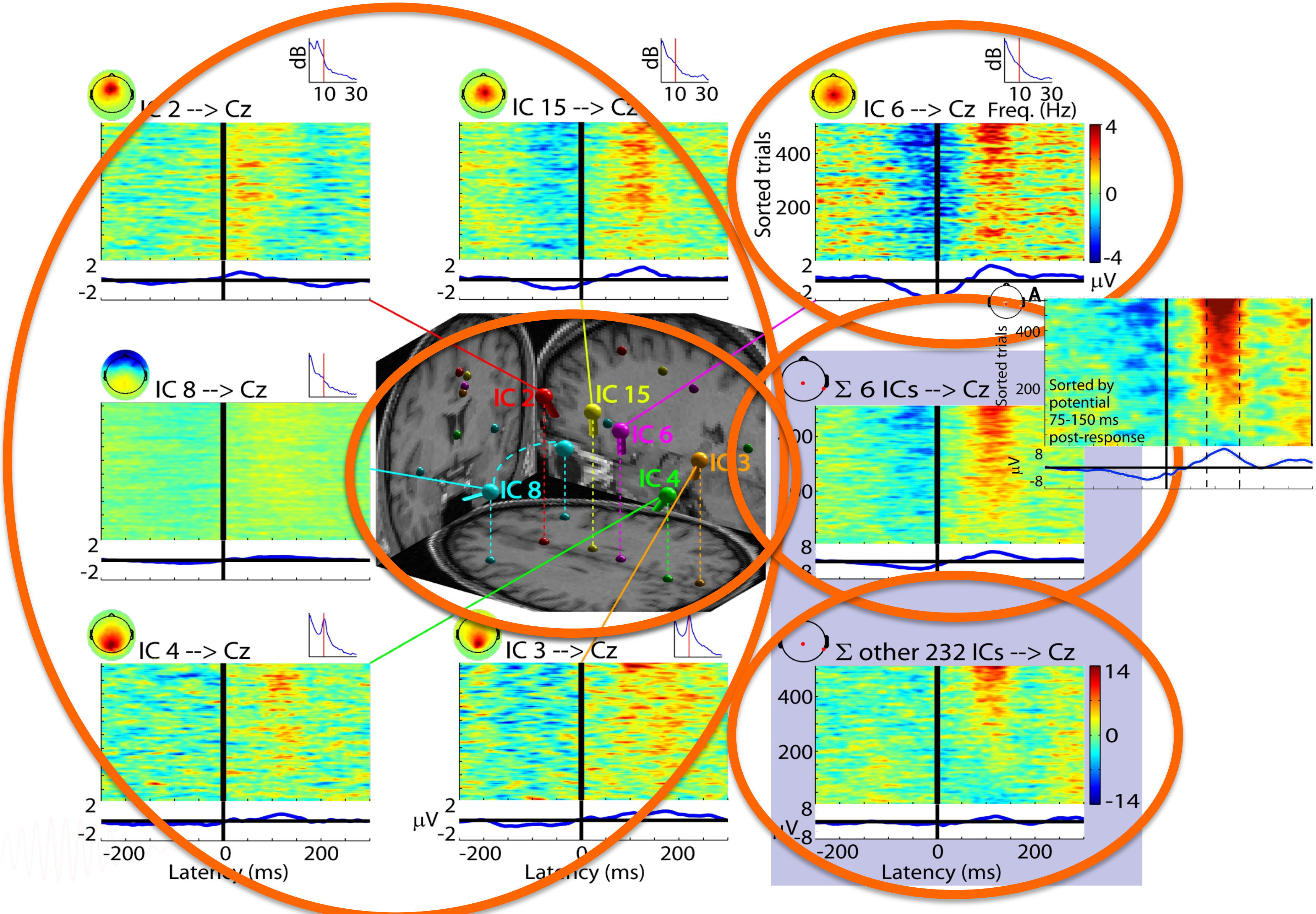


Average ERP

No scalp response in these trials … Why not?

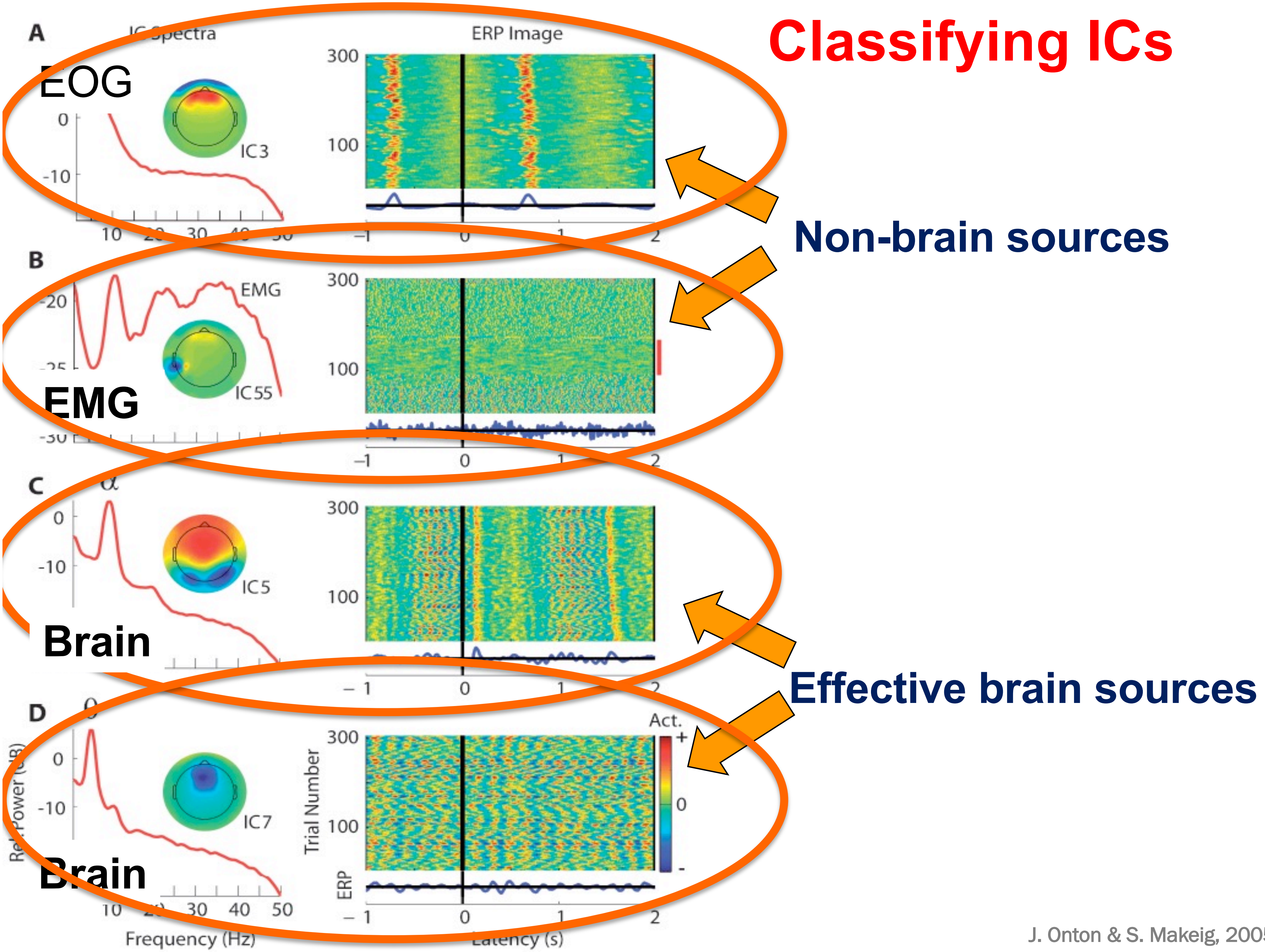
J. Onton & S. Makeig (2009)

The response (at Cz) sums 238 independent sources



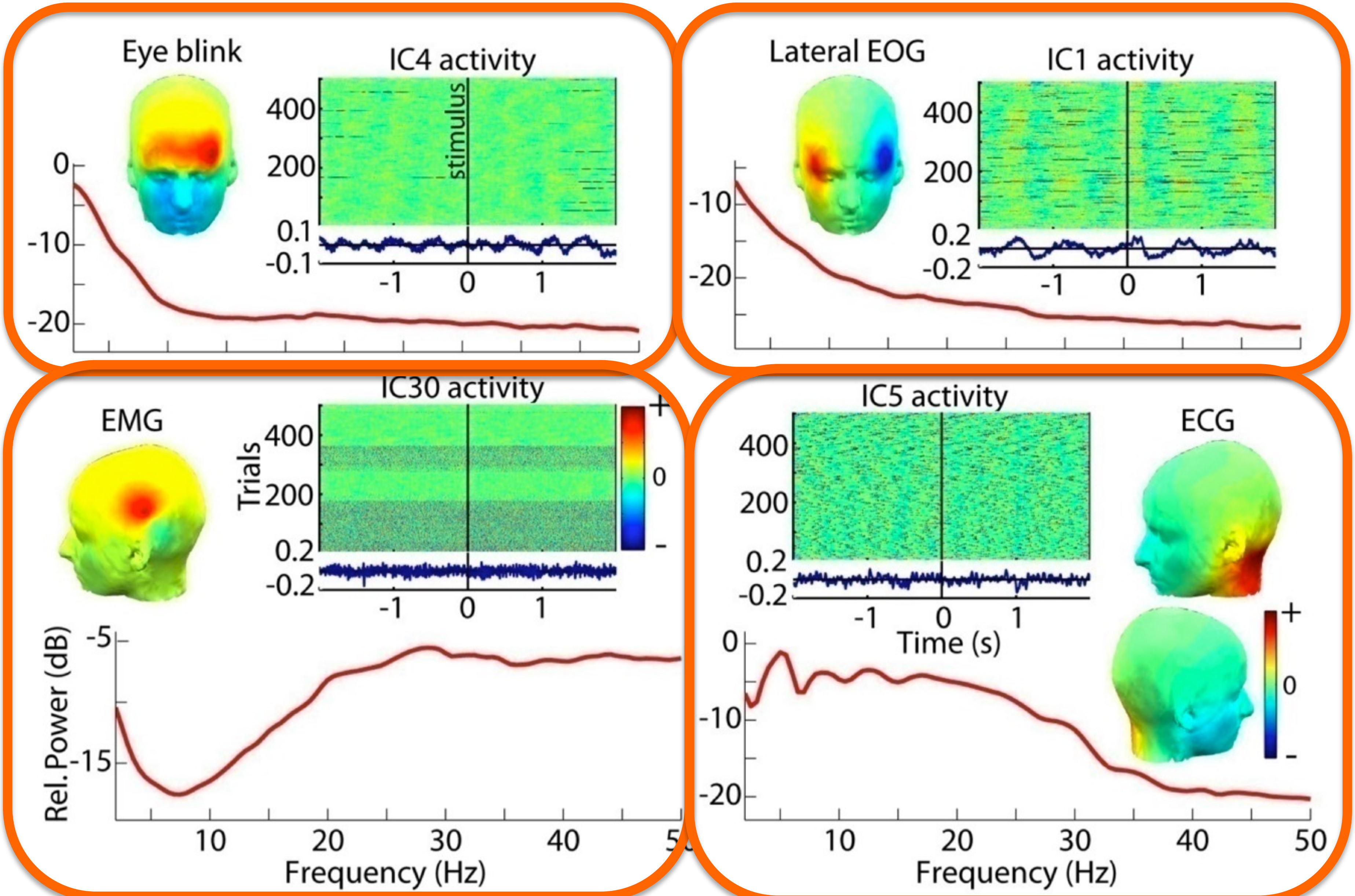
J. Onton & S. Makeig (2009)

Classifying ICs



J. Onton & S. Makeig, 2005

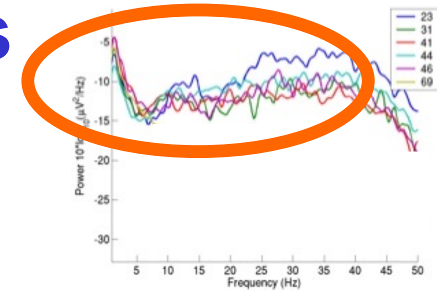
ICA finds Non-Brain Independent Component (IC) Processes …

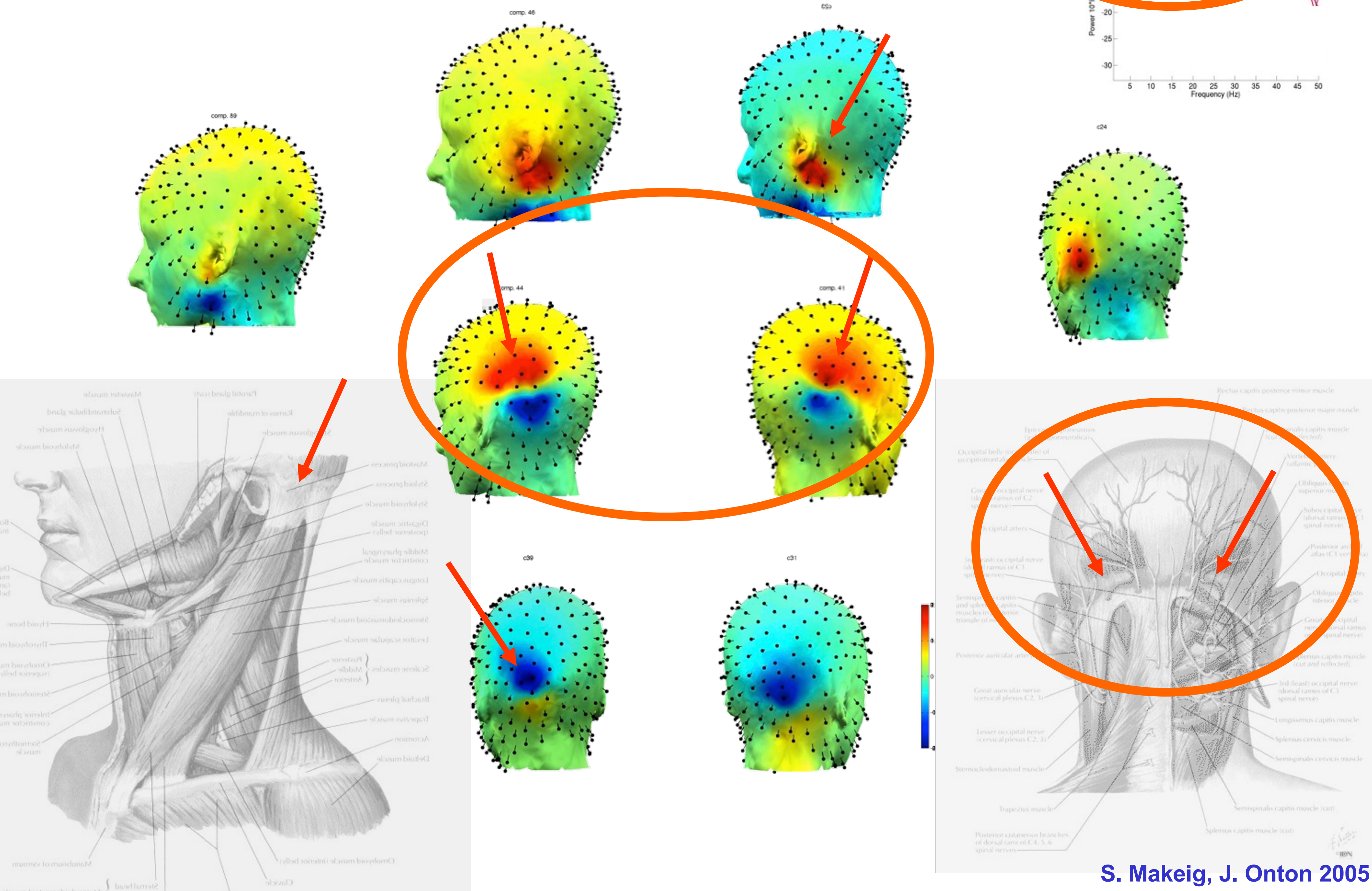


… separates them from the remainder of the data …

J. Onton & S. Makeig

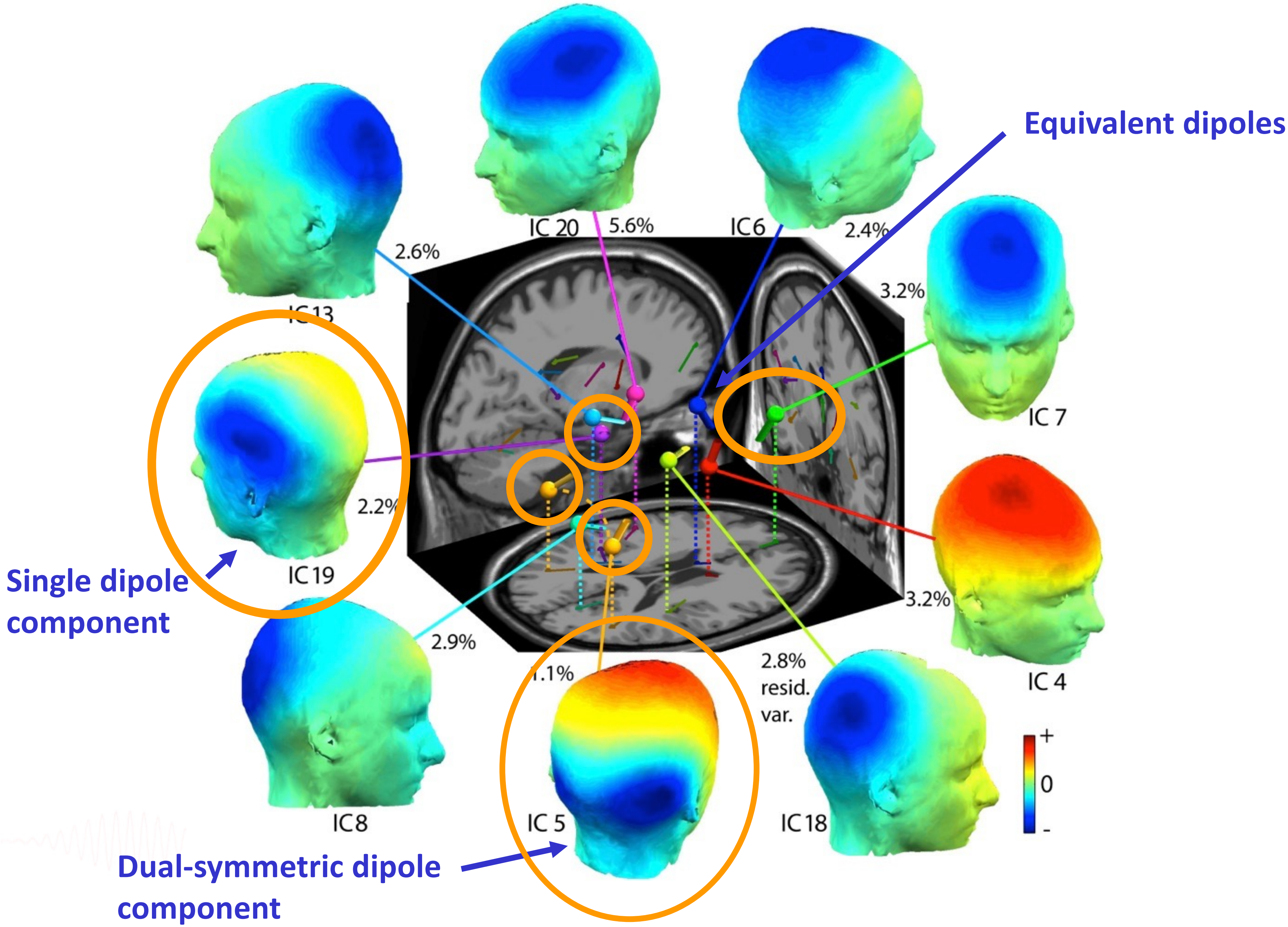
Independent muscle signals





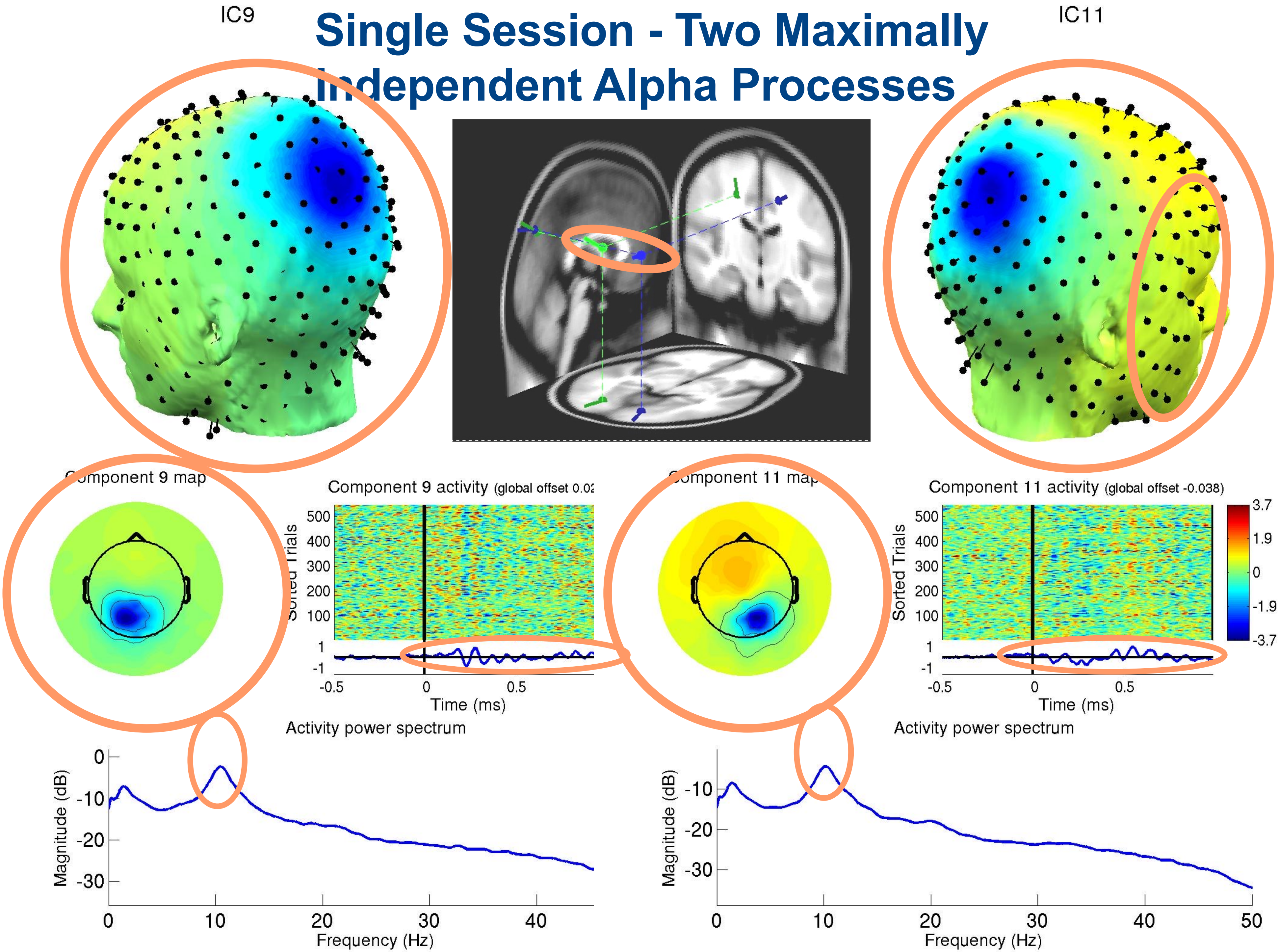
S. Makeig, J. Onton 2005

… and also separates cortical brain IC processes

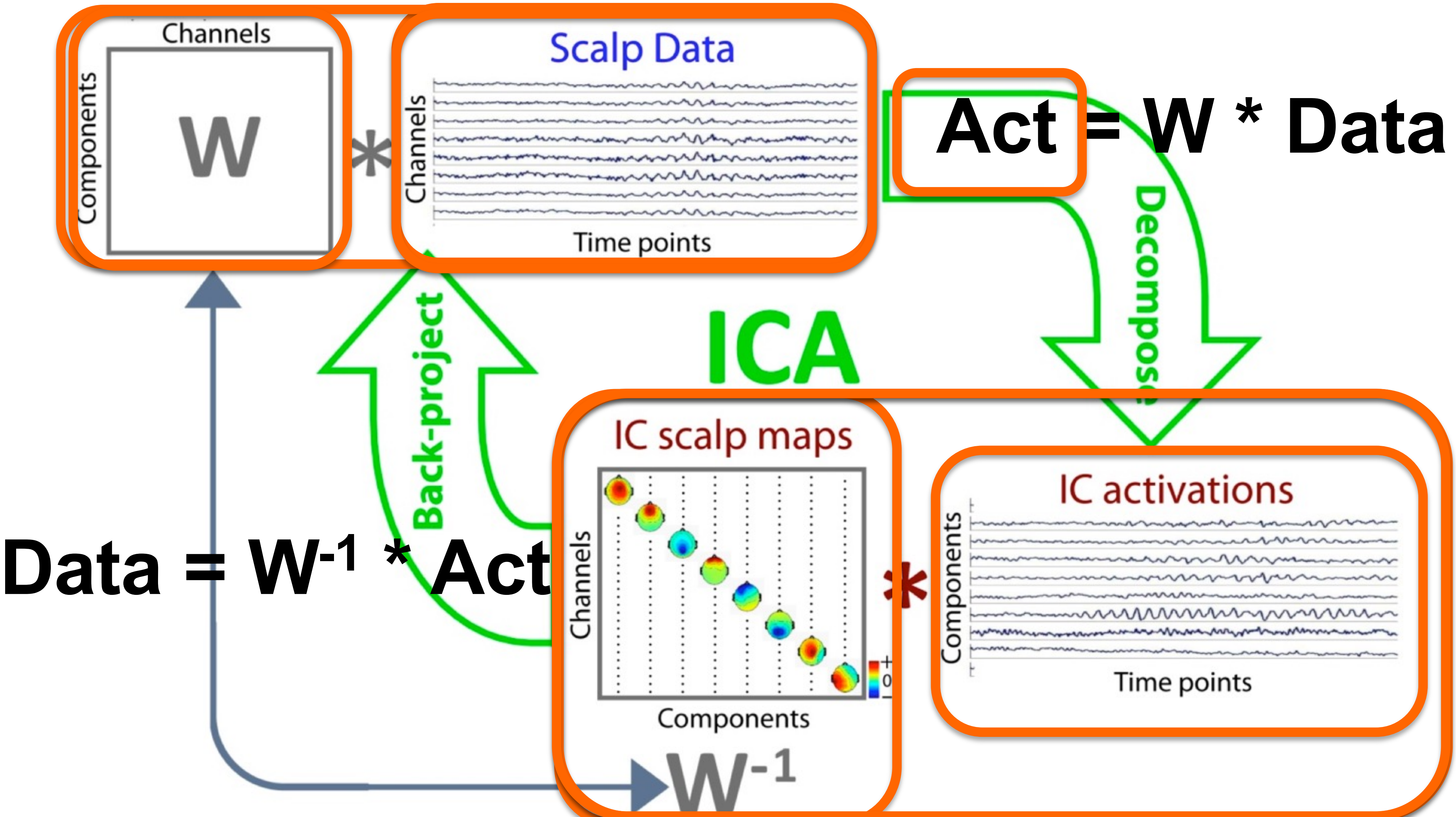


Julie Onton & S. Makeig (2006)

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:

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

Some ICA History

- 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)
- Palmer (2006): Adaptive mixture ICA (AMICA)

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)

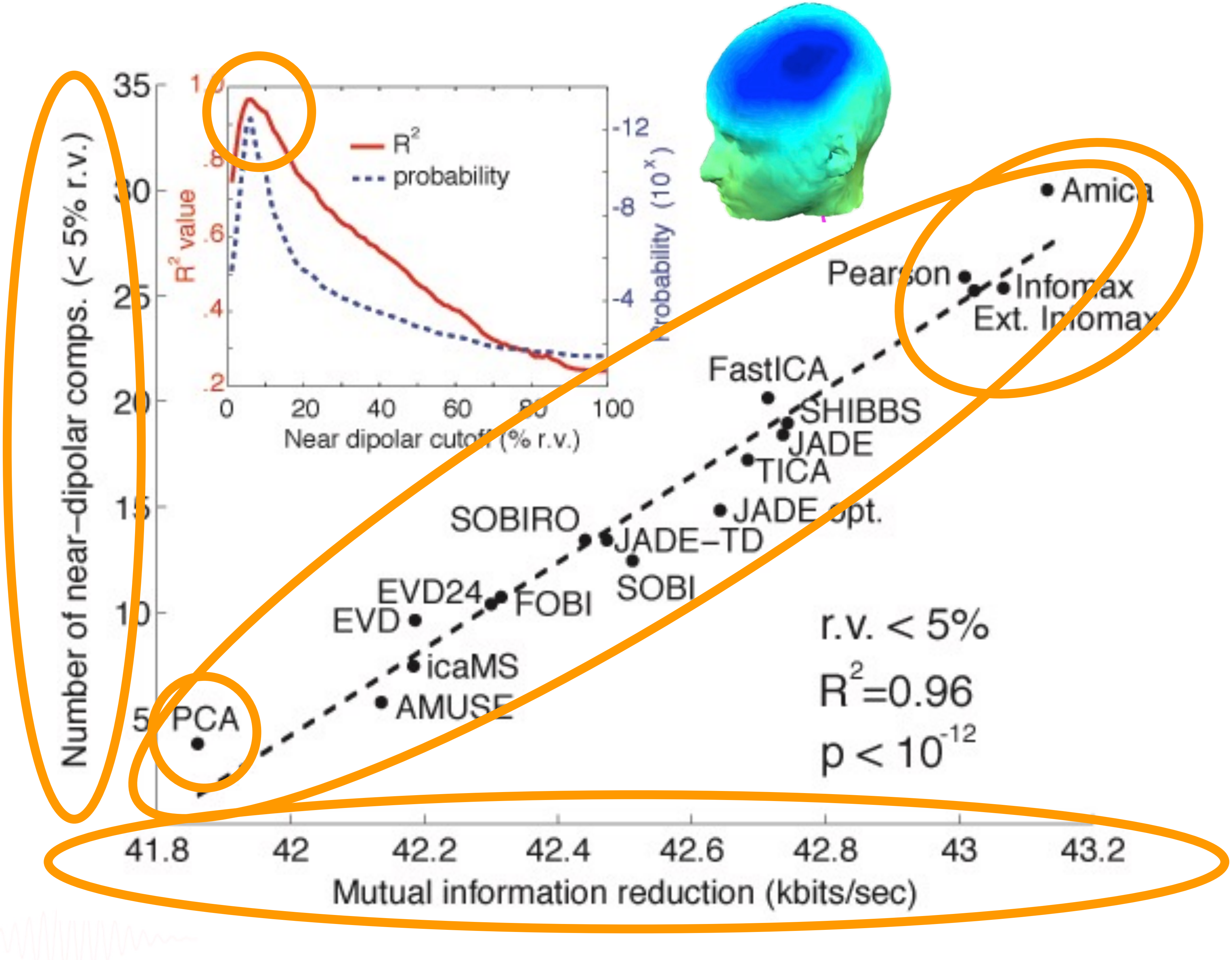
Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent ...**

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

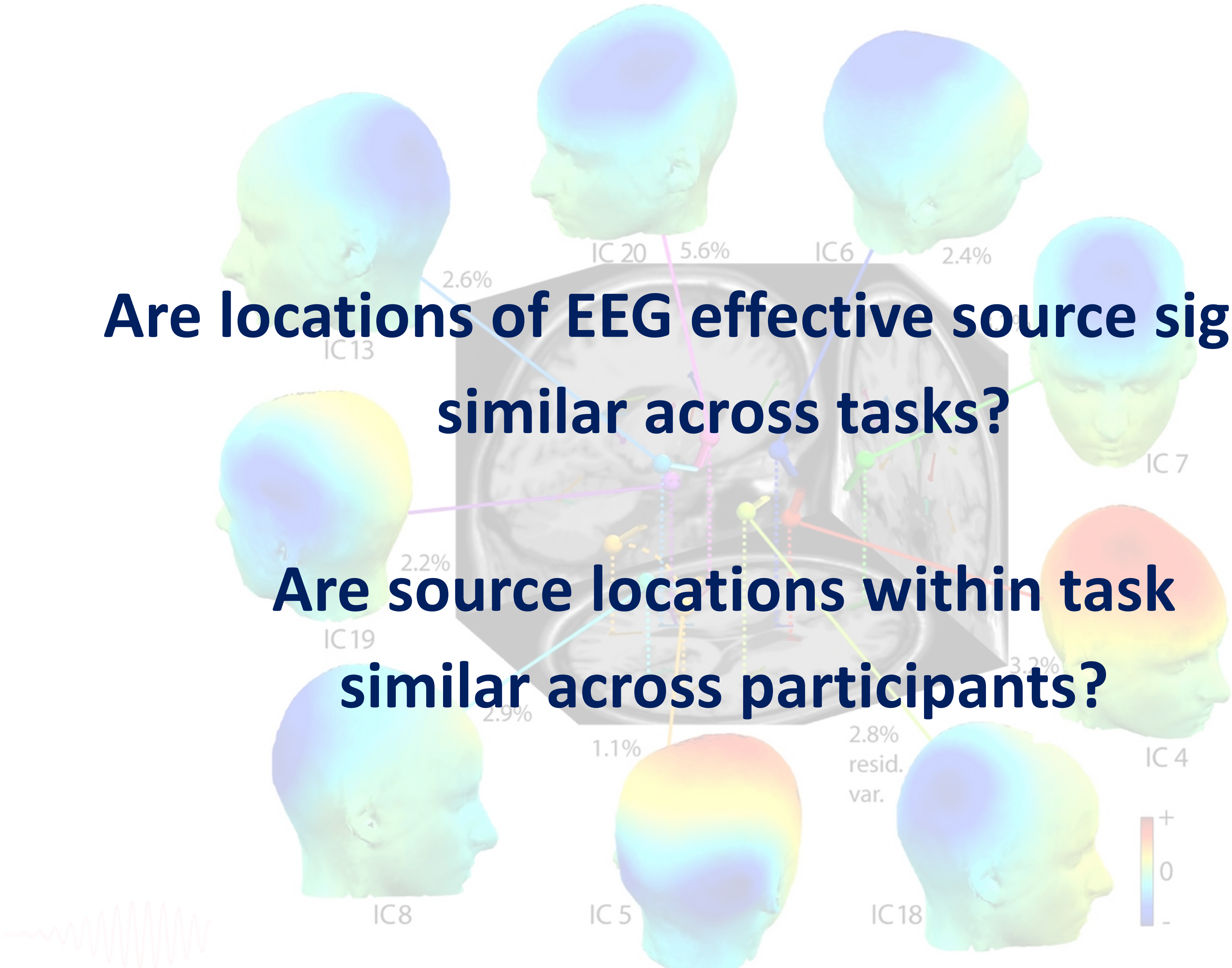
Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent …**

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



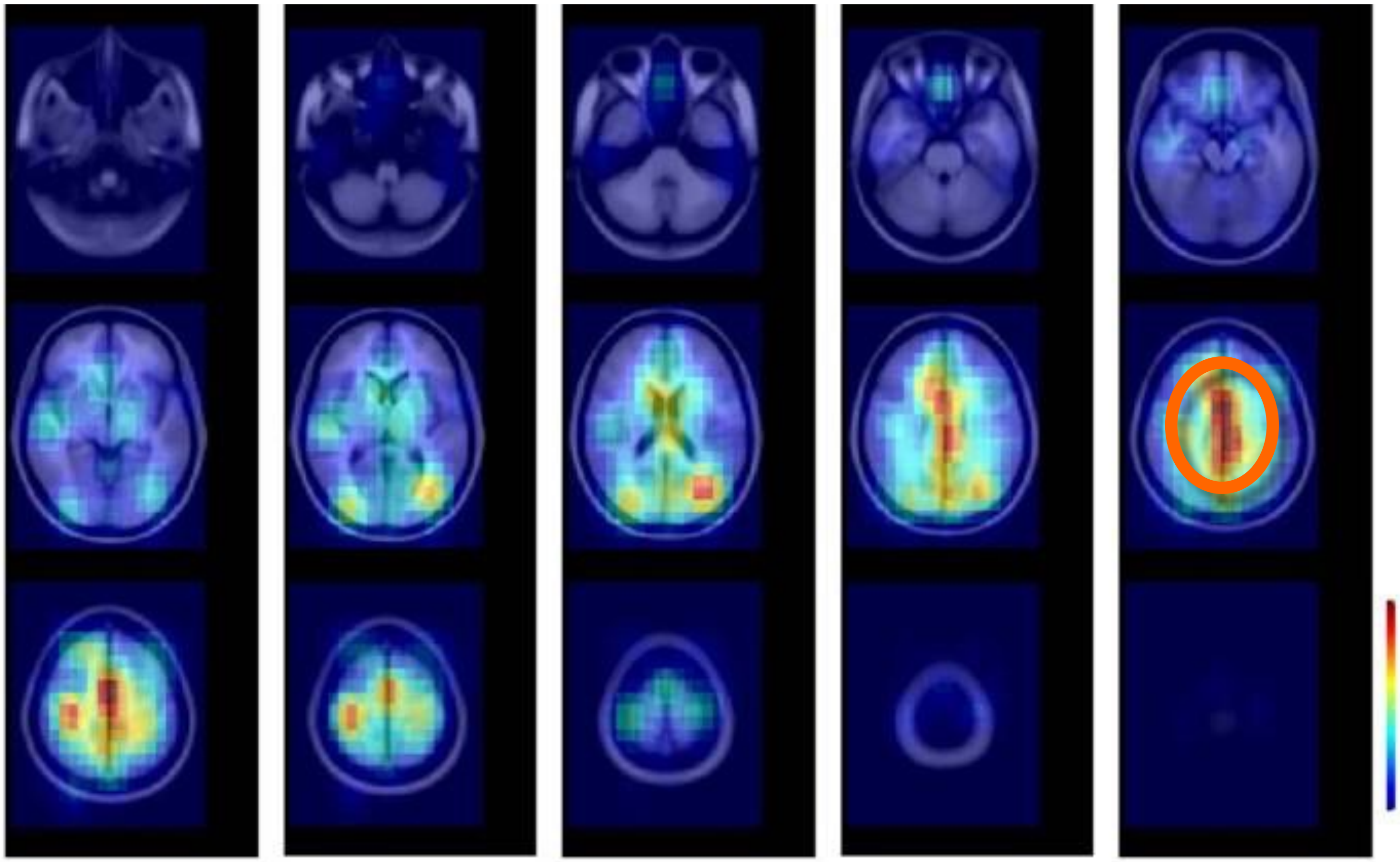
Are locations of EEG effective source signals similar across tasks?

Are source locations within task similar across participants?

S. Makeig (2007)

Effective Source Density

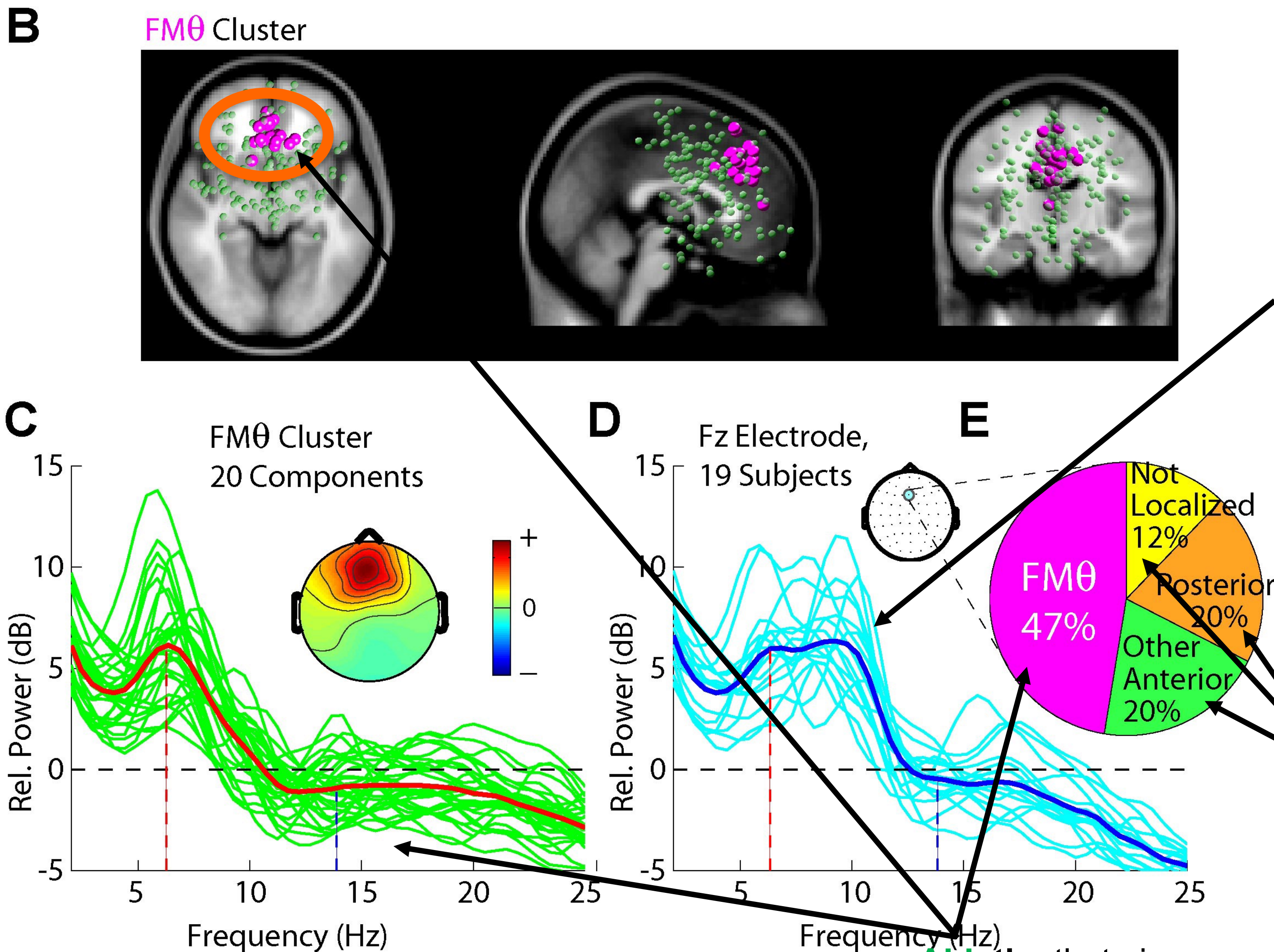
B. Visually cued selective response



Visually cued button press task

Onton & Makeig, 2005

Example: frontal midline theta cluster



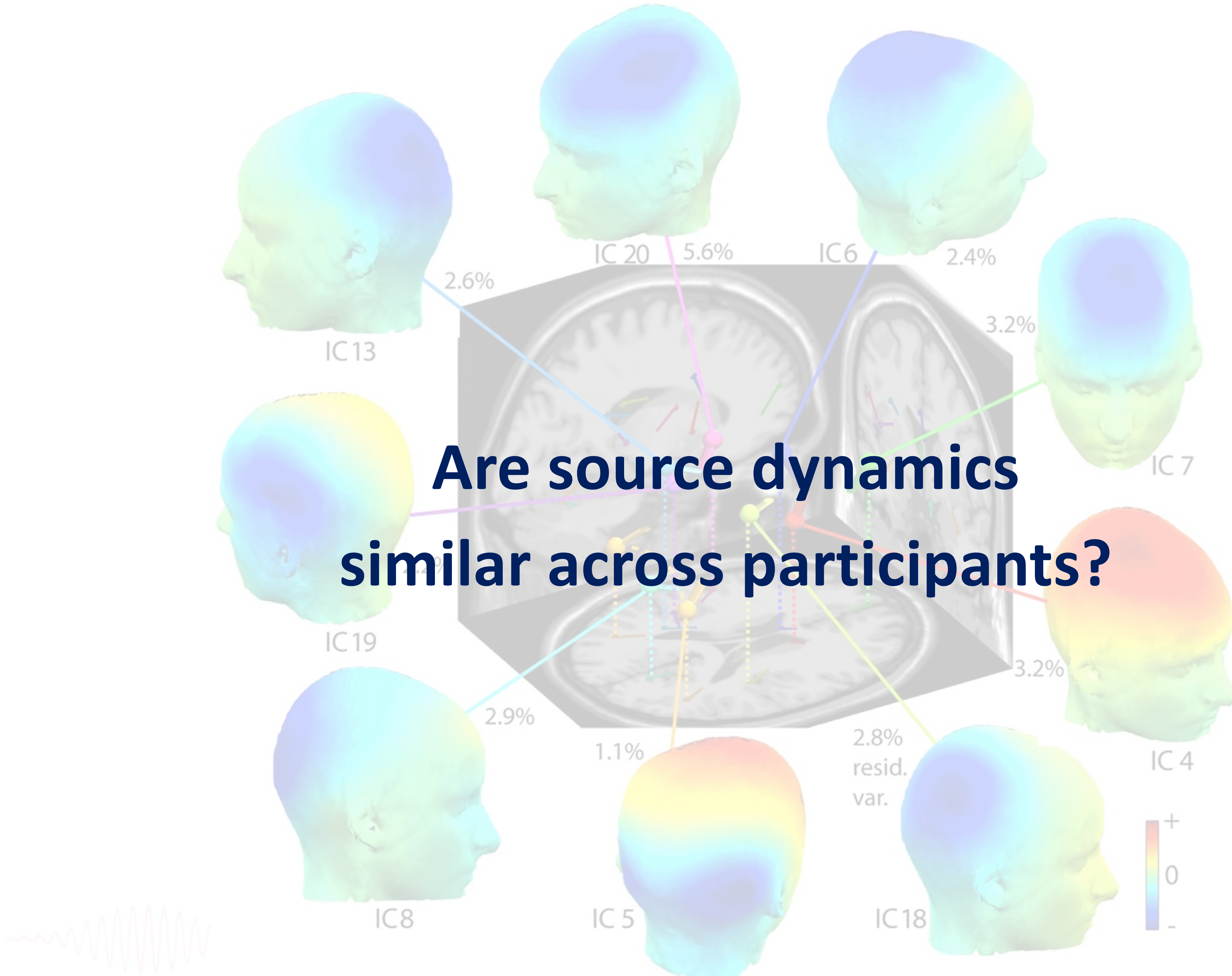
Channel data spactrum for the Fz→Ref channel

NO theta increase with more letters in memory!

ALL the theta increase with more letters in memory ***from this IC cluster!***

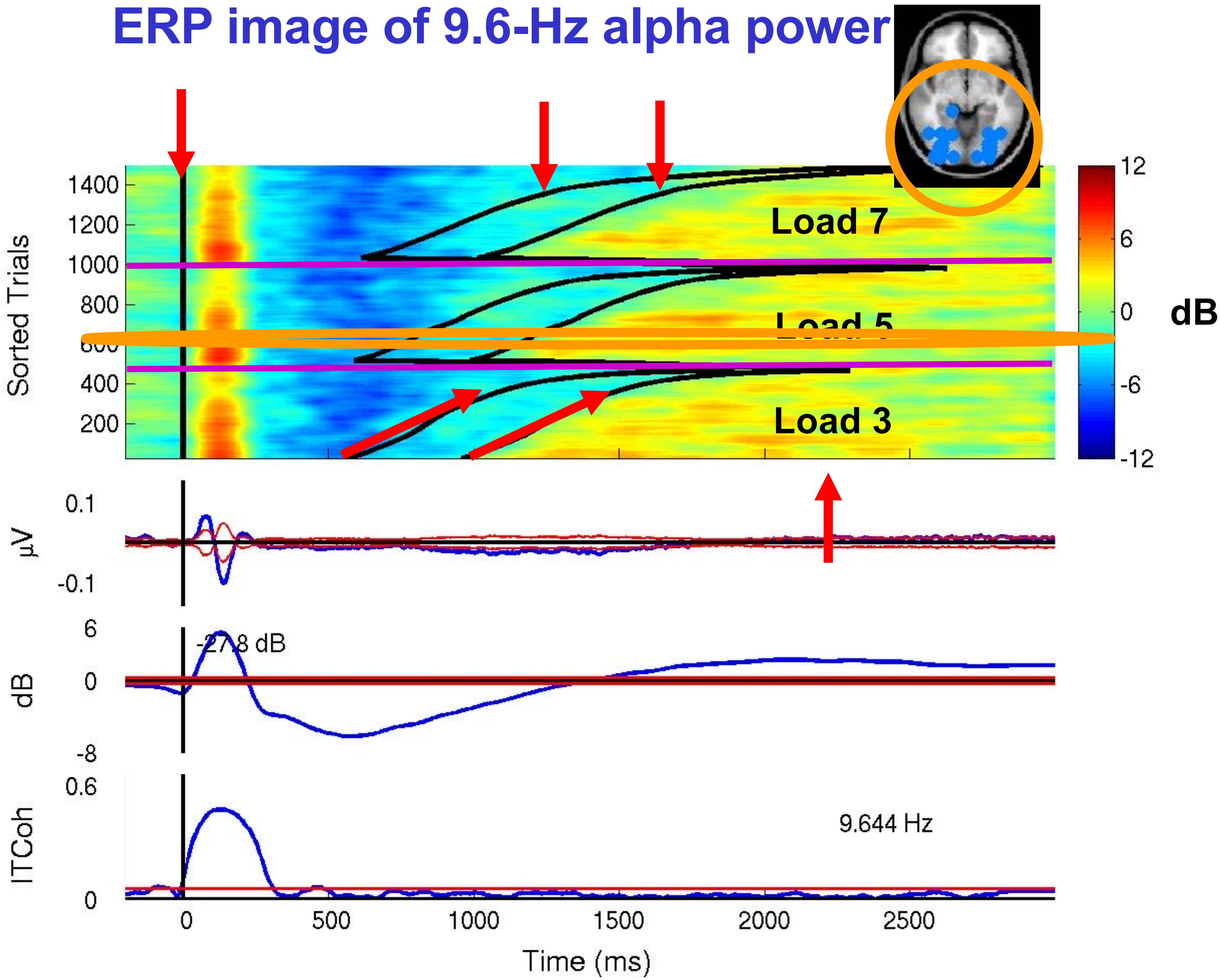
Onton, Delorme & Makeig, *NeuroImage* (2005)

Are source dynamics similar across participants?



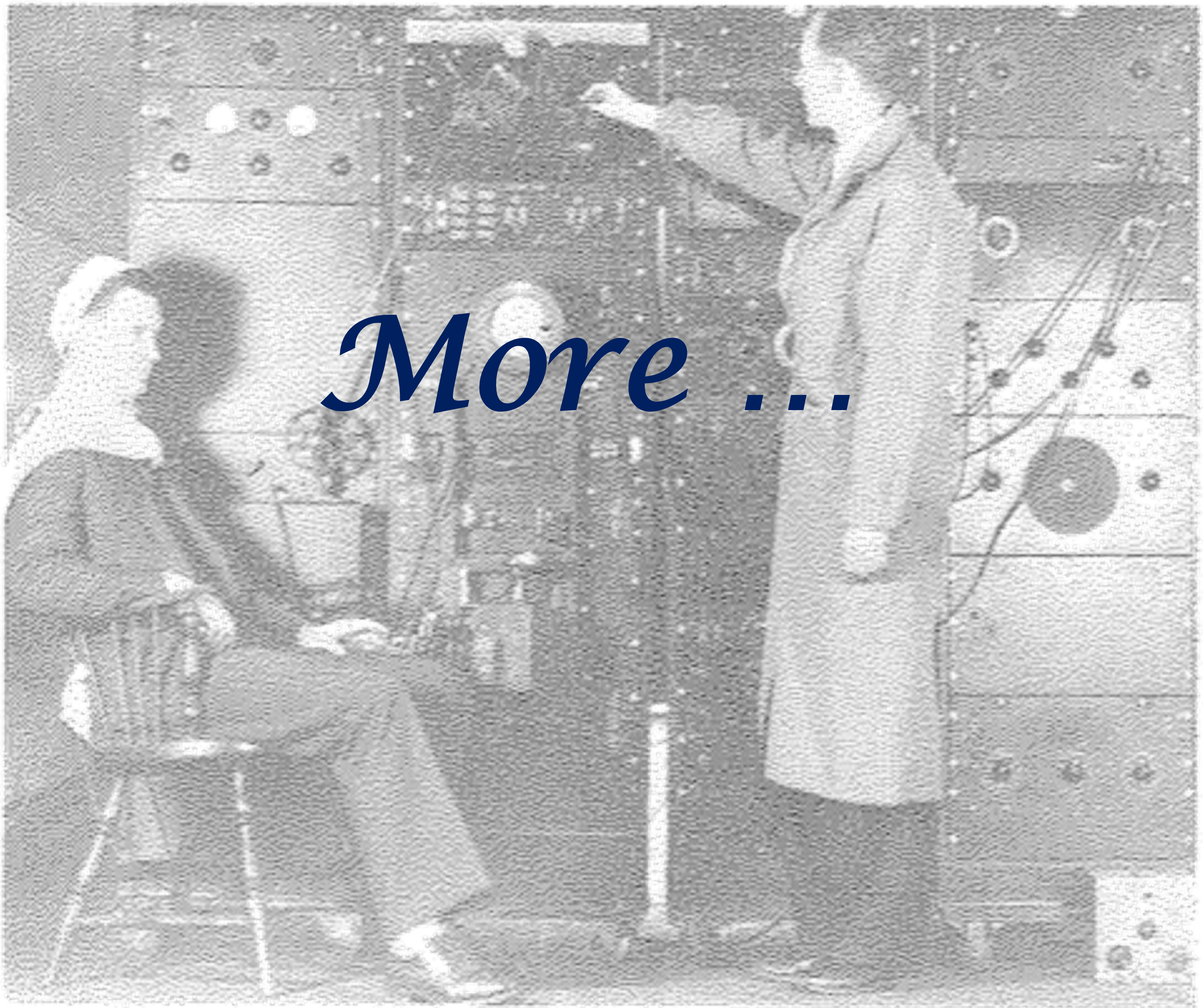
S. Makeig (2007)

ERP image of 9.6-Hz alpha power

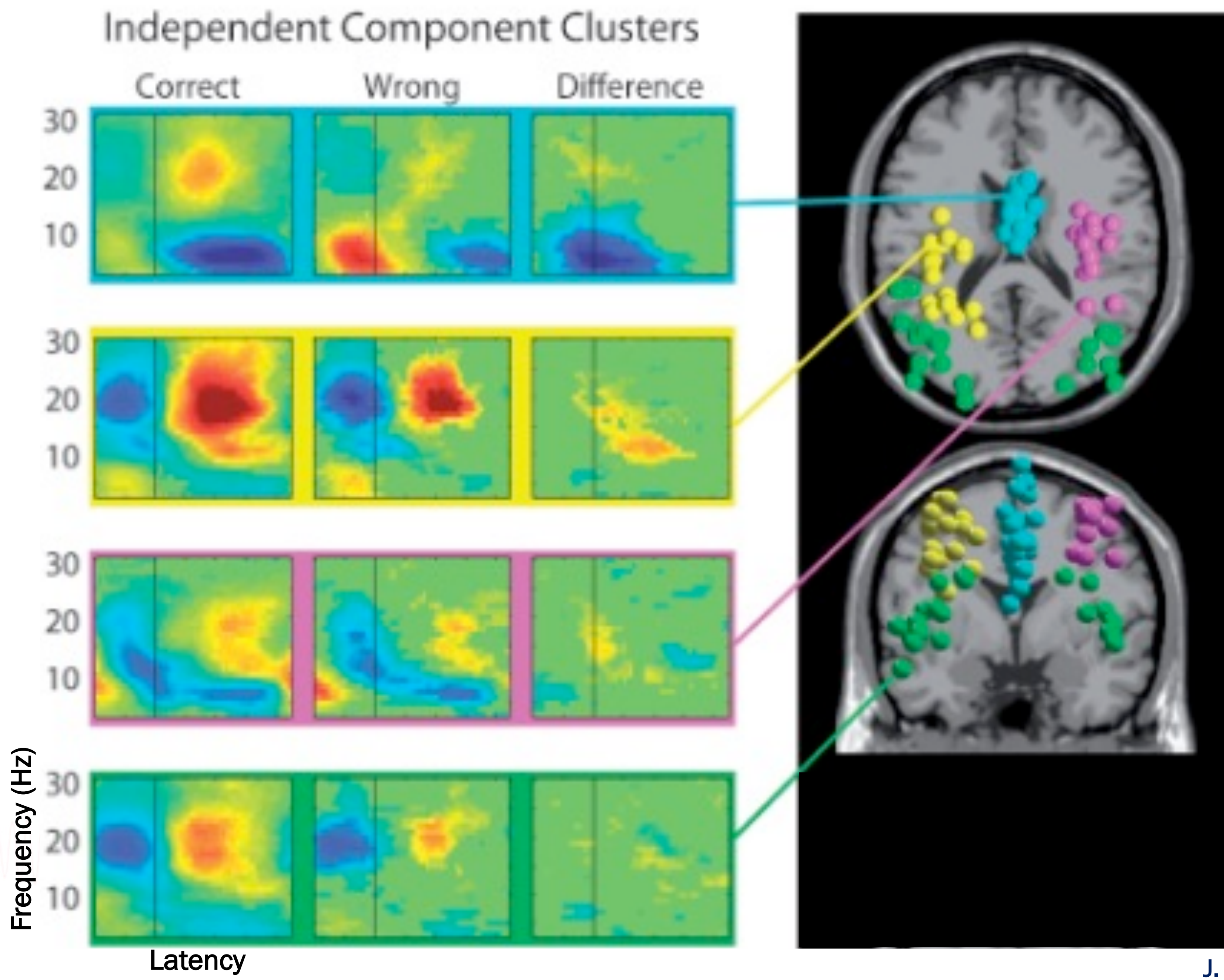


erpimage()

Onton, Delorme & Makeig, 2005.


More ...

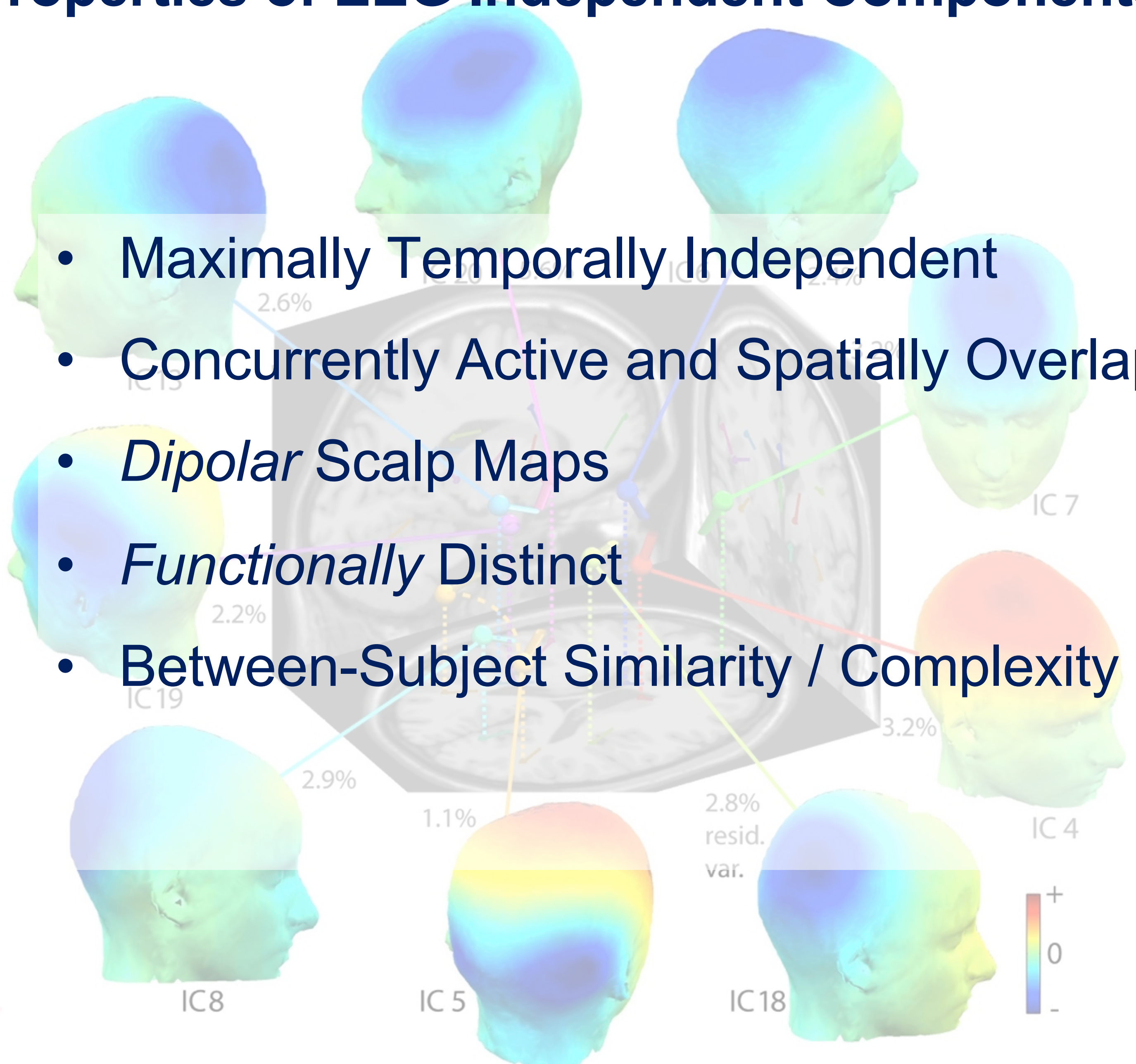
Goal: To cluster equivalent ICs across subjects



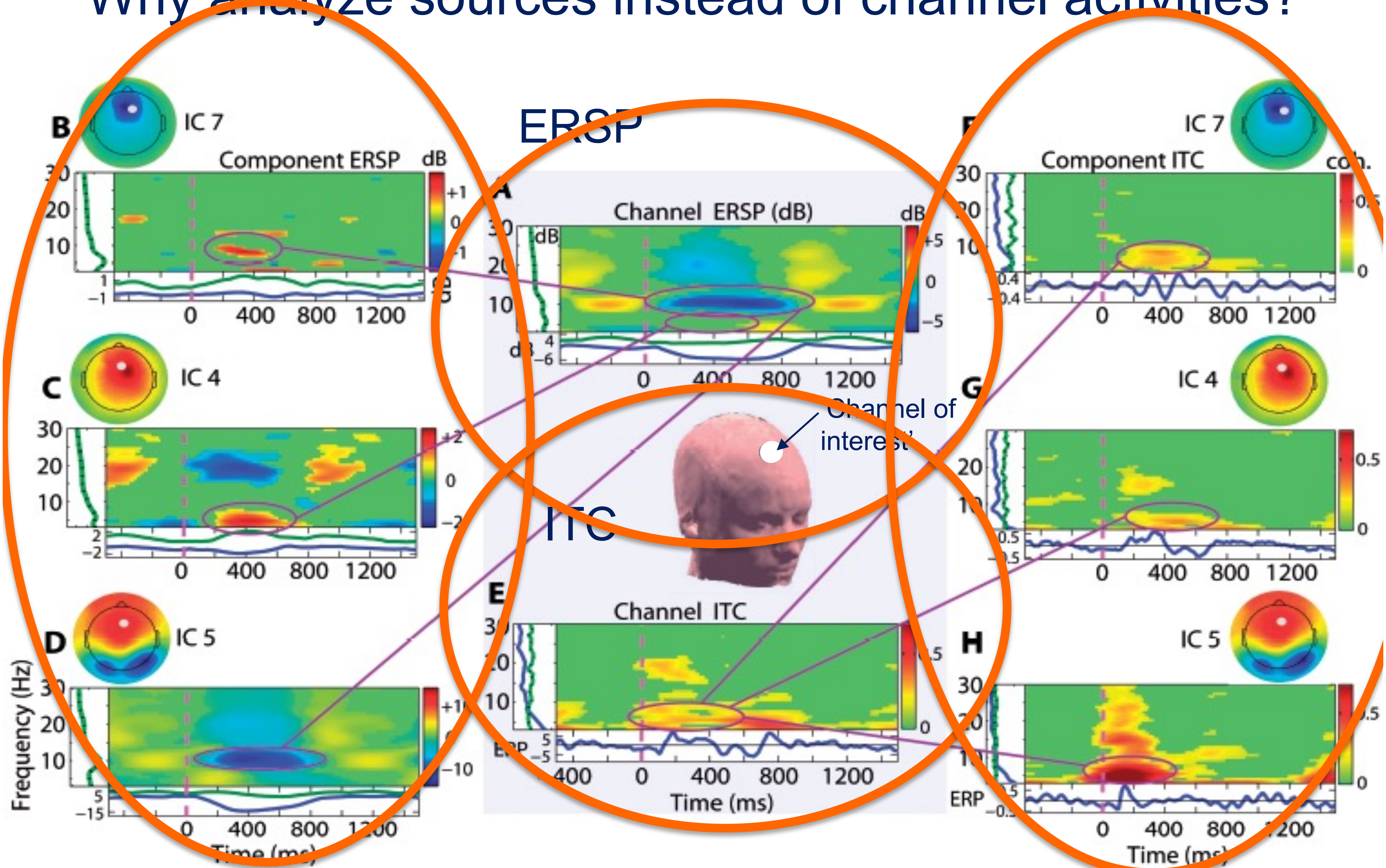
J. Onton, 2005

Properties of EEG Independent Components

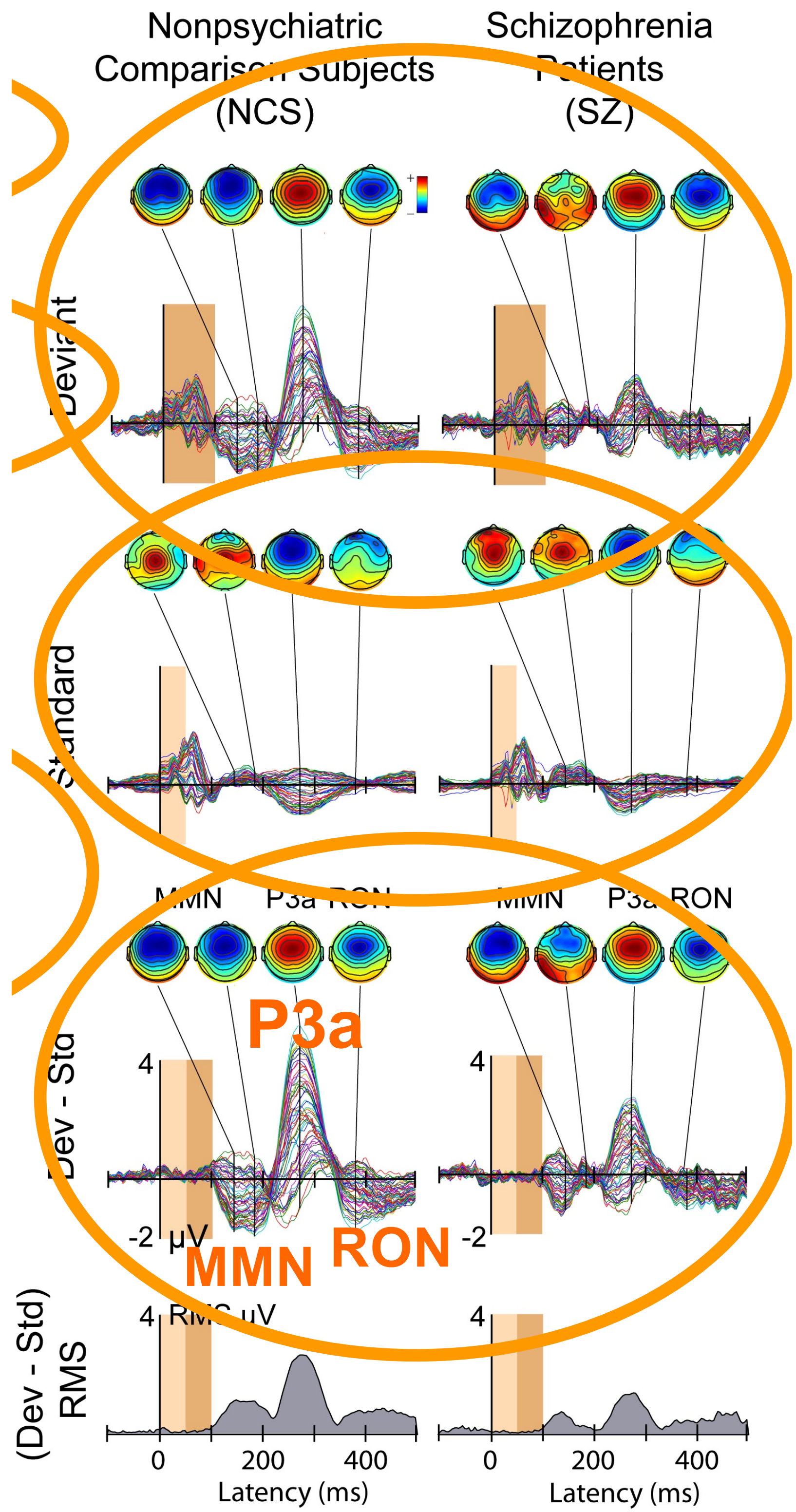
- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar* Scalp Maps
- *Functionally* Distinct
- Between-Subject Similarity / Complexity



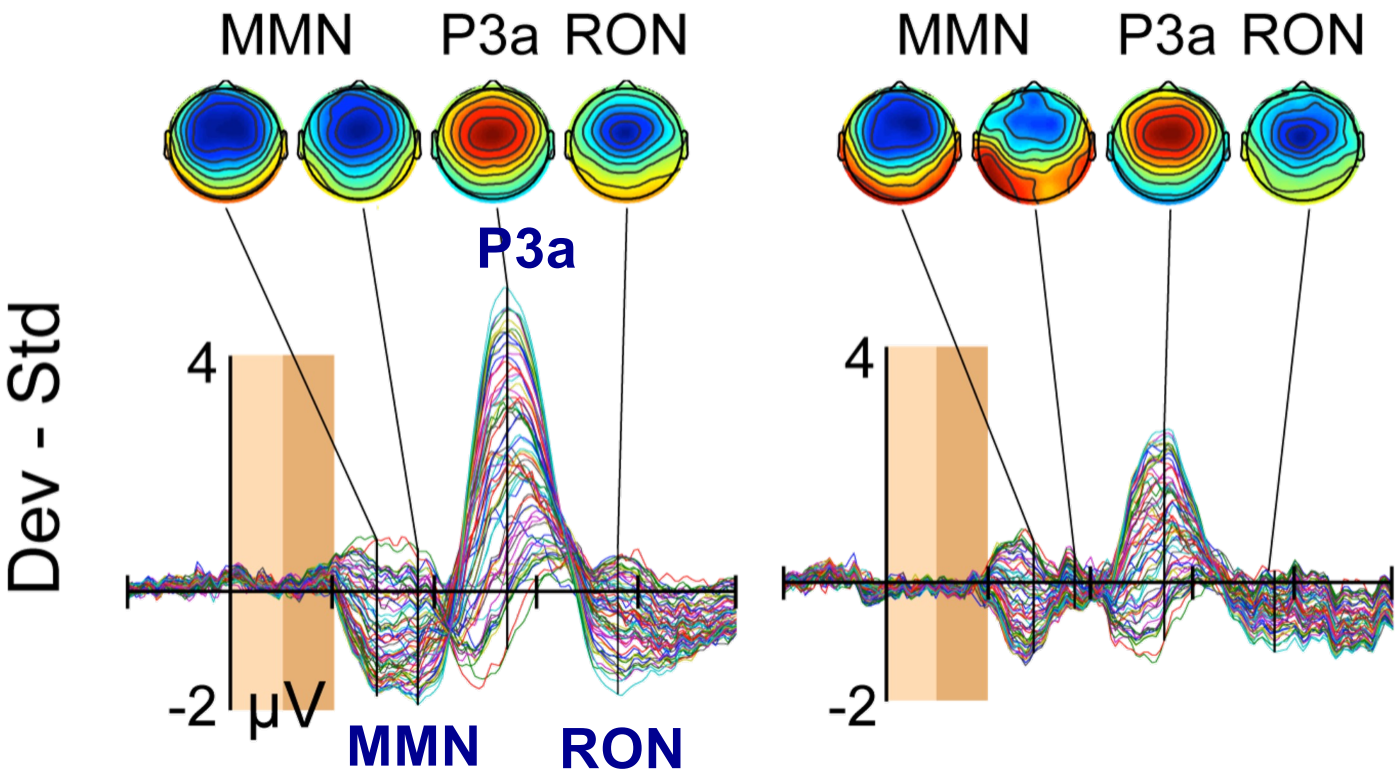
Why analyze sources instead of channel activities?



J. Onton & S. Makeig, 2005

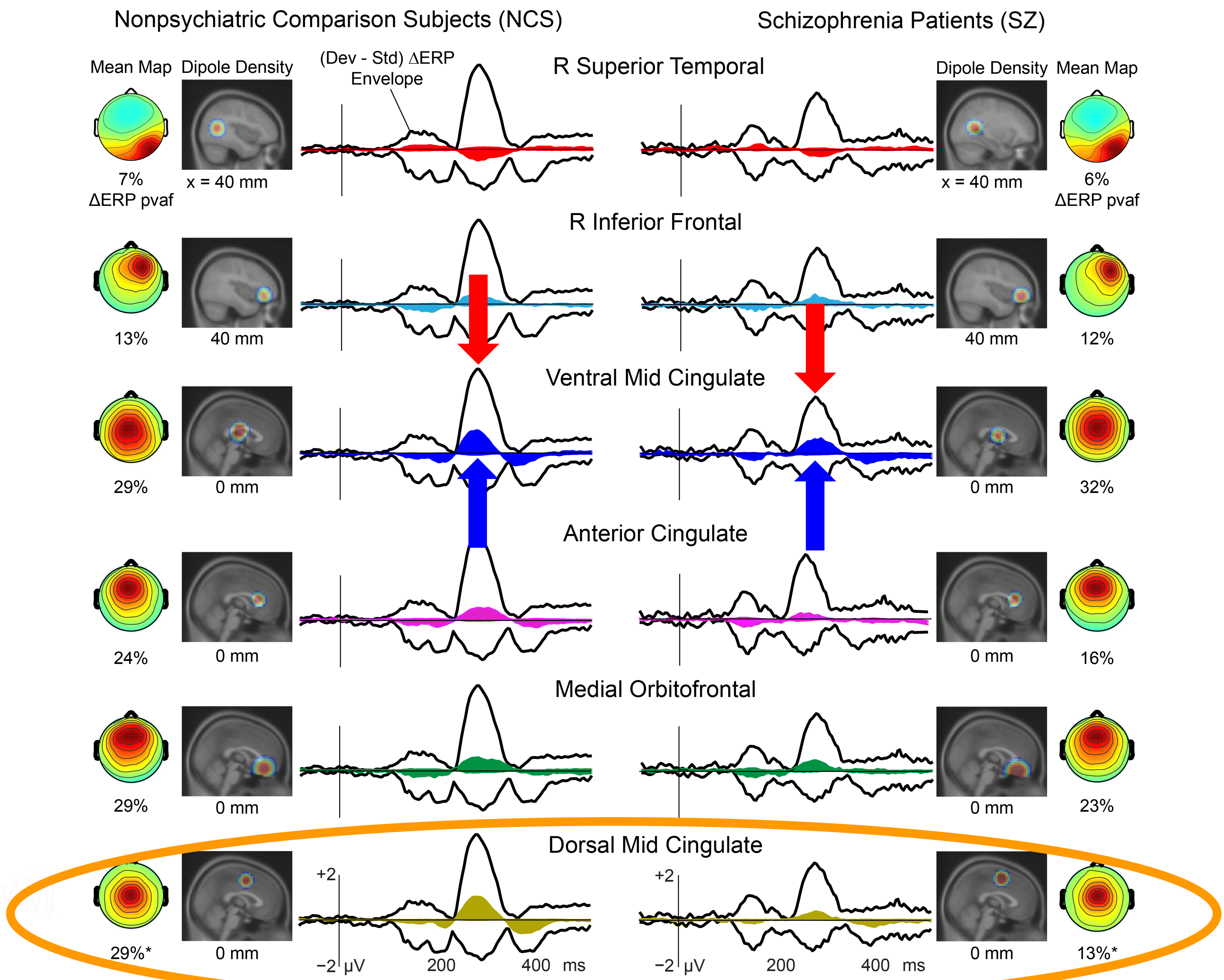

Nonpsychiatric Comparison Subjects (NCS)
Schizophrenia Patients (SZ)
Deviant
Standard
MMN
P3a
RON
Dev - Std
P3a
4
-2 μV
MMN
RON
(Dev - Std) RMS
4 RMS μV
0
200
400
Latency (ms)

Auditory Deviance Response



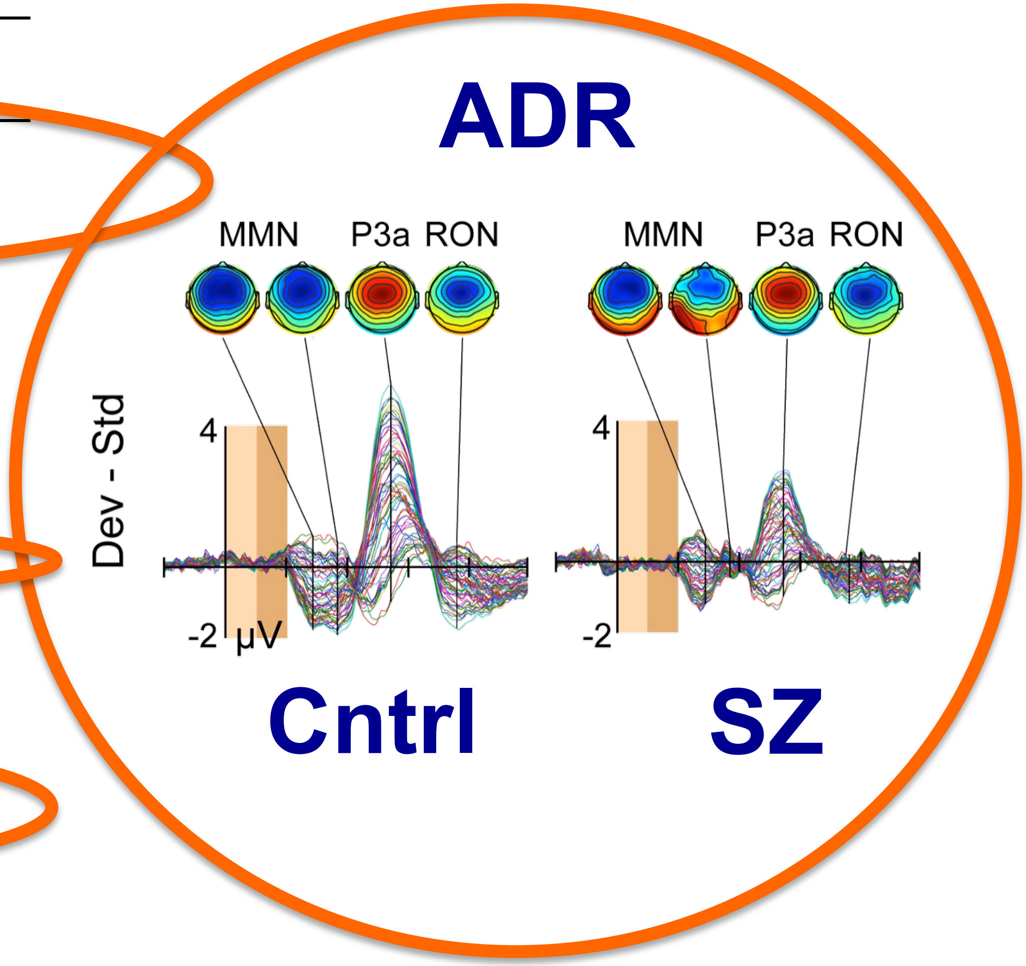
Cntrl **SZ**

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

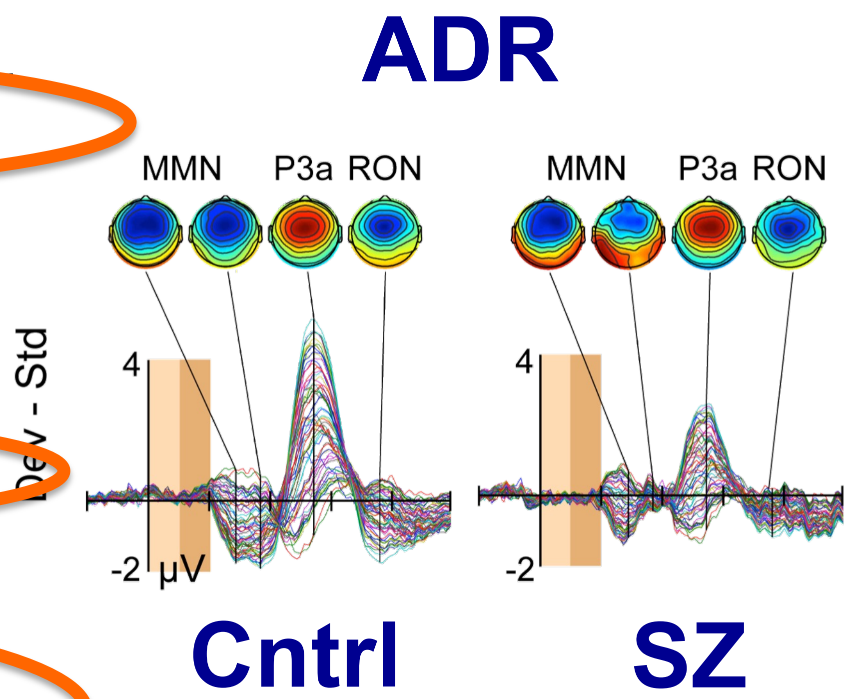


Rissling et al., 2014

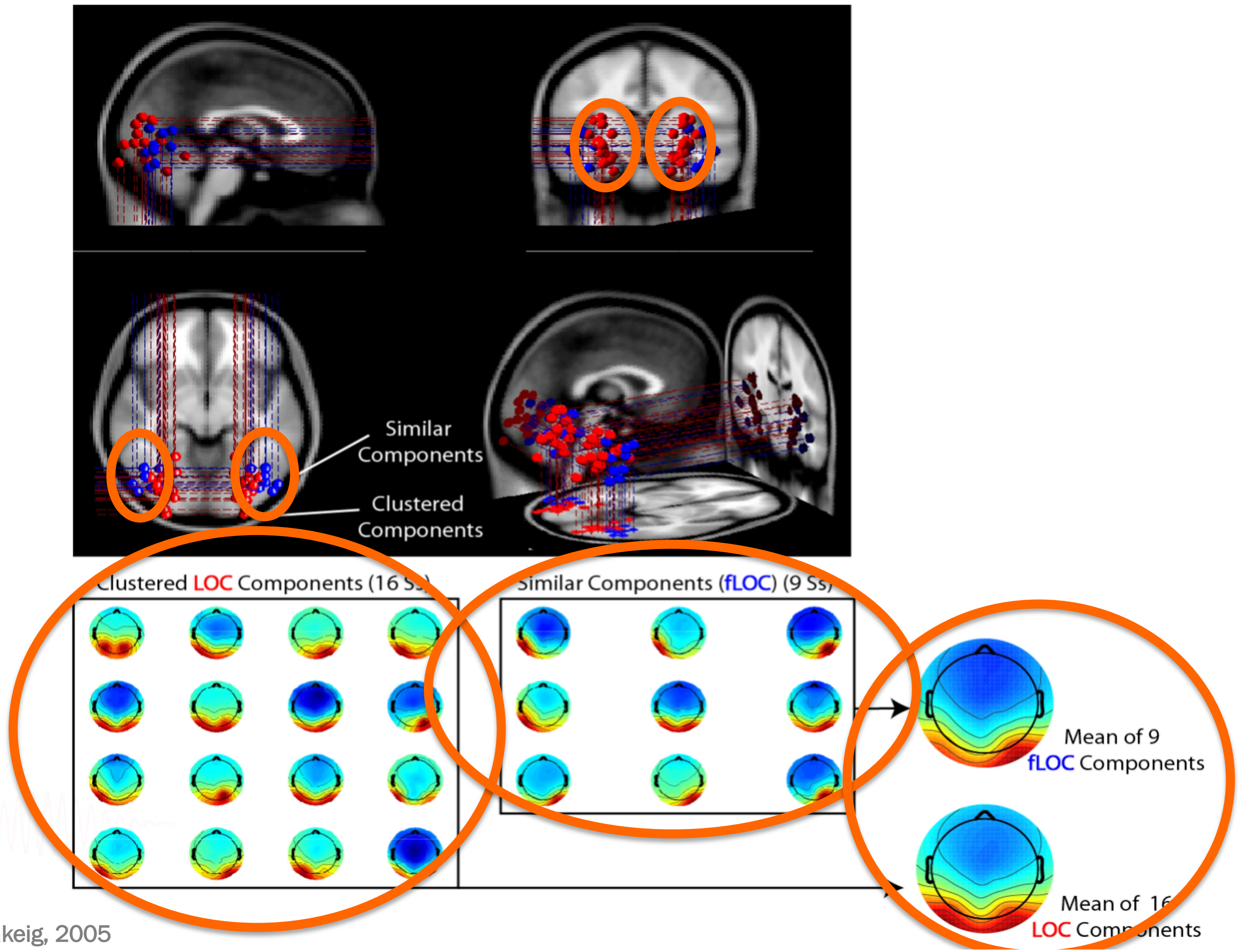
PEAK AMPLITUDES	ERP	r^2
Scalp Electrode (Fz)		
Verbal IQ (WRAT)	P3a	0.11
Functional Capacity (UPSA)	RON	0.12
R Superior Temporal		
Working Memory (LNS Reorder)	RON	0.15
Verbal IQ (WRAT)	RON	0.15
Immediate Verbal Memory (CVLT)	**RON**	**0.28**
Delayed Verbal Memory (CVLT)	RON	0.26
Functional Capacity (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



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

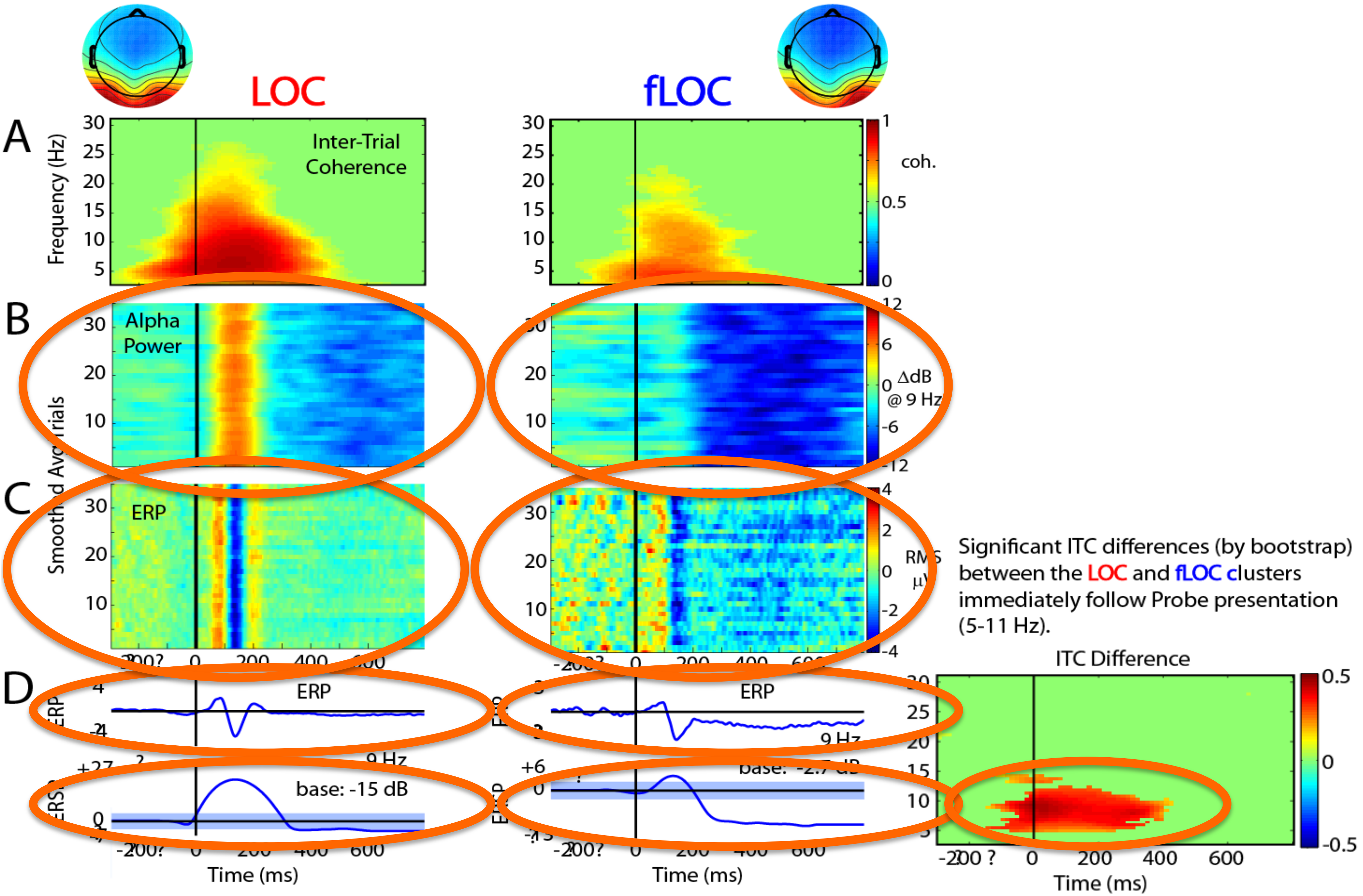


Why don't all subjects contribute to every IC cluster?



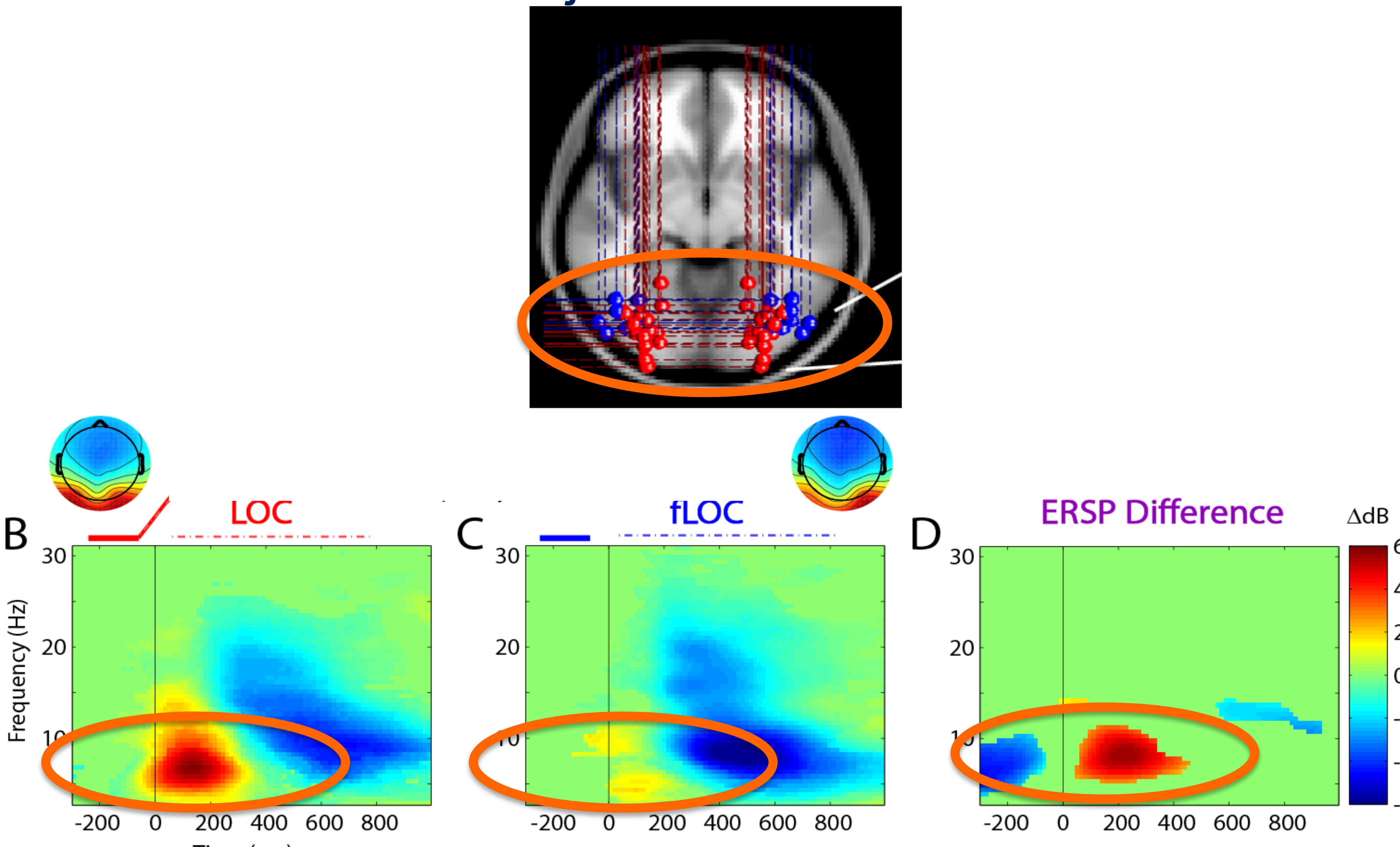
Onton & Makeig, 2005

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?