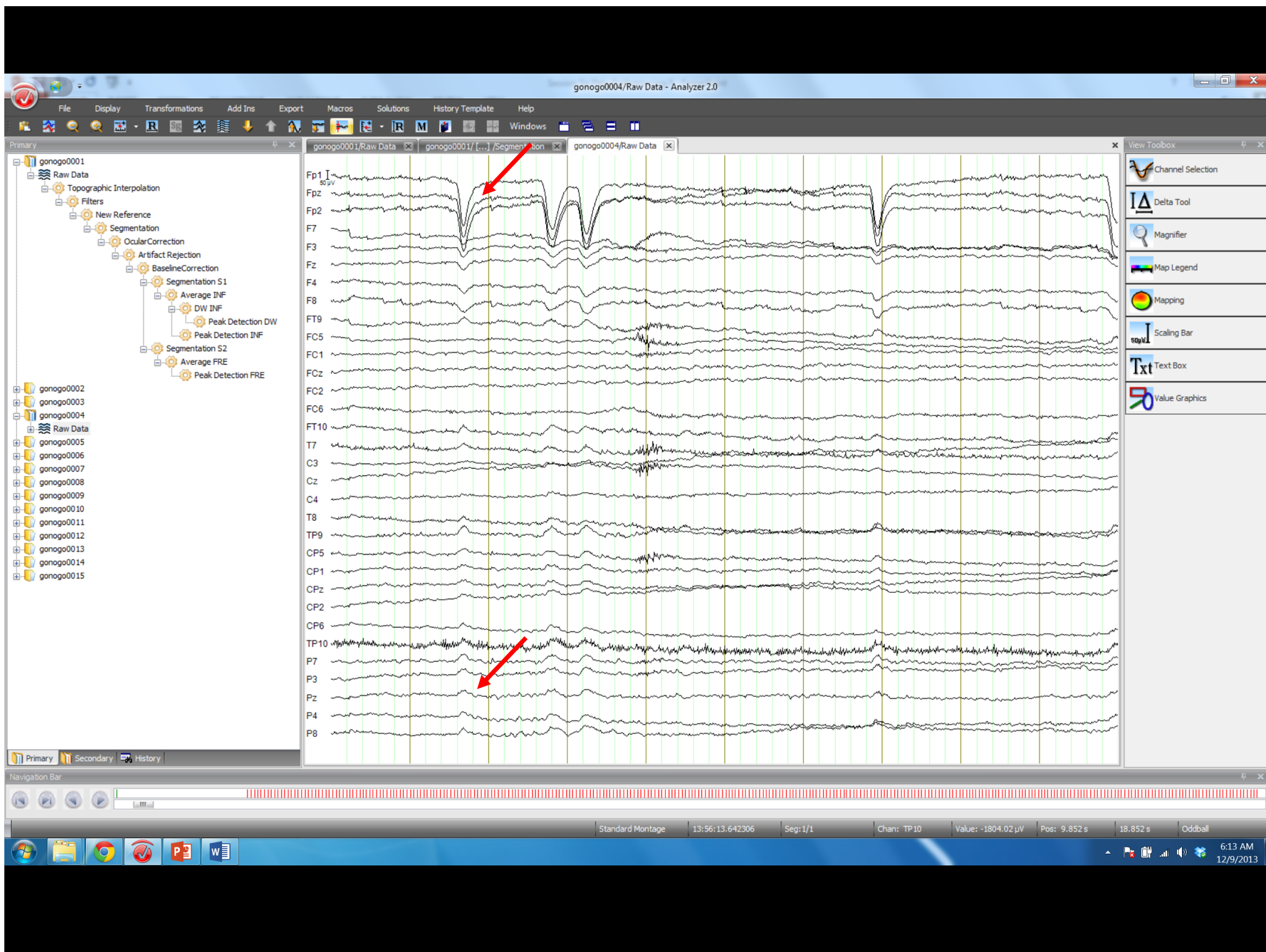
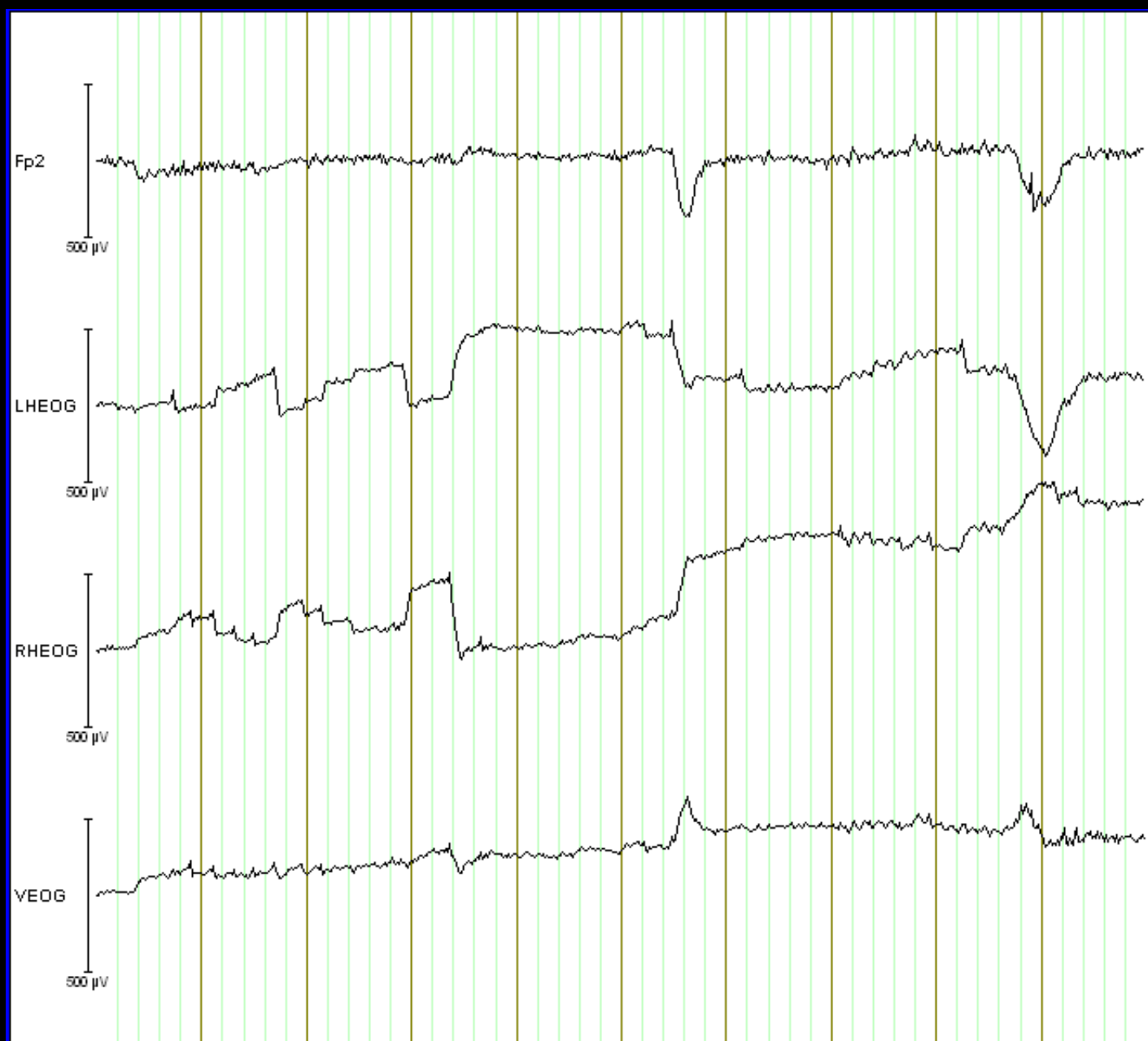


# Ocular Correction

The issue with ocular correction is  
simple...







# Blinks

Blinks impact data across the entire scalp due to signal propagation.

The magnitude of the blink is much greater than the underlying neural activity, hence the signal at that point in time is lost, or at best, blurred.

# What to do?

## 1. Instruction

Instruct participants to try and minimize blinking, or to blink during rest breaks.

Critique:

Is this now a dual task paradigm?

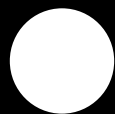
# What to do?

## 2. Environment

Adjust stimuli and lighting to avoid high contrasts.

Critique:

Sleepiness



# What to do?

## 3. Removal

Simply remove all trials from analysis with blinks in them.

Critique:

Data Loss

Systematic Blinking

# What to do?

## 4. Correction

Use a correction method to “remove” the blinks and “interpolate” the missing data.

Critique:

Data Interpolation

Systematic Blinking



With Ocular	Without Ocular
22	44
8	27
5	136
1	18
8	105
24	40
6	91
12	150
22	67
5	68

# Common Correction Techniques

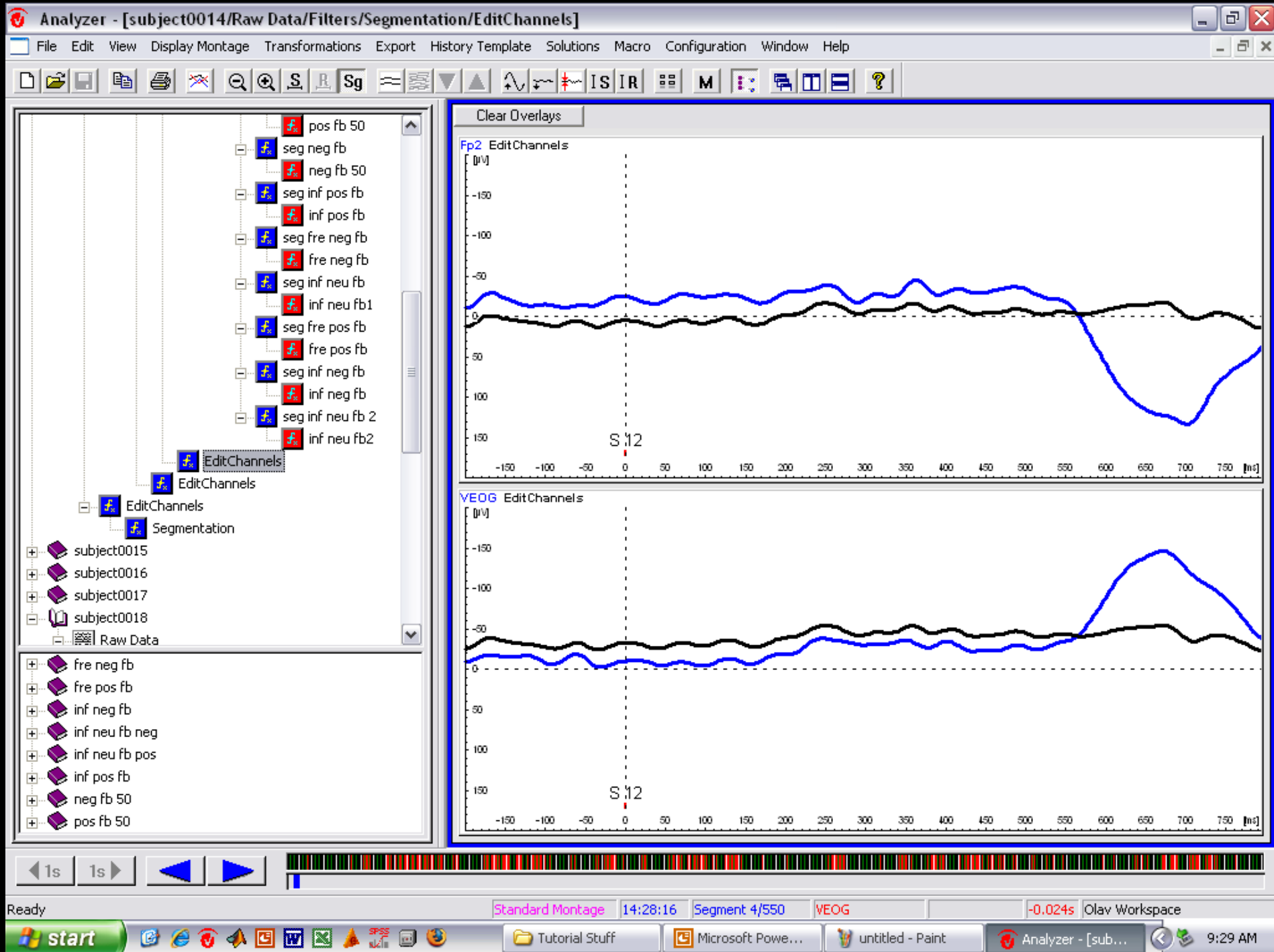
## Gratton and Coles

The GC technique uses EOG channels to identify blinks – based on slope and amplitude criterion. Then, a regression approach is used to estimate the missing data and interpolate it. Scaling factors are used to correct for blink amplitude corrections across the scalp.

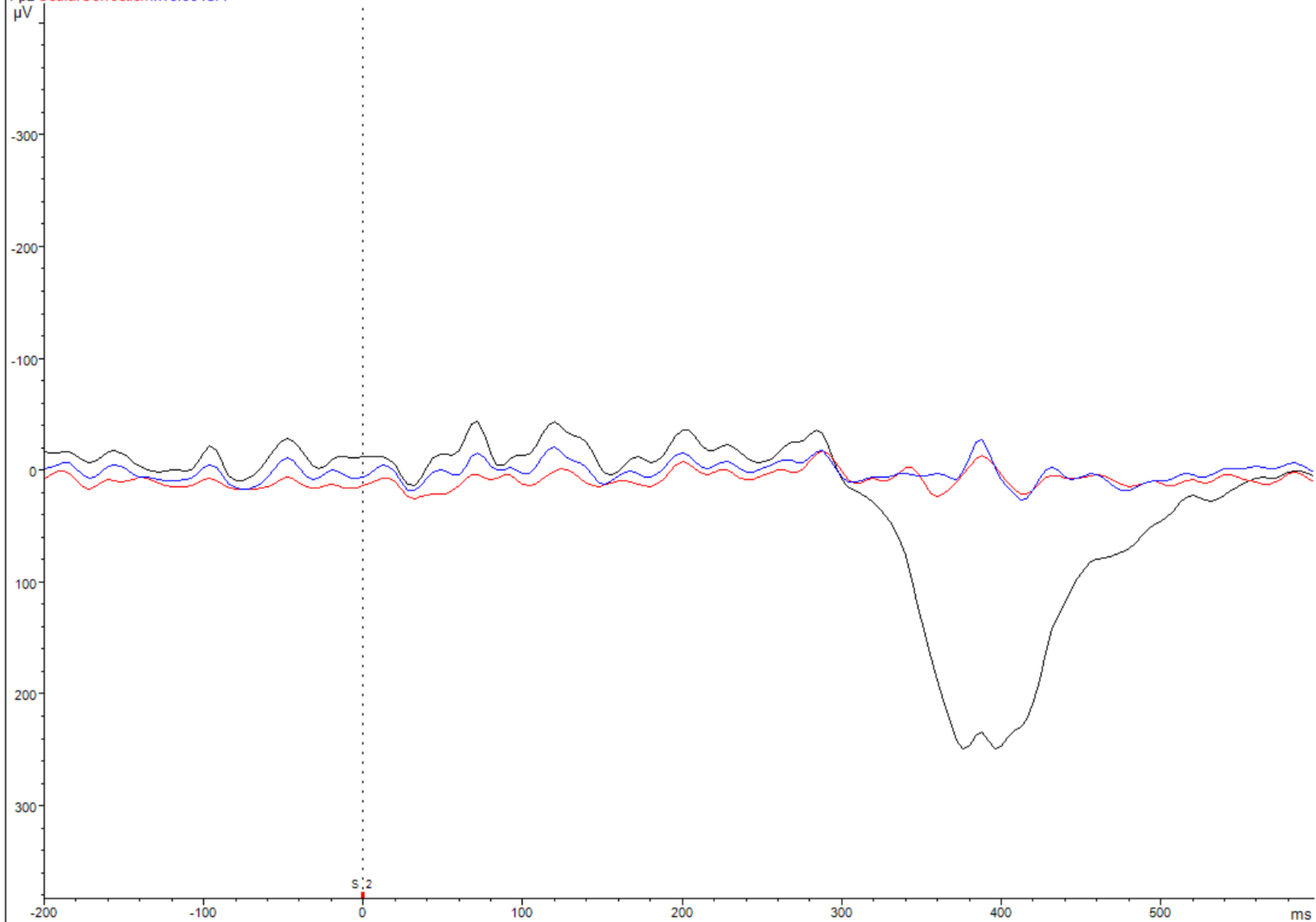
# Common Correction Techniques

## Independent Component Analysis

ICA is a more complex correction procedure (we will discuss it Friday in depth) that uses the ICA solution to identify components that capture variance in the data. Those that capture variance associated with blinks can be subtracted from the data itself to “remove” the blink.



Fp2 OcularCorrectionInverse ICA



# What to do?

## 5. Short Segments

-200 to 400	13
-200 to 600	40
-200 to 800	64
-200 to 1000	84

# Ocular Correction Demonstration

ICA



# ICA

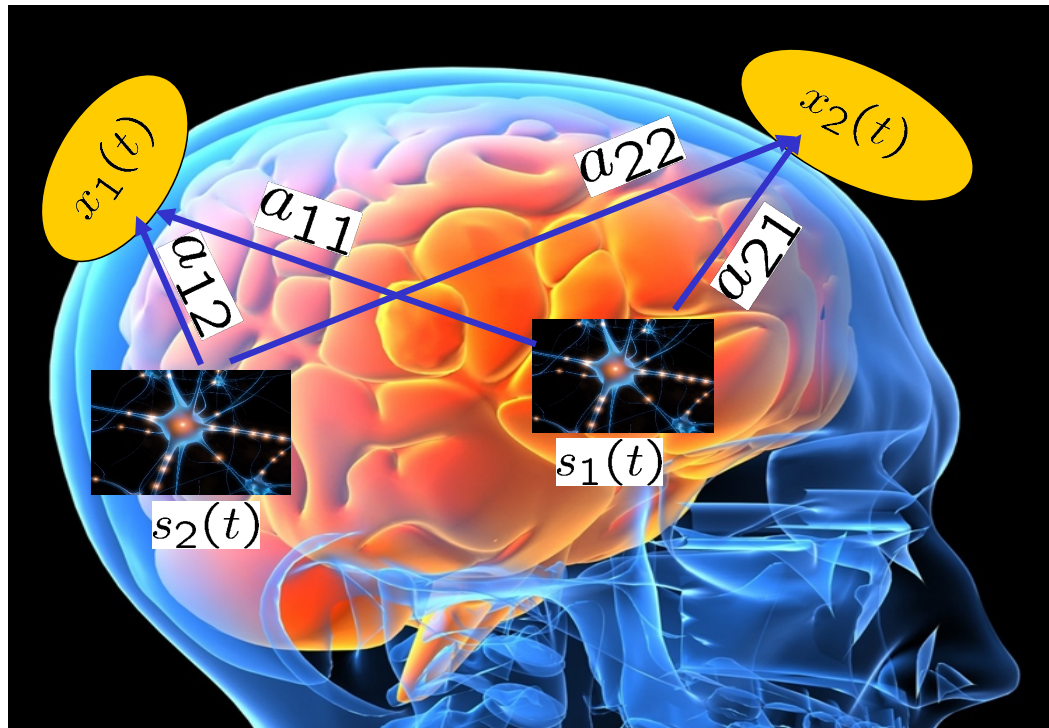
ICA has been primarily used in EEG research to correct ocular artifacts but can be used to isolate spatial components as well.

# Independent Component Analysis

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t)$$

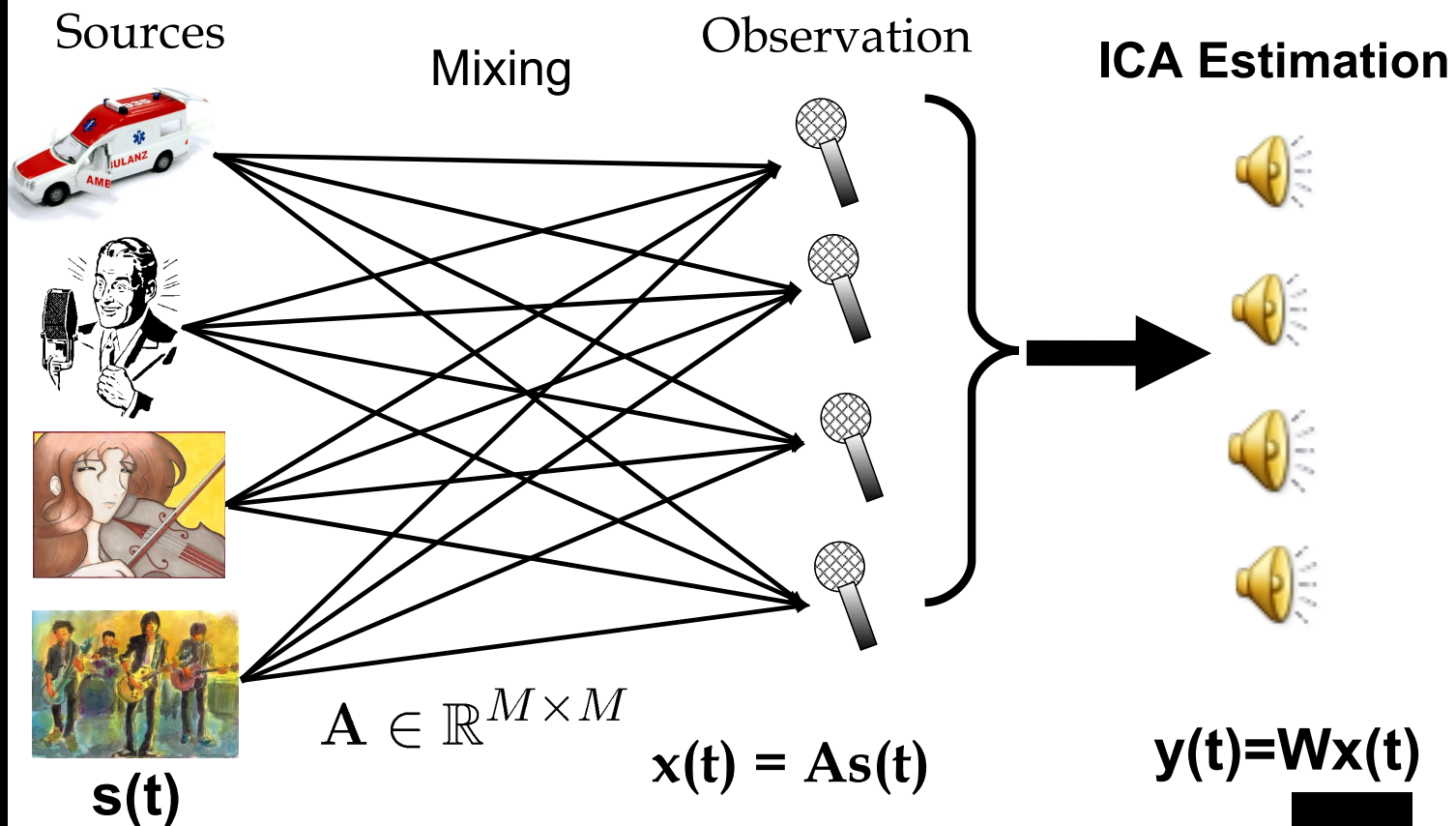
$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t)$$

**Goal:** Estimate  $\{s_i(t)\}$ ,  
(and also  $\{a_{ij}\}$ )

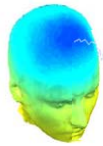


# The Cocktail Party Problem

## **SOLVING WITH ICA**



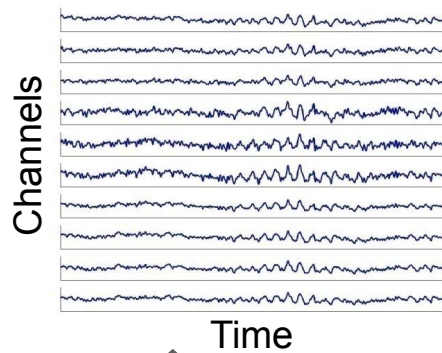
# Independent Component Analysis



$x$  = scalp EEG

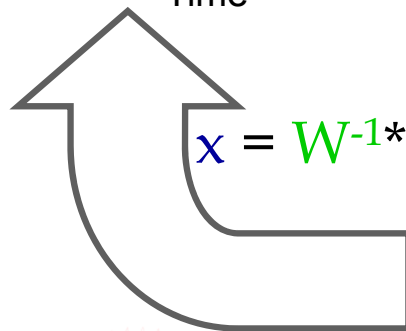
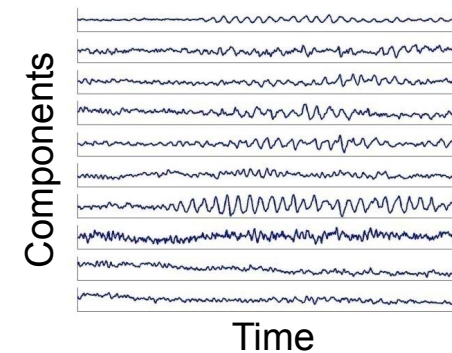
$W$  = unmixing matrix

$u$  = sources



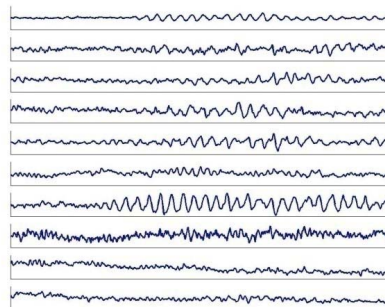
$$W * x = u$$

ICA



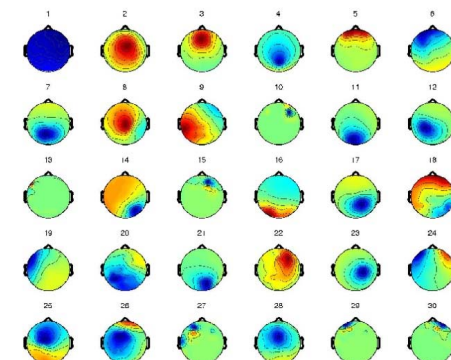
$$x = W^{-1} * u$$

$u$  = sources



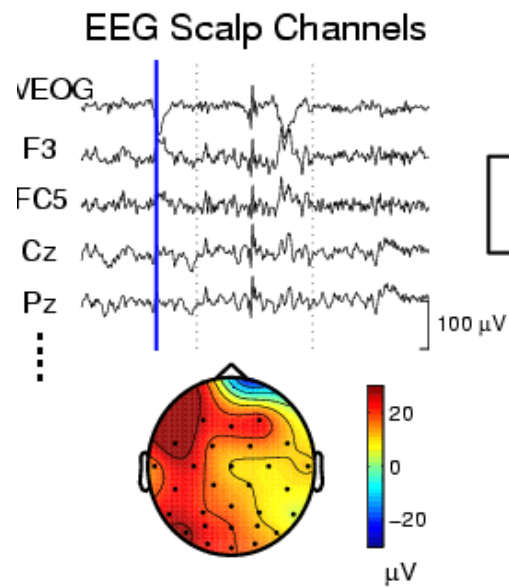
\*

$W^{-1}$  (scalp projections)

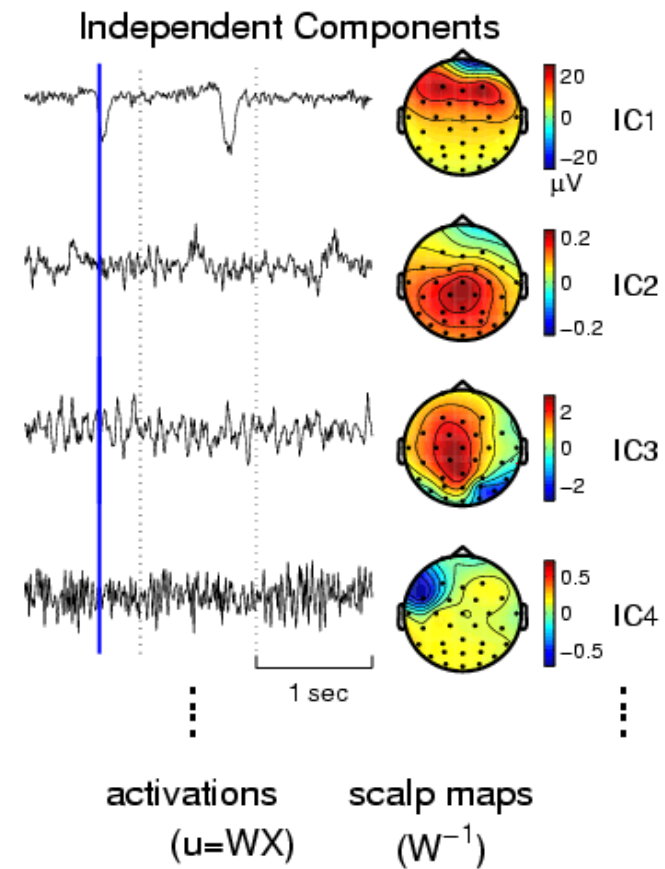




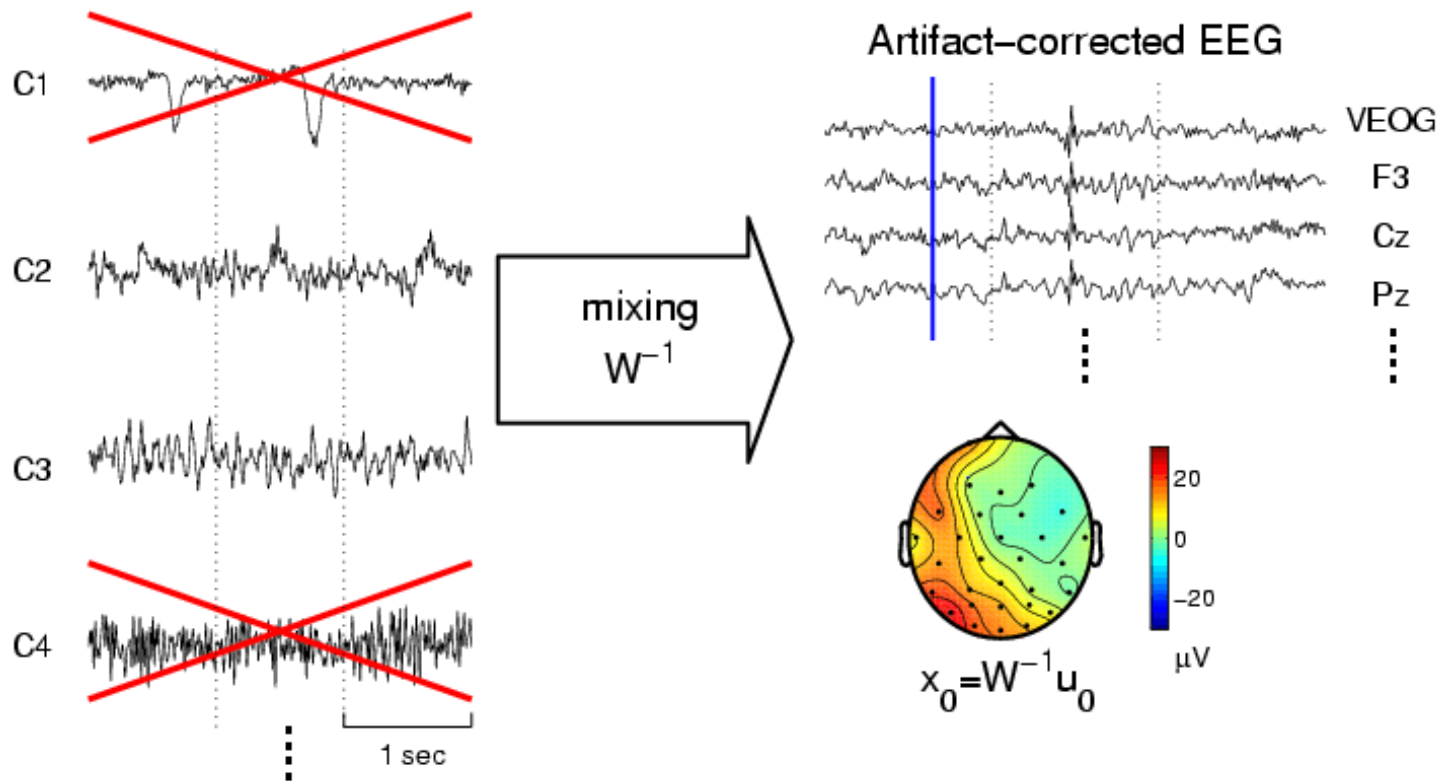
## ICA decomposition



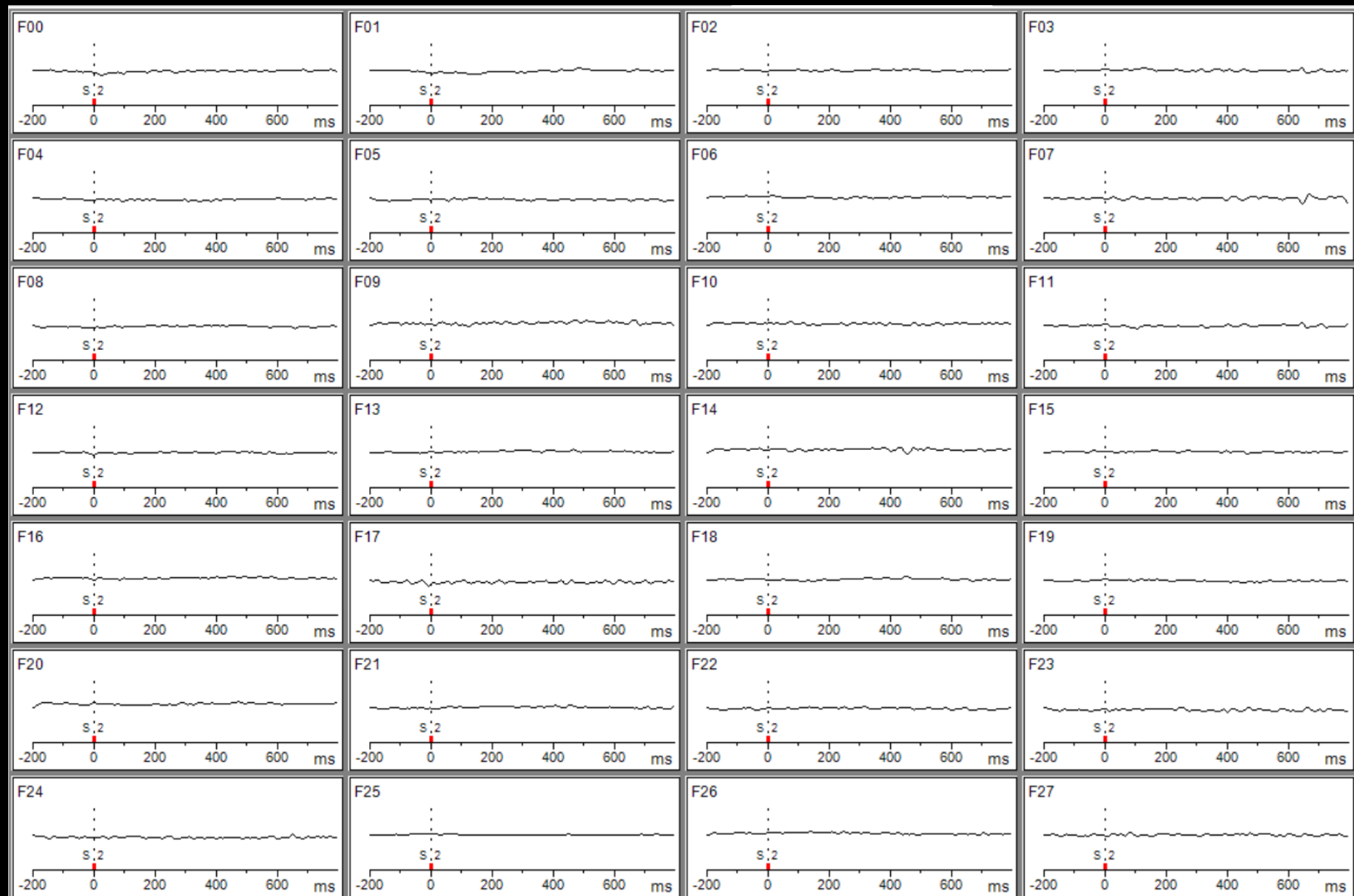
unmixing  
( $W$ )



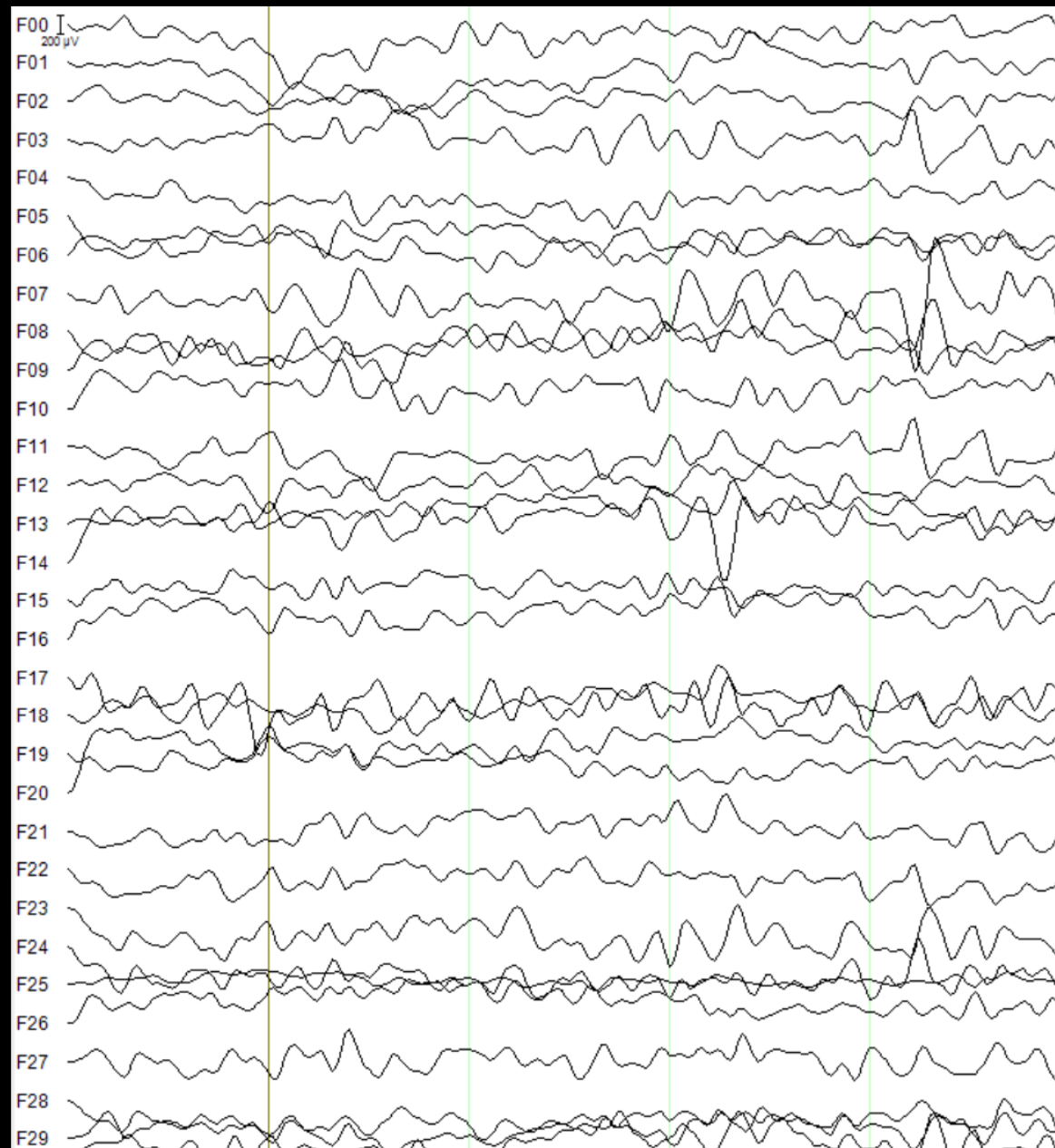
## Summed Projection of Selected Components



What It Looks Like





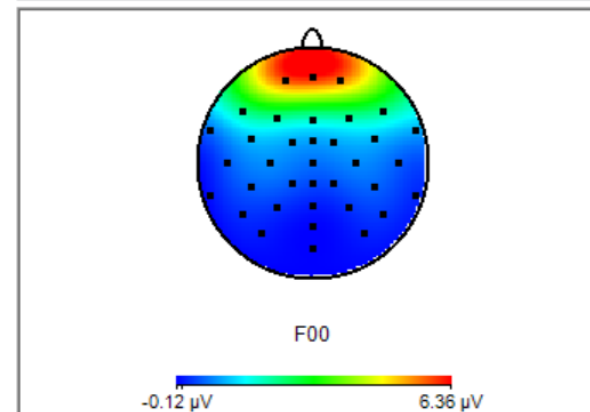


# Inverse ICA - Interactive Mode

Revert

Comments

#	ICA	Fp1	Fpz	Fp2	F7	F3	Fz
#	F00	0.032	0.102	0.035	-0.02	0.012	-0.01
#	F01	-0.047	0.01	0.019	0.034	0.034	-0.01
#	F02	0.009	0.054	-0.042	-0.042	-0.061	0.01
#	F03	0.006	-0.033	0.03	0.005	-0.009	0.01
#	F04	-0.001	-0.017	0.04	-0.045	0.006	-0.01
#	F05	-0.005	0.022	-0.005	-0.035	0.084	0.01
#	F06	0.01	0.011	0.006	-0.055	0.04	-0.01



☐ Show Normed Mappings

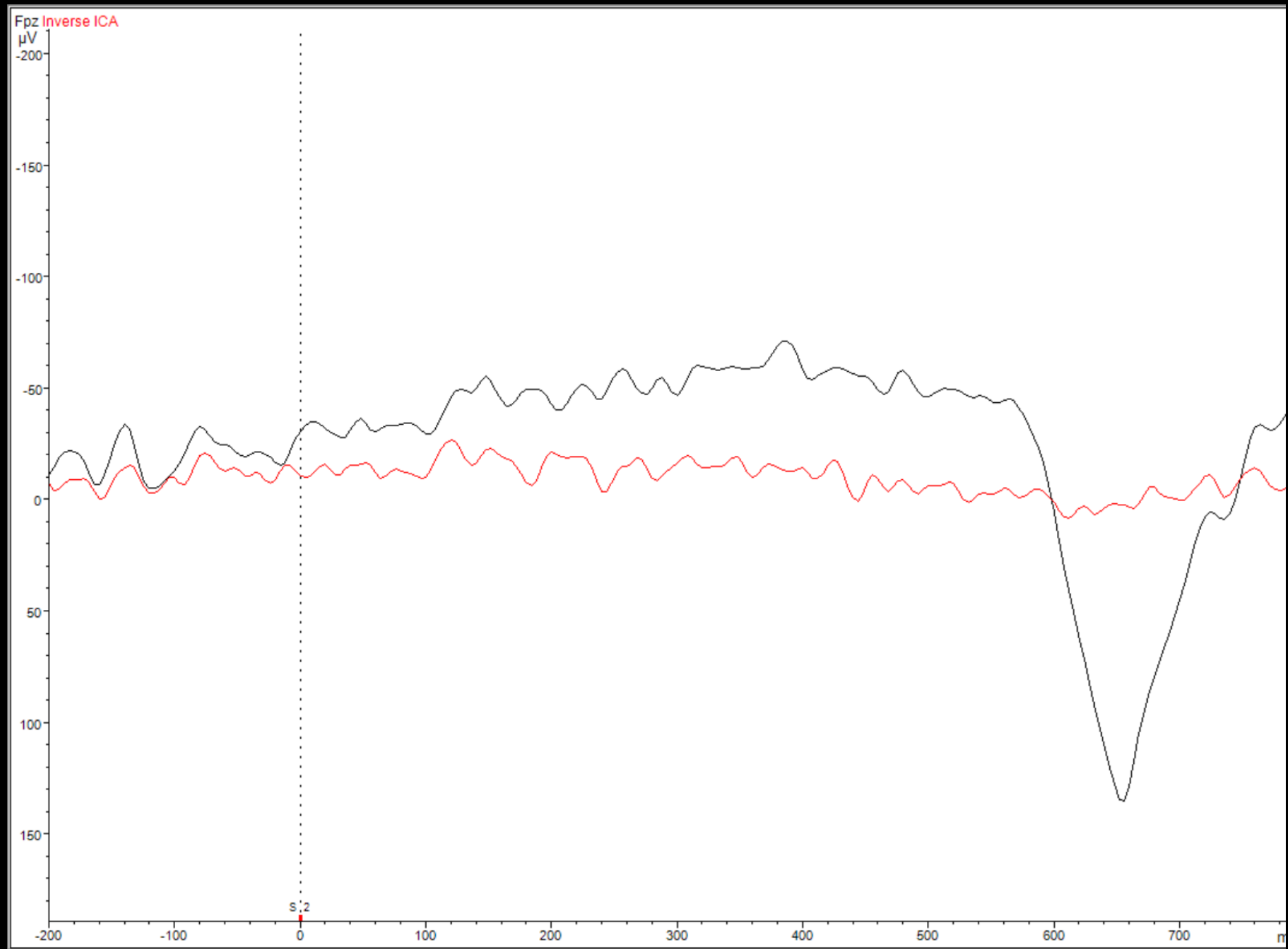
☐ Overlay with Complete Data

ICA Scaling: 100

ICA Components

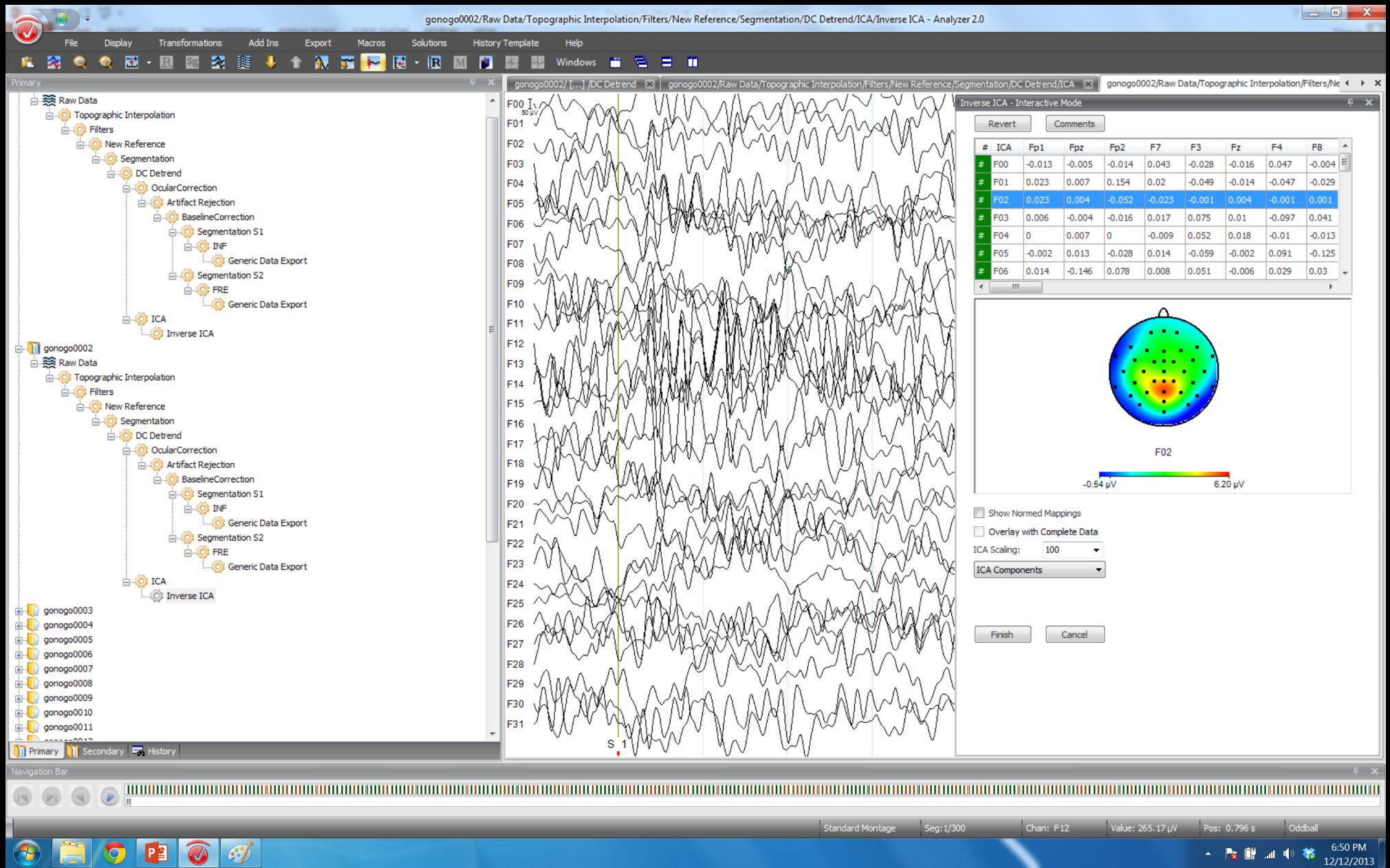
Finish

Cancel



# ICA and EEG

More recently it has been proposed to do the reverse, only keep components with a topography that is of interest.



# Continuous or Segmented Data

ICA will work on continuous or segmented data.

If you want to remove ocular artifacts  
continuous data is fine but you may want to  
train the ICA on a subset of the data.

If you use full data then looking for components  
of interest can be very difficult.

# Continuous or Segmented Data

If you want to use ICA on segmented data you need to ensure you use a sufficient amount of data that does include artifacts.

# ICA Demonstration

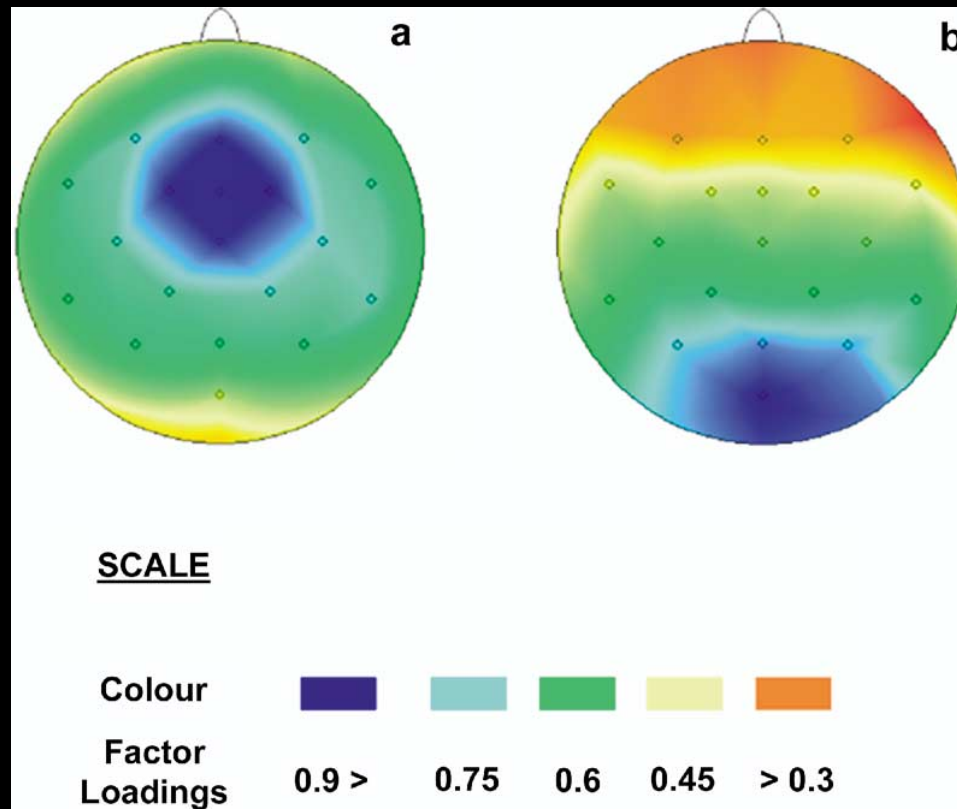
PCA



The data we are working with is very complex and we are making some very broad and simple assumptions to estimate properties of the data and/or correct/adjust/remove bits of the data.

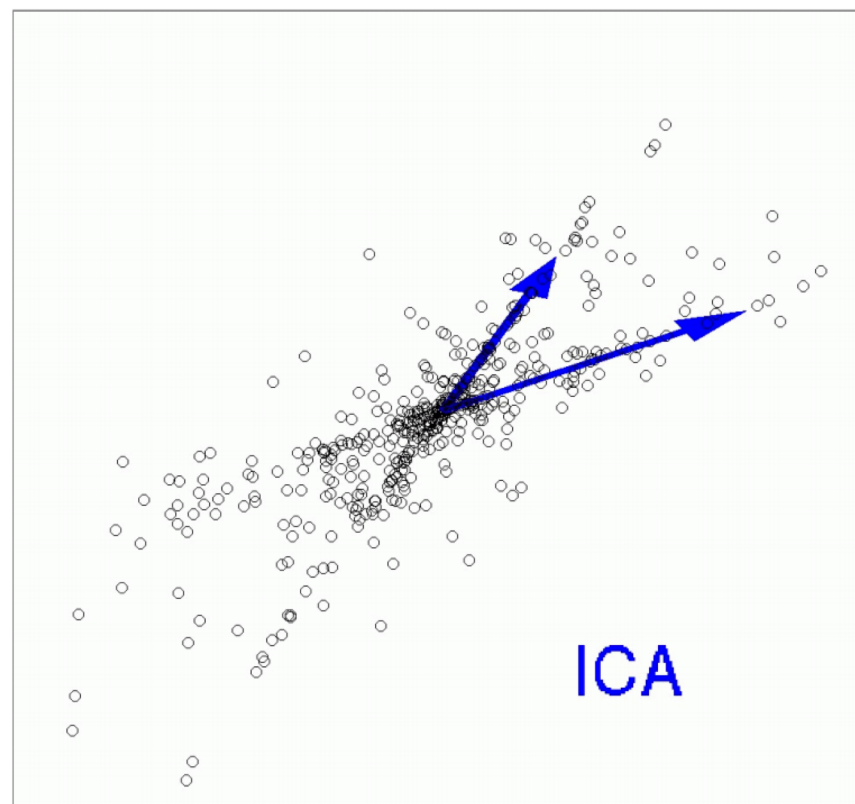
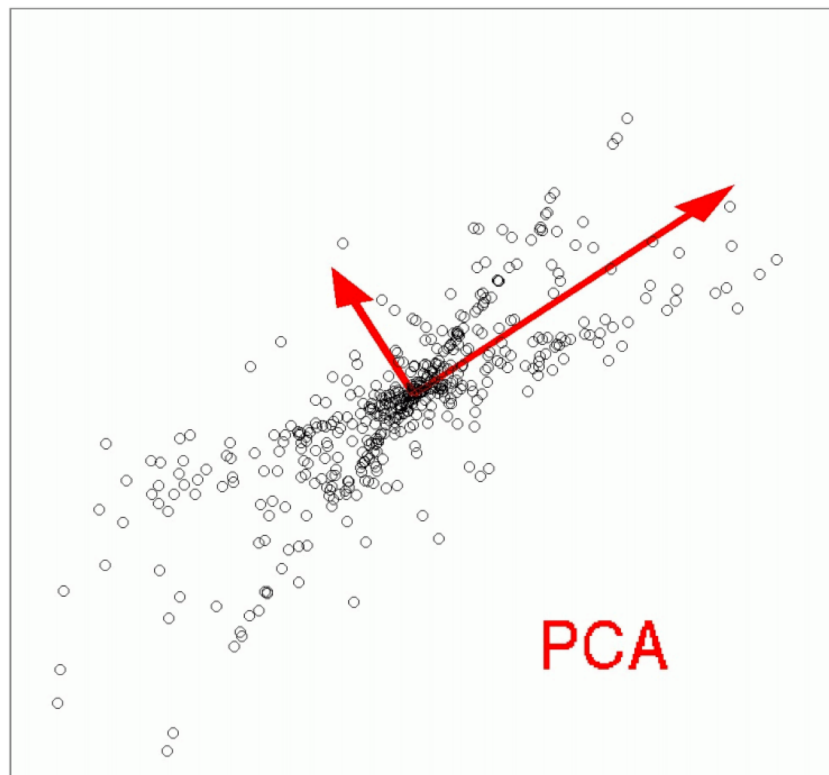
# What do they do?

PCA reduces the data into a series of “components”.



# PCA

Utilizes the first and second moments of the measured data, hence relying heavily on Gaussian features.



# PCA and EEG

PCA is typically used in EEG research to identify spatial, temporal, and/or spatial-temporal components in the data.

At the end of the day...

## DIMENSION REDUCTION

As opposed to having a bunch of channels/time points you have a "spatial component" or a "temporal component"

# Dimension Reduction Before PCA

Channels x Time x Conditions x Participants

4 Dimensions

# Dimension Reduction After Spatial PCA

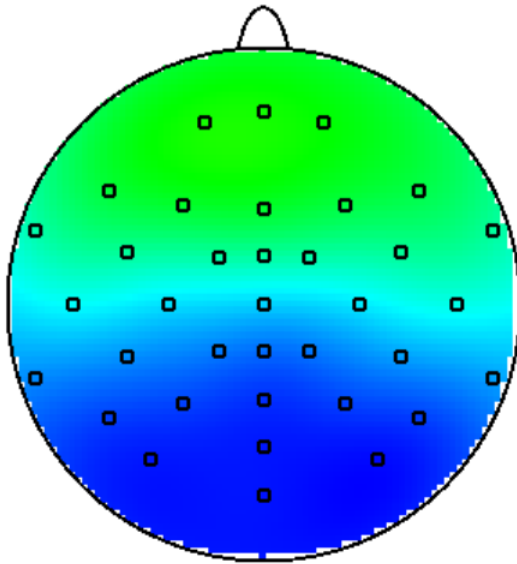
Time x Conditions x Participants

3 Dimensions

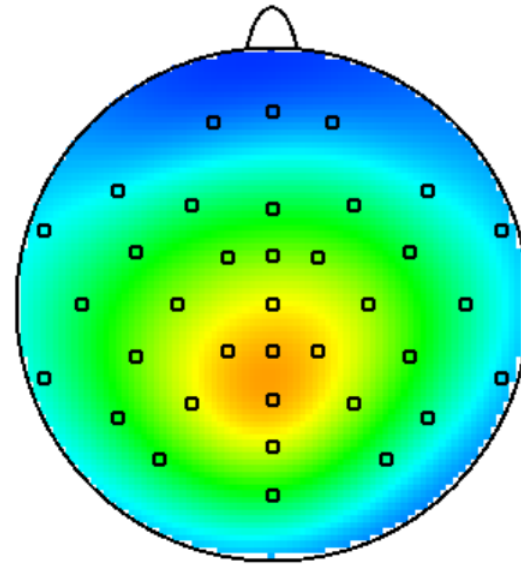
for each Spatial Factor



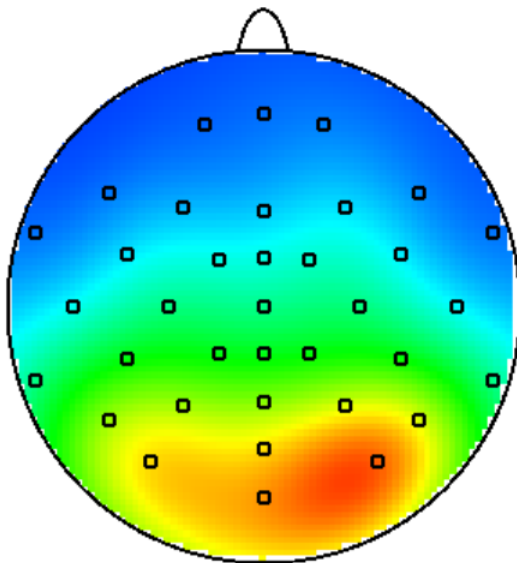
# Spatial Factors



P1



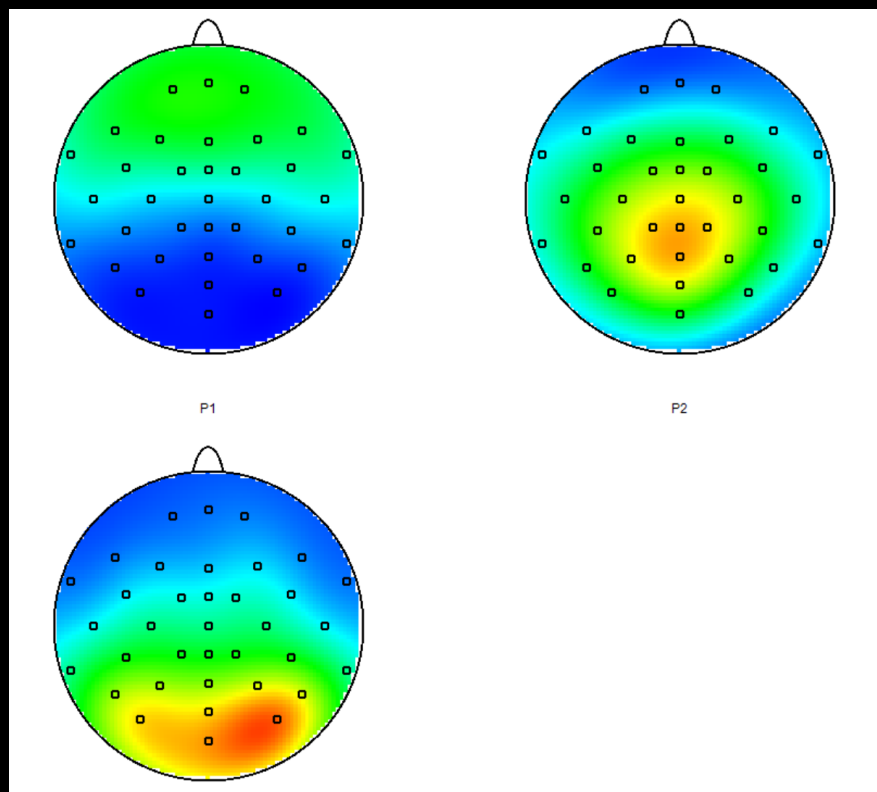
P2



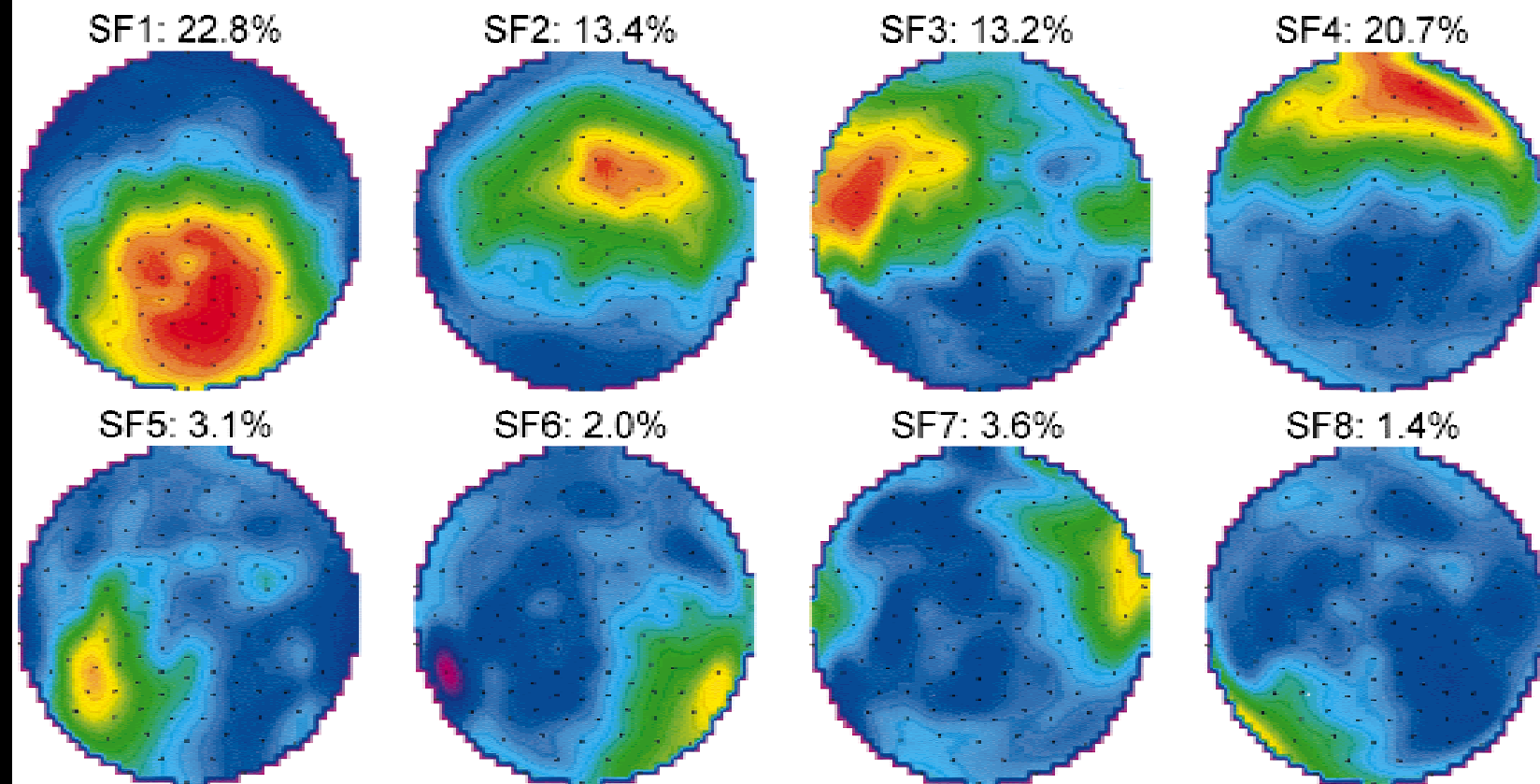
Spatial Factors:  
Each factor will have loadings or weights that when plotted topographically show the spatial components.

# What is a Spatial Component?

Component	Fpz	Fz	FCz	Cz	Cpz	Pz	POz	Oz
1	0.1	0.4	0.9	0.8	0.4	0.3	0.2	0.1
2	0.1	0.1	0.1	0.1	0.4	0.7	0.5	0.2
3	0.9	0.5	0.2	0.1	0.1	0.1	0.1	0.1



# Spatial Factor Loadings (Virtual Electrodes)



# But how is the data reduced?

The component weighting matrix (loadings) are multiplied with the data to create component scores.

Think of it this way, each point in time for each condition for each subject would be weighted by that component relative to the original data value present.

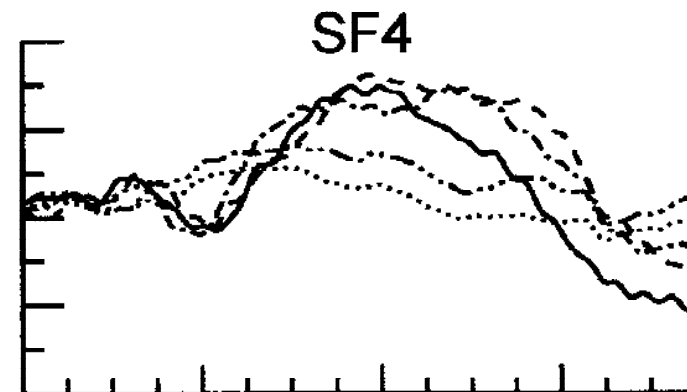
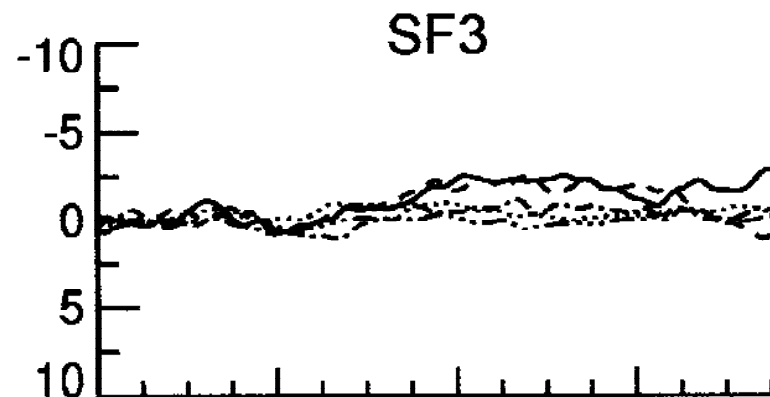
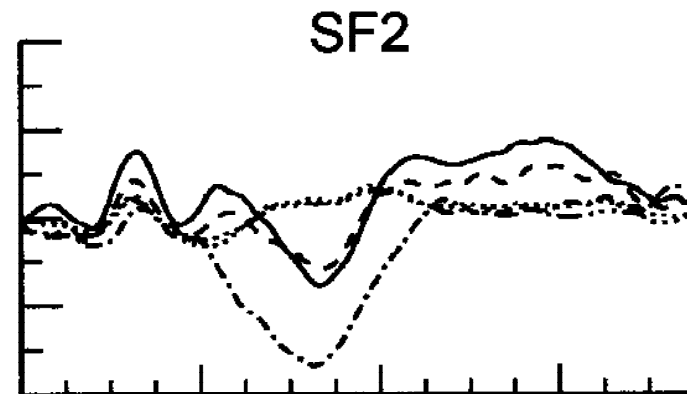
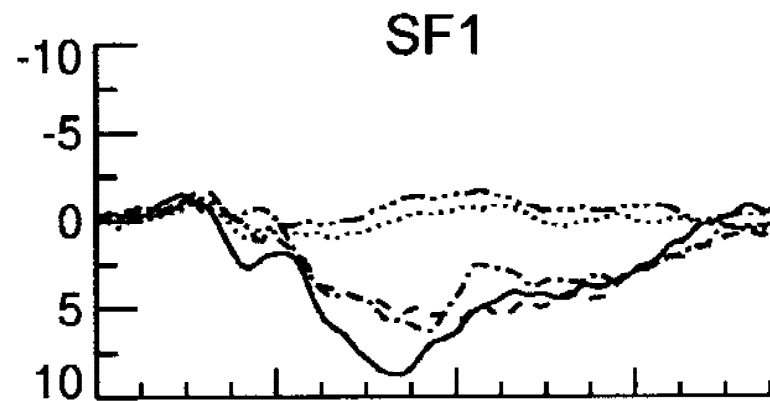
# Virtual ERPs

Once this is done, you can reshape the scores back into the original data format, but with the dimensionality greatly reduced.

## Virtual ERPs: “Attend” Tasks (SF1-8)

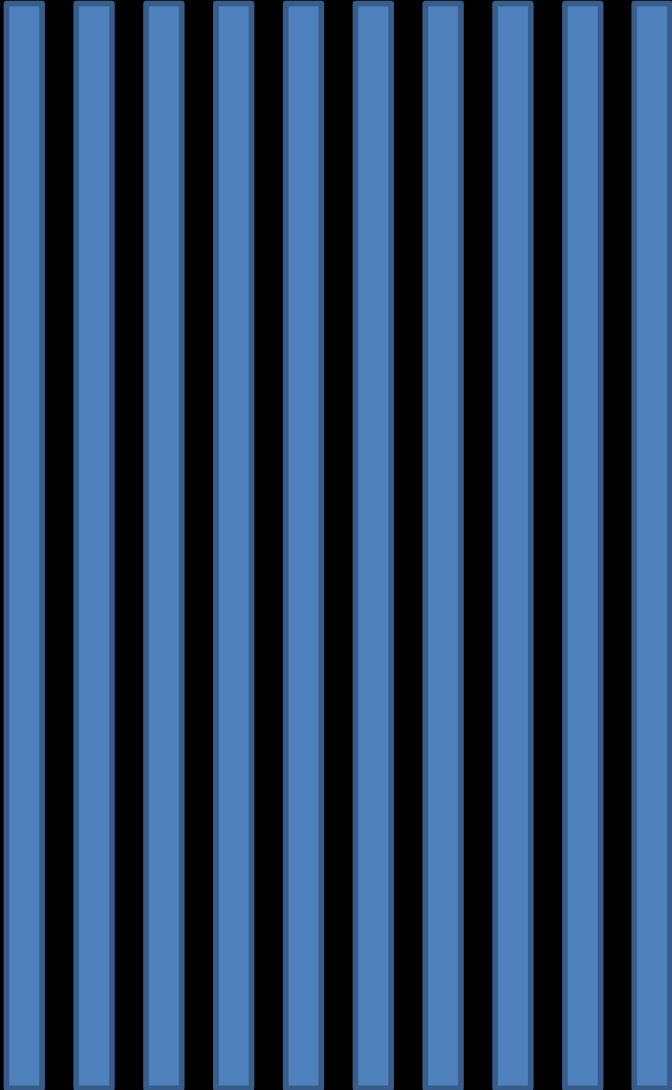
— Classic: Rare  
..... Classic: Freq

--- Novelty: Rare  
- - - Novelty: Novel  
- . - Novelty: Freq



Channels

Time X Conditions X Subjects



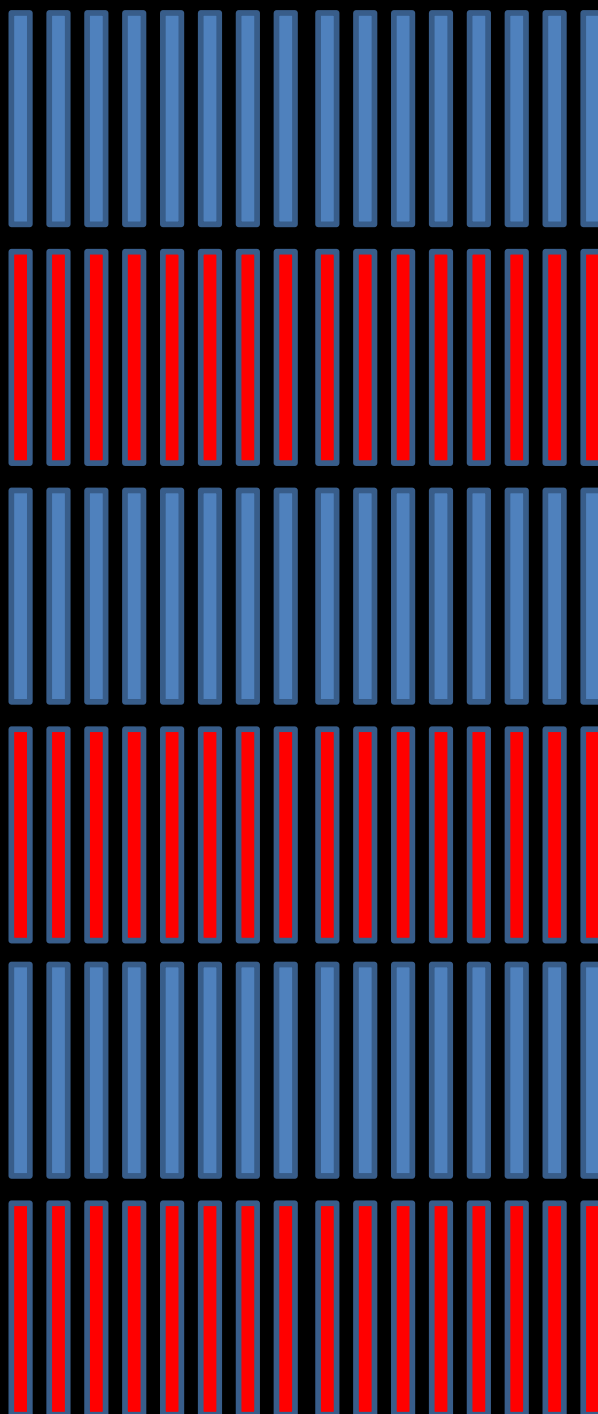
Time X Conditions X Subjects

Channels

Subject 1

Subject 2

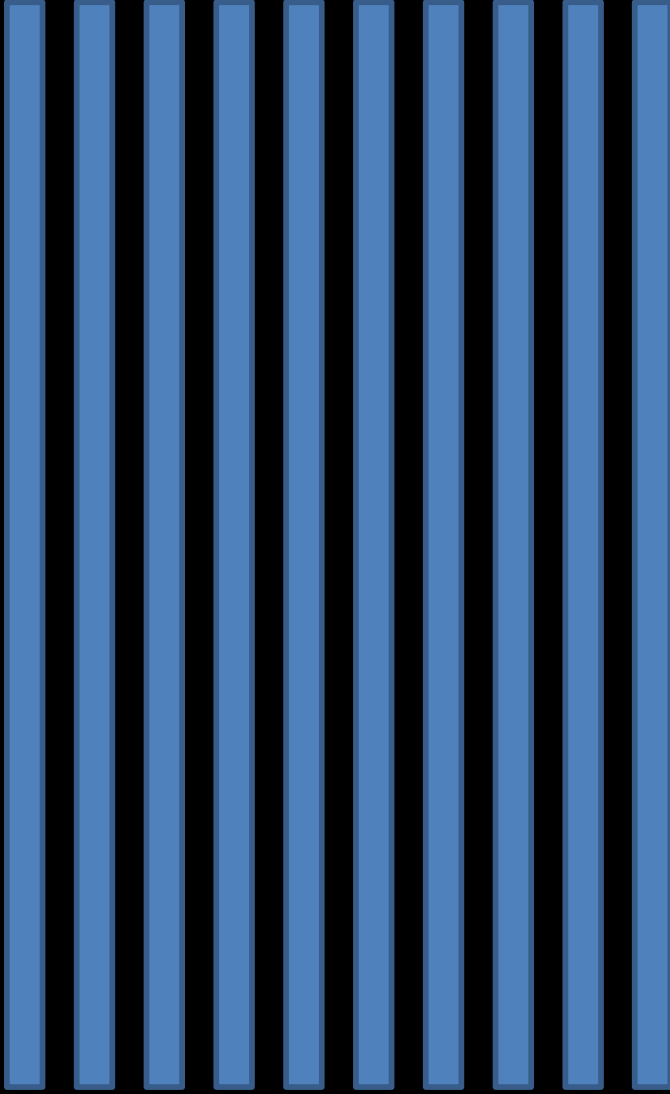
Subject 3





Time X Conditions X Subjects

## Channels



## Components

Time X Conditions X Subjects



Time X Conditions X Subjects

Channels

Subject 1

Subject 2

Subject 3



## Component 1 Scores

Time X Conditions X Subjects

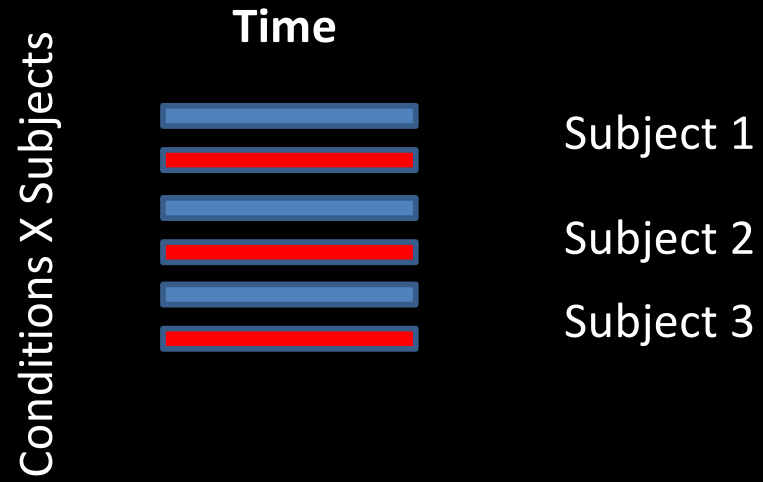


Subject 1

Subject 2

Subject 3

# Virtual ERPs



# Temporal PCA

The same logic as Spatial PCA but with Time Points as the DV.

# Dimension Reduction Before PCA

Channels x Time x Conditions x Participants

4 Dimensions

# Dimension Reduction After Temporal PCA

Channels x Conditions x Participants

3 Dimensions

for each Temporal Factor

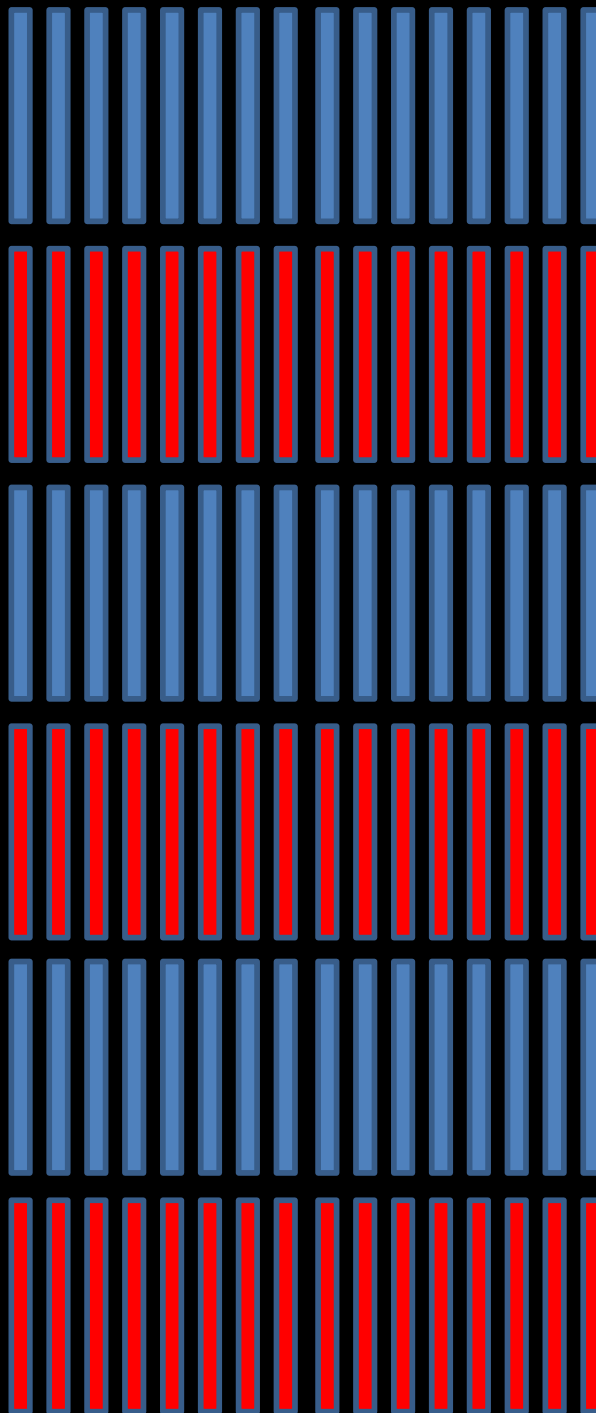
Channels X Conditions X Subjects

Time

Subject 1

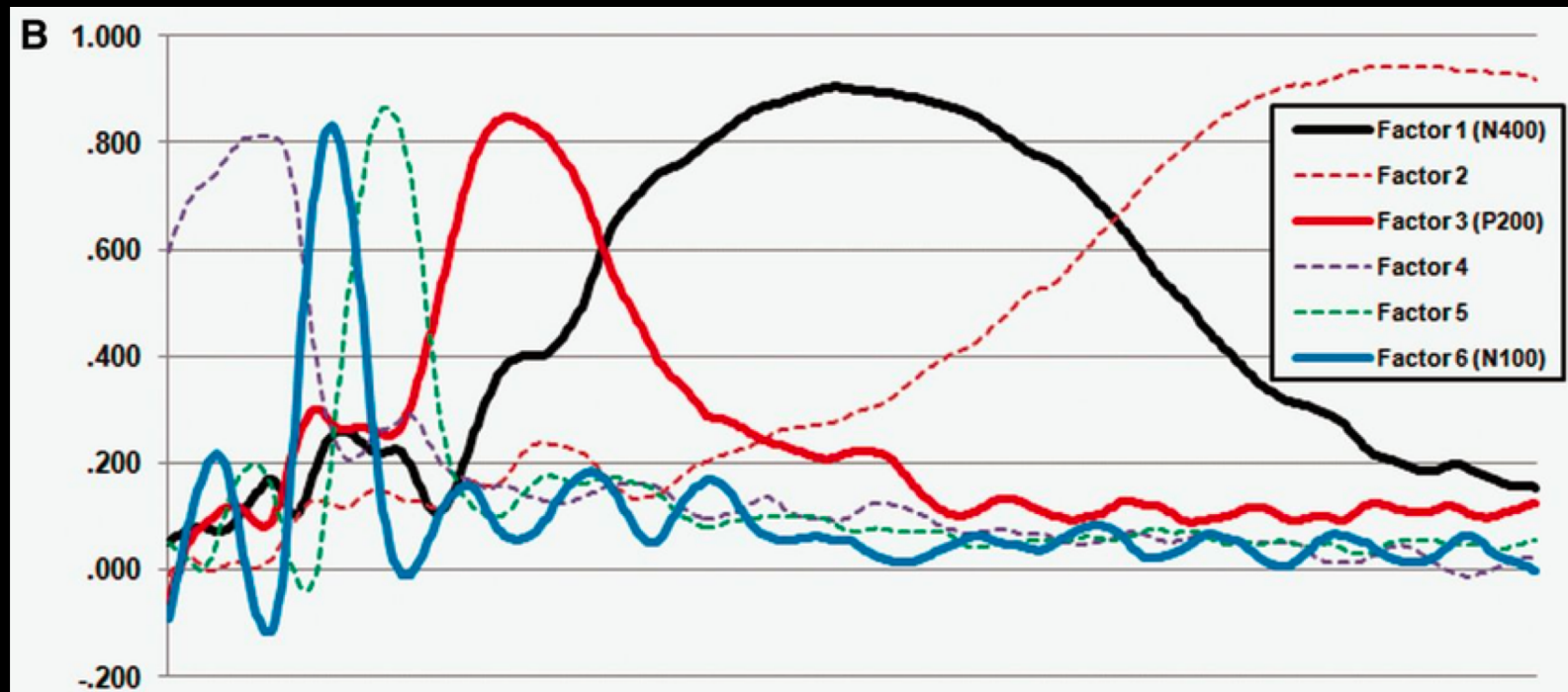
Subject 2

Subject 3





# Temporal Factor Loadings



# Spatial – Temporal or Temporal – Spatial PCA

The logic is simple – you run a PCA on either the spatial or temporal dimension first and then you run a second PCA on the virtual data from one of the factors from the first PCA to reduced the dimensionality further.

## Virtual ERPs



The second PCA would collapse the time dimension to a series of factors.

For each factor you would have a score. For example, a factor might be maximal between 200 to 300 ms.

As such, when you reshape the data one last time, you would have a single score for each participant for each condition. The score reflects the value of a single spatial and temporal component.

## Conditions X Subjects

	Temporal Component 1	Temporal Component 2	Temporal Component 3
Subject 1	■ ■ ■ ■ ■	■ ■ ■ ■ ■	■ ■ ■ ■ ■
Subject 2	■ ■ ■ ■ ■	■ ■ ■ ■ ■	■ ■ ■ ■ ■
Subject 3	■ ■ ■ ■ ■	■ ■ ■ ■ ■	■ ■ ■ ■ ■

Subject 1

Subject 2

Subject 3

## Conditions X Subjects

- Temporal Component 1

Subjects

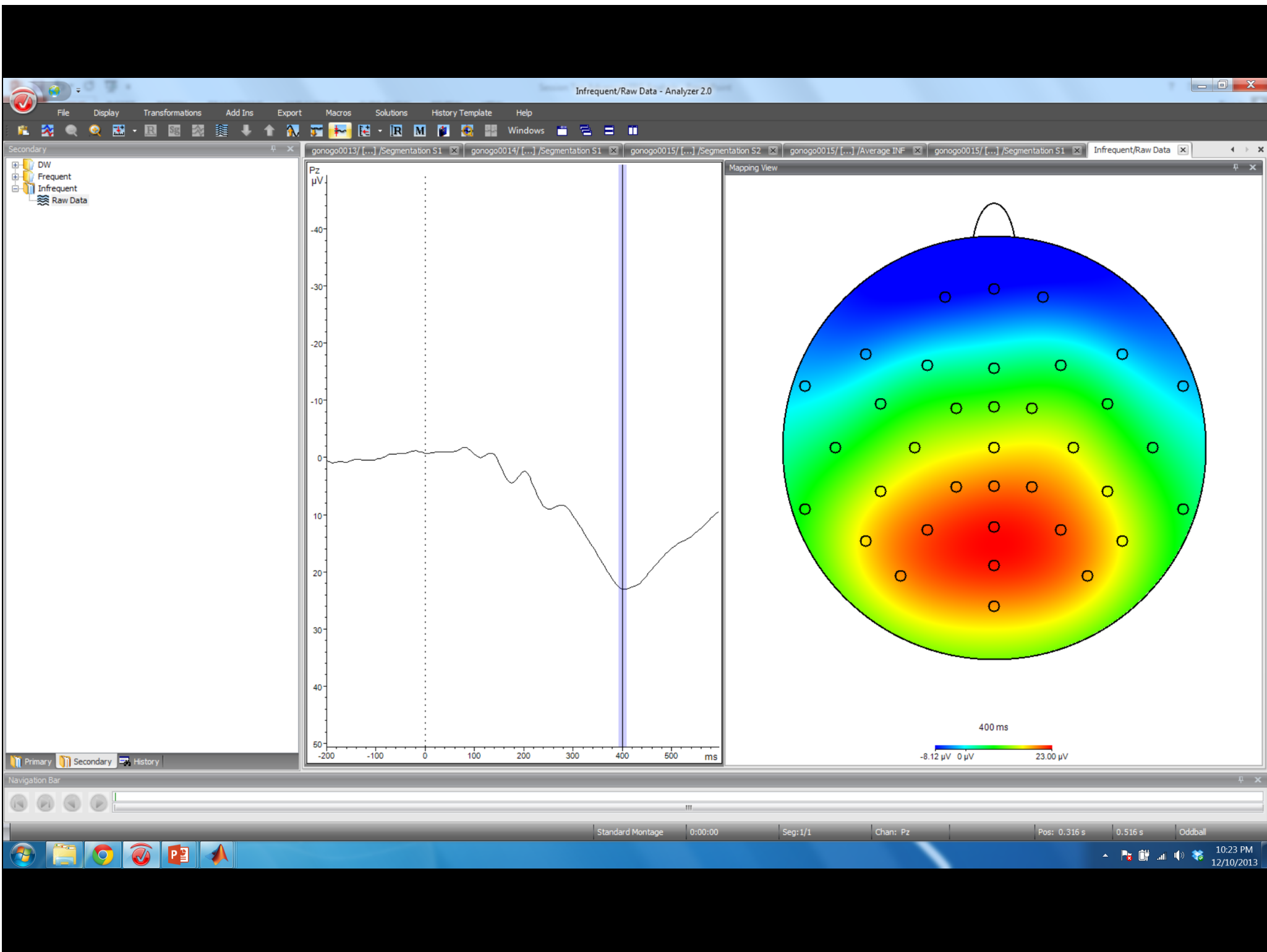
Conditions



# Source Analysis and a Few Other Tricks

# Source Analysis





# The Inverse Problem

Simply put, because we do not know the number of dipoles at any given point in time, there are an infinite number of potential solutions for a given scalp topography.

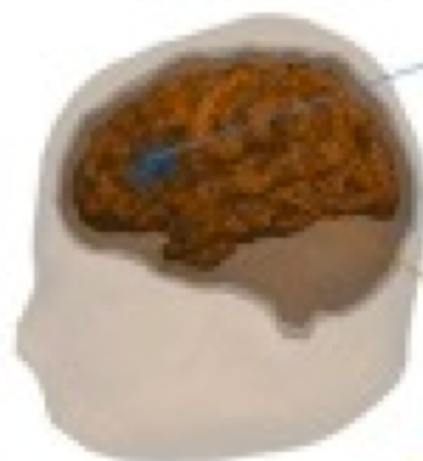
Source localization techniques try to get around this by using certain assumptions.

Forward problem: data generation

$$X = G \cdot S$$

head model

lead field matrix:  $G$



source signals

signal matrix:  $S$

activity of interest

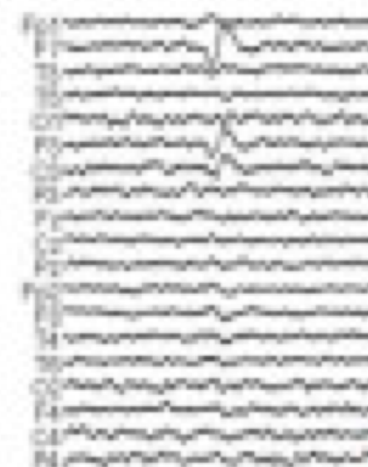


background activity



EEG

data matrix:  $X \in \mathbb{R}^{N \times K}$



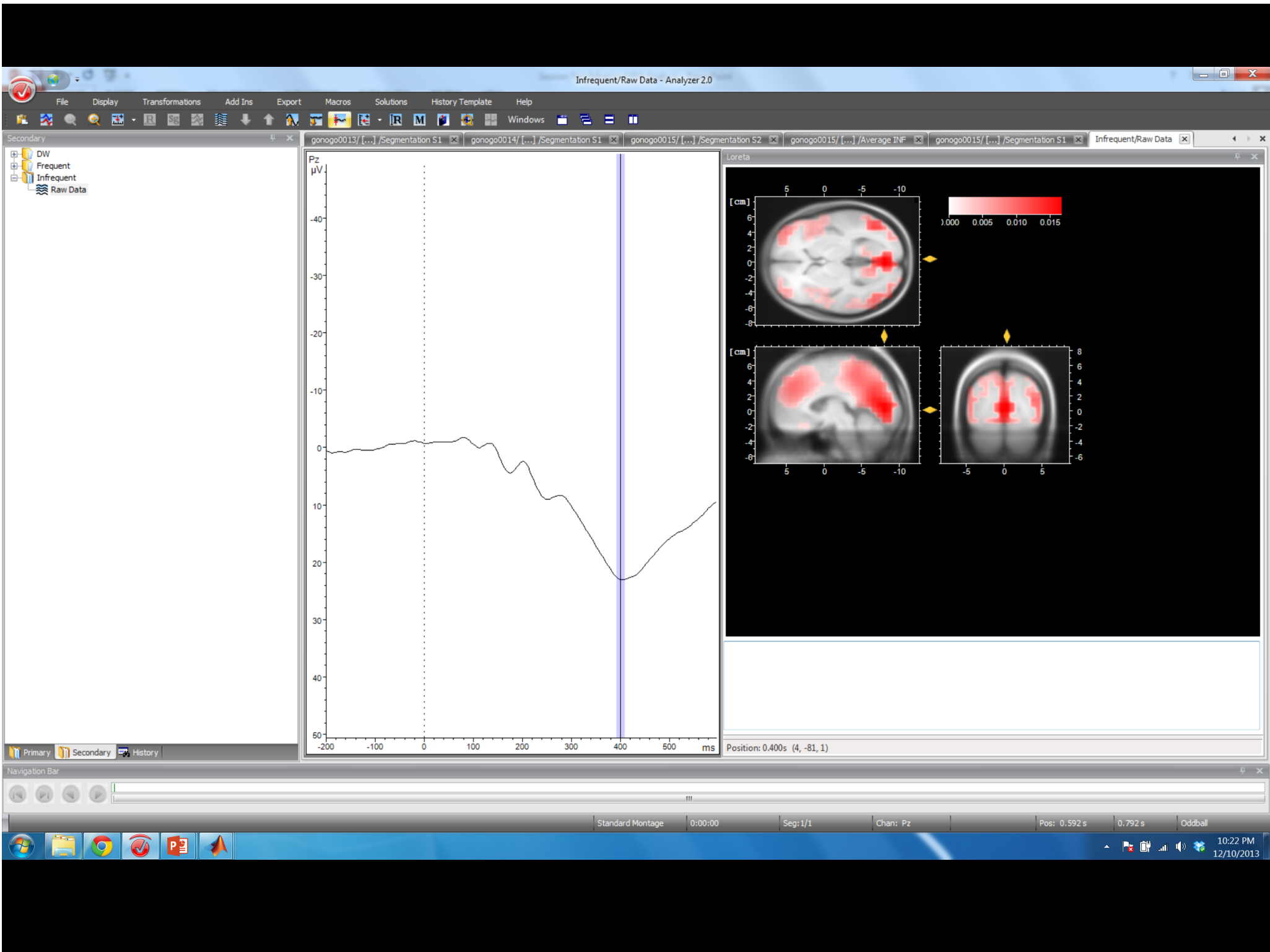
$$S = ?$$

Inverse problem: source imaging

# Discrete vs Distributed Source Solutions

Discrete source solutions focus on dipoles.

Distributed source solutions use dipoles but “consider” whole brain activity.



# Loreta

The mathematics behind these techniques is beyond the scope of what we are covering here.

In a nutshell, Loreta divides the brain into a number of voxels each containing three perpendicular dipoles (X, Y, Z) and then varies the parameters of these dipoles for each voxel to obtain the observed scalp topography.

# Loreta

There are still multiple solutions. So, the determined solution is the one that has the best smoothness – contiguous voxels should have gradual changes in dipole strength.

There are additional constraints such as perpendicular dipoles only, minimum overall source magnitudes, and more...

# Be careful...

Loreta has a strong bias towards the midline and towards the cortical surface.

When using a source technique, do not use it like fMRI. At best, it is a confirmation of something we already know from other lines of research.

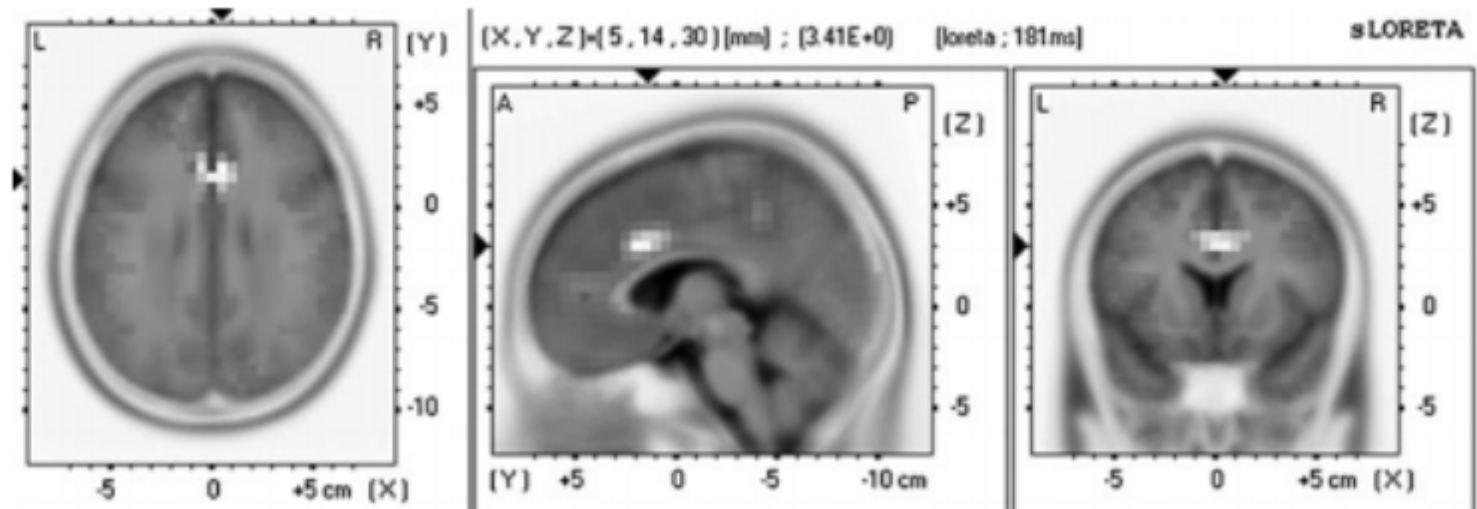


# sLoreta and eLoreta

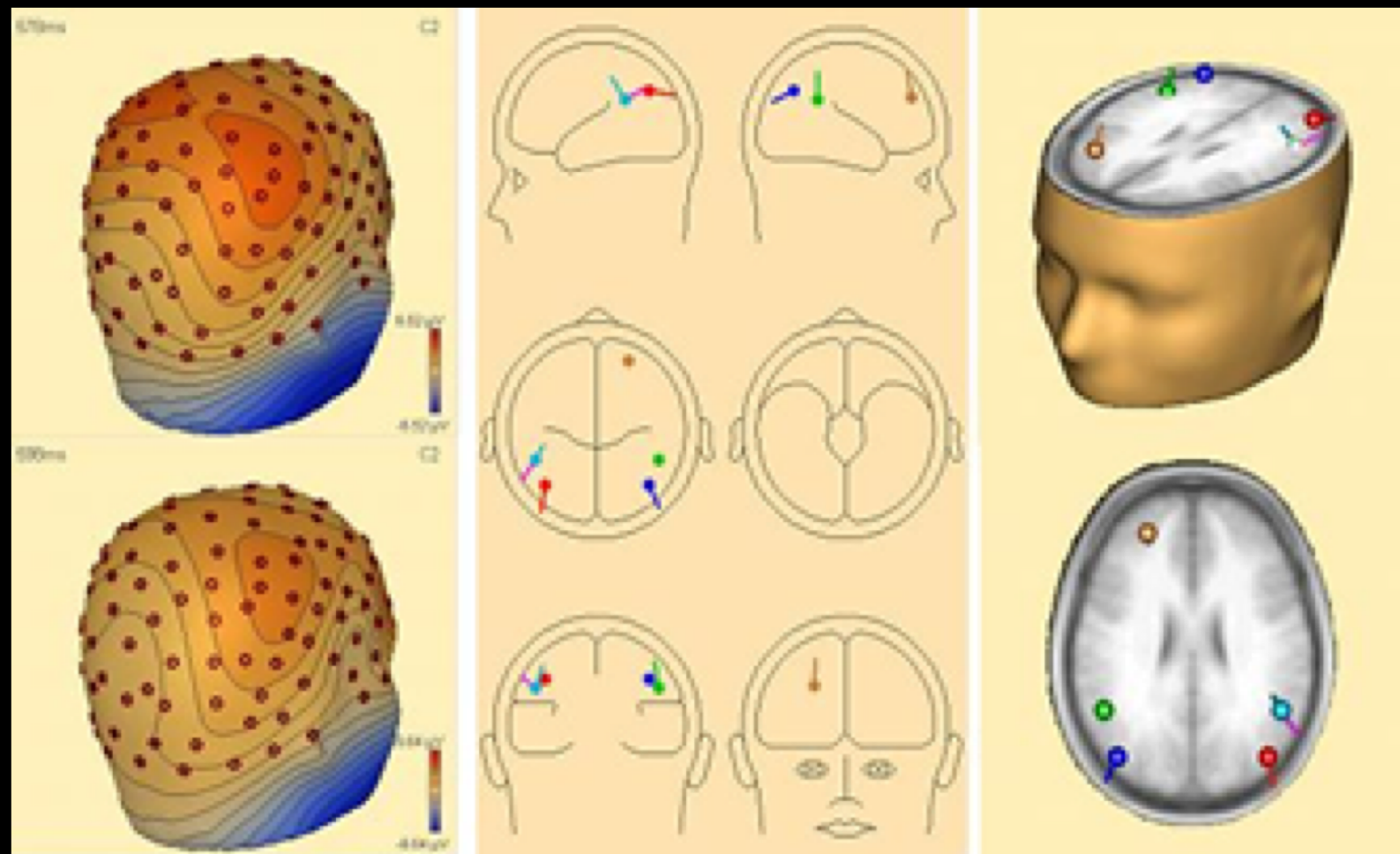
<http://www.uzh.ch/keyinst/loreta.htm>

# Reporting Source Data

**Fig. 4** sLORETA source analysis of self-ownership, as compared with other-ownership, cues at 190–240 ms post-cue-onset. Statistical non-parametric mapping (SnPM) at a significance level of .05 revealed differences localized in Brodmann area 24 (sLORETA value = 34.1) within the cingulate gyrus

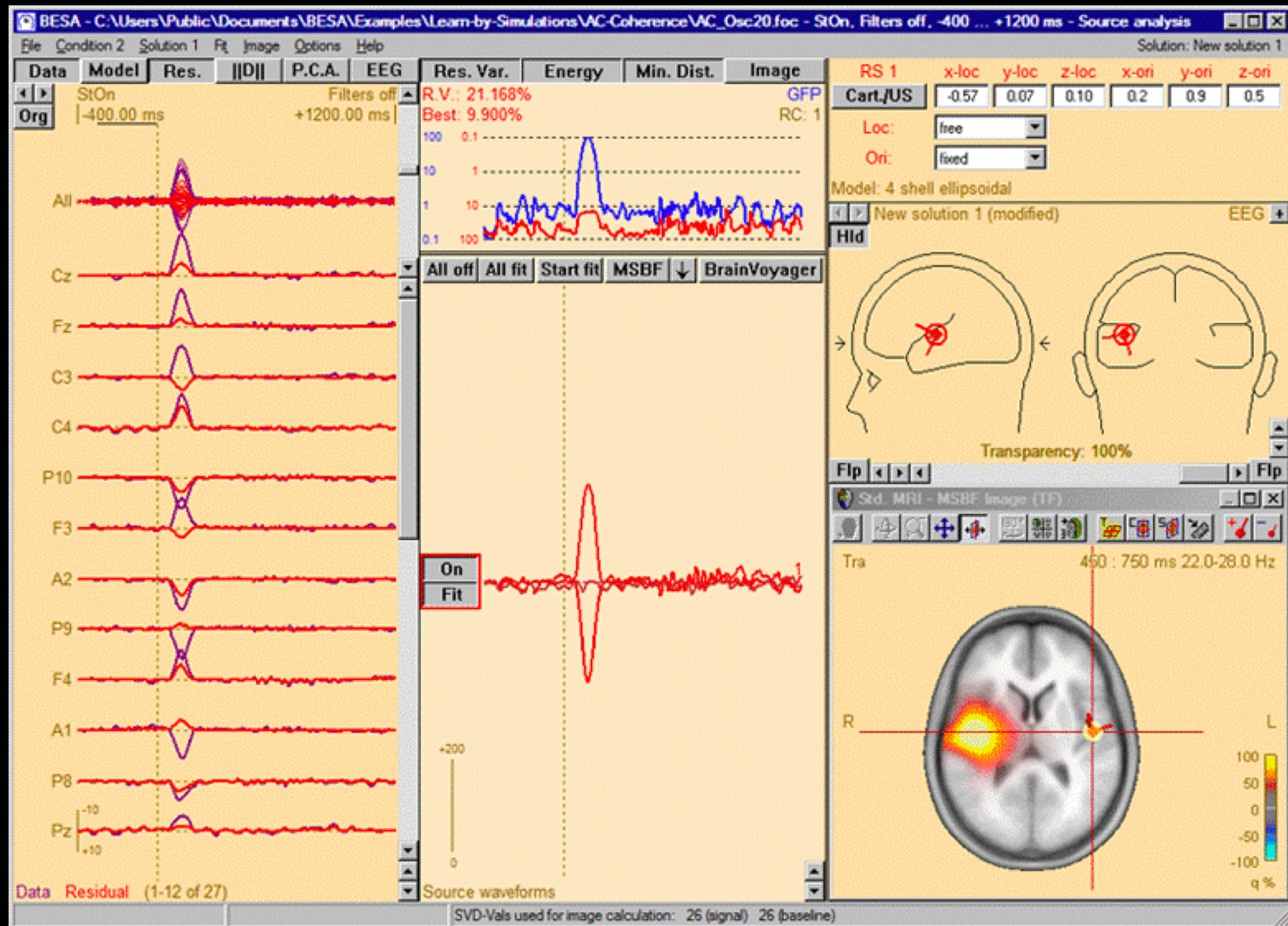


# BESA





# BESA



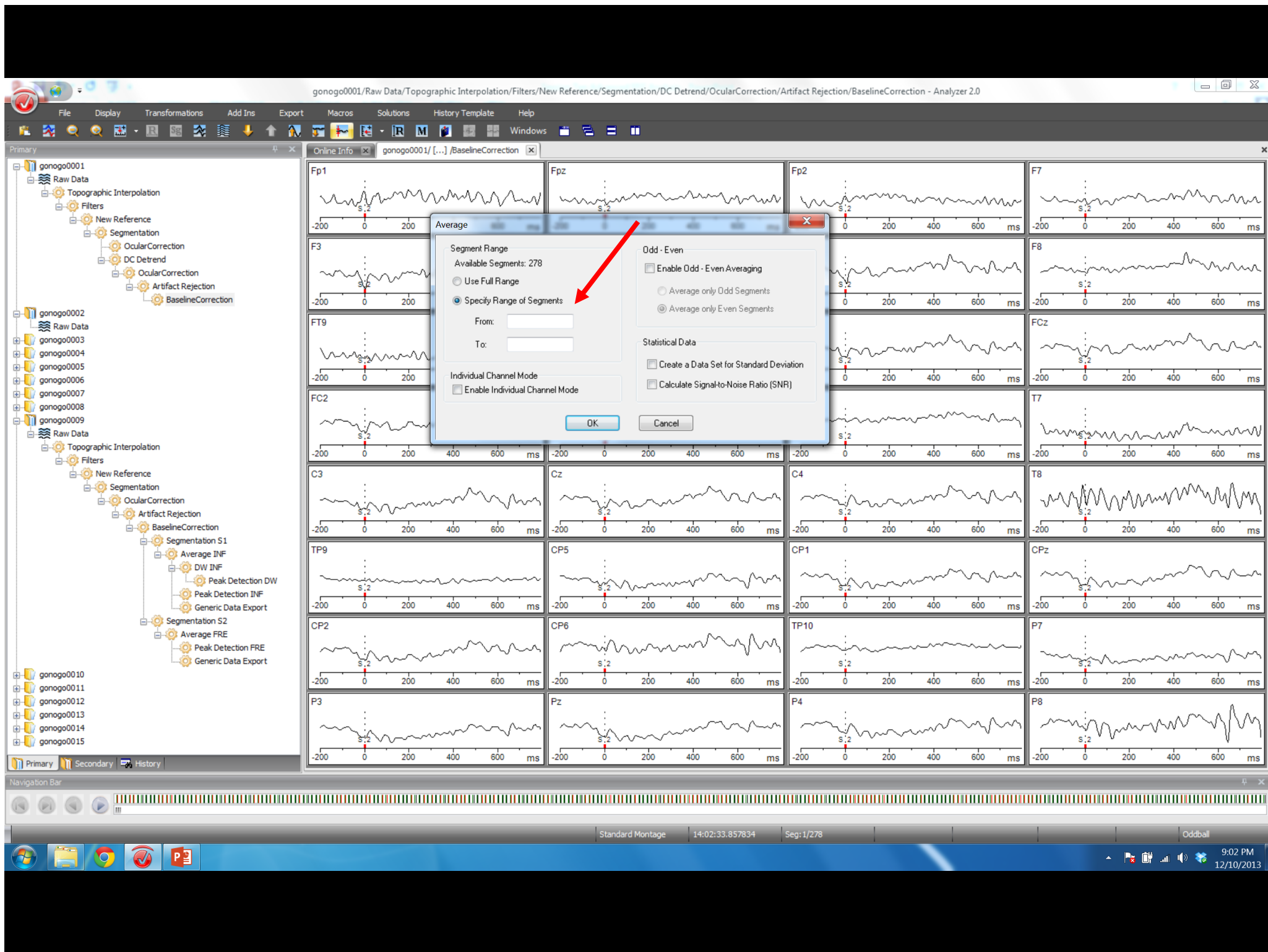
What to do?

# Bins and Running Averages

# Bins

You do not always have to collapse all of your markers across and entire experiment.

You may want to look at learning effects for instance.





# The Oddball Data

For instance, you have 60 infrequent markers, so:

Bin 1 = Segments 1 to 30, Bin 2 = 31 to 60

In a like manner, for the 240 frequent markers:

Bin 1 = Segments 1 to 120, Bin 2 = 121 to 240

You could then do your peak detection to compare the first half of the experiment.

Obviously, the same logic could be applied to any other break down of the average.

# Running Averages

The principle behind a running average is somewhat similar to bins, but somewhat more “smooth”.

As opposed to creating 2 bins (or more bins) you create a series of bins.

If you have 100 segments, you create the following bins:

Average 1 = Segments 1 to 50

Average 2 = Segments 2 to 51

Average 3 = Segments 3 to 52

etc

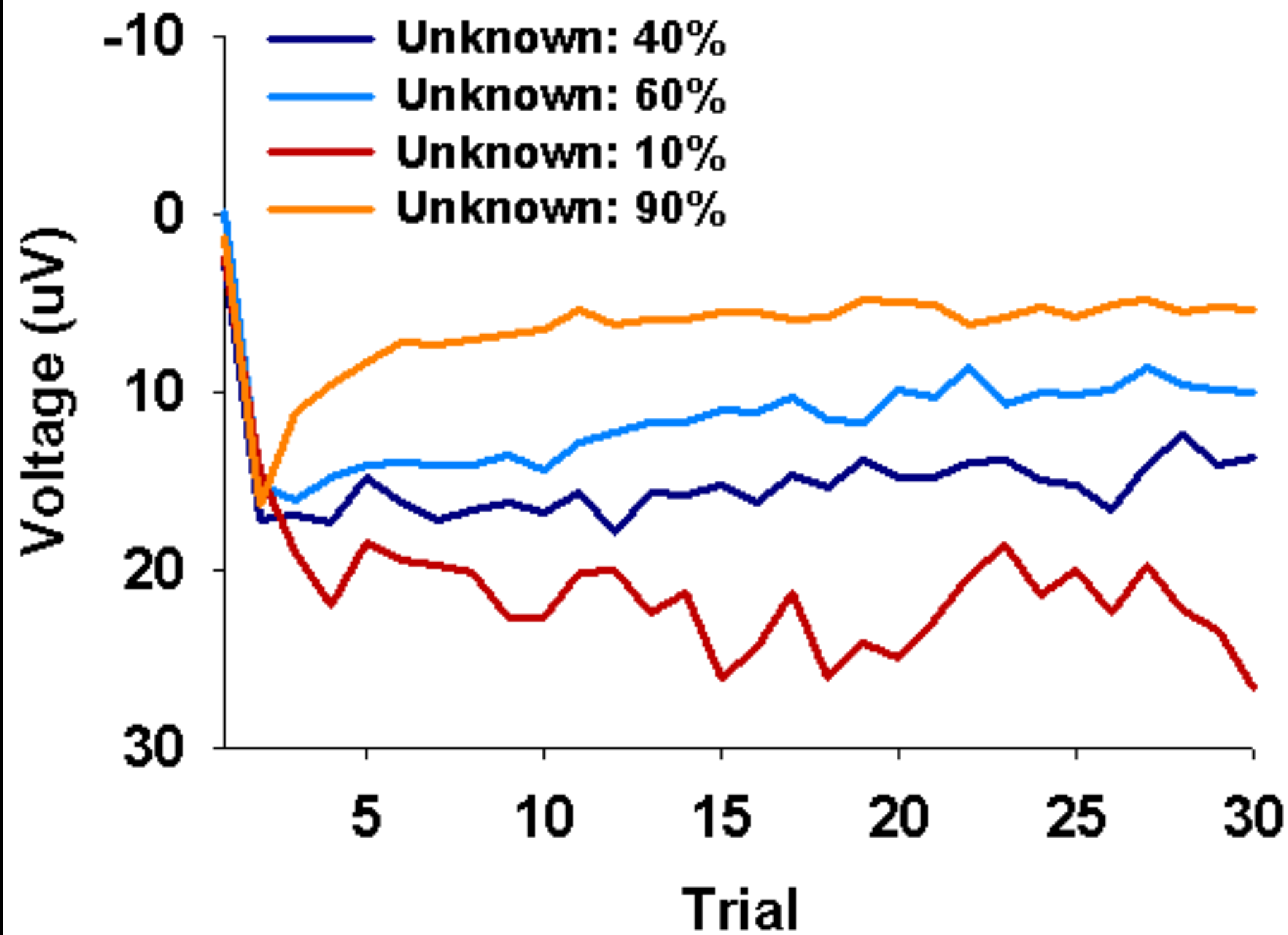
# Running Averages

To be able to do this, you need to work with MATLAB.

But, its not as hard as you think.

Analyzer has a MATLAB interface so you can export directly to MATLAB and then bring the data back in.

# Single Trial Analysis



# So how do you do it?

## Strategy One

Focus on the P300 – its large and easy to see.

Similar to Running Averages, calculate the a peak voltage - the mean – but do this for every trial and not on the average of a subset of trials.

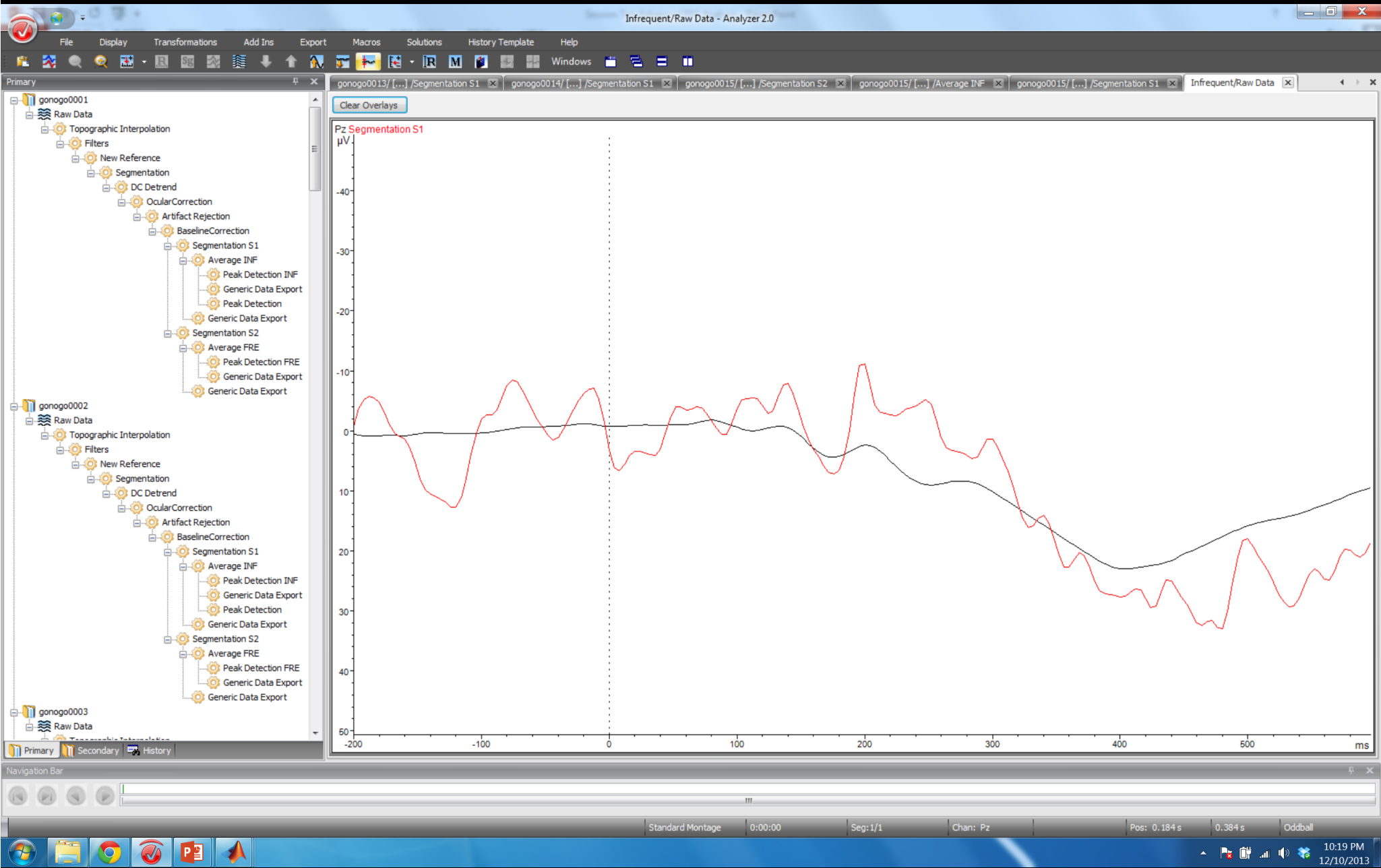
You can then plot this as a function of time.

# So how do you do it?

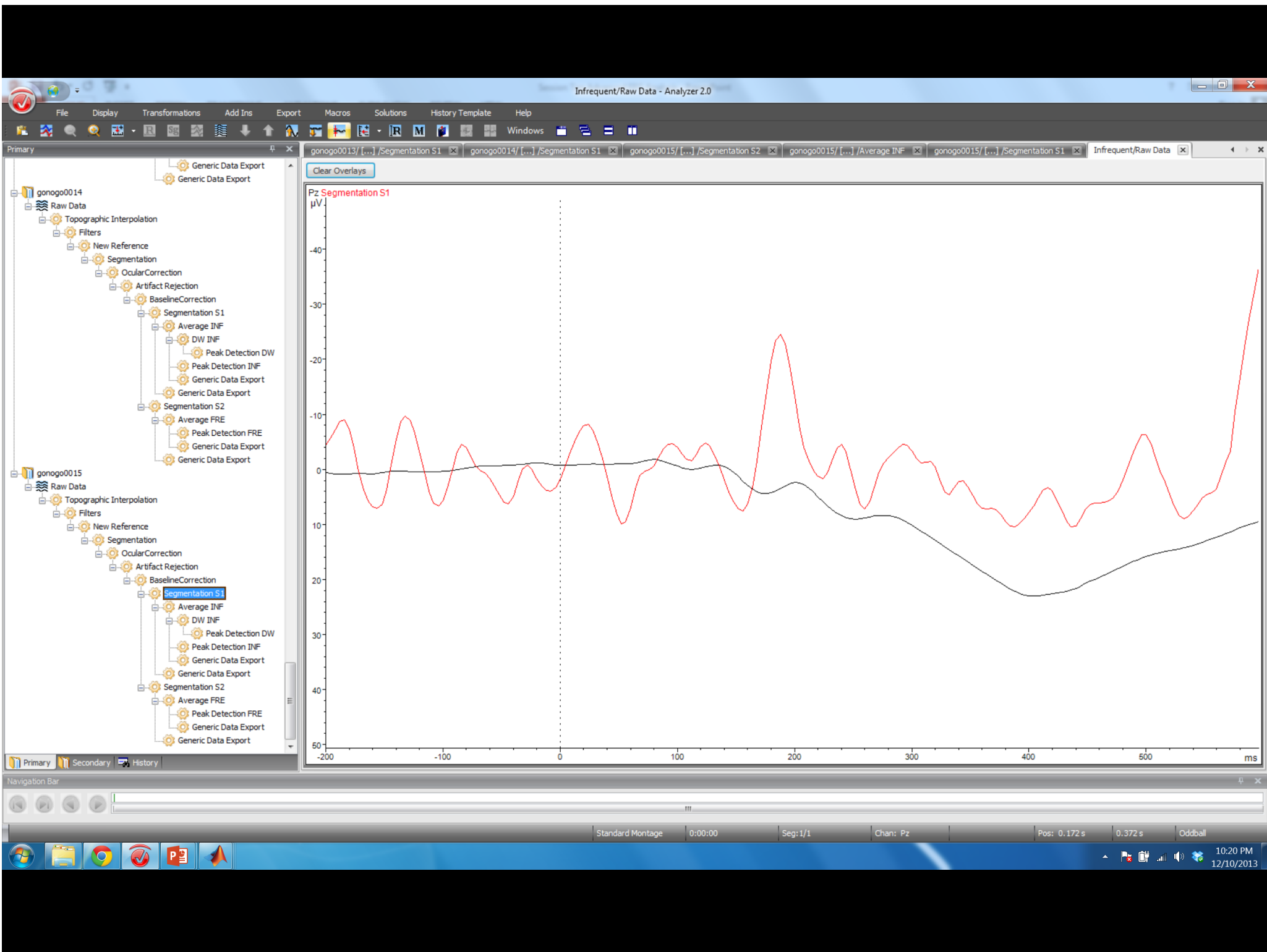
## Strategy Two

Fit a function to the component of interest on the grand average.

Estimate the quality of fit for each individual trial.







# Component Latency and Onset Analysis

# Component Latency

It is relatively easy to do a component latency analysis.

All you need to do is use a maxima / minima peak detection approach and get a latency value for the peak for each subject for each condition.

Following that, traditional null hypothesis testing can be used.

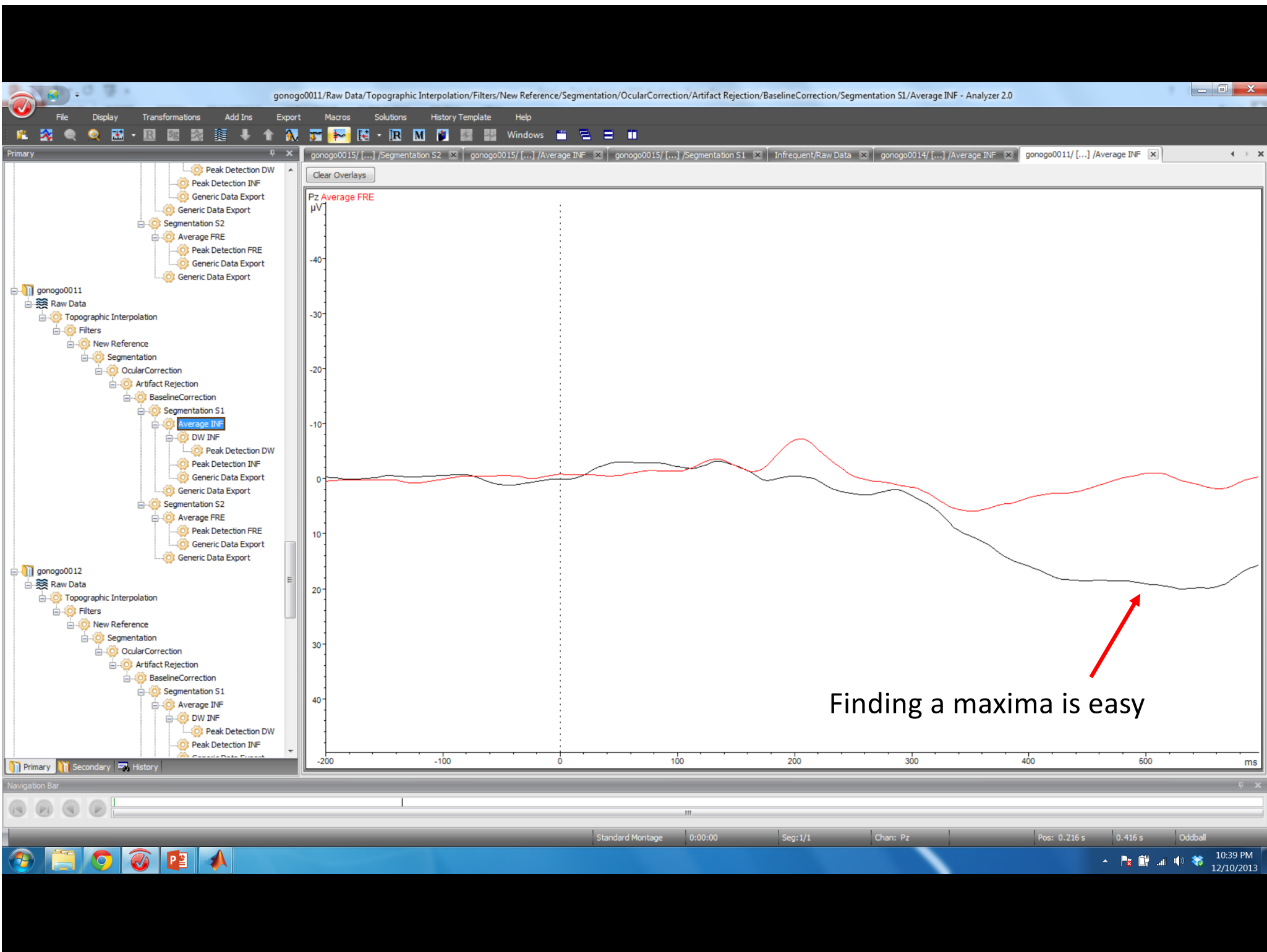
Condition A			Condition B	
12.4 $\mu\text{V}$	324 ms		11.3 $\mu\text{V}$	378 ms
9.8 $\mu\text{V}$	333 ms		10.4 $\mu\text{V}$	402 ms
10.3 $\mu\text{V}$	356 ms		8.8 $\mu\text{V}$	444 ms
4.2 $\mu\text{V}$	401 ms		7.6 $\mu\text{V}$	424ms
7.8 $\mu\text{V}$	367 ms		6.4 $\mu\text{V}$	401 ms

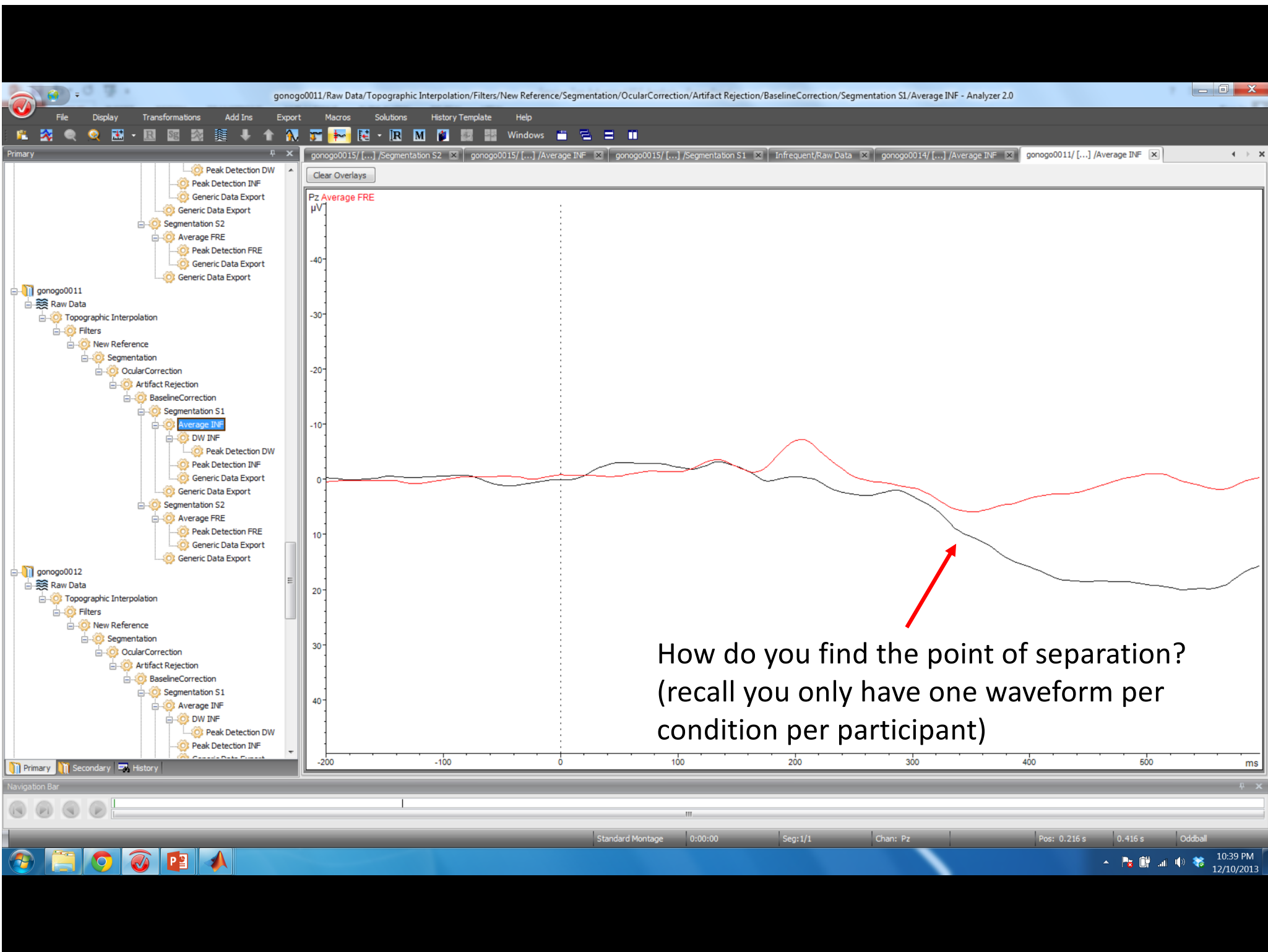
# Component Onset

Component onset is a bit trickier.

Why?

Because we cannot do single trial analysis very effectively, we cannot get an onset value for each subject based on their own data. Why is that?





# One Solution

Rodrigues – Fornells

1. Get the average waveform for each participant for the channel of interest for each condition.
2. Use a running average to smooth these, say  $\pm 5$  time points on either side.
3. Sequentially t-test every point of the smoothed waveforms till there is a significant difference using a stricter alpha (0.001)



# The Problem

This approach will tell you when two conditions diverge, it does not allow you to compare the difference of onset between A and B with the difference in onset between C and D.

# Pooling Electrodes

