Robust Linear Modelling of EEG data: the LIMO EEG plug-in

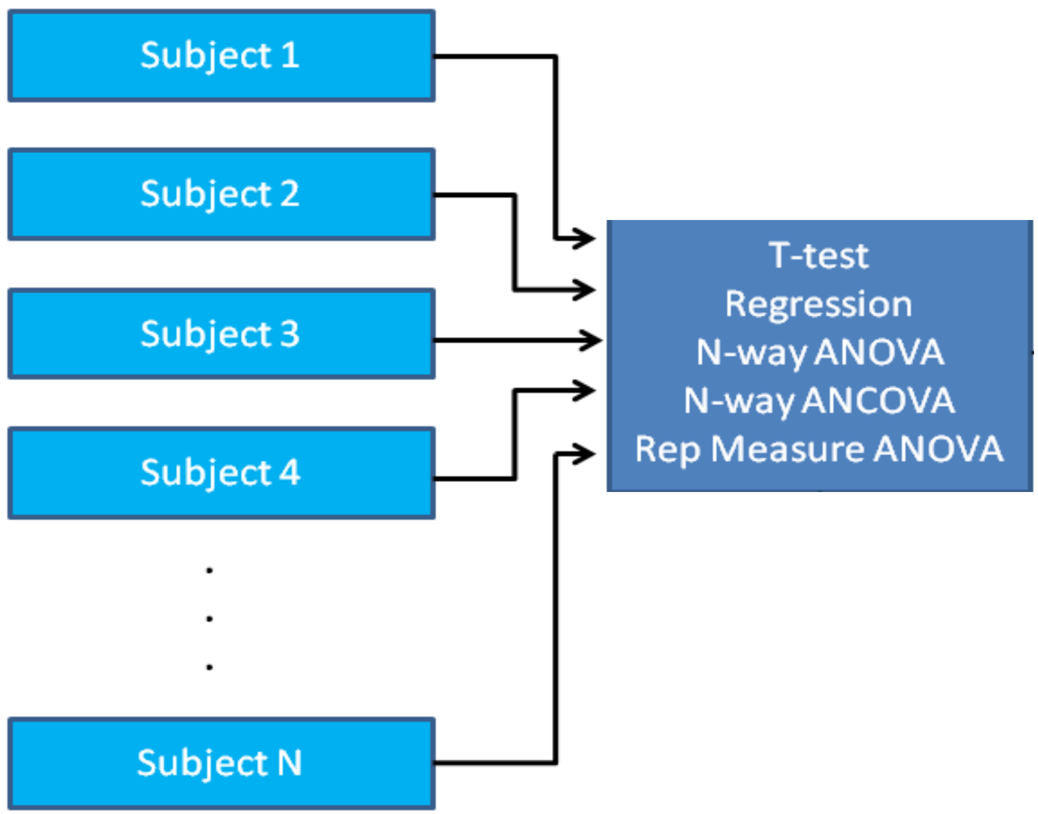
Arnaud Delorme

with Cyril Pernet, PhD, Edinburgh Imaging & Centre for Clinical Brain Sciences

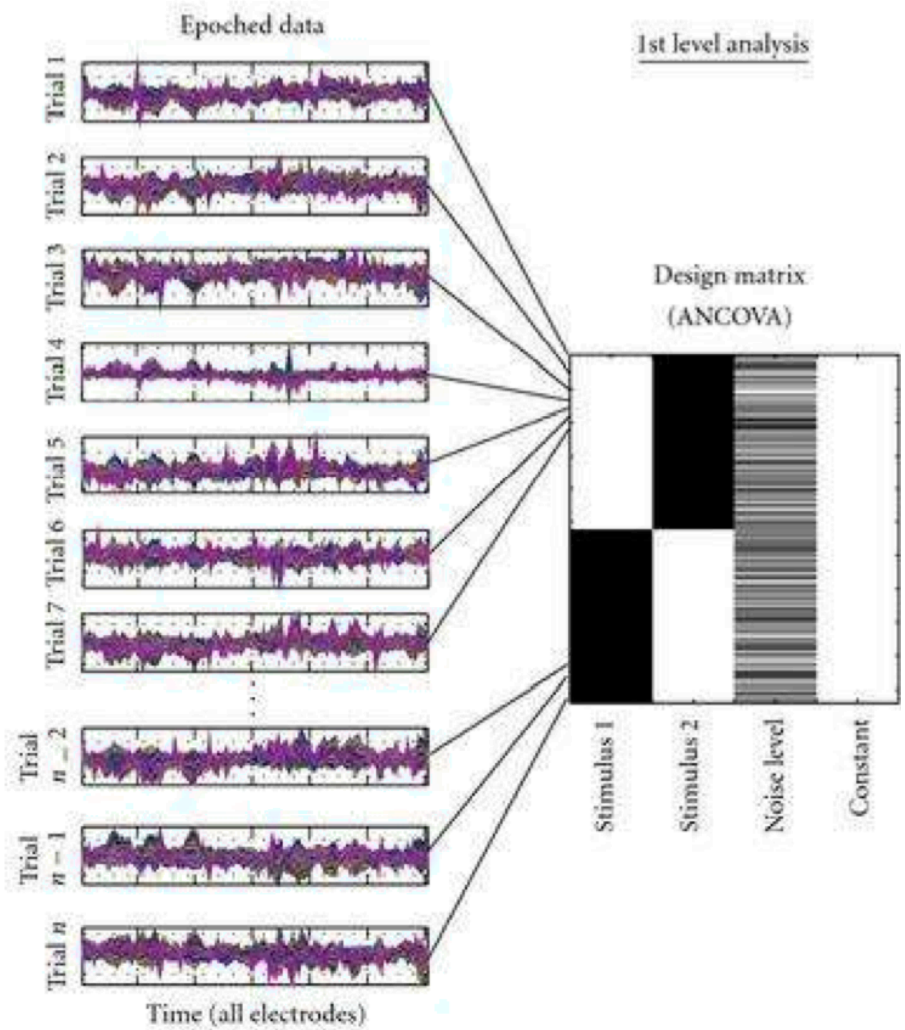
Hierarchical Linear Model

1st level analysis:
GLM: $Y = X\beta + \varepsilon$
→1 β per column of X
(= within subject effects)

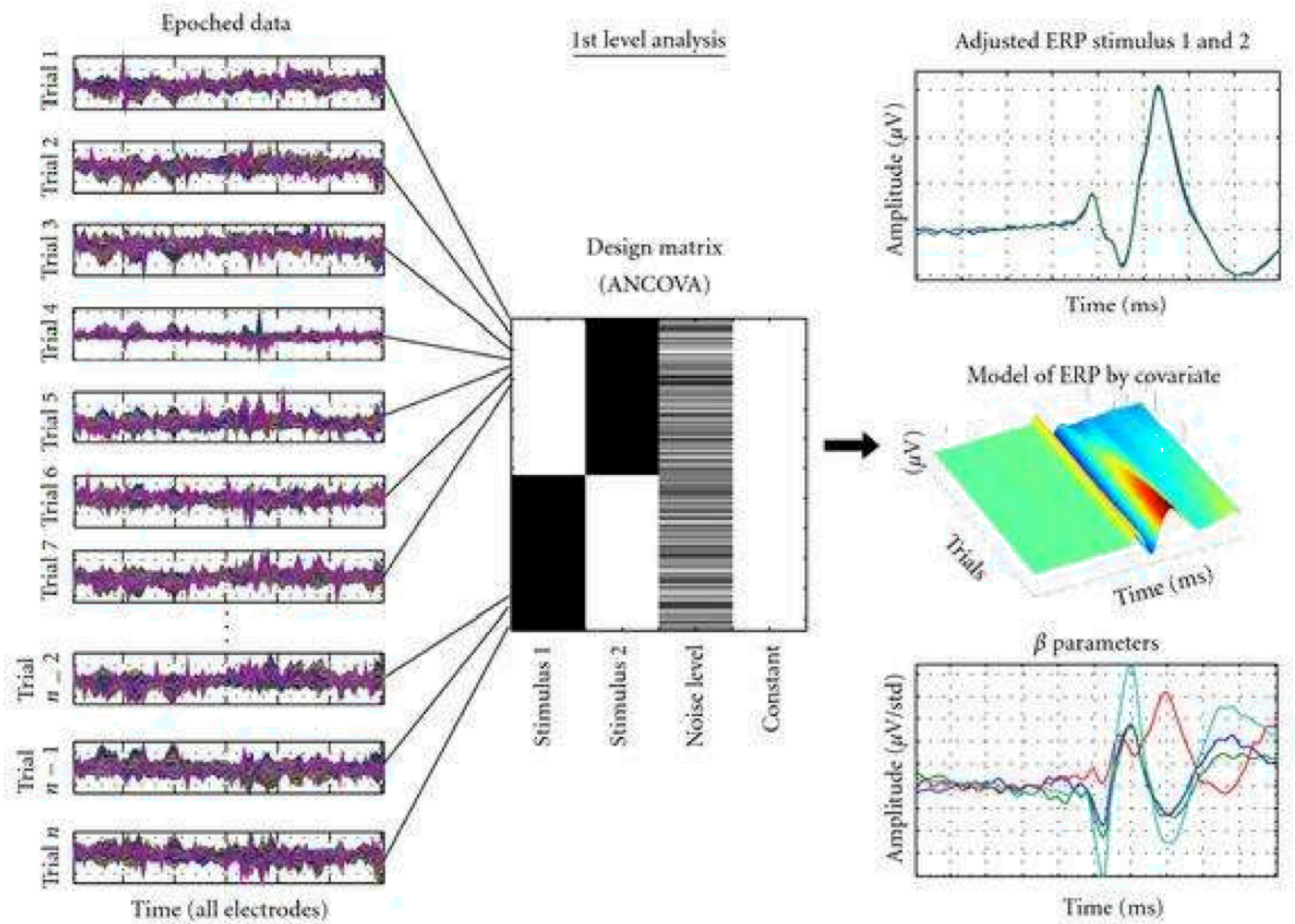
2nd level analysis:
Robust stats (Yuen t-tests, robust GLM, robust Hotelling T^2)



Linear Modeling of EEG data



Linear Modeling of EEG data

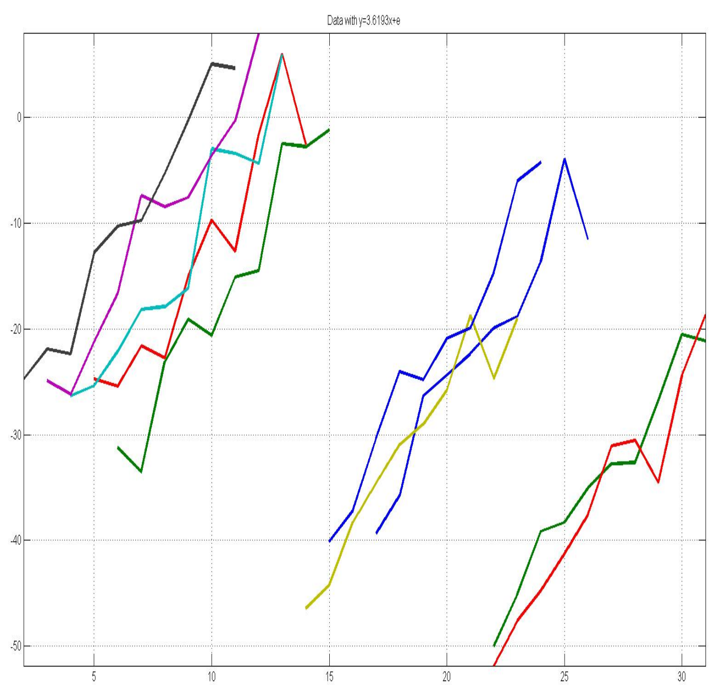


Random Effect Model

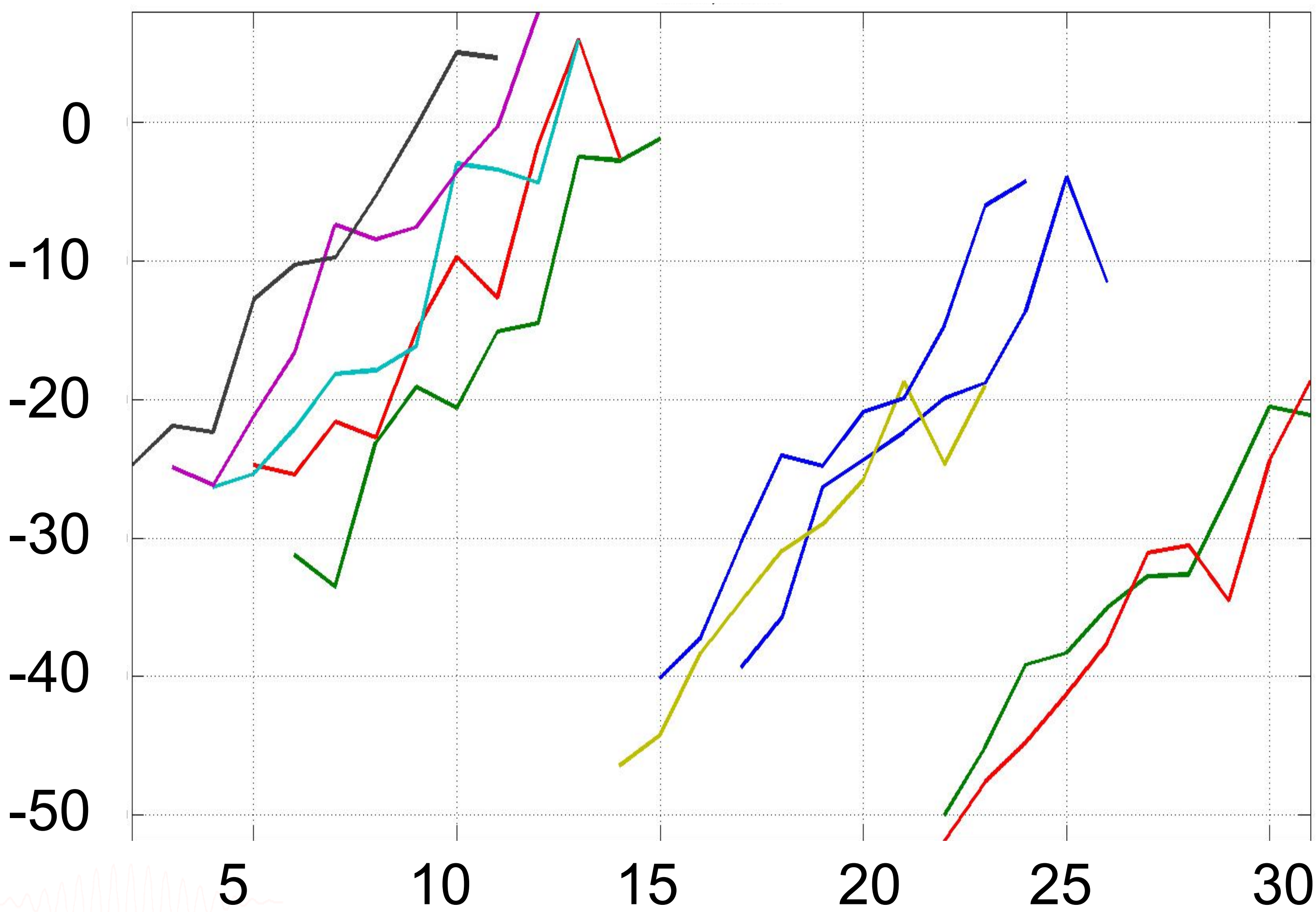
Model the data with fixed effects (the experimental conditions) and a random effect (subjects are allowed to have different overall values – considering subjects as a random variable)

<u>Example</u>: present stimuli from intensity -5 units to +5 units around the subject perceptual threshold and measure RT

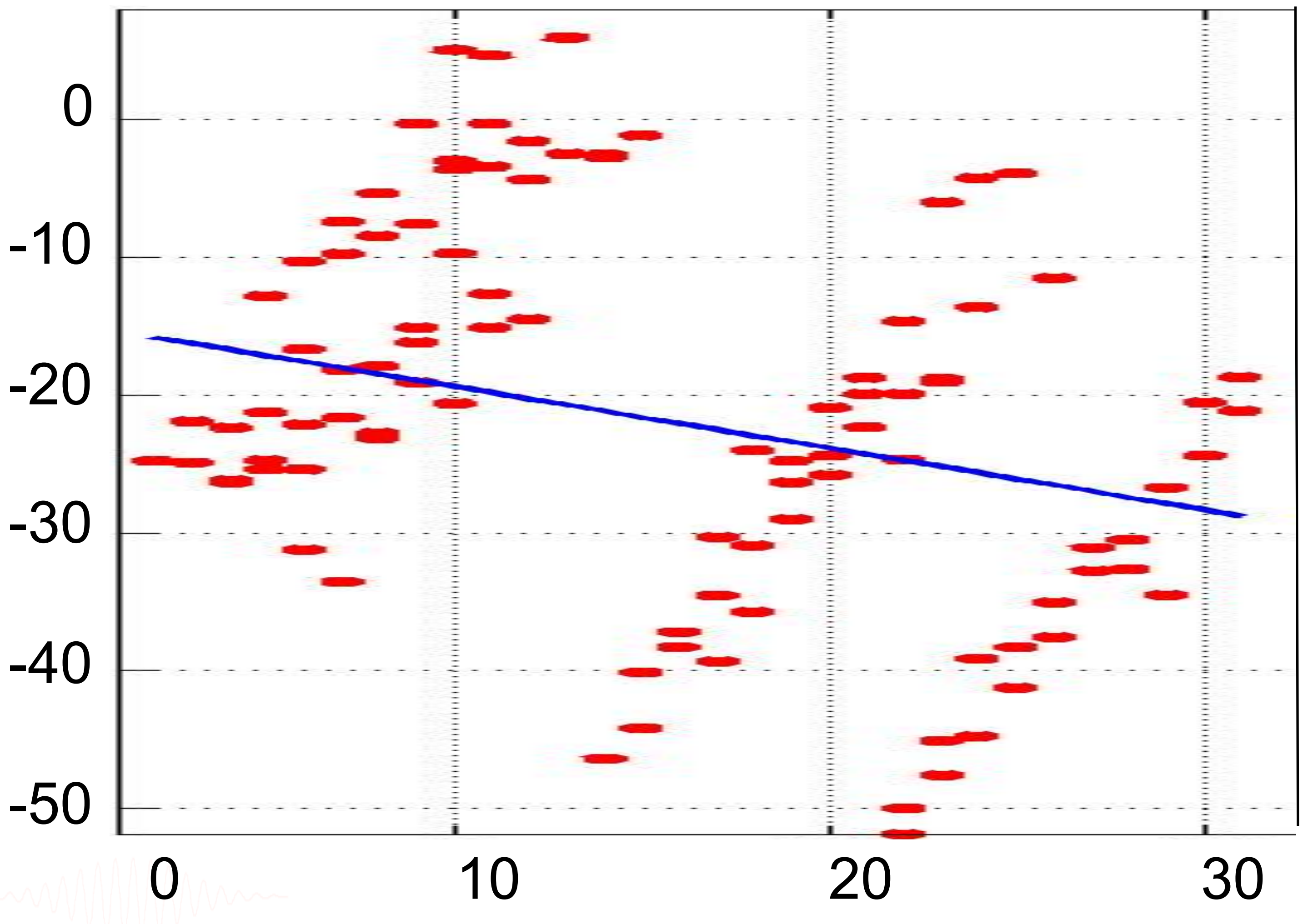
→ Plot the data per intensity



Fixed Effect



Fixed Effect

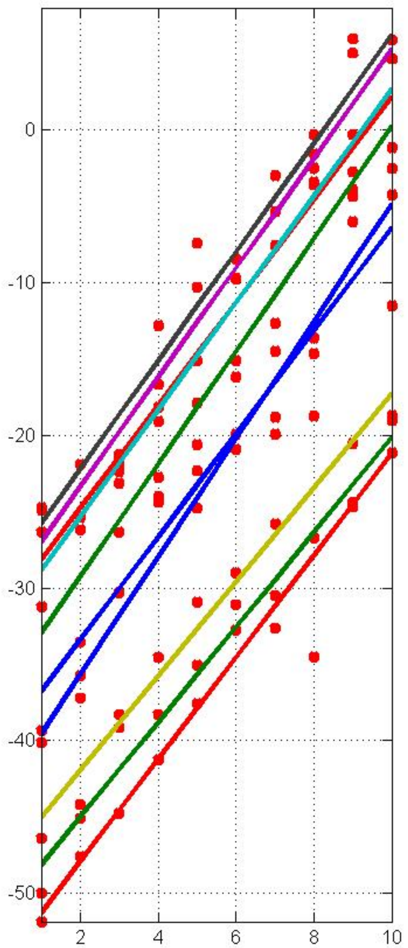


$R^2 = 0.16$
$P<0.03$

Fixed effect = average across subjects → negative correlation?

Random Effect Model

1st level



2nd level

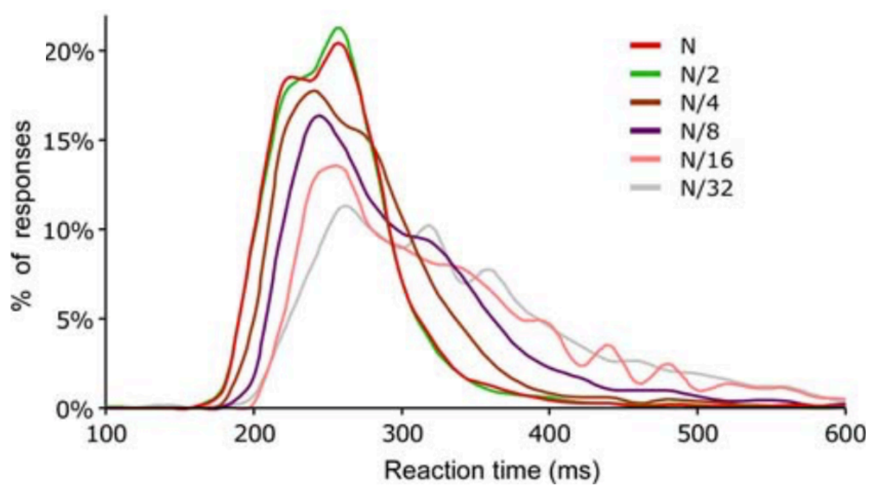
Slope: [3.43 3.51 3.39 3.41 3.62 3.55 …]

$p < 10^{-12}$

A regression is a linear model

Varying factor: Contrast of image

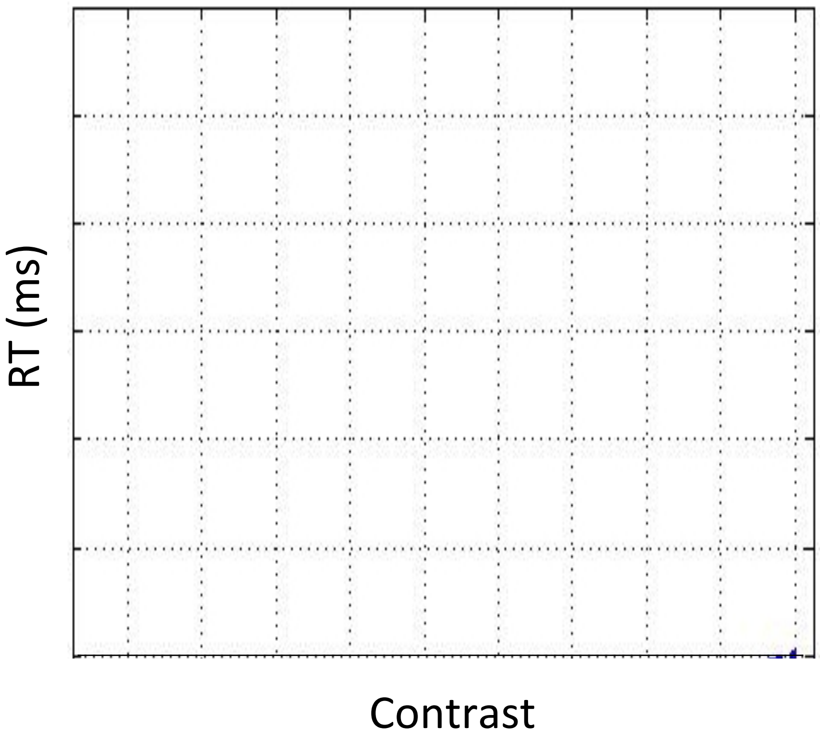
Outcome: Reaction time



Mace, M., Delorme, A., Richard, G., Fabre-Thorpe, M. (2010) Spotting animals in natural scenes: efficiency of humans and monkeys at very low contrasts. *Animal Cognition*, 13(3):405-18.

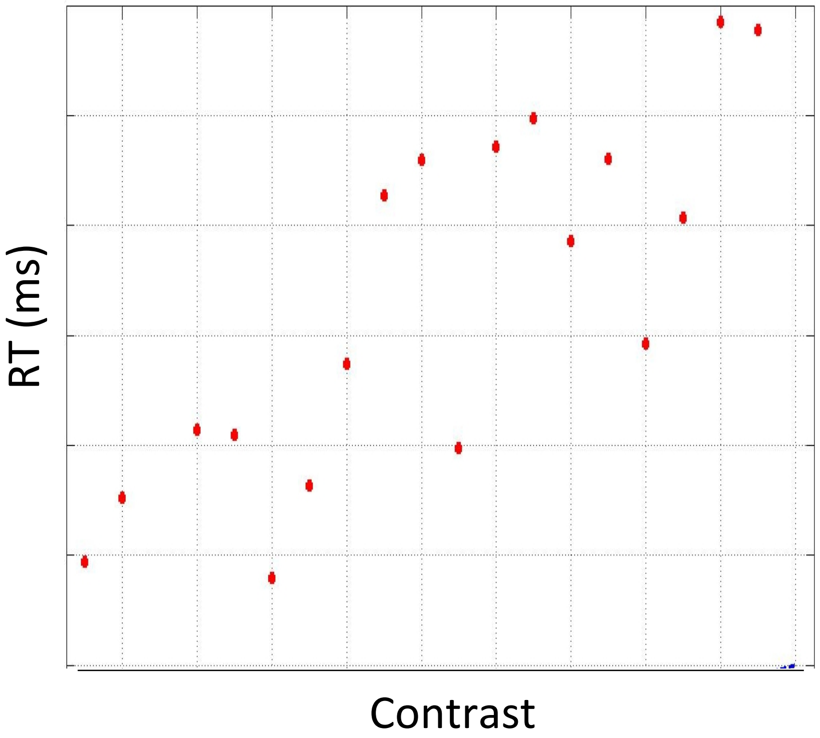
A regression is a linear model

- We have an experimental measure x (e.g. contrast)



A regression is a linear model

- We have an experimental measure x (e.g. contrast)
- We then do the expe and collect data RT (e.g. reaction time)



A regression is a linear model

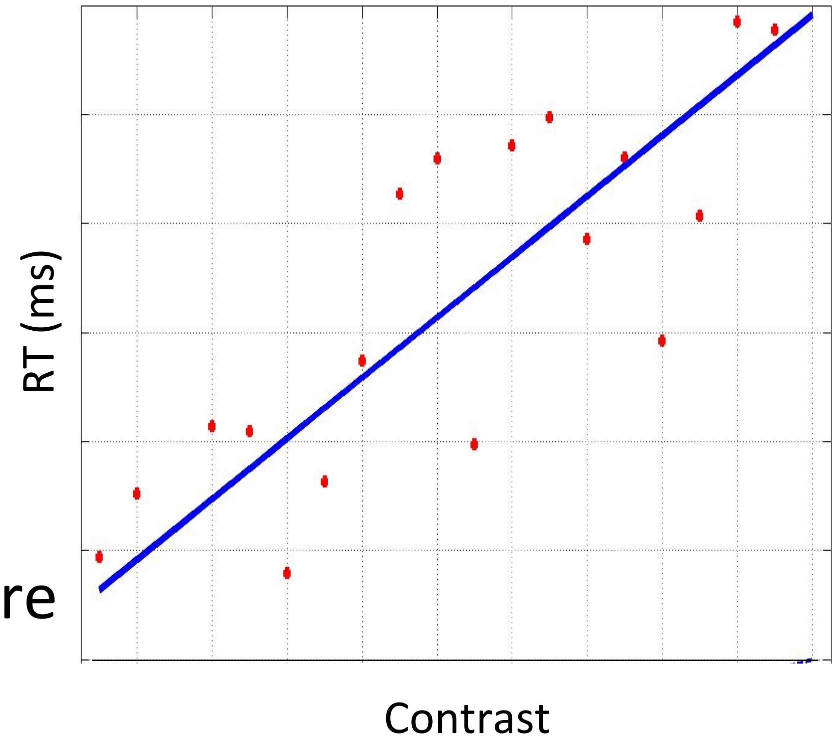
- We have an experimental measure x (e.g. contrast)

- We then do the expe and collect data RT (e.g. reaction time)

- Model: RT = $\beta_0 + x\beta_1 + \varepsilon$

- Do some maths / run a software to find β_1 and β_0

- RT^ = 23.6 + 2.7x



A regression is a linear model

For each trial

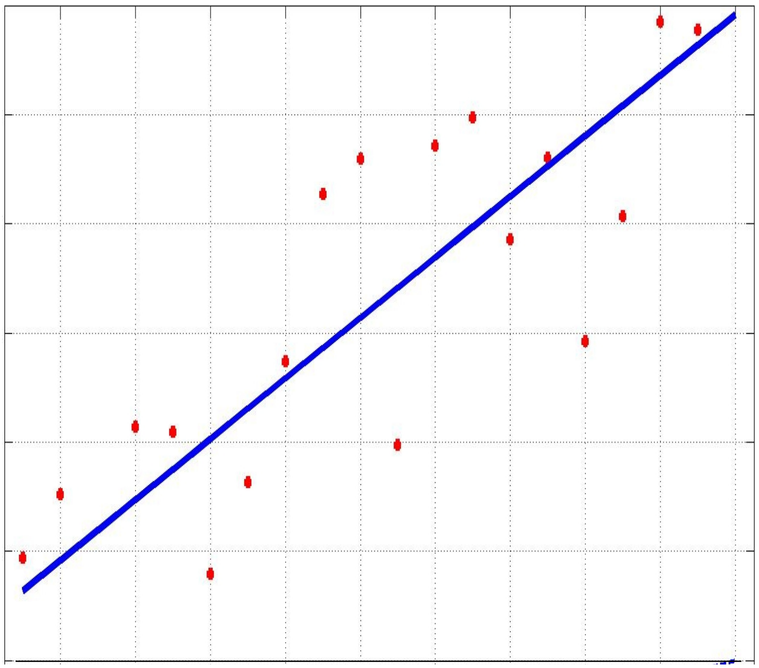
Contrast level

$RT_1 = \beta_0 + 10*\beta_1 + \varepsilon_1$

$RT_2 = \beta_0 + 5*\beta_1 + \varepsilon_2$

$RT_3 = \beta_0 + 7*\beta_1 + \varepsilon_3$

...



A regression is a linear model

For each trial

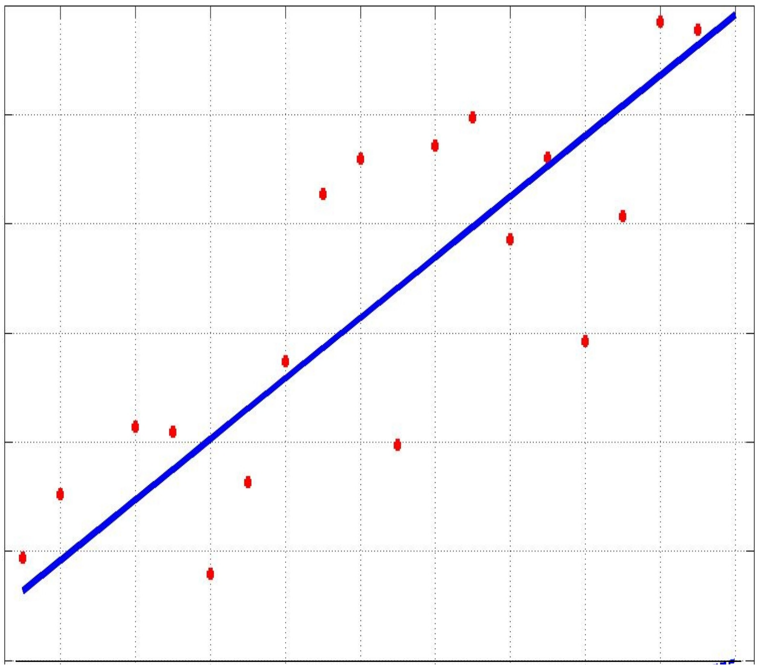
Contrast level

$RT_1 = \beta_0 + 10*\beta_1 + \varepsilon_1$

$RT_2 = \beta_0 + 5*\beta_1 + \varepsilon_2$

$RT_3 = \beta_0 + 7*\beta_1 + \varepsilon_3$

...



To test for significance compare the original regression model $RT_i = \beta_0 + c_i*\beta_1 + \varepsilon_i$ with the simplified model $RT_i = \beta_0 + \varepsilon_i$

Compare these errors

An ANOVA is a linear model

Varying factor: Type of image

Outcome: Reaction time (go/no-go)

No-animal

Delorme, A., Richard, G., Fabre-Thorpe, M. (2010). Key visual features for rapid categorization of animals in natural scenes. *Frontier in psychology*, 1:21

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

For trial 13 (for example second trial of birds) we have

$$RT_{2,2} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,2}$$

el

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

For trial 13 (for example second trial of birds) we have

$$RT_{2,2} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,2}$$

Statistics: if there is an effect of treatment then error of the simplified model $RT_{i,j} = \beta_0 + \varepsilon_{i,j}$ should be lower than the original model $RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$

Compare these errors

This is a GLM that is also strictly equivalent to running an ANOVA

The GLM can do both a Regression and an ANOVA

Varying factor: Type of image **AND** contrast
Outcome: Reaction time (go/no-go)

For example, for trial (first bird with contrast $c_{2,1}$) we have

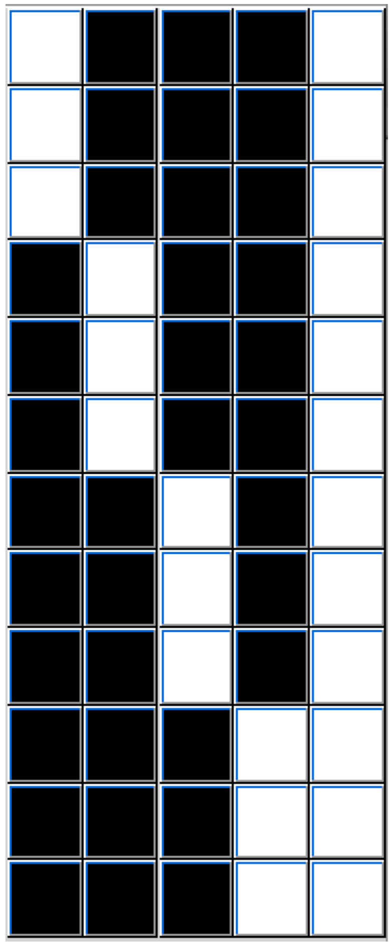
$$RT_{2,1} = \beta_0 + \underbrace{0*\beta_1 + 1*\beta_2 + 0*\beta_3 + 0*\beta_3}_{\text{Categorical var. ANOVA}} + \underbrace{c_{2,1}*\beta_4}_{\text{Continous var. REGRESSION}} + \varepsilon_{2,1}$$

The design matrix

Y	Gp
8	1
9	1
7	1
5	2
7	2
3	2
3	3
4	3
1	3
6	4
4	4
9	4

Design matrix

G_1 G_2 G_3 G_4 C

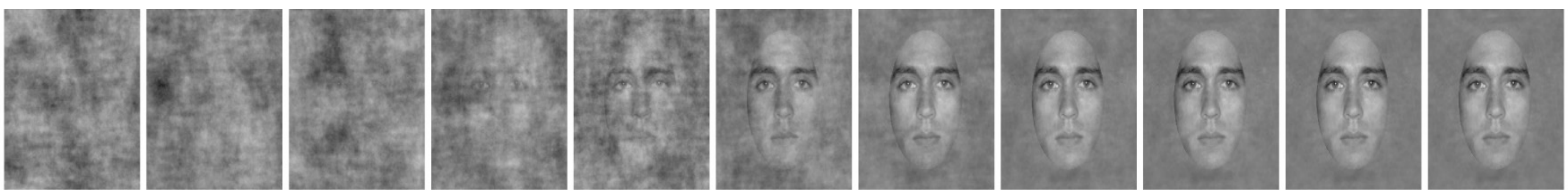


y(1..3) = 1xβ1+0xβ2+0xβ3+0xβ4+c+error
y(4..6) = 0xβ1+1xβ2+0xβ3+0xβ4+c+error
y(7..9) = 0xβ1+0xβ2+1xβ3+0xβ4+c+error
y(10..12) = 0xβ1+0xβ2+0xβ3+1xβ4+c+error
...

$$\begin{Bmatrix} 8 \\ 9 \\ 7 \\ 5 \\ 7 \\ 3 \\ 3 \\ 4 \\ 1 \\ 6 \\ 4 \\ 9 \end{Bmatrix} = \begin{Bmatrix} 1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1 \end{Bmatrix} * \begin{Bmatrix} \beta 1 \\ \beta 2 \\ \beta 3 \\ \beta 4 \\ c \end{Bmatrix} + \begin{Bmatrix} e1 \\ \\ \vdots \\ \\ e13 \end{Bmatrix}$$

Design considerations

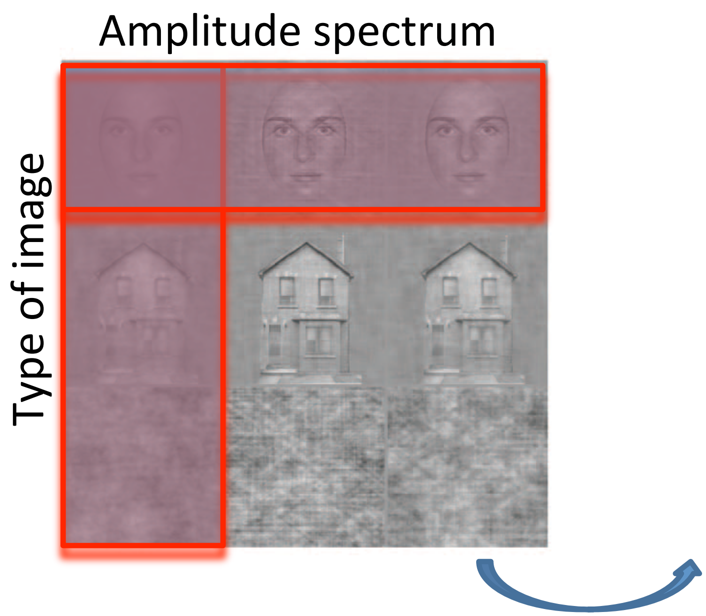
Illustration with a set of studies looking at the effect of stimulus amplitude and phase information

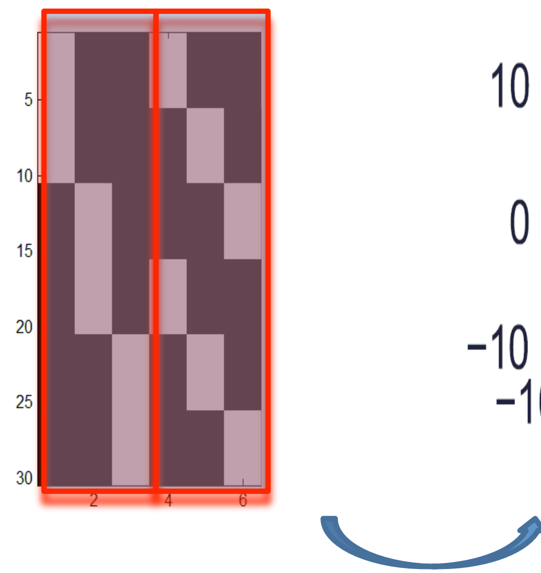


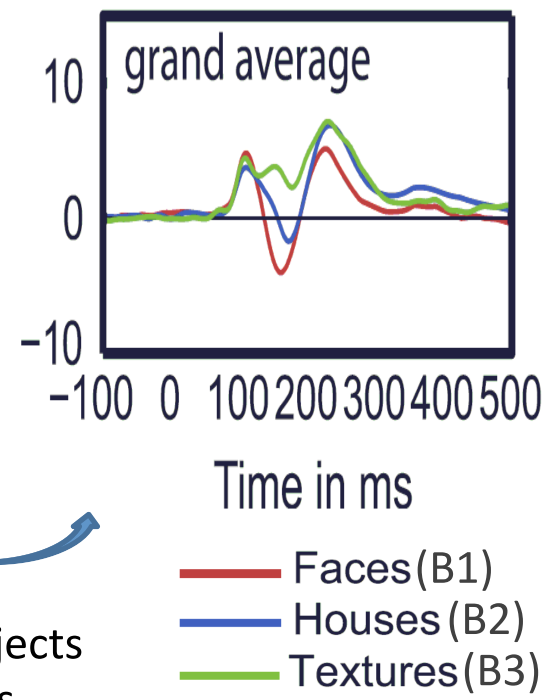
Rousselet, Pernet, Bennet, Sekuler (2008). Face phase processing. *BMC Neuroscience 9:98*

Factorial Designs: N*N*N*…

Categorical designs: Group level analyses of course but also Individual analyses with bootstrap



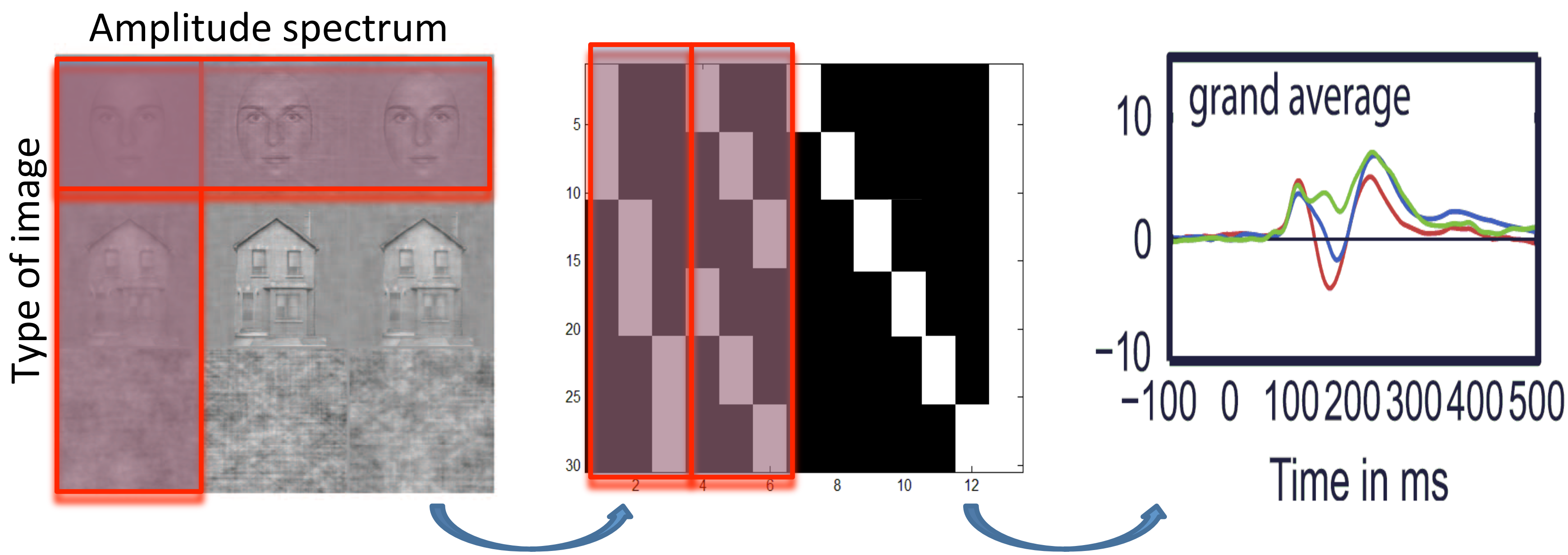




B1→B3 represents type of image, 2nd level ANOVA across subjects
B4→B6 represents amplitude, 2nd level ANOVA across subjects
Interaction between phase and amplitude can be assessed using 2-way ANOVA

Factorial Designs: N*N*N*…

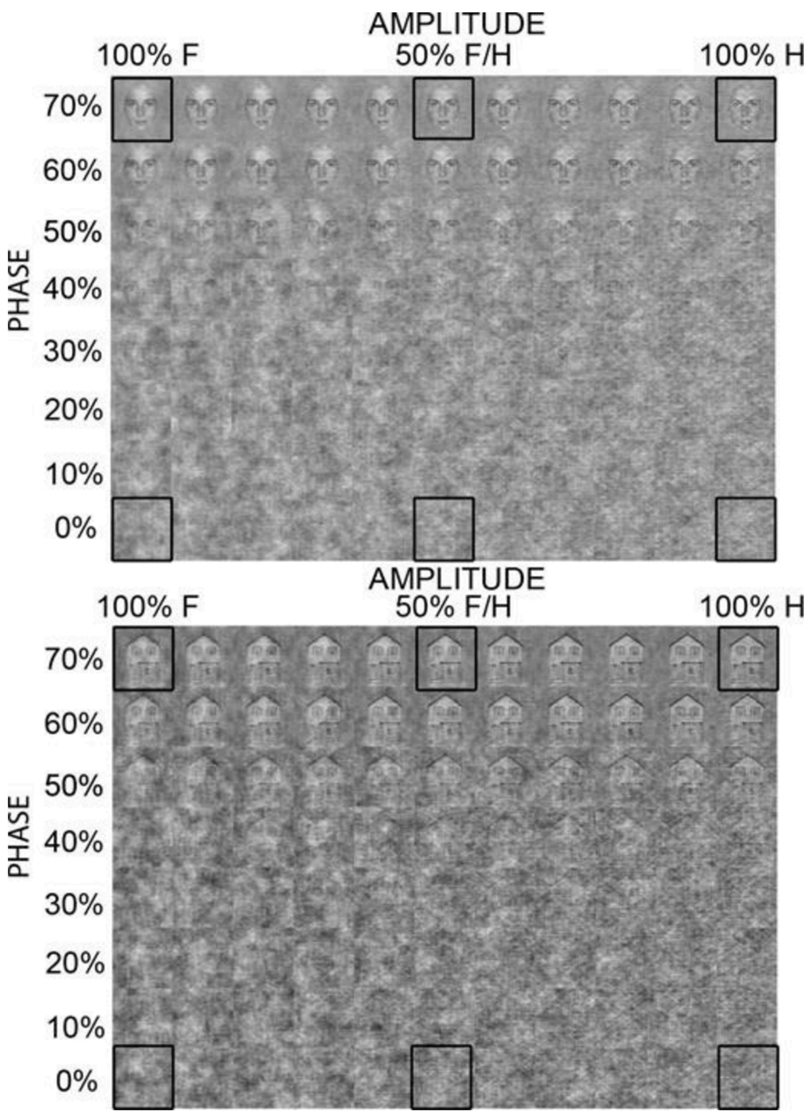
Categorical designs: Group level analyses of course but also Individual analyses with bootstrap

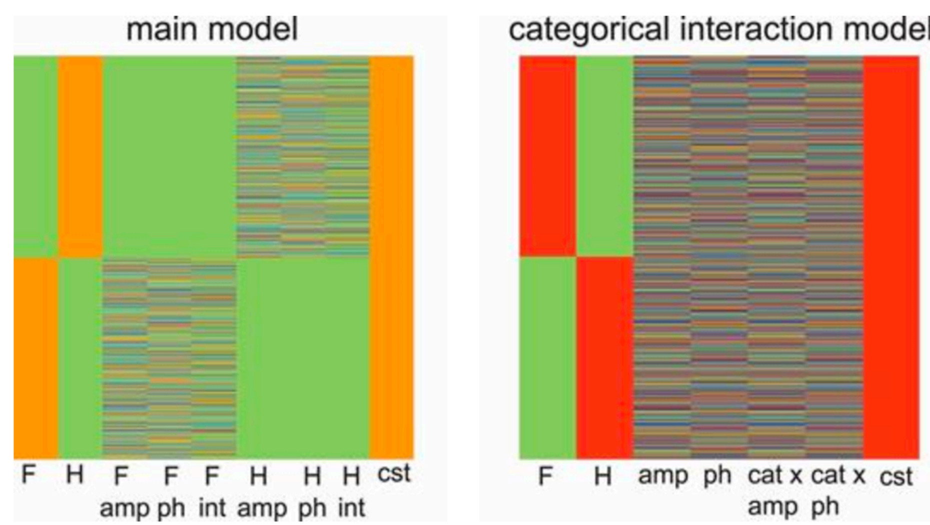


Interaction between phase and amplitude can be assessed using 1-way ANOVA on B7 to B12. There is no interaction left between B1-B3 and B4-B6

[Bienek, Pernet, Rousselet (2012). Phase vs Amplitude Spectrum. Journal of Vision 12(13), 1–24]

Continuous designs

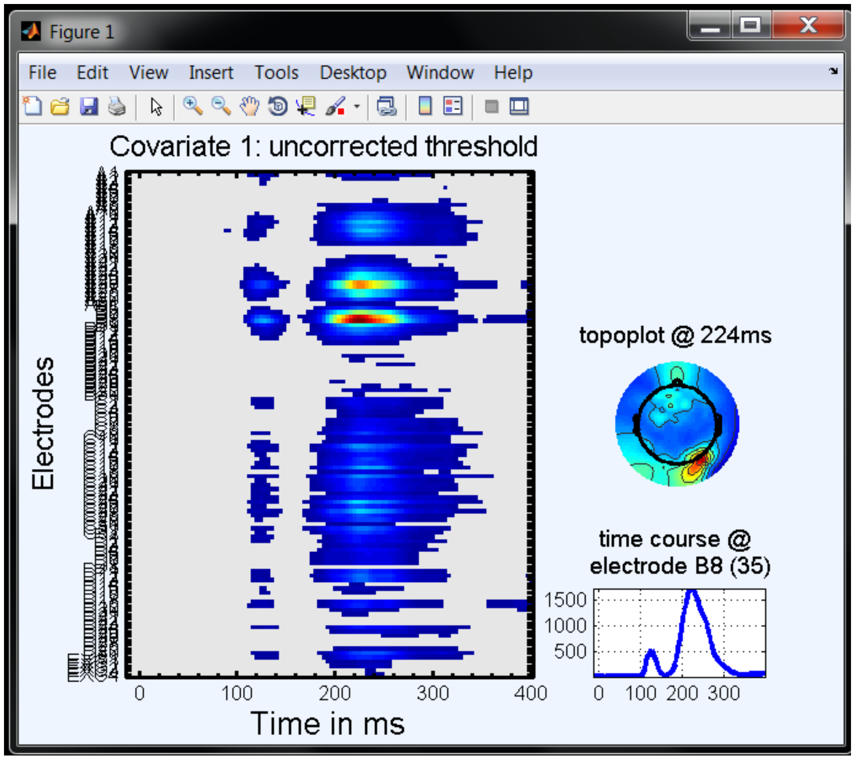


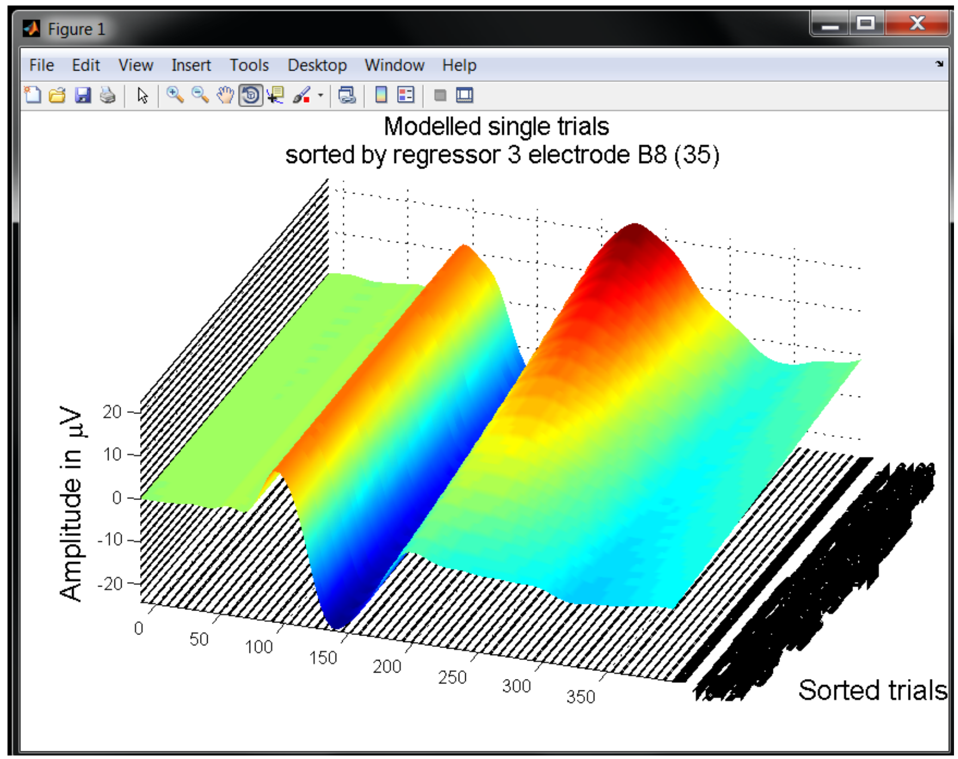


Note: just face and houses (no noise here)

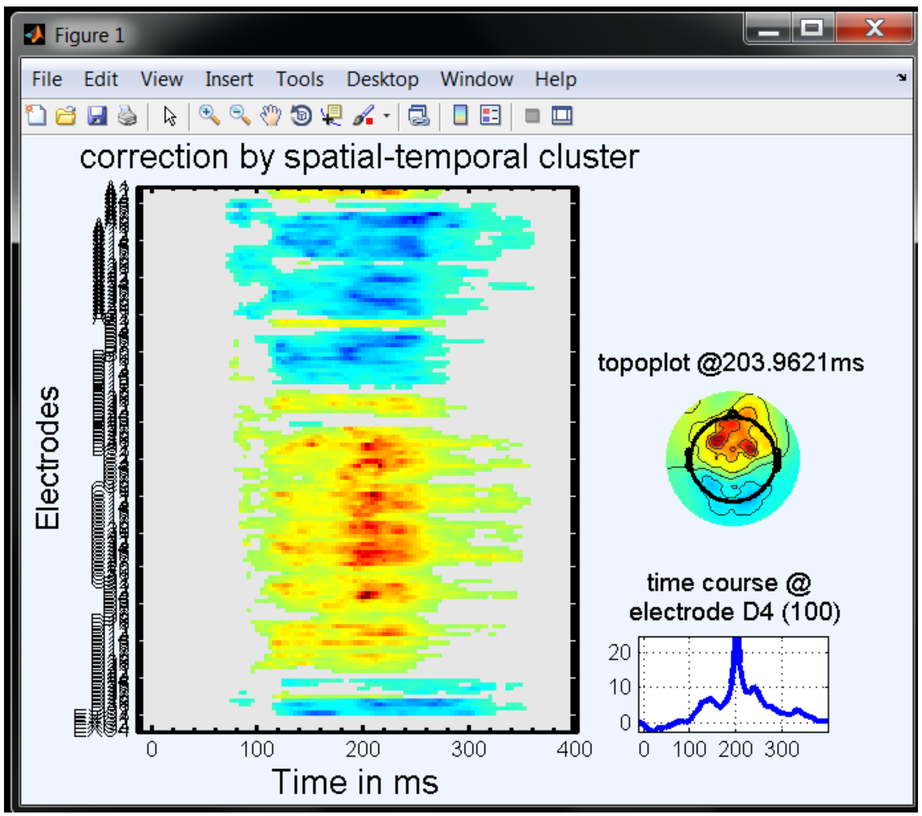
What have we done: results

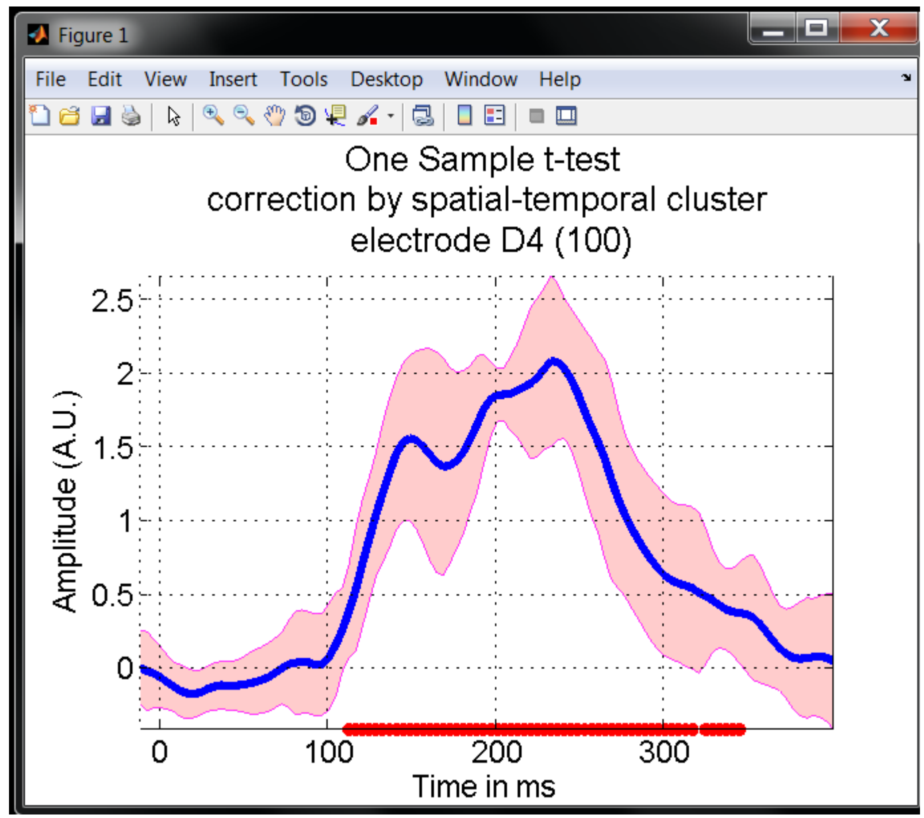
- Image all (R2, condition, covariate)
- Course plots, topoplots





Review group level results





Design questions!

- Let's think how to analyse your data!
- Nb of conditions / covariates
- contrasts
- 1st level covariates
- 2nd level covariates