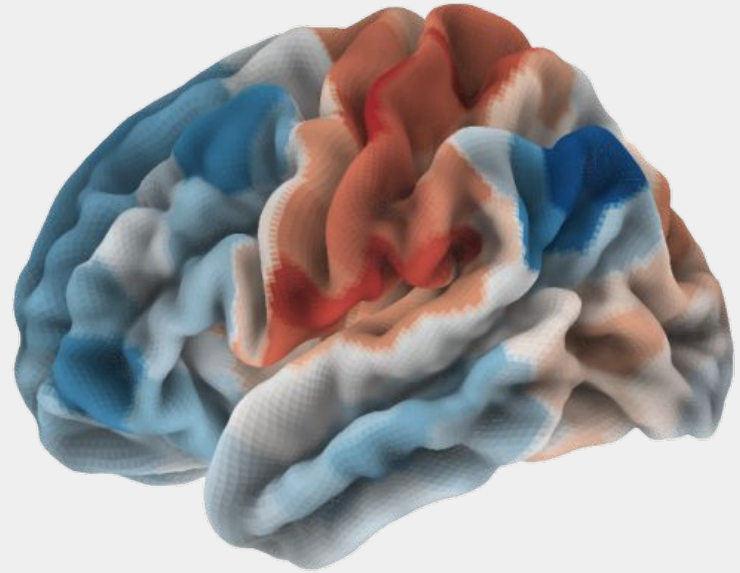


Fundamentals of fMRI data analysis

Karolina Finc

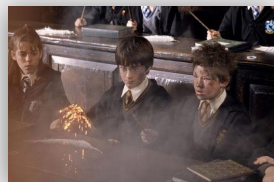
Centre for Modern Interdisciplinary
Technologies Nicolaus Copernicus University
in Toruń

PART #5: General Linear Model 2



Study plan

Open science &
neuroimaging



BEFORE

fMRI data
preprocessing



fMRI data manipulation
in python



Functional
connectivity



General
Linear Model



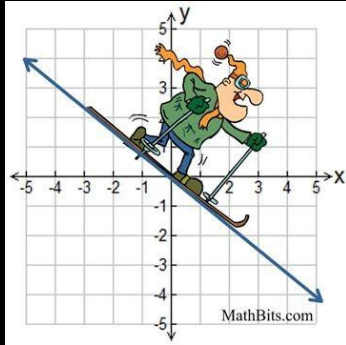
AFTER



Machine Learning
on fMRI data



Slope-intercept form of linear function



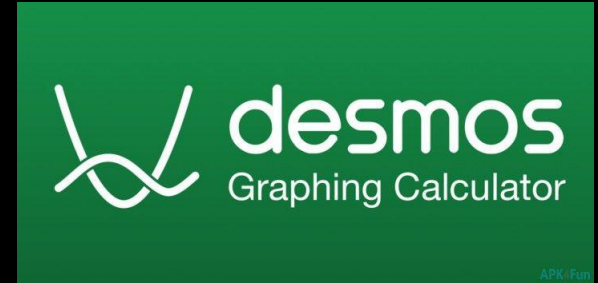
slope

intercept

$$y = mx + b$$

coefficient
s

The diagram illustrates the components of the slope-intercept form equation $y = mx + b$. Arrows point from the word "slope" to the variable m (labeled "coefficient s") and from the word "intercept" to the variable b .



<https://www.desmos.com/calculator>

Linear regression

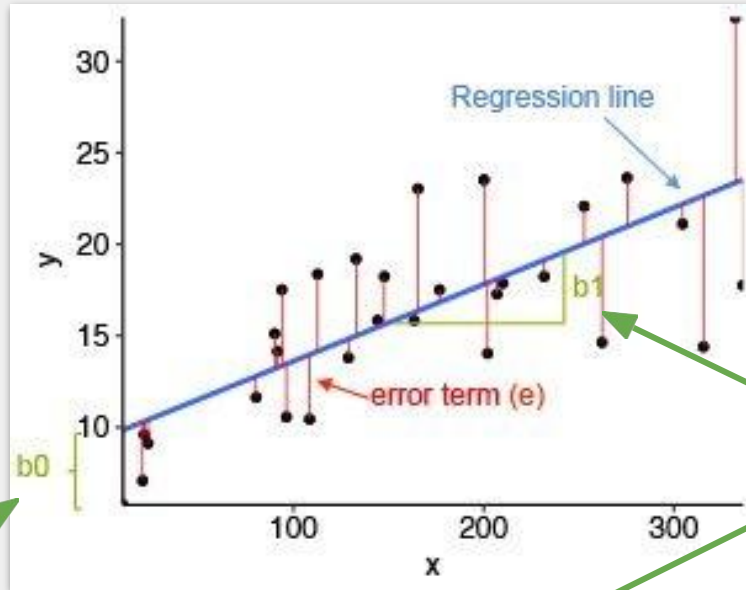


Regression line provides a **model** of the data

Regression problem:
predict real-valued output

Regression is an example of **supervised learning**
(answers are given)

Fitting regression line



Find such β_0 and β_1 that minimize cost function: **sum of squared errors** function

intercept

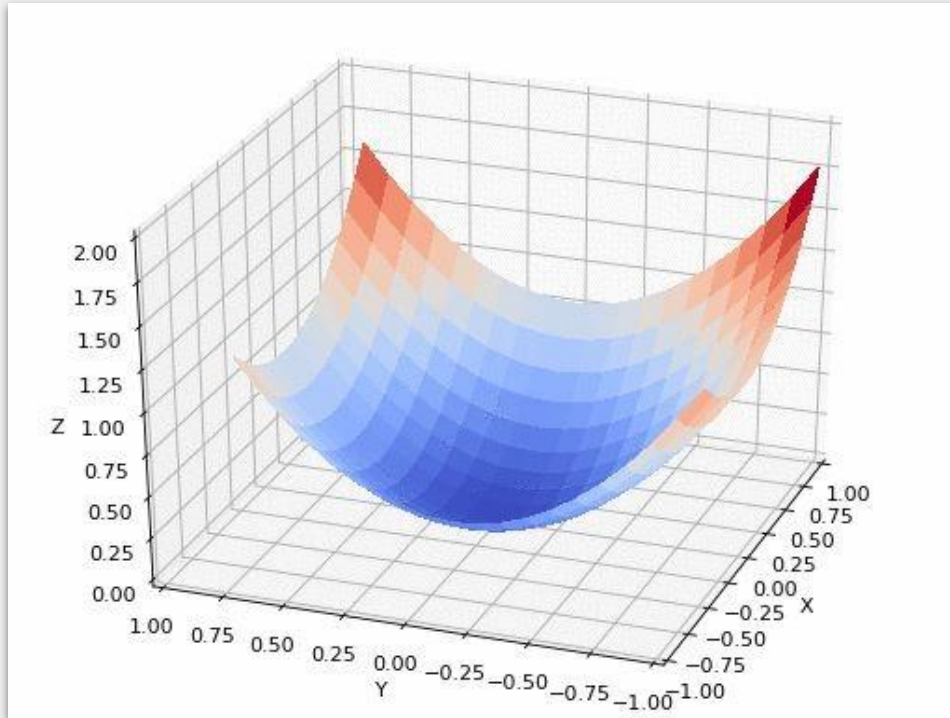
slope

Linear Regression: Single Variable

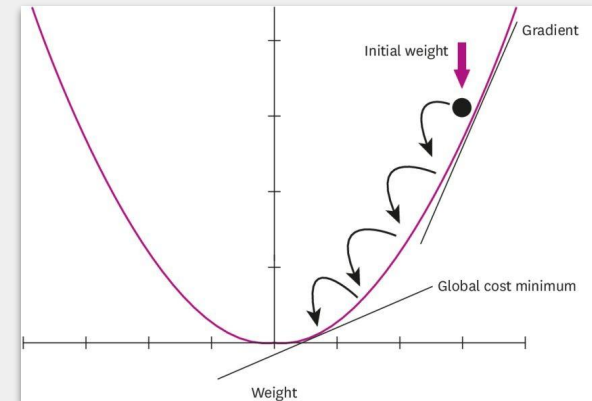
$$\hat{y} = \beta_0 + \beta_1 x + \epsilon$$

Predicted output Coefficients Input Error

Gradient descent



- Algorithm for minimizing cost function
- Is used not only in linear regression



Linear combination



= a



+

b



+ ϵ

What **combination** of Lily & James gives a better **prediction** of harry?

Multiple linear regression

Linear Regression: Single Variable

$$\boxed{\hat{y}} = \beta_0 + \beta_1 \boxed{x} + \boxed{\epsilon}$$

Predicted output Coefficients Input Error

Linear Regression: Multiple Variables

$$\boxed{\hat{y}} = \beta_0 + \beta_1 \boxed{x_1} + \dots + \beta_p \boxed{x_p} + \boxed{\epsilon}$$

Each parameter β_i is interpreted as the effect of x_i controlling for all other variables in the model.

Matrix notation

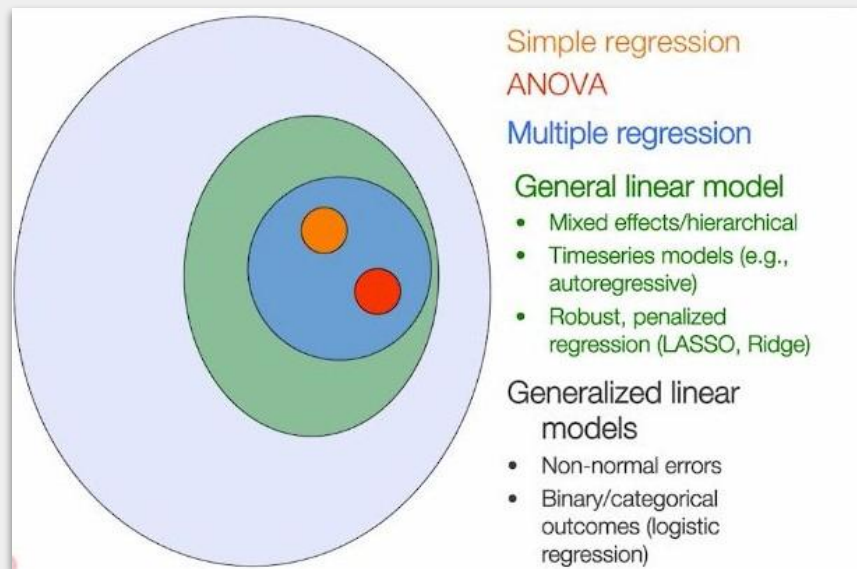
$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{np} & \cdots & X_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Observed Data Design matrix Model parameters Residuals


Coursera, *Principles of fMRI*

Generalized Linear Model



The general linear model (GLM) approach treats the fMRI data as a linear combination of model functions, predictors, plus noise, or error.

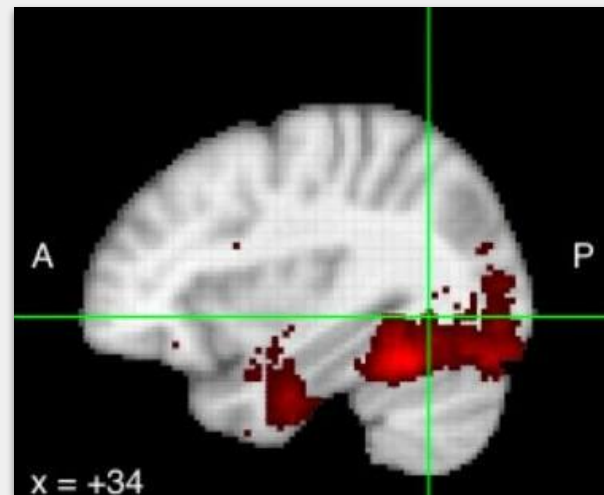




What we want to achieve
when asking participant to
perform a task in the fMRI
scanner?

Goals of task-based fMRI

1. Induce in a study participant to do actions or experience cognitive states you're interested in.
2. You want to detect brain signals that are related to this cognitive states or actions.



Presentation method

Equipment

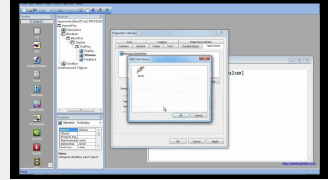
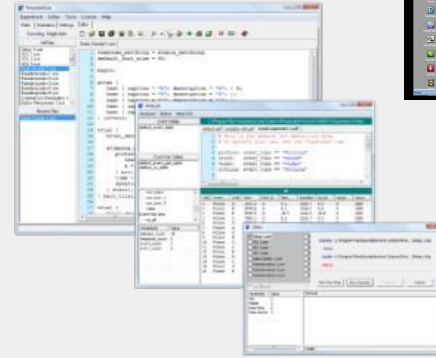
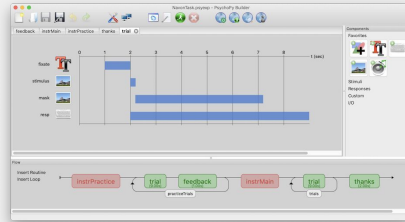


Presentation method

Equipment

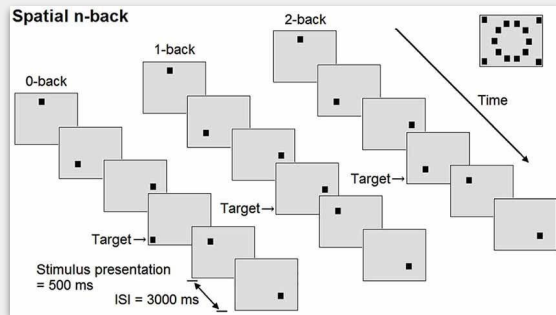


Software



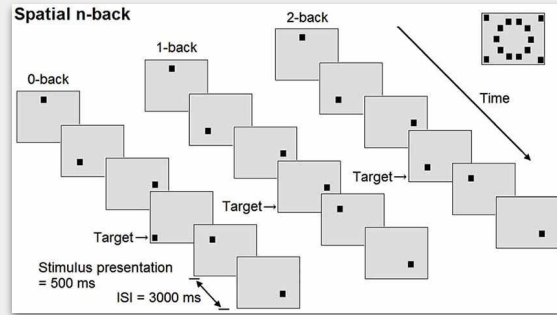
What to present?

Simple stimuli

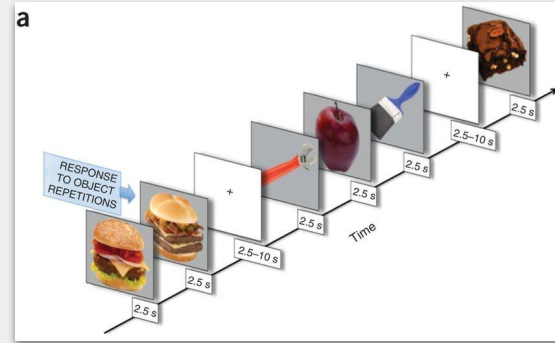


What to present?

Simple stimuli

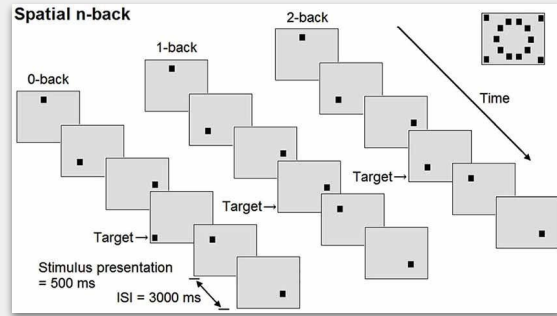


Photos



What to present?

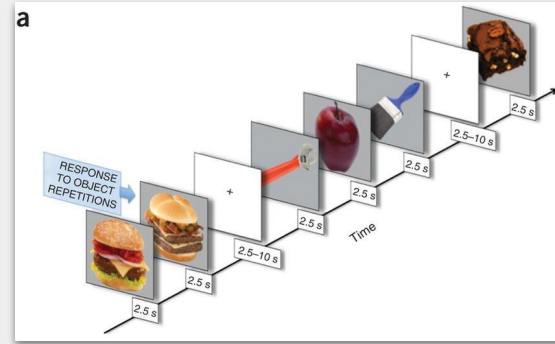
Simple stimuli



Movie

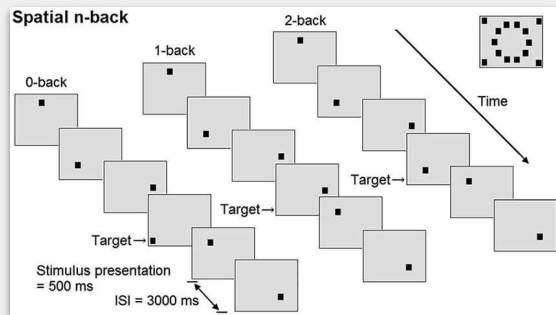


Photos

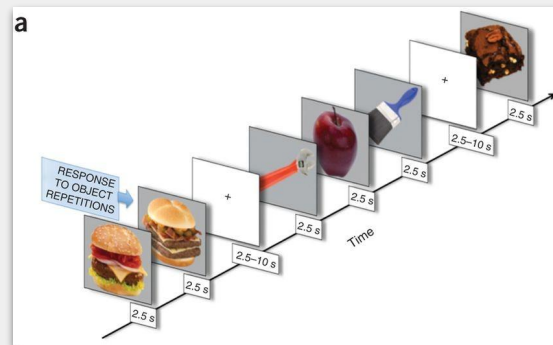


What to present?

Simple stimuli



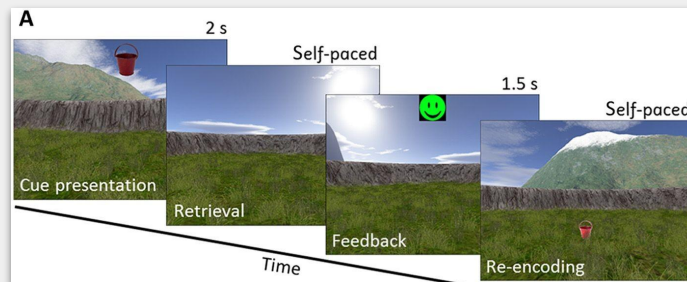
Photos



Movie

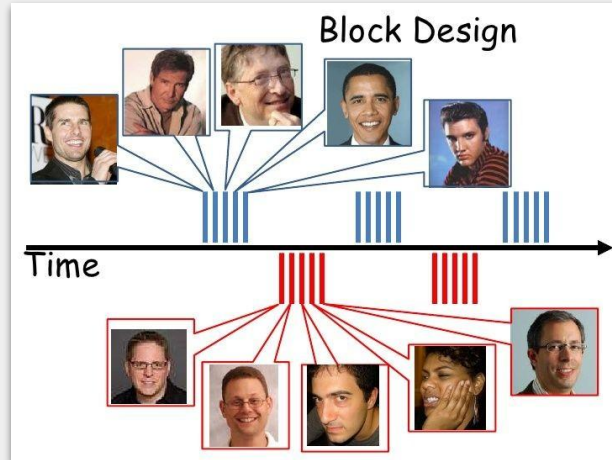


Virtual reality



Task designs

Famous
people

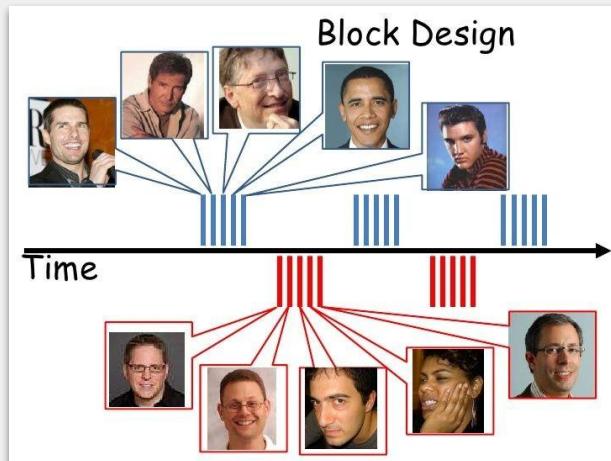


Non-famous
people

Block design
similar events are
grouped

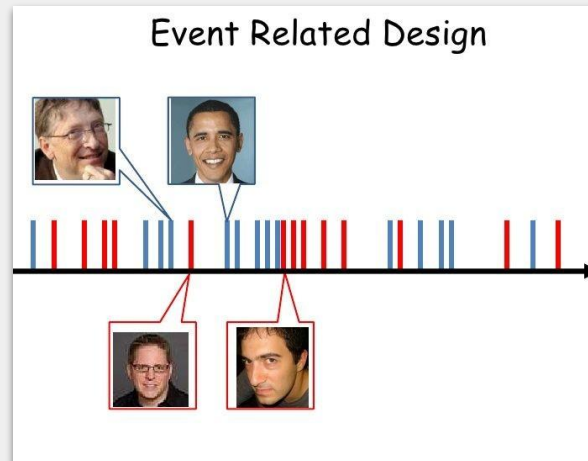
Task designs

Famous
people



Non-famous
people

Block design
similar events are
grouped



Event-related design
events are mixed

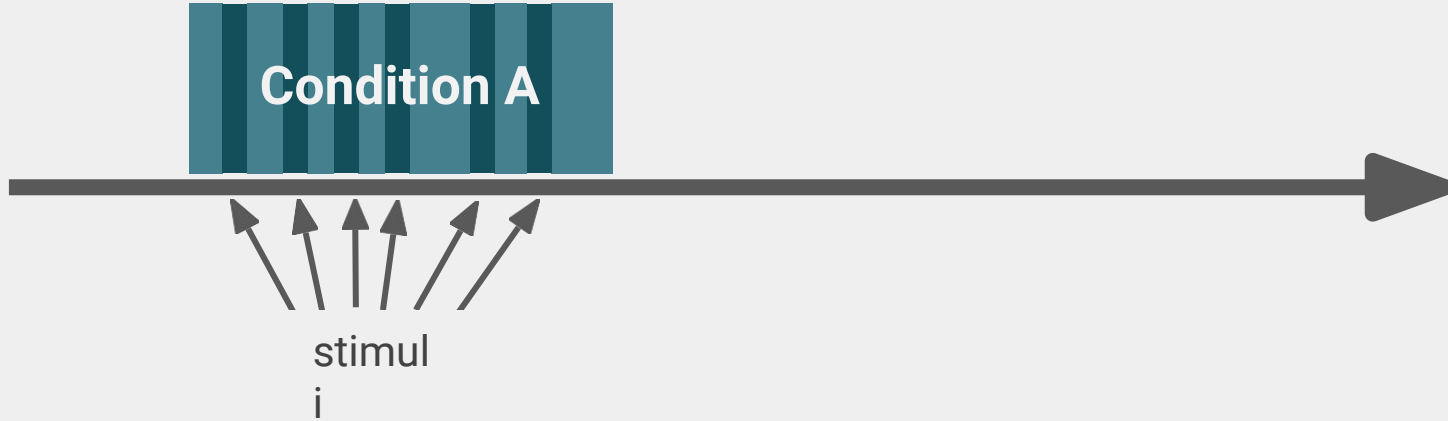
Events parameters



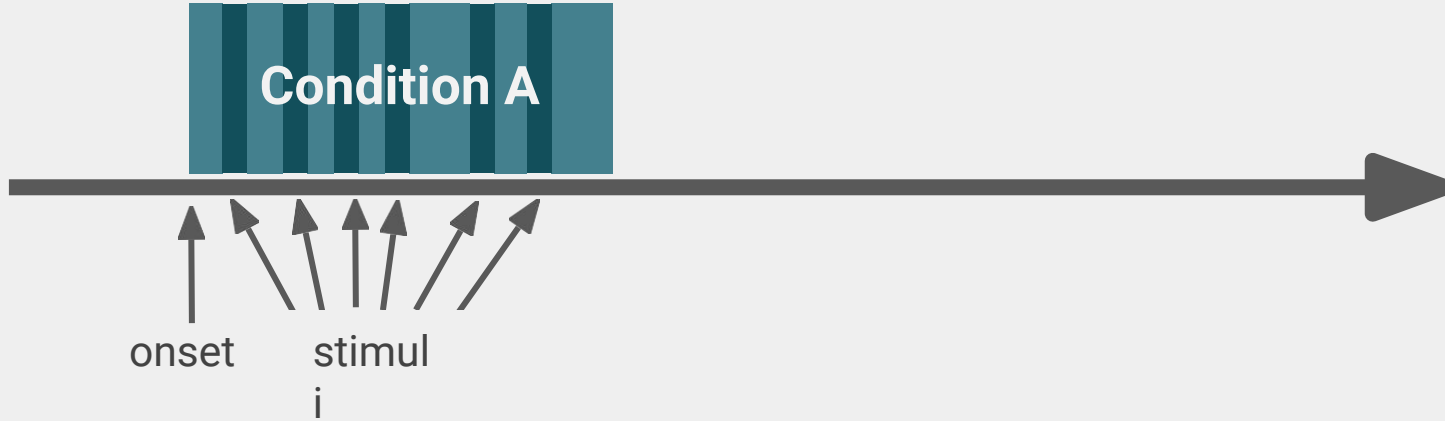
Events parameters



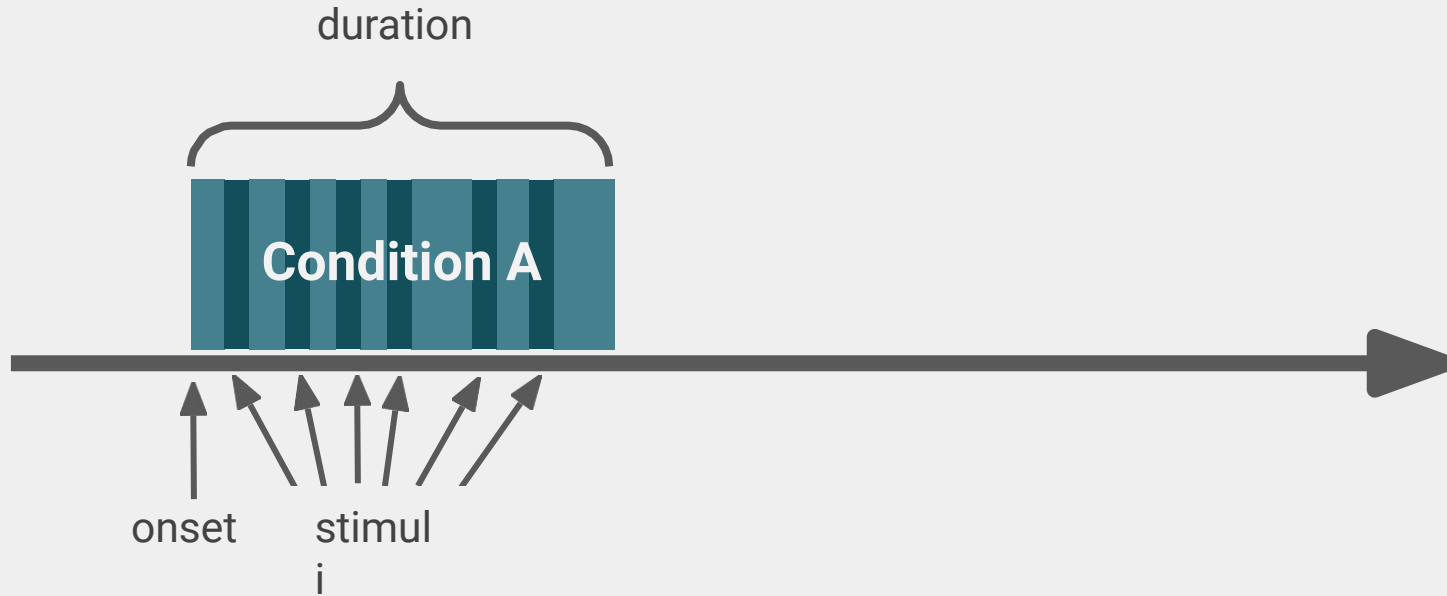
Events parameters



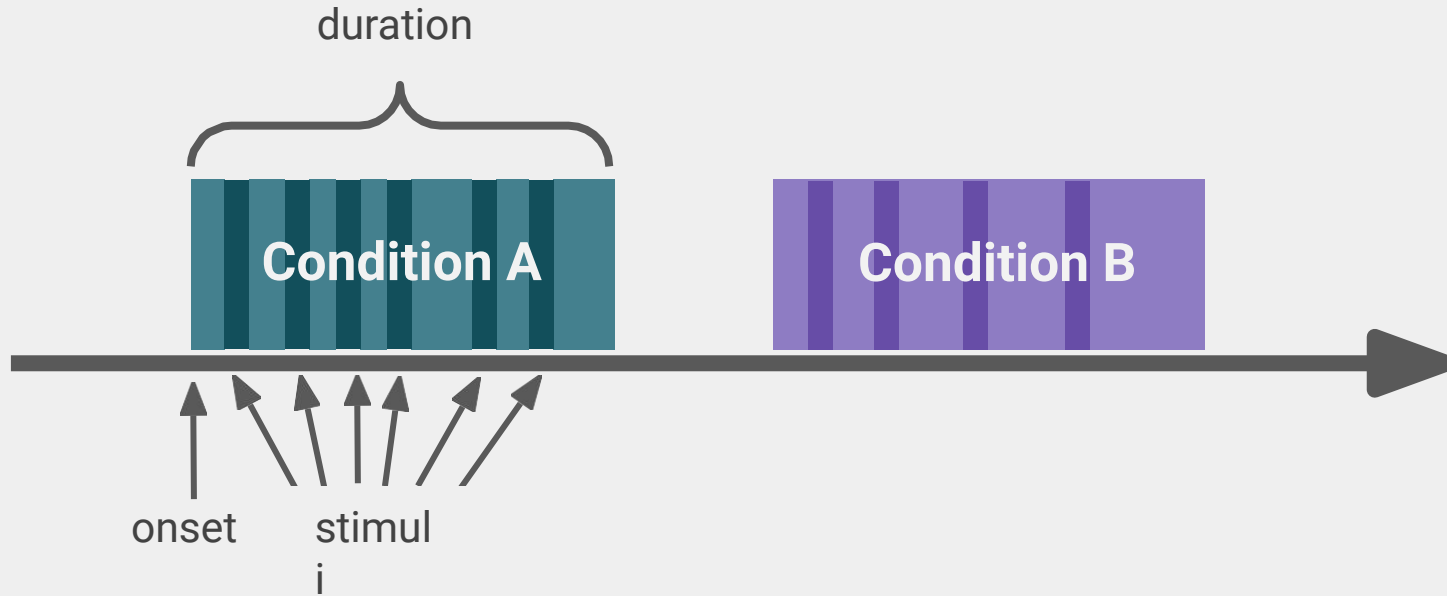
Events parameters



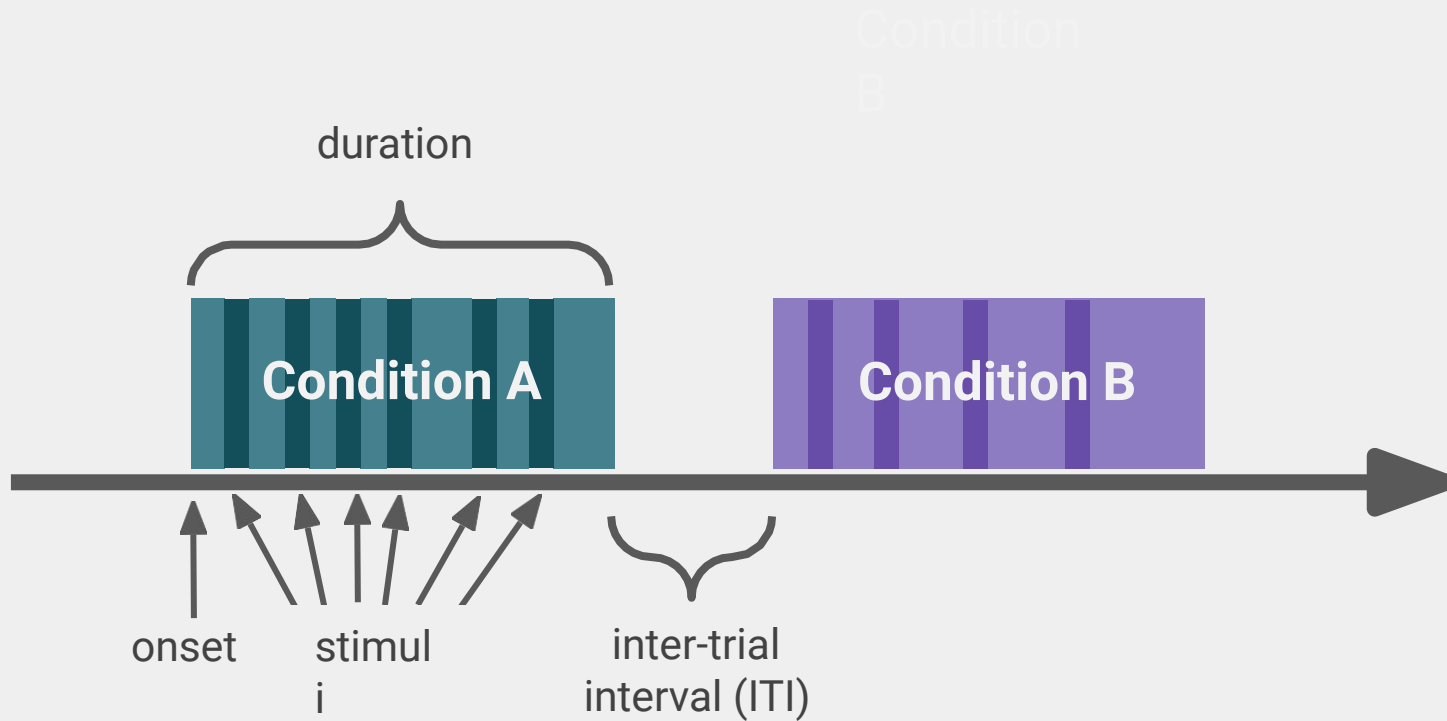
Events parameters



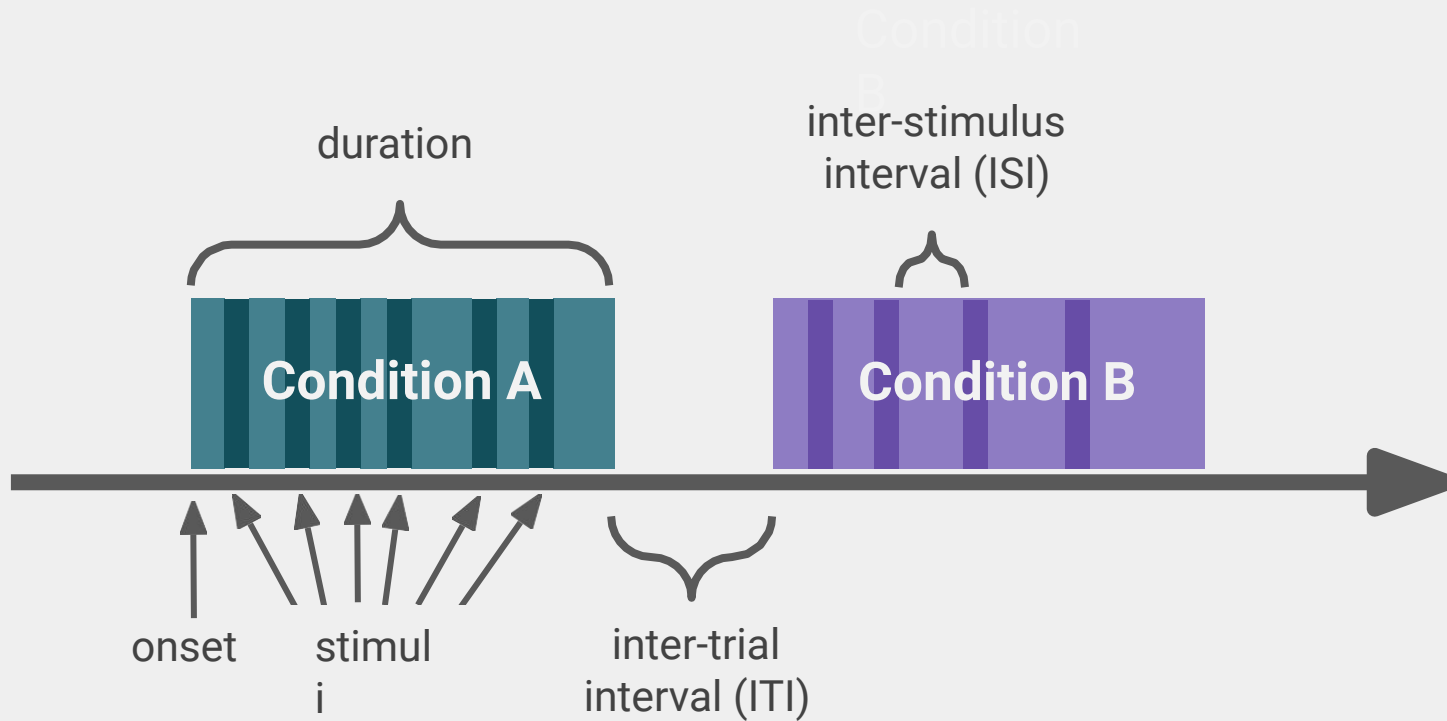
Events parameters



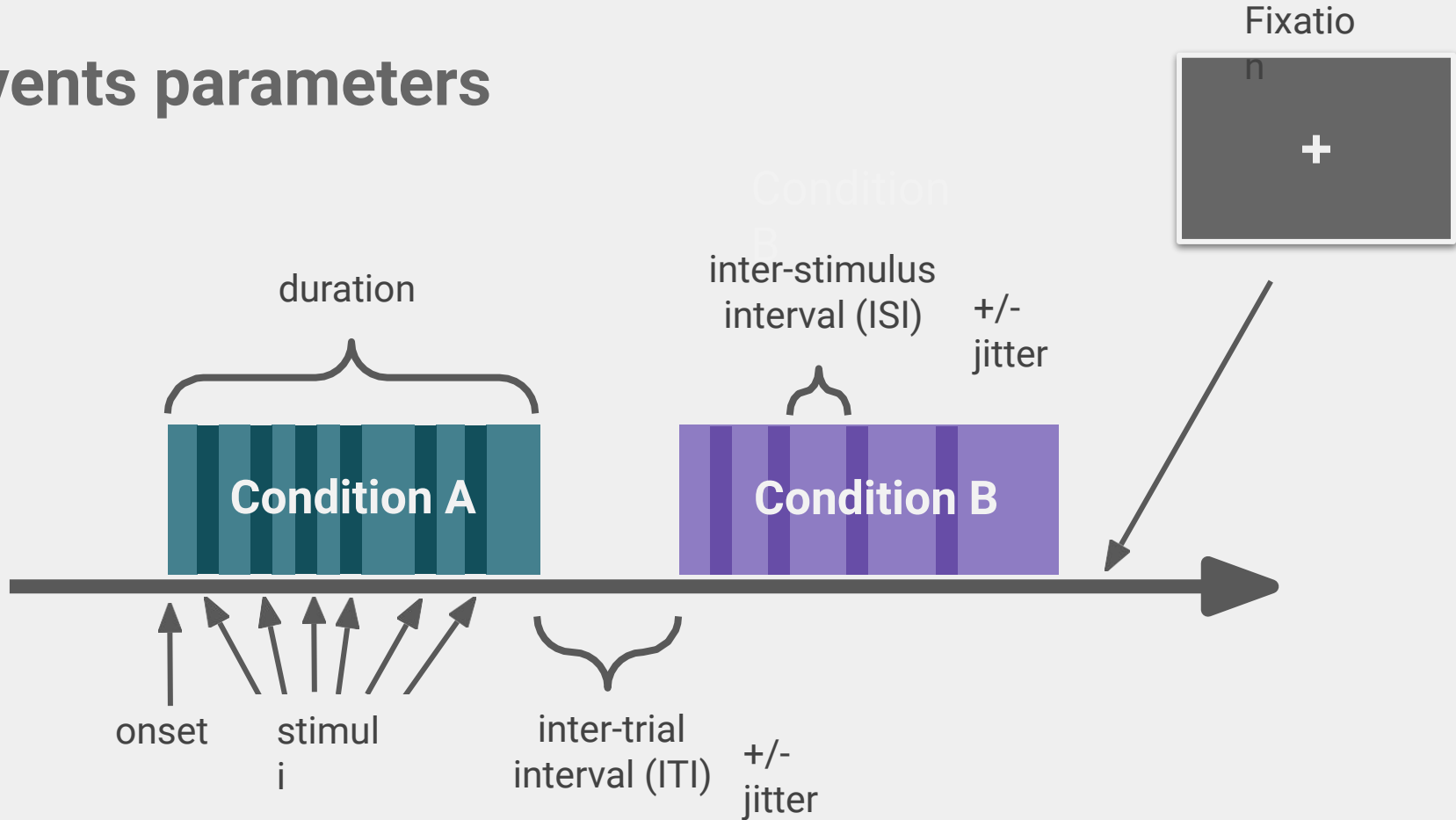
Events parameters



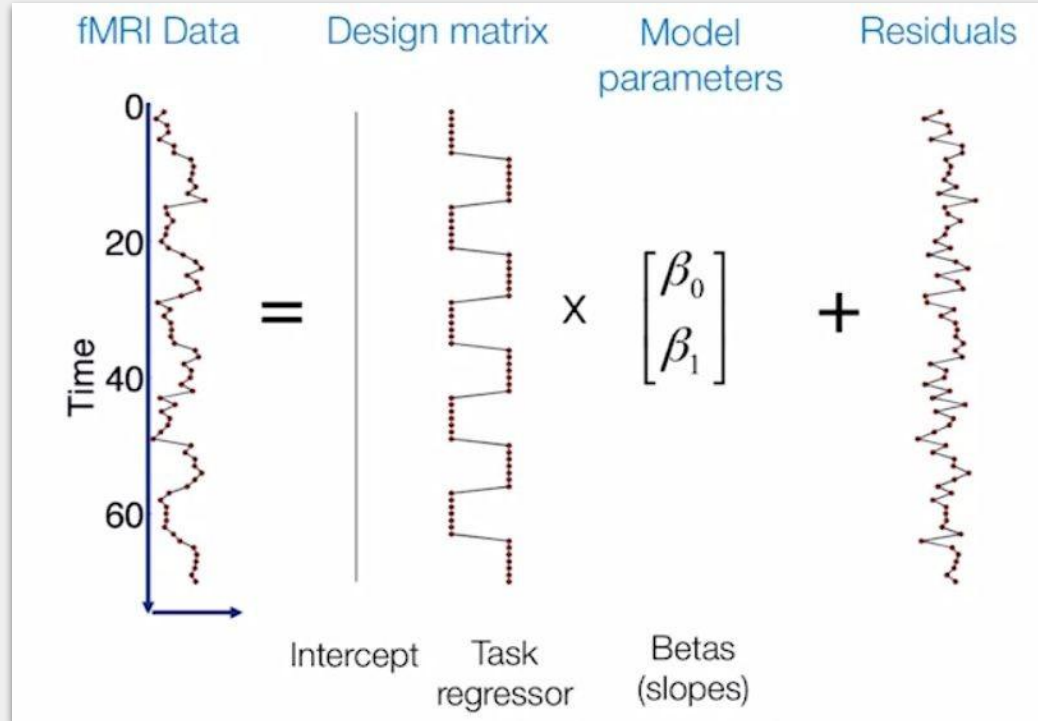
Events parameters



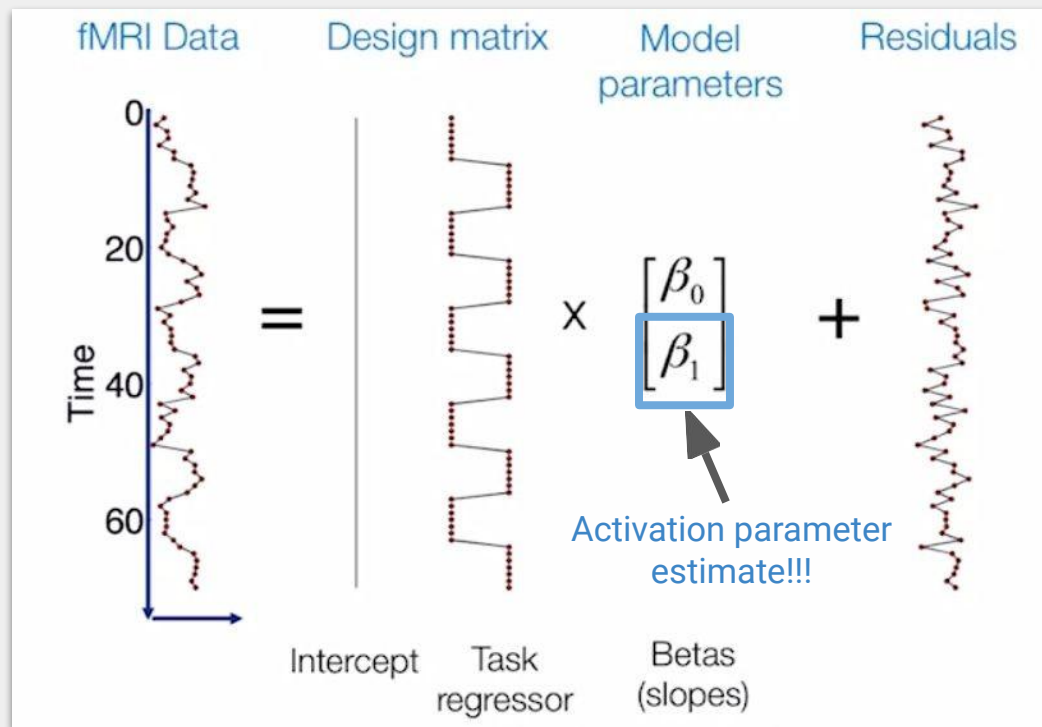
Events parameters



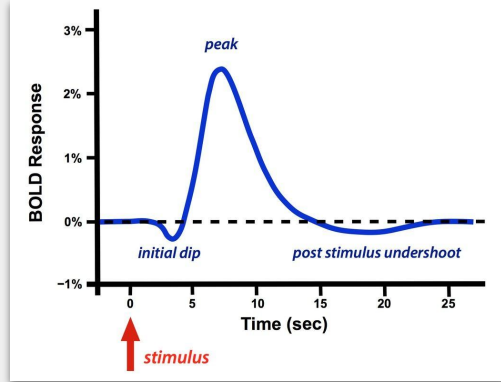
First level GLM



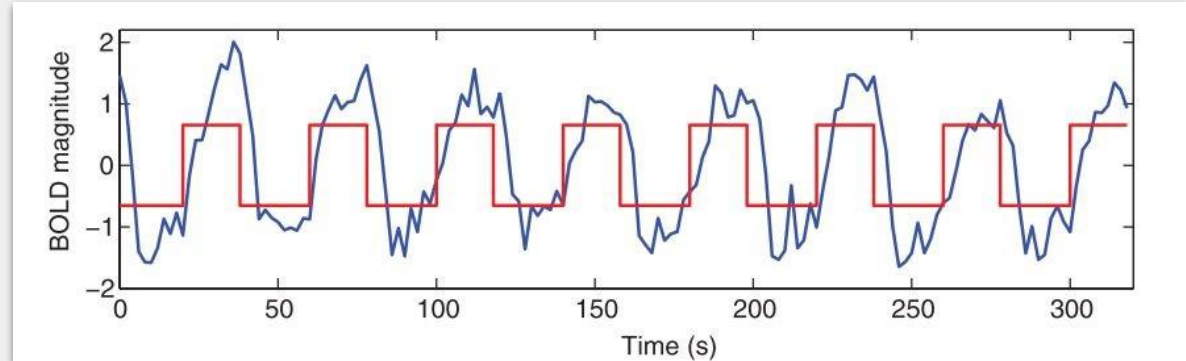
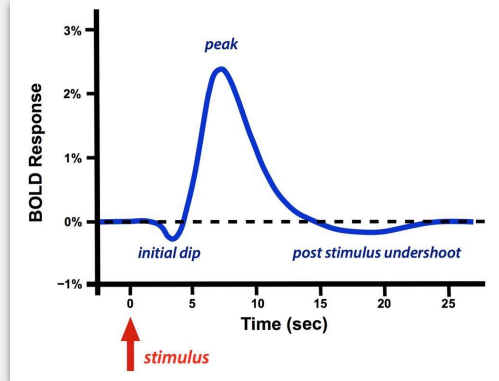
First level GLM



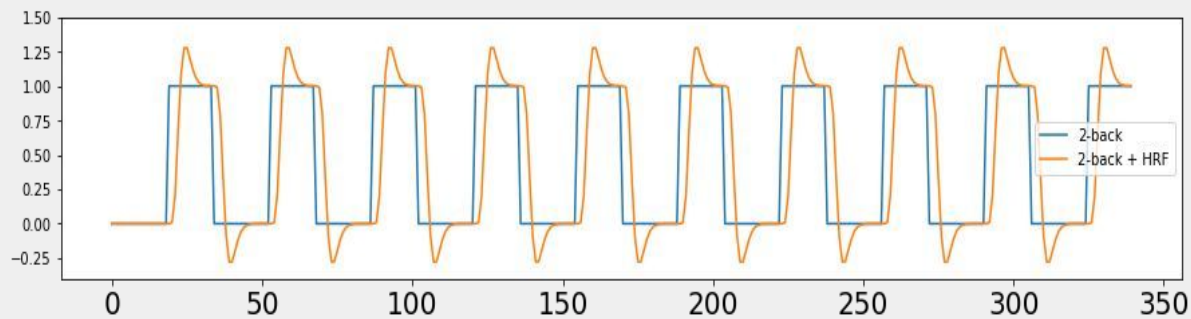
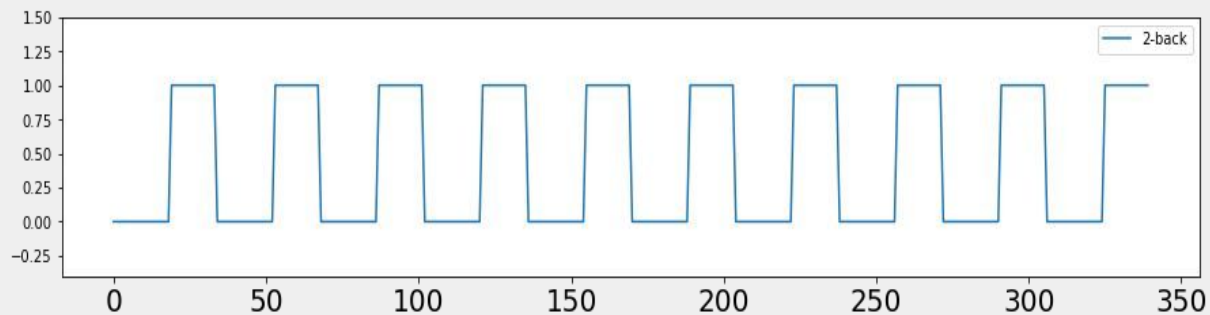
Haemodynamic delay!



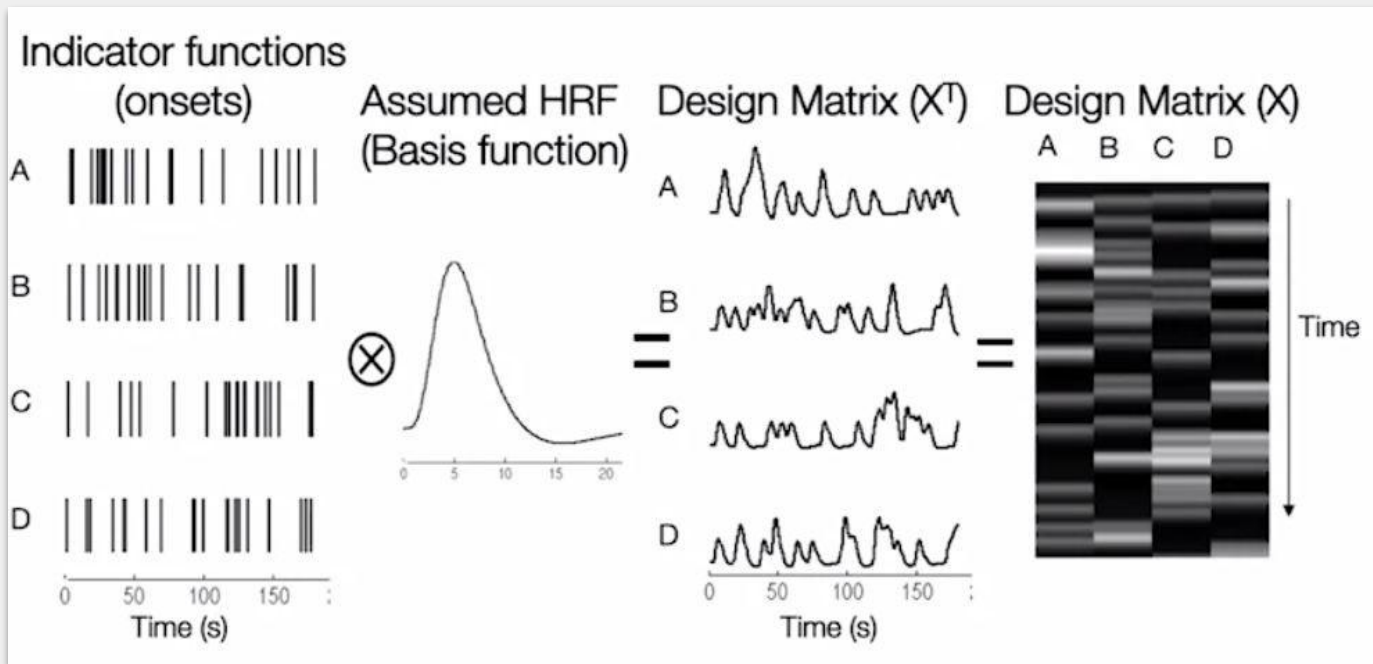
Haemodynamic delay!



Convolution



Building design matrix from events data

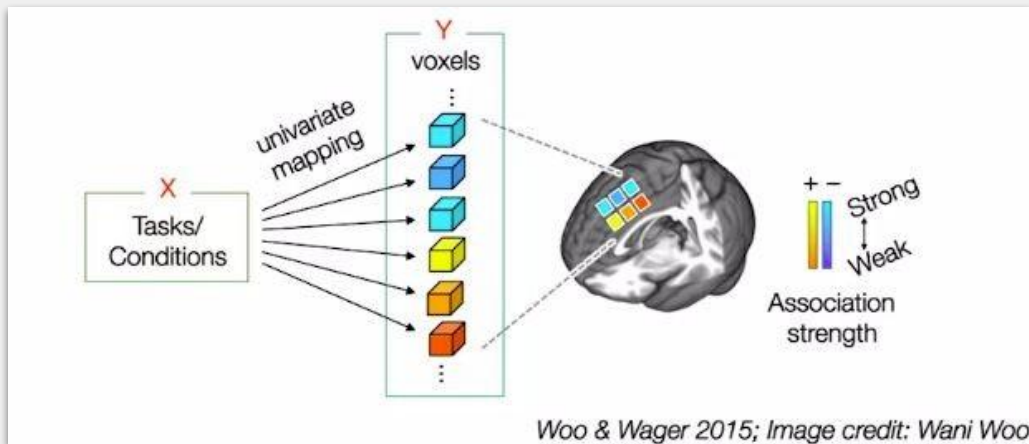




How to do this on the whole brain?

Mass univariate approach

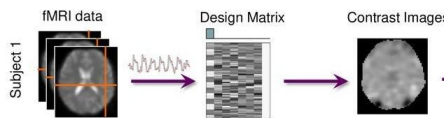
Assumes voxels are independent, each owns a separate model.



Analysis steps

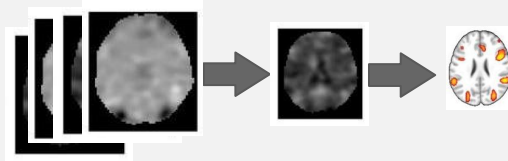
1-level analysis (within-subject; individual)

How was the brain active in
a one particular brain?

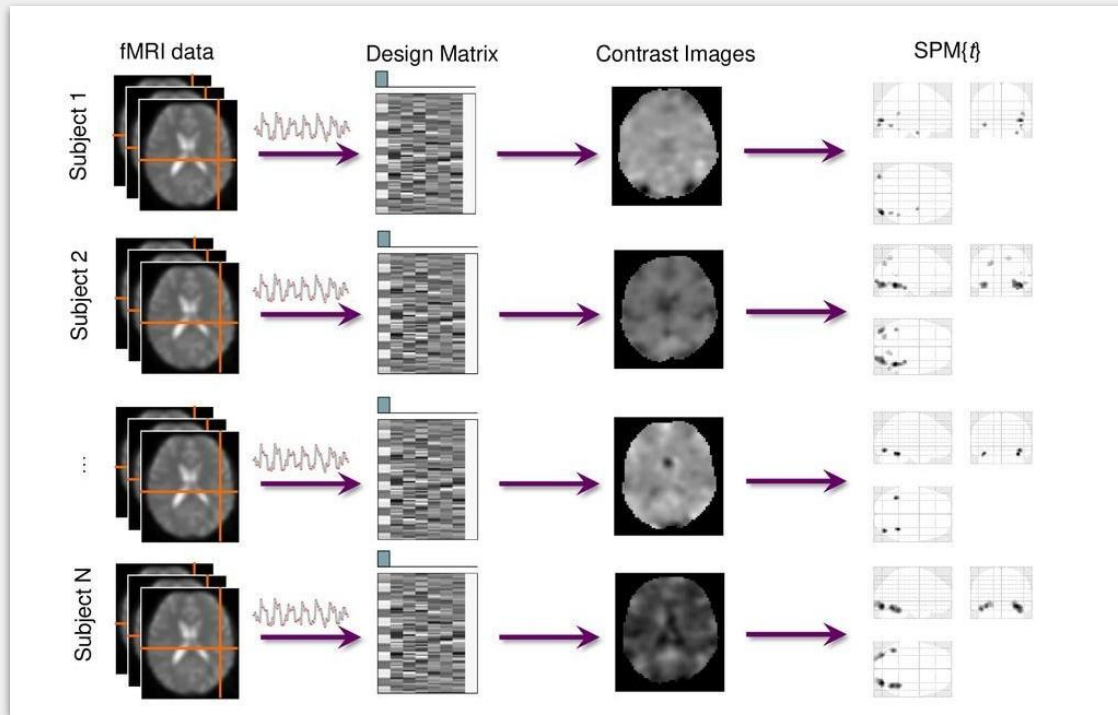


2-level analysis (across-subject; group)

How brain activations look
in general (for the whole
group)?
Differences between groups?

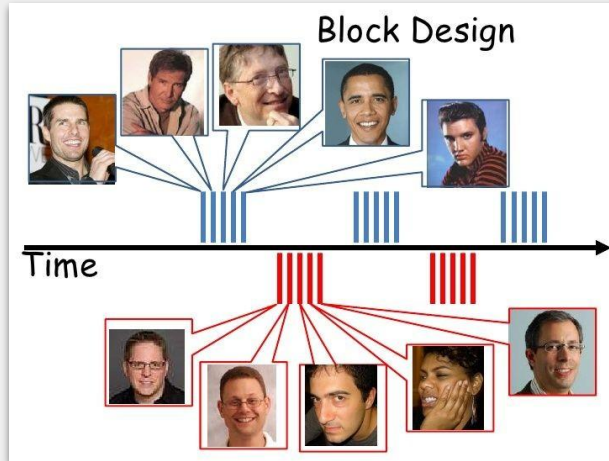


1-level analysis



Contrasts

Famous
people



Non-famous
people

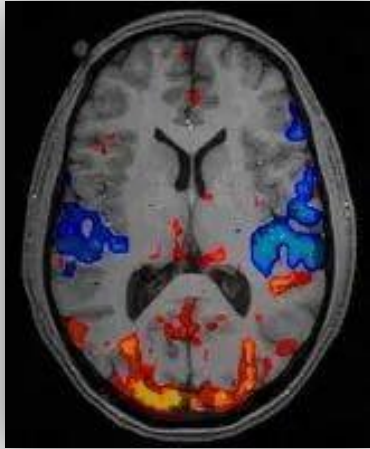
- Difference between conditions
- Each condition separately
- Average of two conditions

This functions can be assessed with different linear **contrasts**.

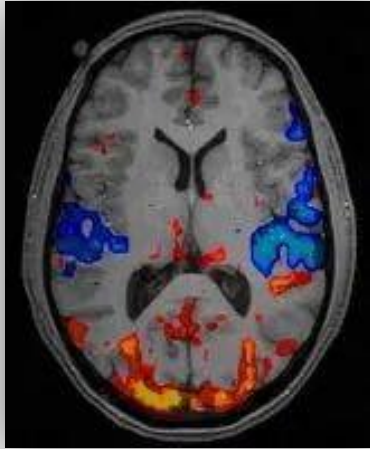
Contrast - a *linear combination* of GLM parameters.

Is this this combination **significantly** different from zero (T, F statistical testing)?

Multiple comparison correction

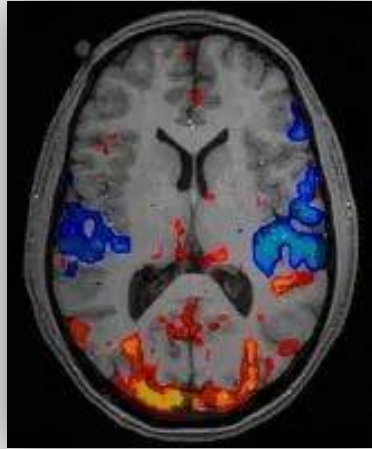


Multiple comparison correction



For 100 000 voxels
~ 5 000 false
positives!!!

Multiple comparison correction



For 100 000 voxels
~ 5 000 false
positives!!!

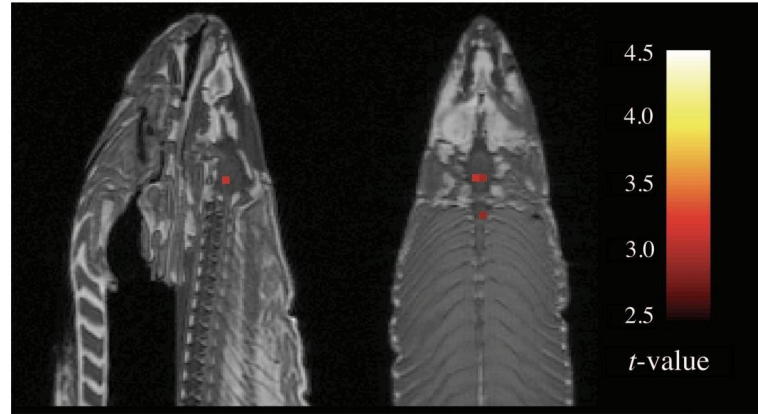


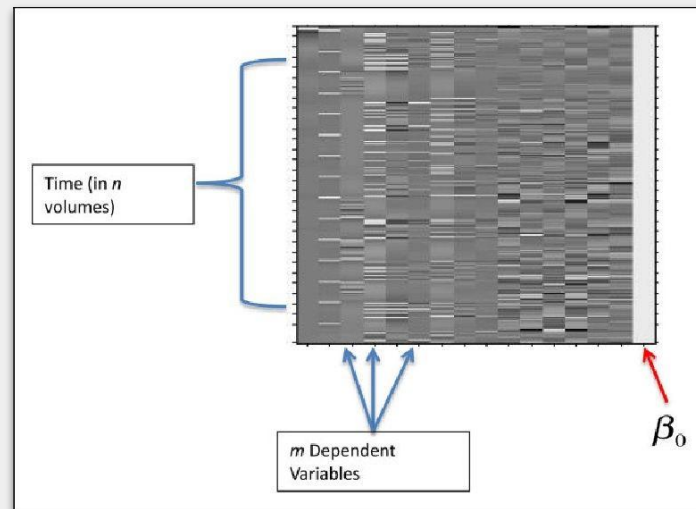
Fig. 1. Sagittal and axial images of significant brain voxels in the task > rest contrast. The parameters for this comparison were $t(131) > 3.15$, $p(\text{uncorrected}) < 0.001$, 3 voxel extent threshold. Two clusters were observed in the salmon central nervous system. One cluster was observed in the medial brain cavity and another was observed in the upper spinal column.

Important questions about task design

- Block vs. event-related?
- How many conditions?
- How many runs?
- What's repetition time (TR)?
- How stimuli were organized in time (onsets, durations)?



Nilearn:
Machine learning for Neuro-Imaging in Python

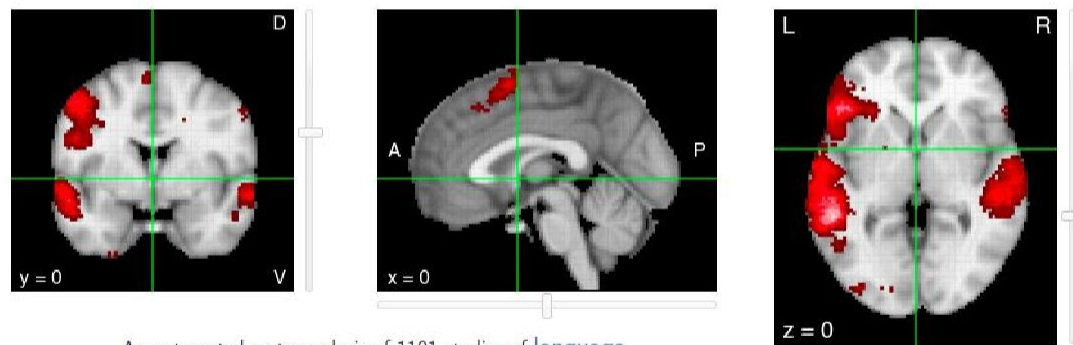


neurosynth.org



Neurosynth is a platform for large-scale, automated synthesis of functional magnetic resonance imaging (fMRI) data.

It takes thousands of published articles reporting the results of fMRI studies, chews on them for a bit, and then spits out images that look like this:



An automated meta-analysis of 1101 studies of [language](#)

Database Status

507891 activations reported in [14371 studies](#)

Interactive, downloadable meta-analyses of [1335 terms](#)

[150,000 brain locations](#)

Homework

1. GitHub Classroom

GLM analysis & fMRI results
plotting

Classroom for **GitHub**

Your course assignments on GitHub



Next



Functional connectivity