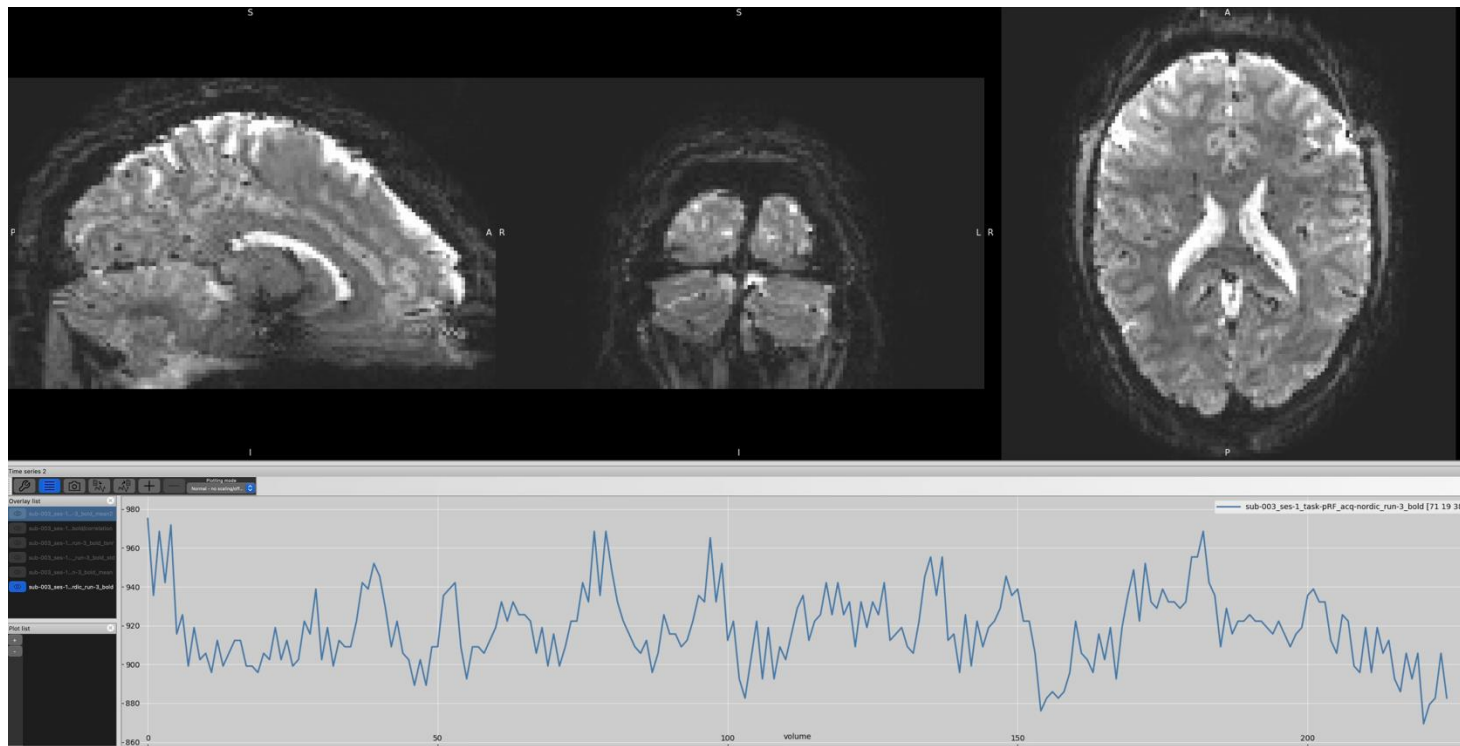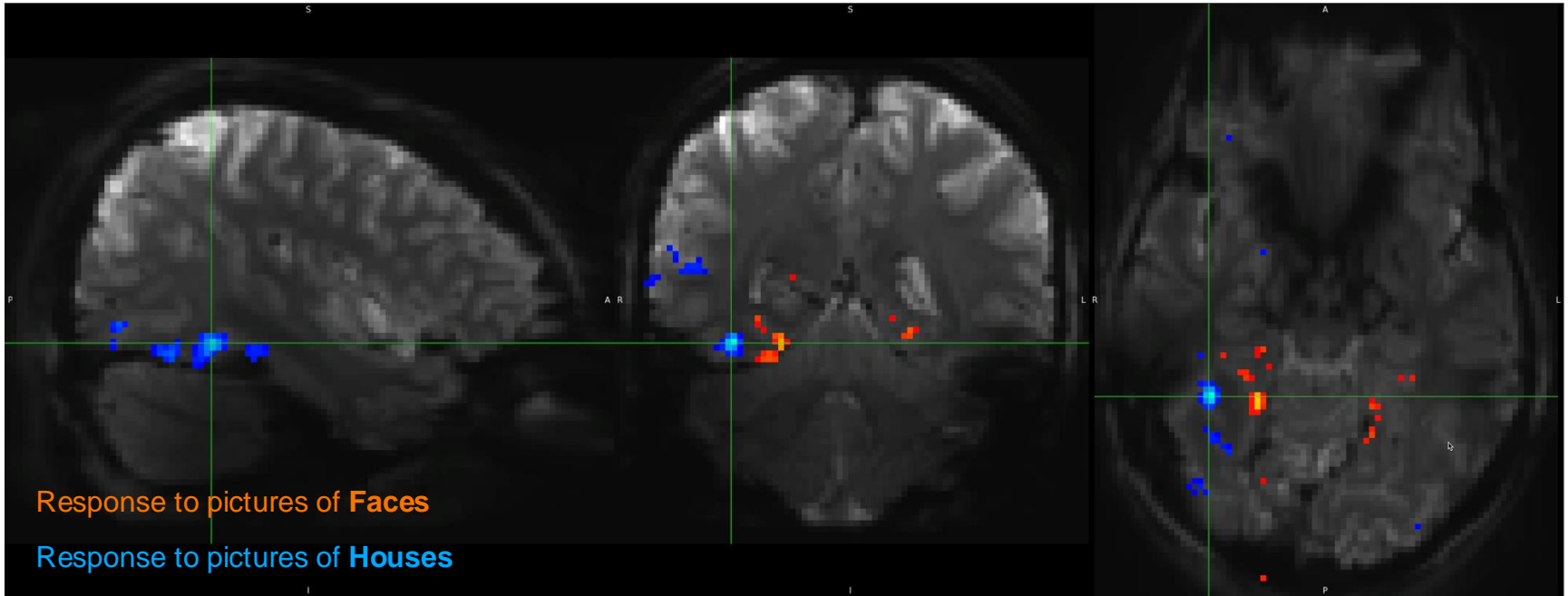# General linear model

Raw fMRI data

Each voxel time series is different, and many factors contribute to this difference
(experimental factors, local blood flow, breathing etc)

Statistical output

Response to pictures of **Faces**

Response to pictures of **Houses**

Despite the variable contributions of different factors, we can make robust **statistical inferences**

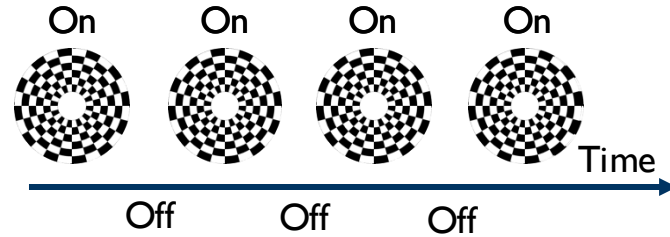→ **How do we link the experiment to the fMRI data?**

Linking the experiment to the data

Linking the experiment to the data

1) Experiment
Participants perform a paradigm
in which experimental components
are manipulated (stimuli, instructions,
prompted actions).
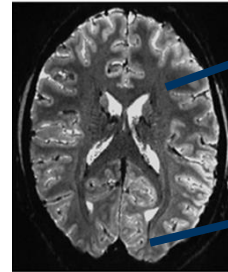
Example On-Off visual paradigm

2) Data
Recording MRI signals while participants
perform the experiment in the scanner.

Certain parts of the observed fluctuations
(not all) will be related to the experiment.

Raw fMRI data

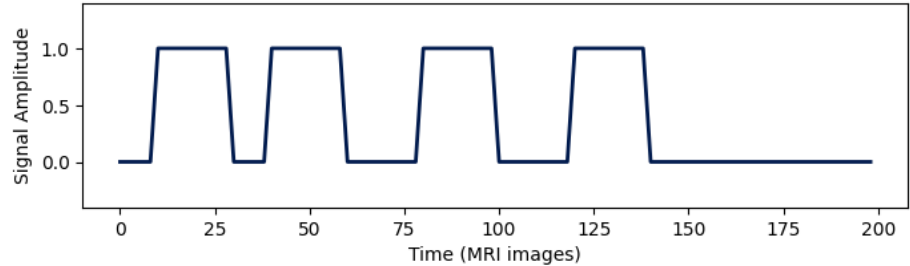Voxel time series 1

Voxel time series 2

Linking the experiment to the data

3) Regressor design
Regressors (or predictors) are **predicted time series** that are based on your experiment. For example, stimulus on = 1, stimulus off = 0.
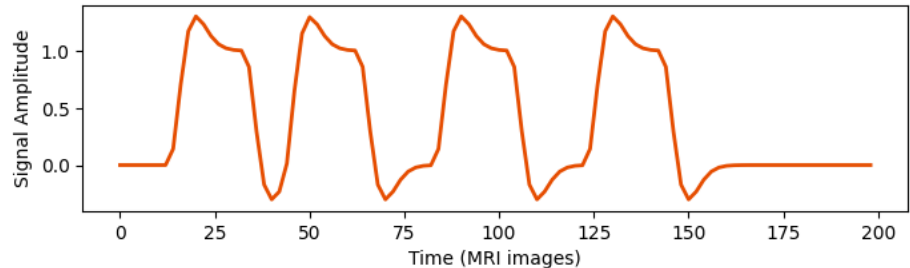
Simple regressor modeling stimulus presence

4) HRF Convolution
When linking the experiment to the data, we need to account for the **hemodynamic lag**.

We do this by **convolving the regressors** with the hemodynamic response function.

**Same regressor as above, convolved with the HRF**

Side quest: **What is convolution?**

Convolution is one of the most important techniques in **signal processing**.
It is a mathematical operation on **two functions** that produces a **third function.**

It expresses how one function modifies another by computing their overlap as one shifts over the other. Essentially, each point in **Function 3** is a **weighted sum of the points in Function 1**, with the **weighting coming from Function 2 (**called the "**Kernel**").
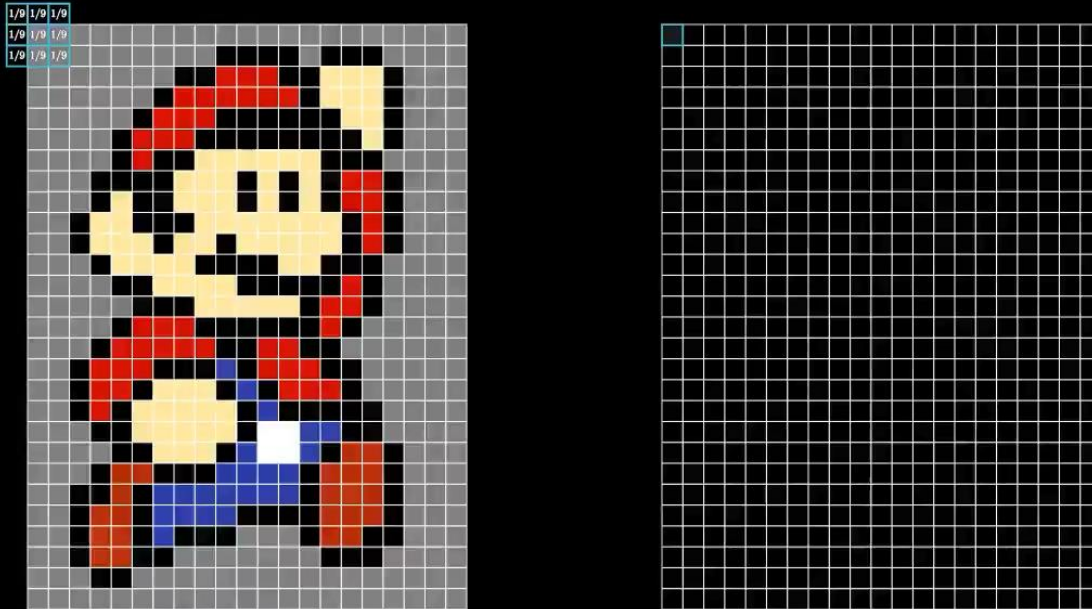
HRF Convolution Example
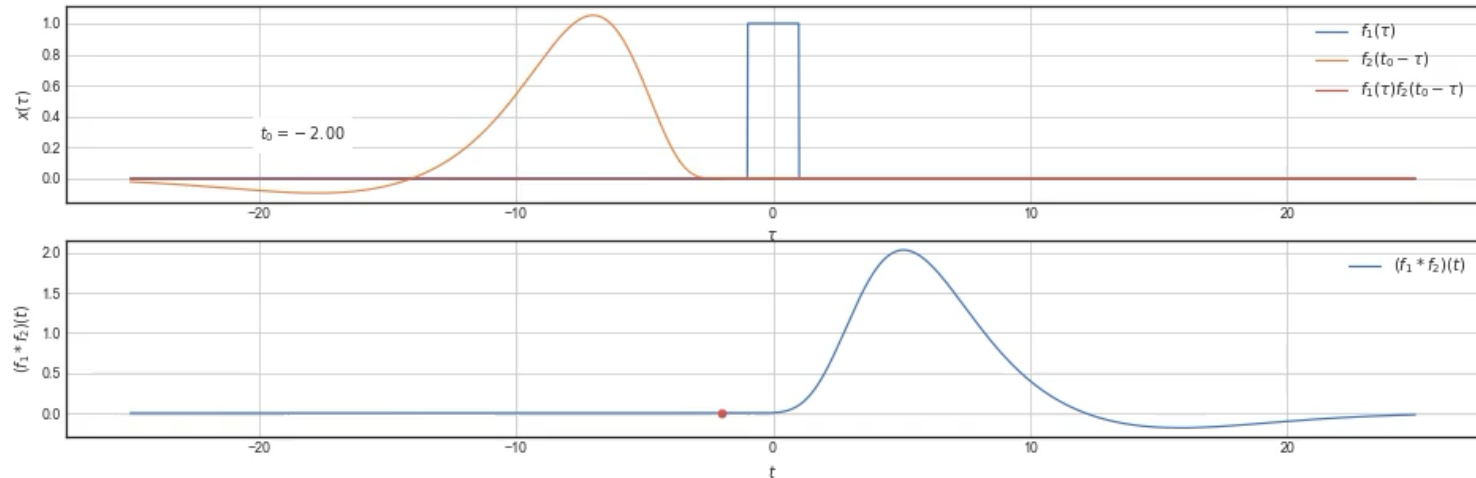
Side quest: **What is convolution?**

Convolution can be performed in **multiple dimensions** (1D for time series, 2D for images, 3D for volumentric data…). In **fMRI preprocessing and analysis**, convolution is used all the time!

Example of 2D Convolution

Video from https://www.3blue1brown.com

Side quest: **What is convolution?**

Convolution can be performed in **multiple dimensions** (1D for time series, 2D for images, 3D for volumentric data…). In **fMRI preprocessing and analysis**, convolution is used all the time!
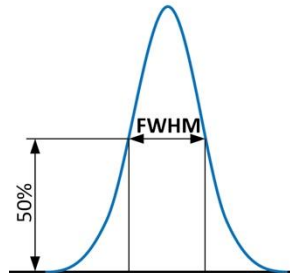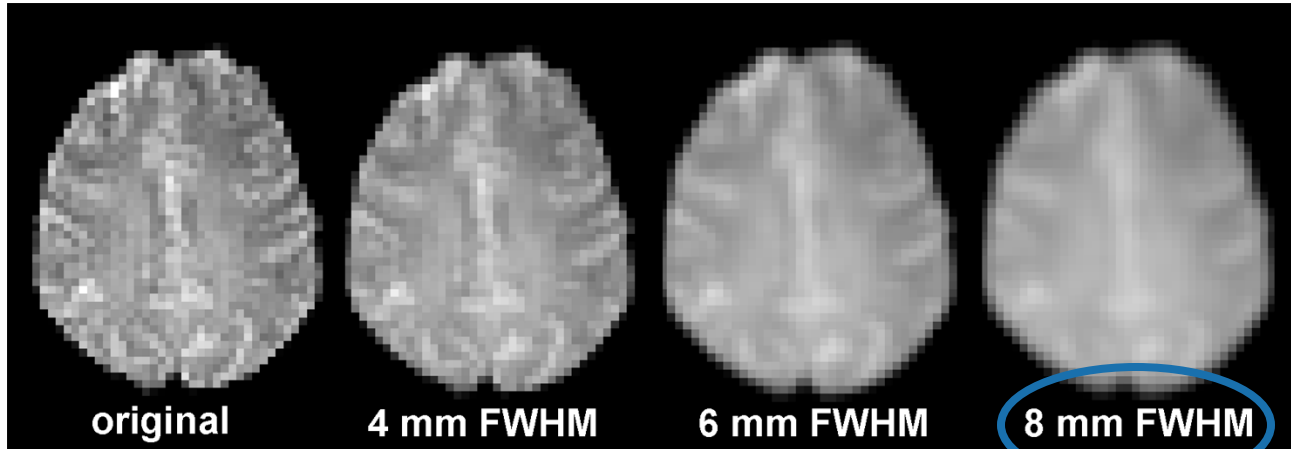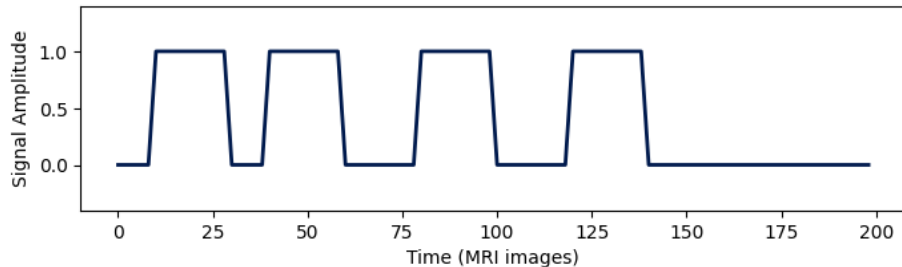
Example of 3D convolution (Spatial Smoothing)

original   4 mm FWHM   6 mm FWHM   8 mm FWHM

Gaussian smoothing kernel with 8mm Full-Width-at-Half-Maximum (FWHM)

Note that the smoothing kernel is 3D, unlike the 1D graph on the left.

Image: https://www.cnblogs.com/minks/p/5964403.html

Hemodynamic response function (HRF) convolution

Simple regressor modeling stimulus presence

HRF convolution

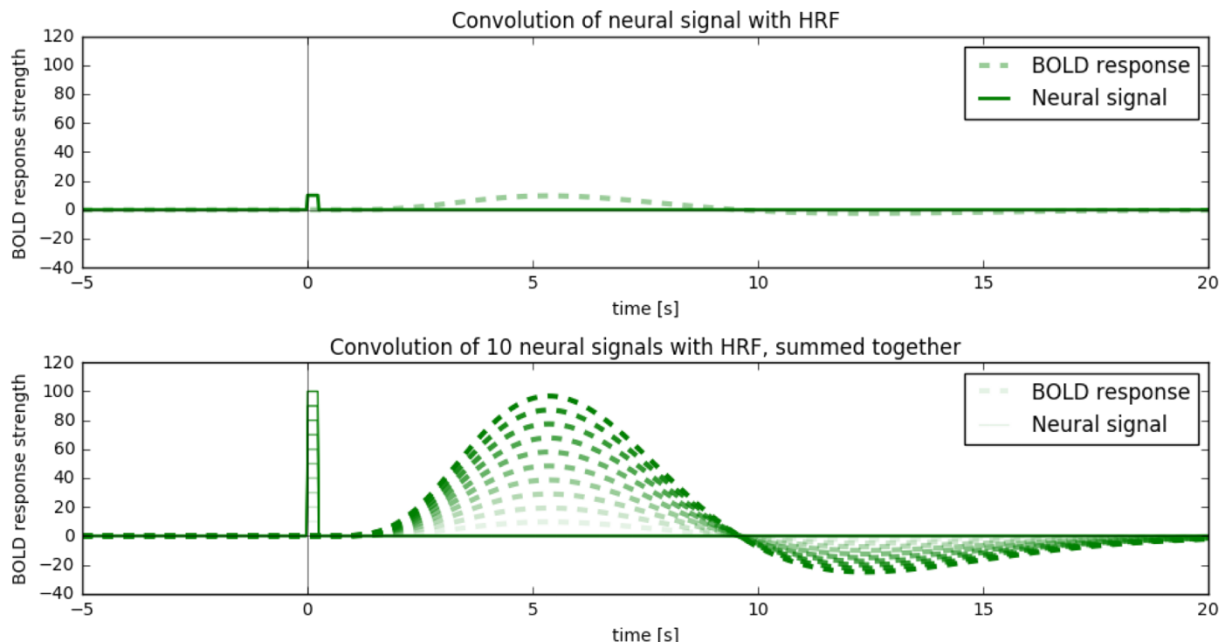**Predicted MRI signal (Convolved regressor)**
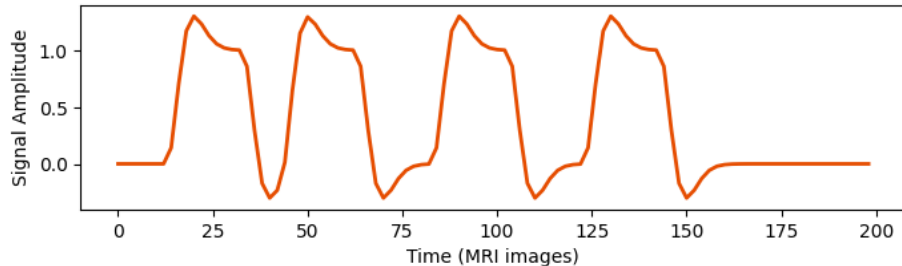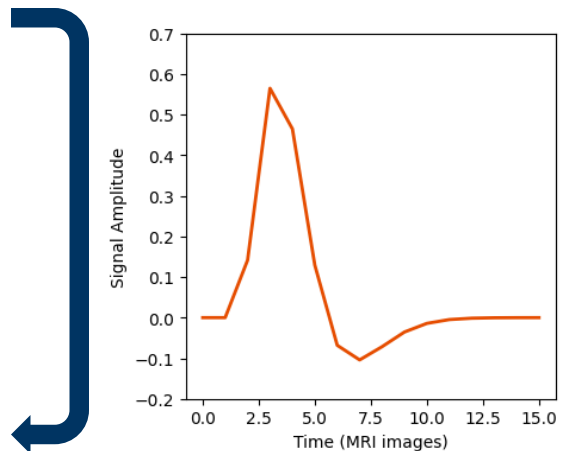
The convolved regressor reflects the predicted voxel time series

Hemodynamic response function (HRF) convolution

Assumption 1:
Fixed HRF shape.

Assumption 2:
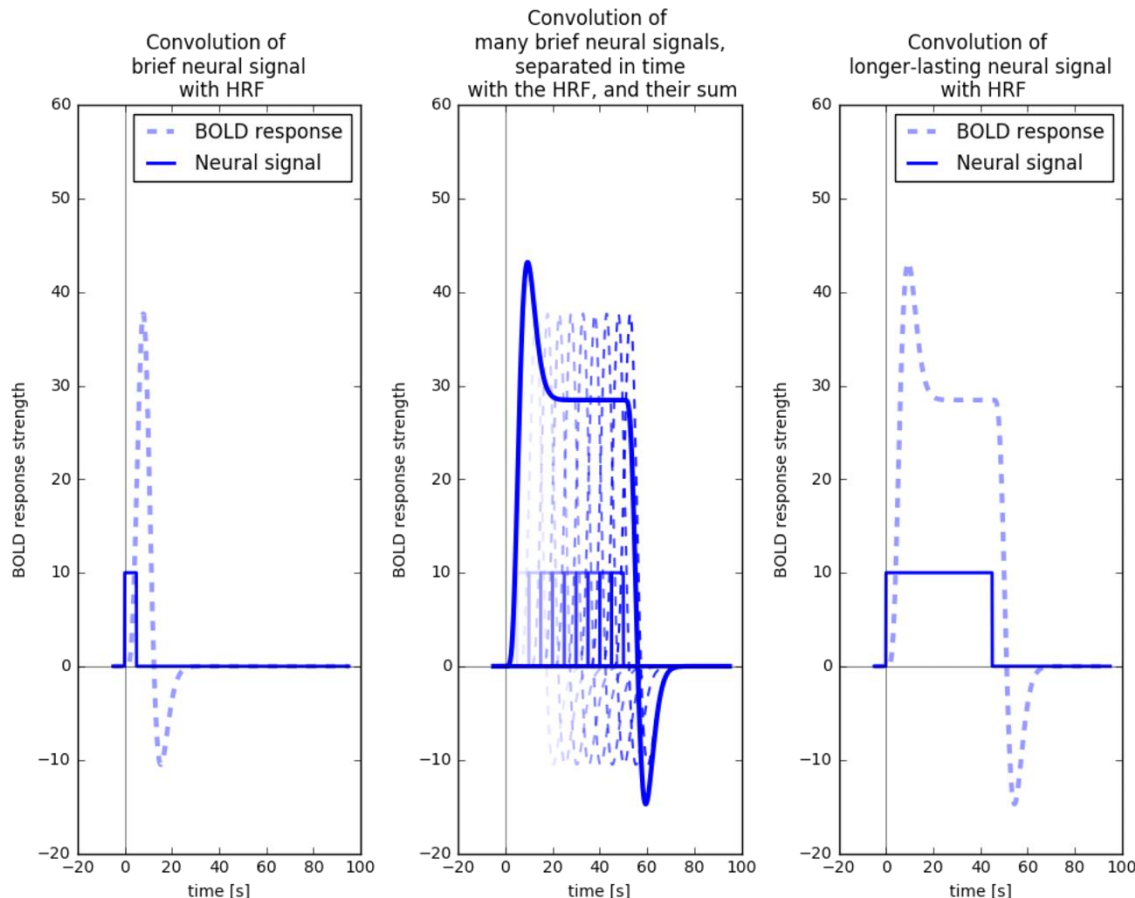BOLD Responses scale linearly with amplitude of the neural event.

Hemodynamic response function (HRF) convolution

Assumption 1:
Fixed HRF shape.

Assumption 2:
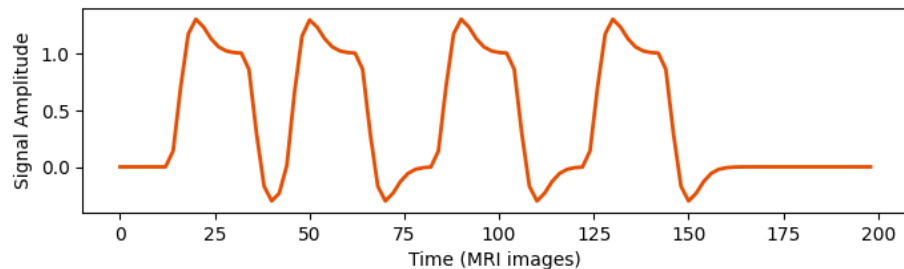BOLD Responses scale linearly with amplitude of the neural event.

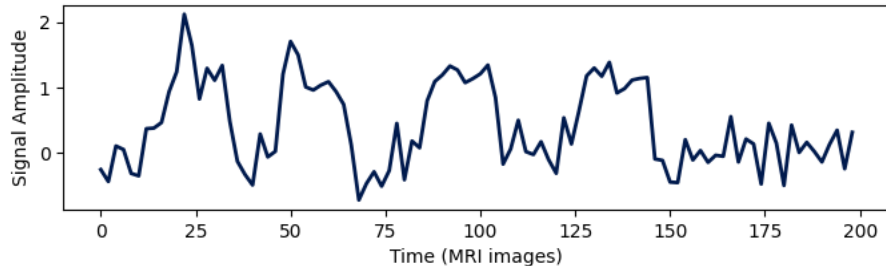**Assumption 3:**
Responses add up linearly in time.

Going from one regressor to many regressors
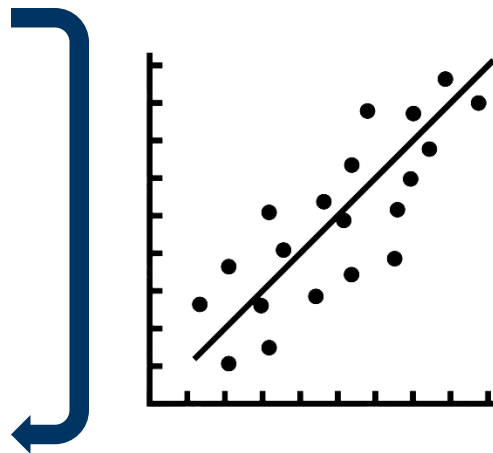
One regressor

Predicted MRI signal (Convolved regressor)

**Observed MRI time series of one voxel (data)**

Linear regression with 1 regressor

Estimates how well a regressor explains the changes in the observed time series

Many regressors

Following the introduced regressor-design logic, we build not one but many regressors for our experiment, all of which are then combined in one **Design Matrix**.

Time (MRI images)

Design Matrix

Regressors of interest | Nuisance regressors | Constant
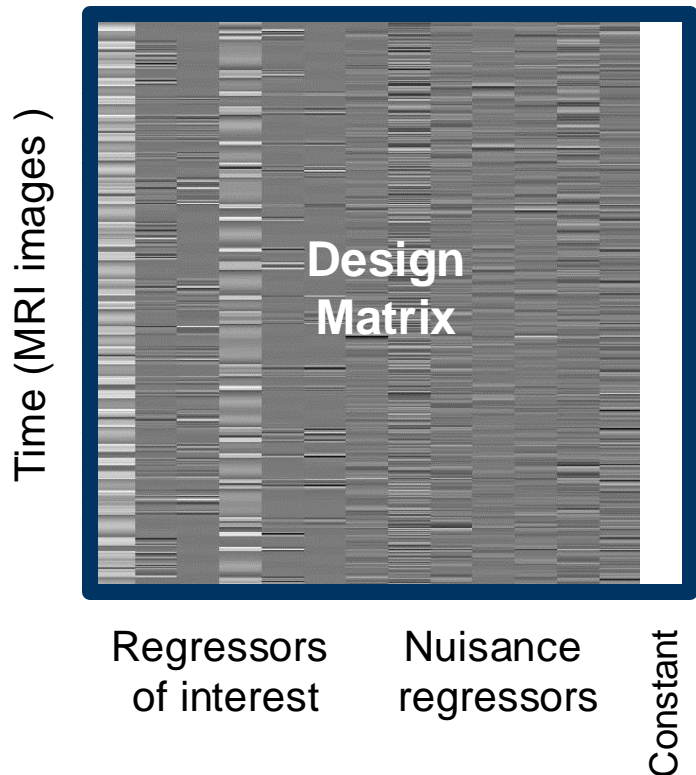
Regressors of interest
- Regressors you are interested in analyzing.
- Typically represent relevant task conditions (e.g., faces, houses…)

Nuisance regressors
- Model variance you are NOT interested in and that you want to account for.
- Always includes head motion.
- Other typical ones include breathing, heartbeat, button presses, noise components

Constant term
- Models the mean of the voxel time series.

To fit many regressors simultaneously, we use a **general linear model.**

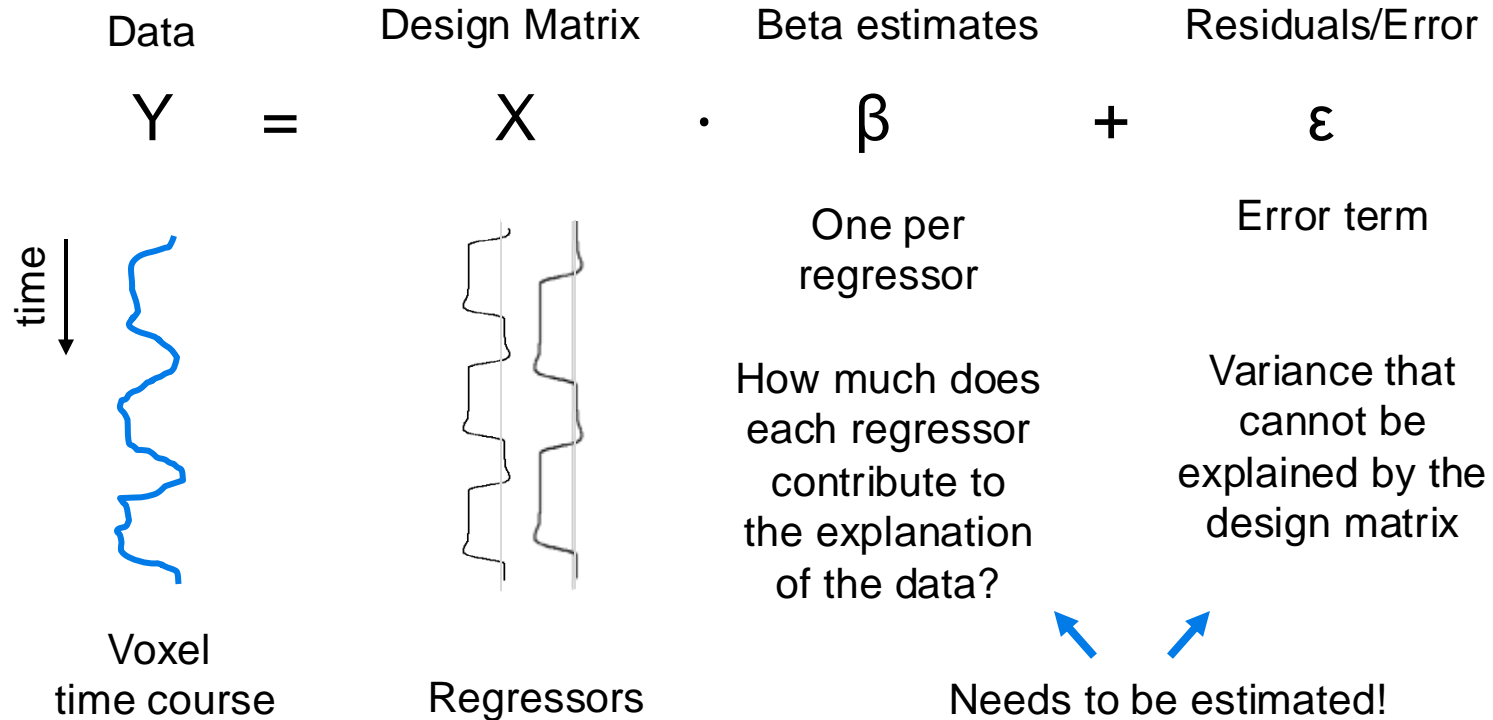# Mass-univariate General Linear Model (GLM)

Statistical method that models **dependent variables** (i.e. data) as weighted combinations of **independent regressors** (i.e. design matrix).

By **fitting all regressors simultaneously**, we (hope to) disentangle the regressors' **individual explanatory power.**

It is called "mass-univariate" because it is **fit to every single voxel** separately.

General Linear Model (GLM)

Data		Design Matrix		Beta estimates		Residuals/Error

$$Y = X \cdot \beta + \varepsilon$$

One per regressor

Error term

How much does each regressor contribute to the explanation of the data?

Variance that cannot be explained by the design matrix

time

Voxel time course
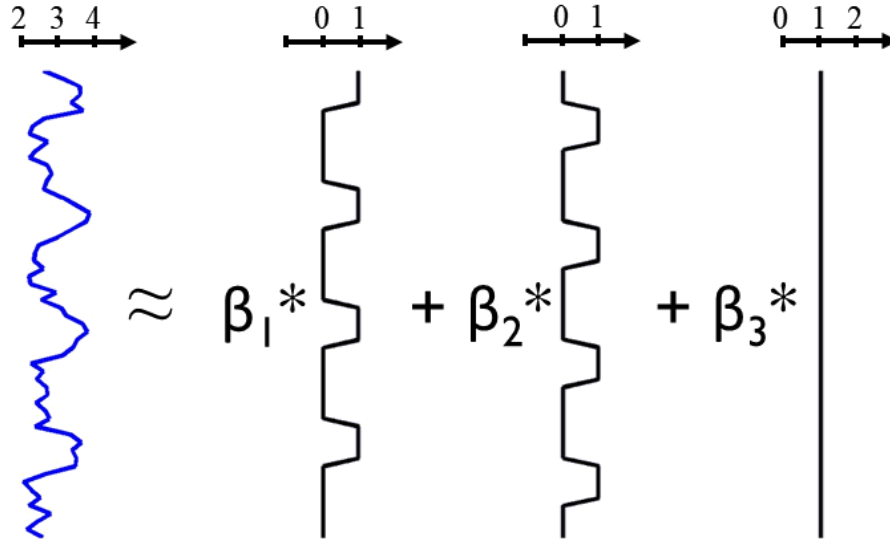
Regressors

Needs to be estimated!

General Linear Model (GLM)

The estimation entails finding the beta estimates (β) such that the linear combination of the regressors 'best' matches the data.
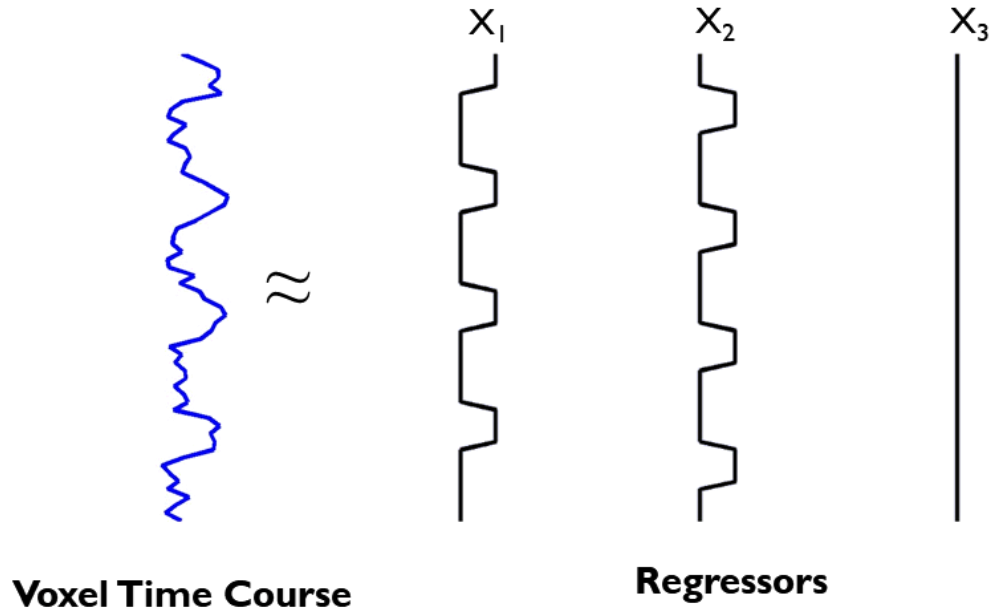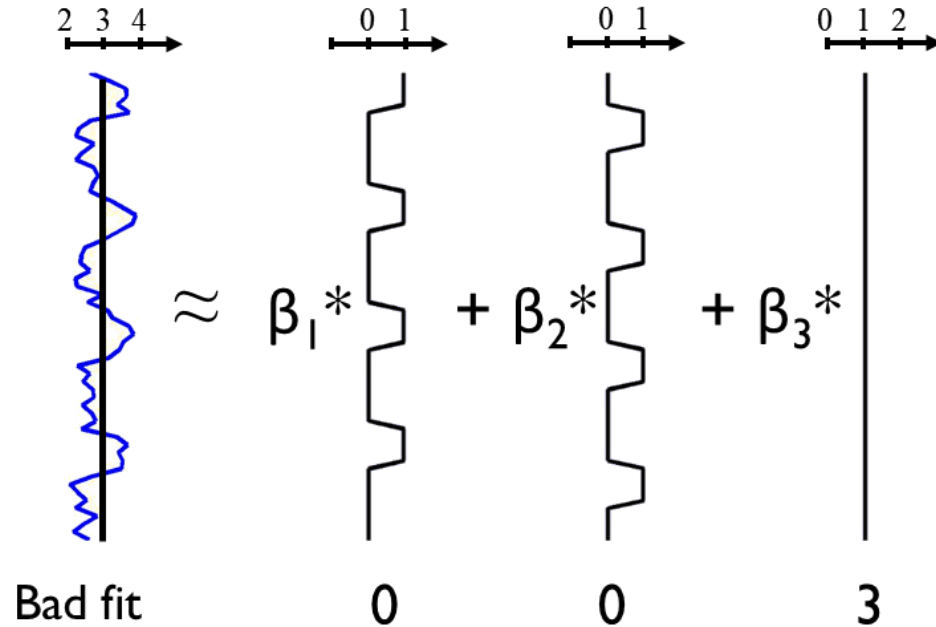

Voxel Time Course   **Regressors**

General Linear Model (GLM)

Modeling voxel time courses with a linear combination of hypothetical time-series (regressors). Same model for each voxel → One beta estimate per regressor per voxel.

General Linear Model (GLM)

Modeling voxel time courses with a linear combination of hypothetical time-series (regressors).
Same model for each voxel → One beta estimate per regressor per voxel.

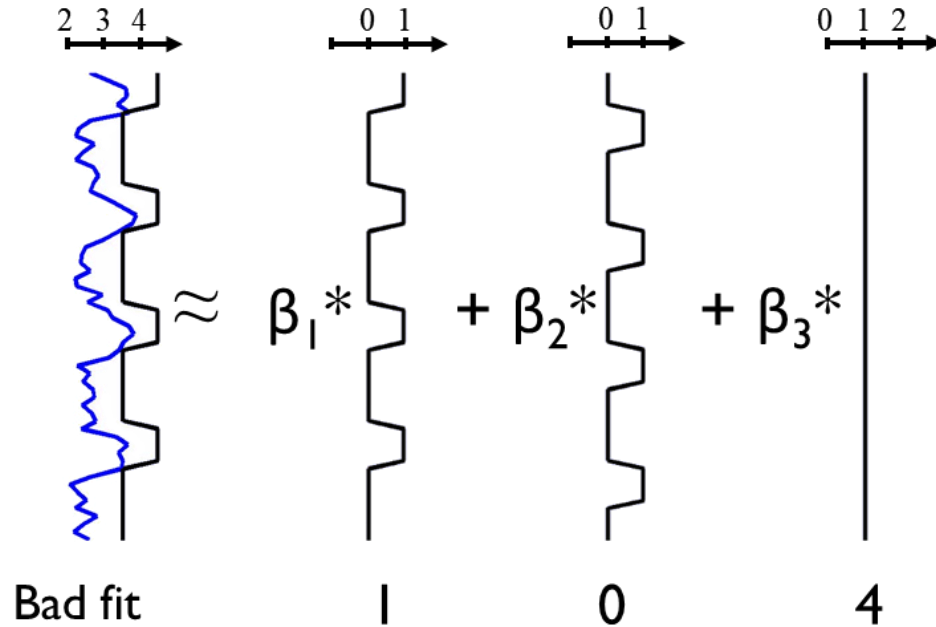Bad fit   0   0   3

General Linear Model (GLM)

Modeling voxel time courses with a linear combination of hypothetical time-series (regressors).
Same model for each voxel → One beta estimate per regressor per voxel.

Bad fit ≈ $\beta_1 *$ [ ] + $\beta_2 *$ [ ] + $\beta_3 *$ [ ]
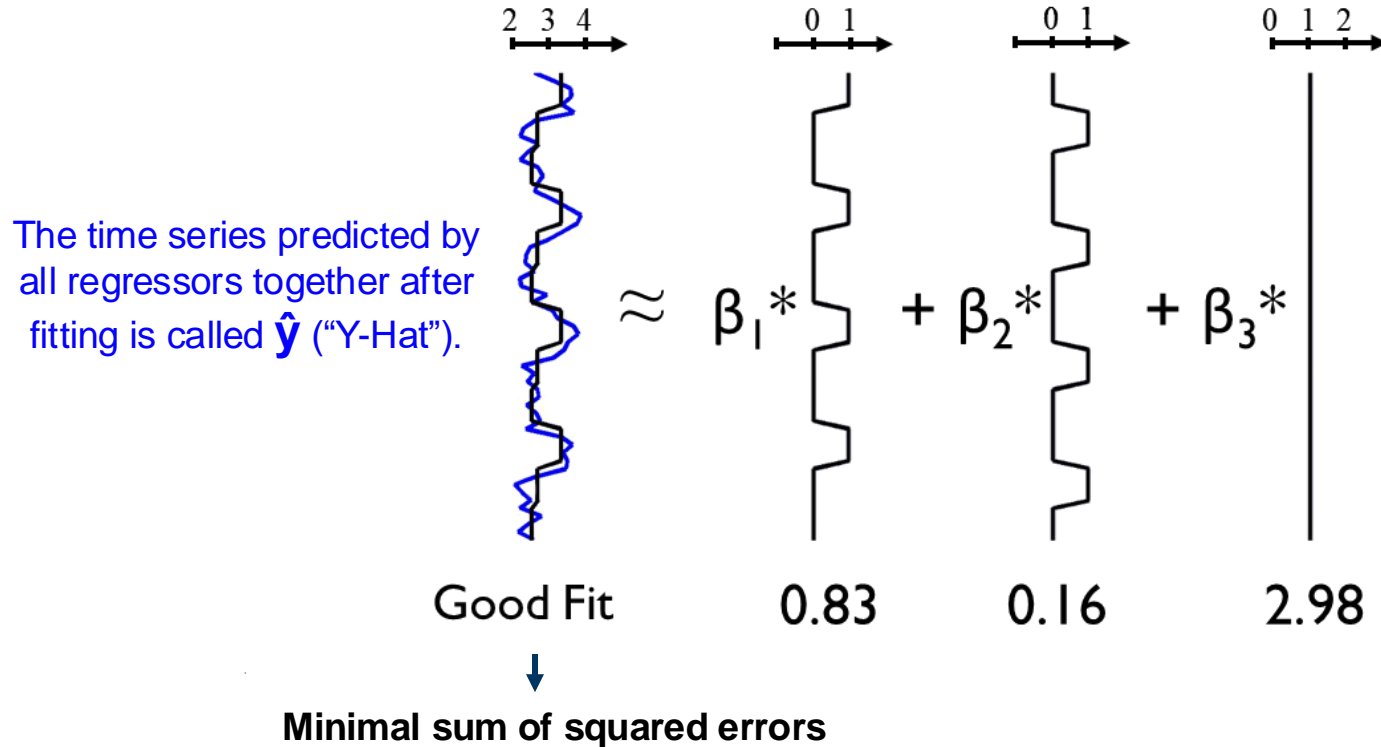
| | 1 | 0 | 4 |

General Linear Model (GLM)

Modeling voxel time courses with a linear combination of hypothetical time-series (regressors).
Same model for each voxel → One beta estimate per regressor per voxel.

The time series predicted by all regressors together after fitting is called $\hat{\mathbf{y}}$ ("Y-Hat").

$$\approx \beta_1 * \quad + \beta_2 * \quad + \beta_3 *$$

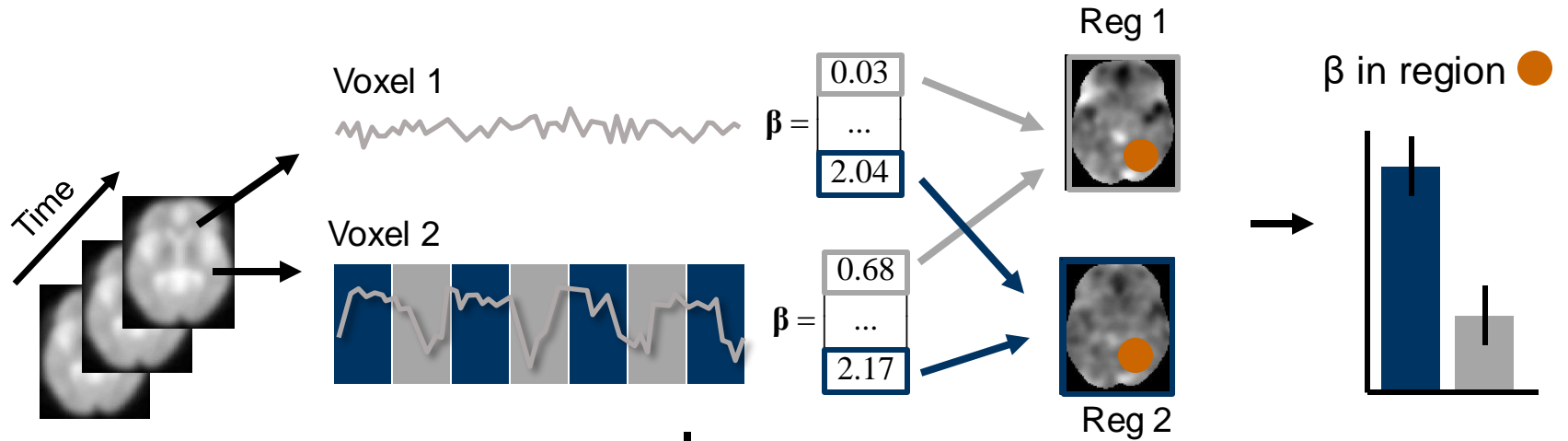Good Fit   0.83   0.16   2.98

↓

Minimal sum of squared errors

General Linear Model (GLM)

We can perform this analysis for all voxels simultaneously using matrix-based linear algebra, meaning that **it is fast**!

tldr: The GLM gives you a matrix containing one beta estimate per voxel and regressor

GLM - Statistical inference

Voxel 1

Time

Voxel 2

$\beta =$ | 0.03 |
| ... |
| 2.04 |

$\beta =$ | 0.68 |
| ... |
| 2.17 |

Reg 1

Reg 2

β in region ●

The **t-Statistic** tells us how much an estimated effect (beta) differs from zero relative to the data variability.
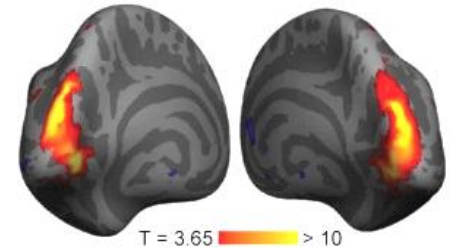
We often differentiate between **Regions of interest (ROI) analyses** and **Voxel-wise analyses.**

t-test

■ > ▧

Overlay statistical map on T1-scan

Thresholded at e.g. p<0.001.

T = 3.65 ▬ > 10

GLM - Statistical inference

How is the t-Statistic computed?

t-Score   **Standard error of the model**   **Degrees of freedom**

$$t = \frac{\hat{\beta}}{\hat{\sigma}}$$

$$\hat{\sigma} = \sqrt{\frac{\Sigma_{i=1}^{N}(\hat{y}_i - y_i)^2}{df}}$$

$$df = N - P$$

β = Beta estimates   ŷ = Model prediction   N = number of time points
σ = Standard Error   y = Data   P = number of regressors
                            df = Degrees of freedom

tldr: You get a t-value and associated p-value for each voxel and regressor
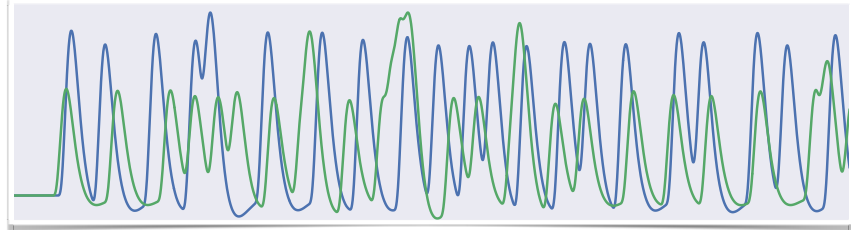
Example analysis:
Identifying voxels with preference for faces or houses
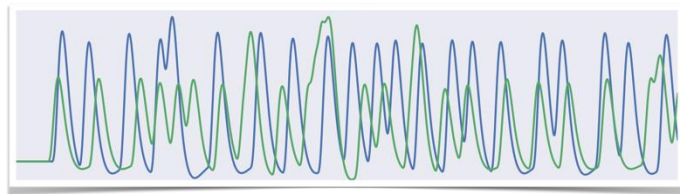
Example analysis: Faces vs. Houses

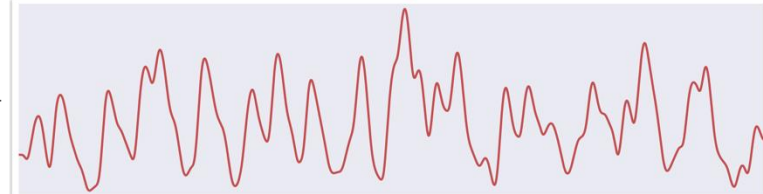**Face and house stimuli**

Regressors modeling faces and houses

Fitting all regressors to the data using a GLM
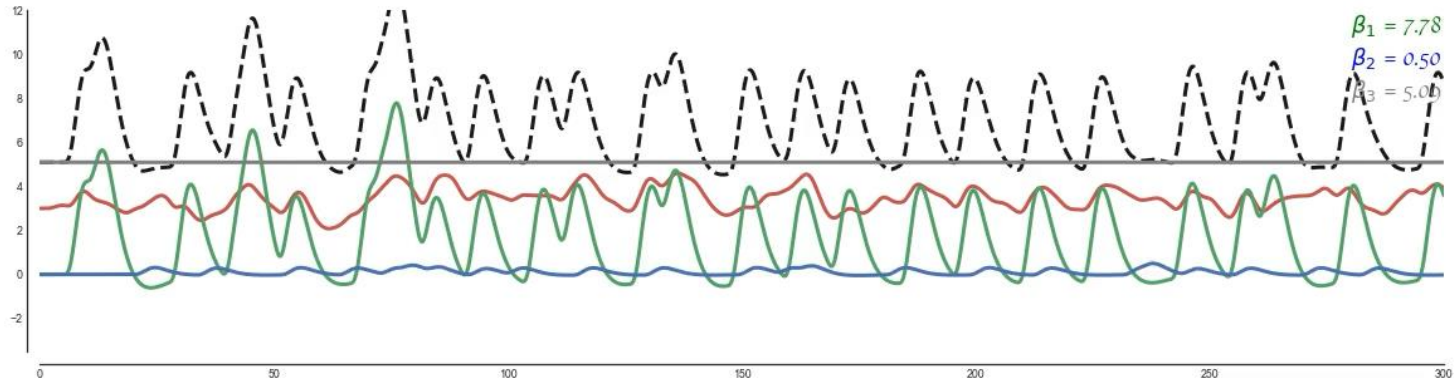

x_1
x_2

Model Regressors

y_1

Data (single voxel)

$$y_i = a_i \cdot x_1 + b_i \cdot x_2 + c_i + N(0,\sigma)$$

Example analysis: Faces vs. Houses

Visualizing different model fits

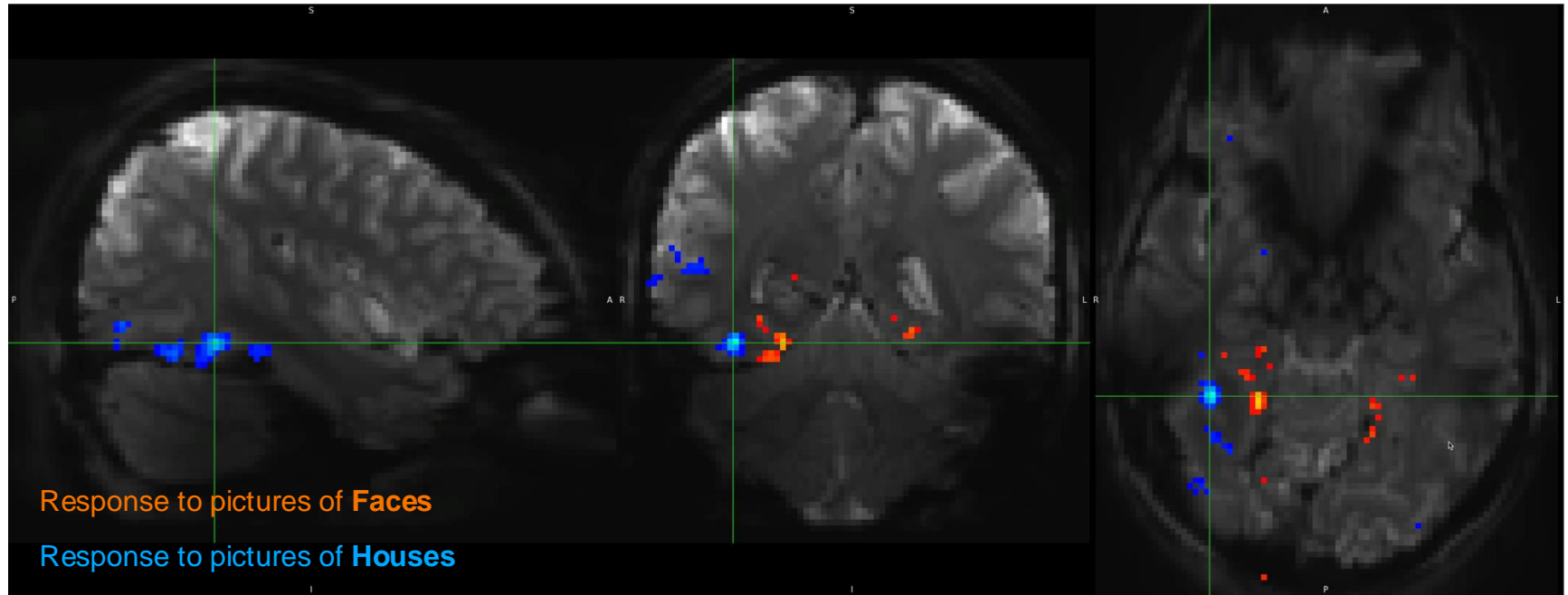Dashed line = Linear combination of all fitted regressors (ŷ)

Red = Data

Green = Face regressor

Blue = House regressor

Quiz: Which regressor explained the data best?

Example analysis: Faces vs. Houses

Response to pictures of **Faces**

Response to pictures of **Houses**

The colored blobs are the t-statistic estimated for the two regressors thresholded at a certain p-value (p<0.001)

fMRI contrasts

fMRI contrasts

We often want to know whether the BOLD signal was stronger in one condition vs. another.
(i.e. compare betas estimated for different regressors).

Most fMRI software packages allow you to do this by defining "**Contrast vectors**" that encode the comparison.
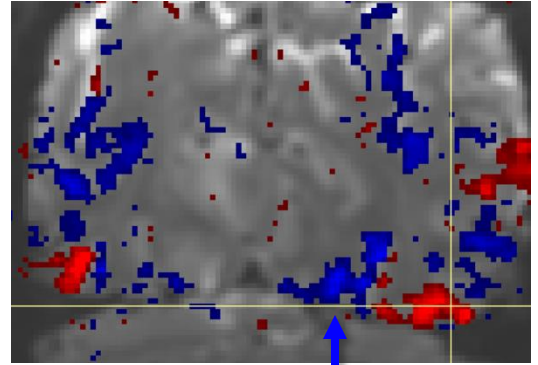
Faces > [Houses, Objects]

Negative contrast? Yes!

Test: **Contrast vector:**

Faces > Houses   Faces = 1, Houses = -1

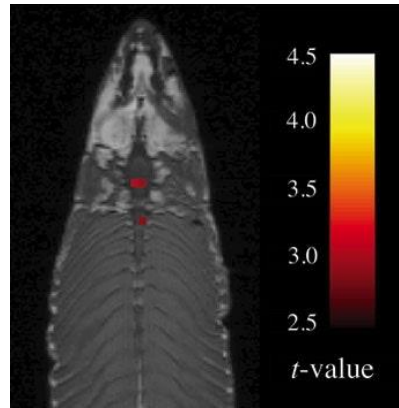Faces > [Houses, Objects]   Faces = 2, Houses = -1, Objects = -1

Contrasts must always be balanced such that the contrast vectors average to zero.

tldr: Contrasts make it easy to compare beta estimates obtained for different conditions

Bennet et al. 2009

Multiple comparison correction

Multiple comparison correction

What is a p-value?

The probability of obtaining a t-score as extreme as, or more extreme than, the observed result, assuming that the null hypothesis is true (i.e., the data is just noise).
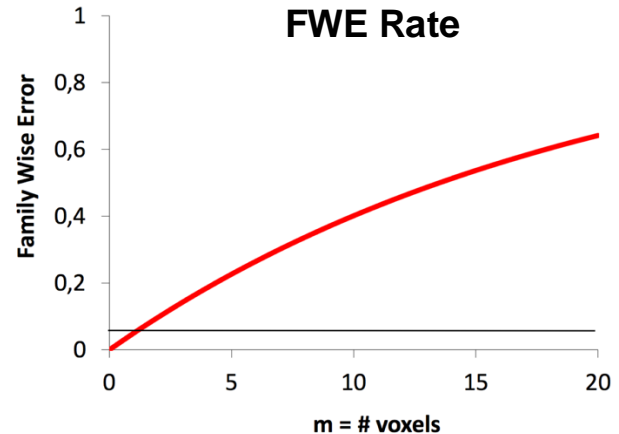
What is the multiple-comparison problem?

When fitting the same GLM to >100.000 voxels, many will show a significant effect even though there is none.

These spurious results are called **False Positives.**
The rate at which they occur is called the
Family-wise error rate (FWE).

FWE Rate

Family Wise Error

m = # voxels

Multiple comparison correction

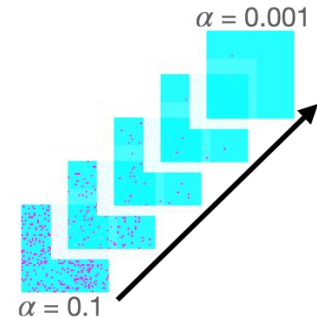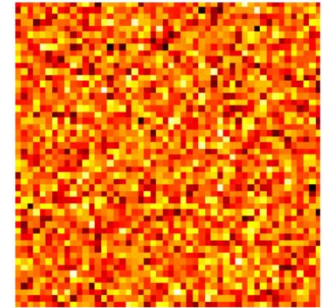## How do you correct for multiple comparisons?

Bonferroni correction

- Divide the significance level by the number of tests.
- Very conservative, few **False Positives**.
- Risks high number of **False Negatives**.
- Assumes independence across voxels.

False Discovery Rate correction

- Controls the proportion of false positives among significant results.
- Less strict and more sensitive than Bonferroni.
- Reasonable trade-off between false positives and false negatives.
- Assumes that tests are statistically dependent.

P-values in pure noise

$\alpha = 0.001$

$\alpha = 0.1$

Multiple comparison correction
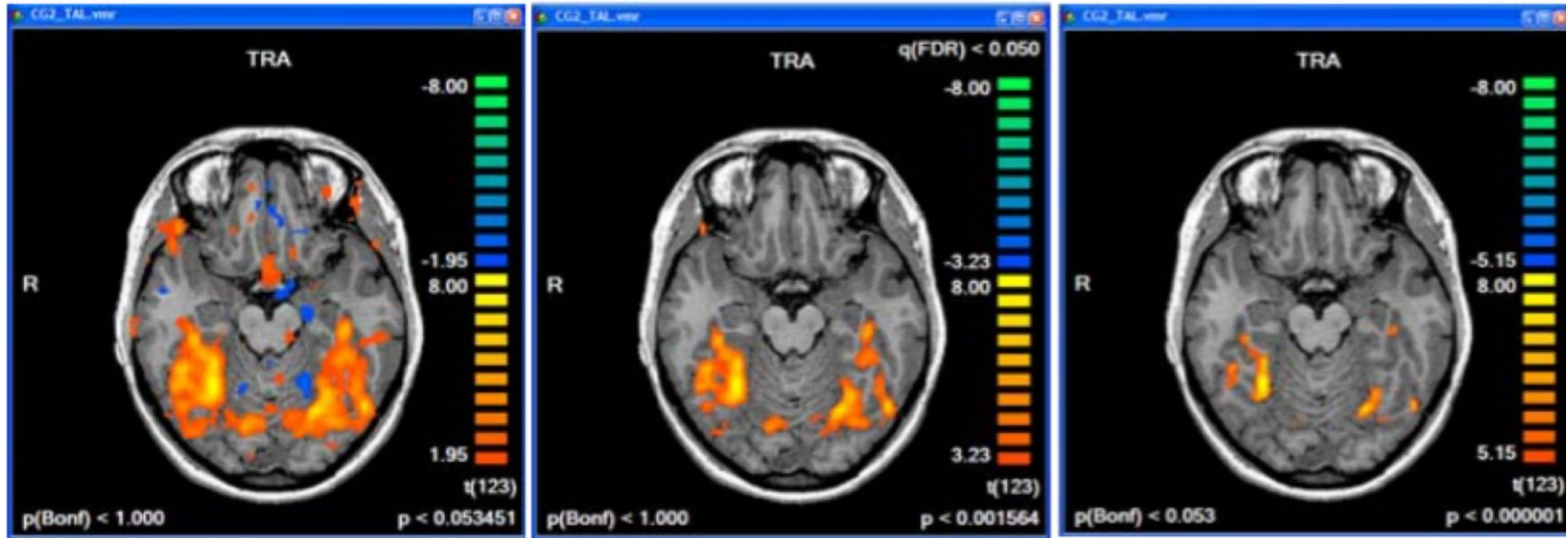
No MCC
p = 0.05
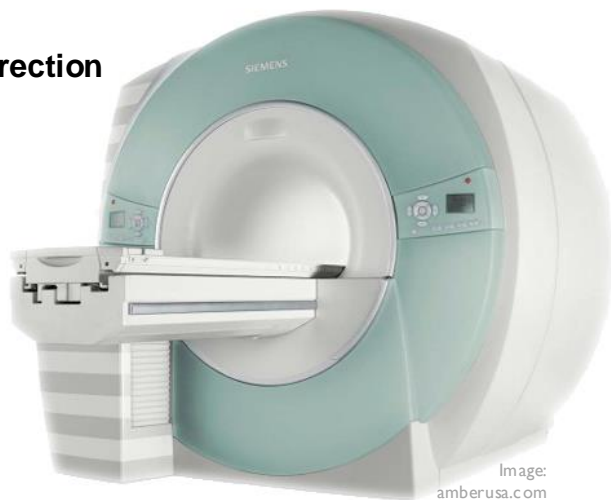
FDR MCC
q = 0.05

Bonf. MCC
p(corr) = 0.05

tldr: When running many statistical tests, you need to correct for multiple comparisons, and there are multiple ways of doing this with different pro's & con's.

Key terms to remember

- **General linear model**
- **On-Off paradigm**
- **Voxel time series**
- **Regressor**
- **Design matrix**
- **Convolution**
- **HRF-Convolution**
- **Kernel**
- **Linear regression**
- **Mass-univariate**
- **Regressors of interest**
- **Nuisance regressors**
- **Constant term (in a GLM)**

- **Minimal sum of squared errors**
- **Beta estimate**
- **Error term, Residuals**
- **Y-Hat**
- **t-Statistic**
- **Regions of interest analysis**
- **Voxel-wise analysis**
- **Multiple comparison correction**
- **fMRI contrasts**

Happy data analysis!