FMRI Data analysis

Analysis? What Analysis?

Different ways to analyze the signal contained in fMRI data:

- Resting state data (search for correlations between voxels in a given brain)

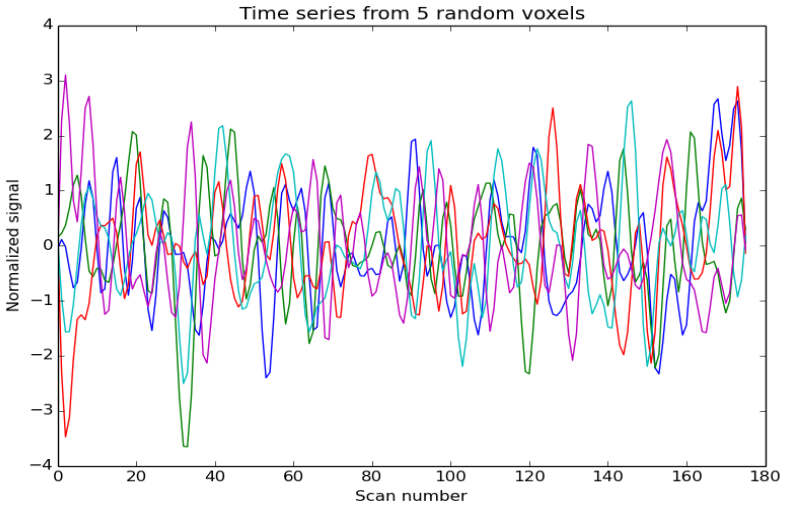
- Inter subject correlations (correlations between voxels across brains)

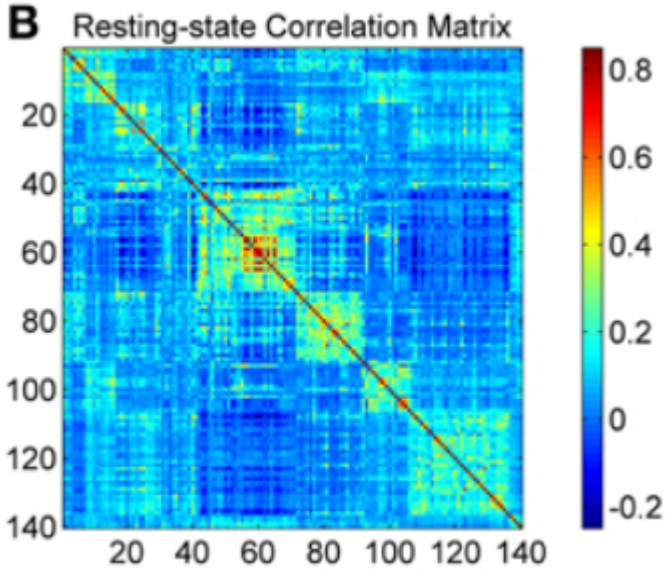
- Detection of task related activations in each voxel (search for correlation between behavior/stimulation and voxels' time courses (univariate analyses))

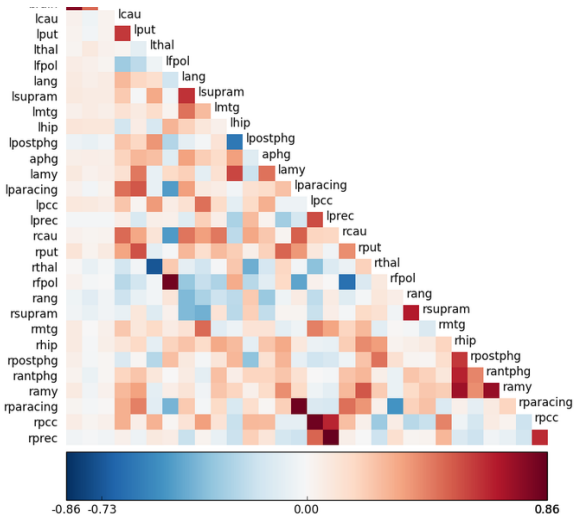
- Representational Similarity Analysis & Pattern Classification (Brain decoding)

Resting state data

In the absence of task, measure ‘spontaneous fluctuations of brain activity’. For two voxels, or two brain regions, one can compute the correlation of their temporal time courses.







Using Independent Component Analysis (ICA), to identify networks in resting state data

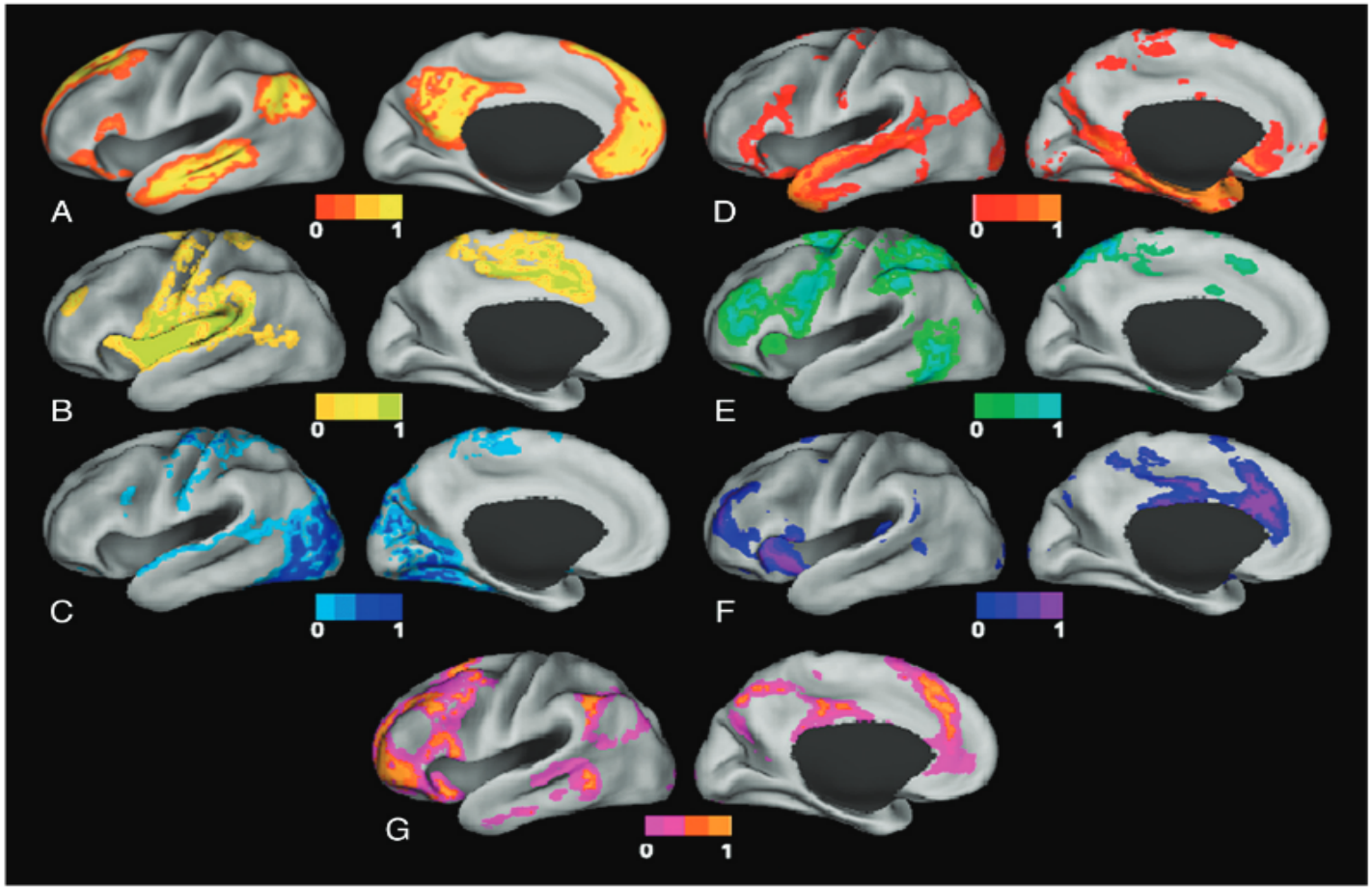
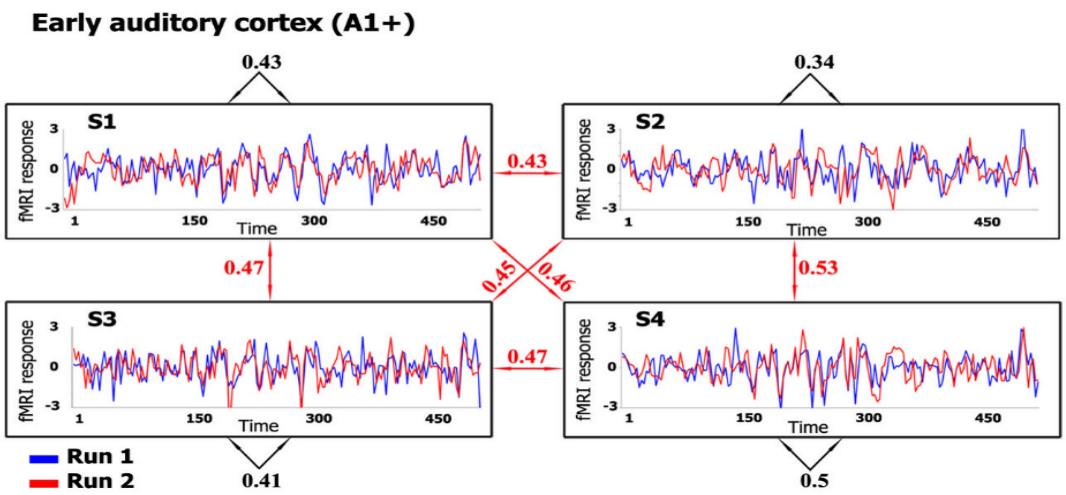


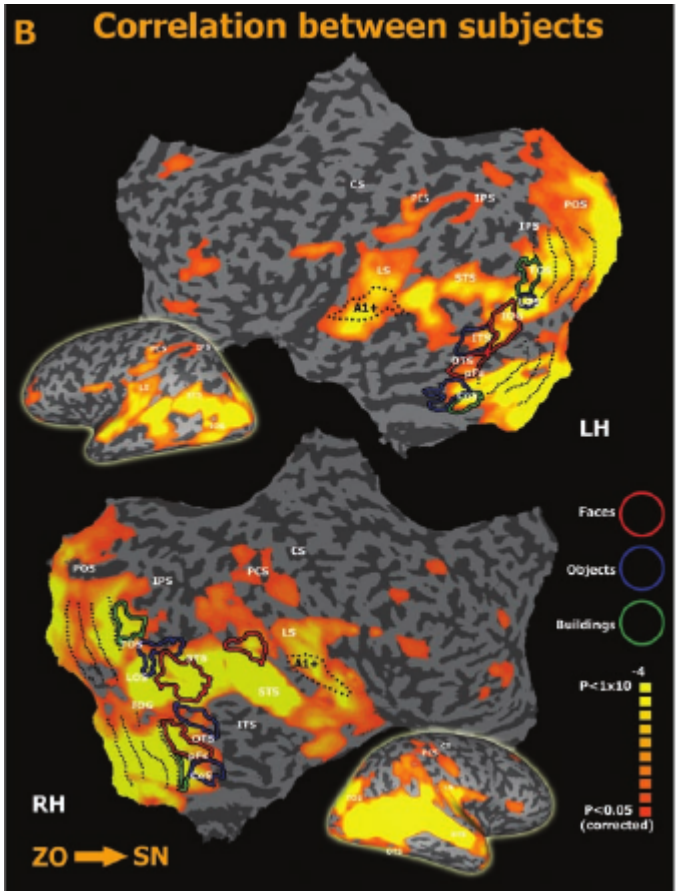
FIG 1. Surface plots of RSNs. *A*, Default mode network. *B*, Somatomotor network. *C*, Visual network. *D*, Language network. *E*, Dorsal attention network. *F*, ventral attention network. *G*, Frontoparietal control network.

See Damoiseaux et al. (2006) *PNAS*

Inter subjects Correlations (ISC)



Hasson, U. 2004. “Intersubject Synchronization of Cortical Activity During Natural Vision.” Science 303 (5664): 1634–40.



Beware: these are not “activations”

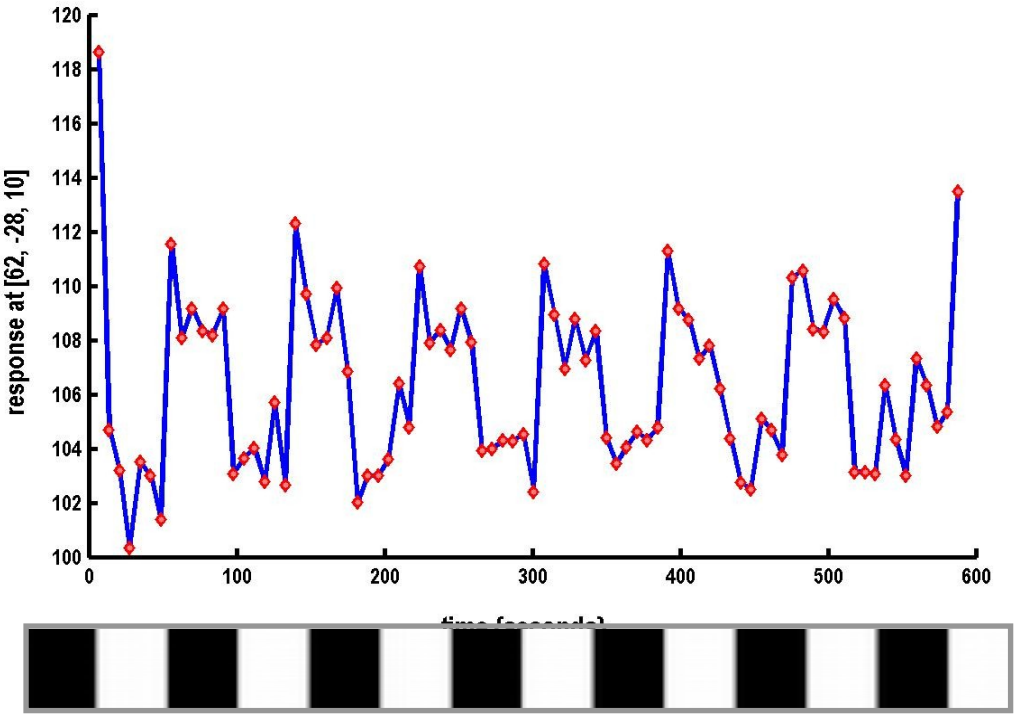
Detection of task-related activations

(this represents the vast majority of the studies)

Example where participants passively listened to words or silence

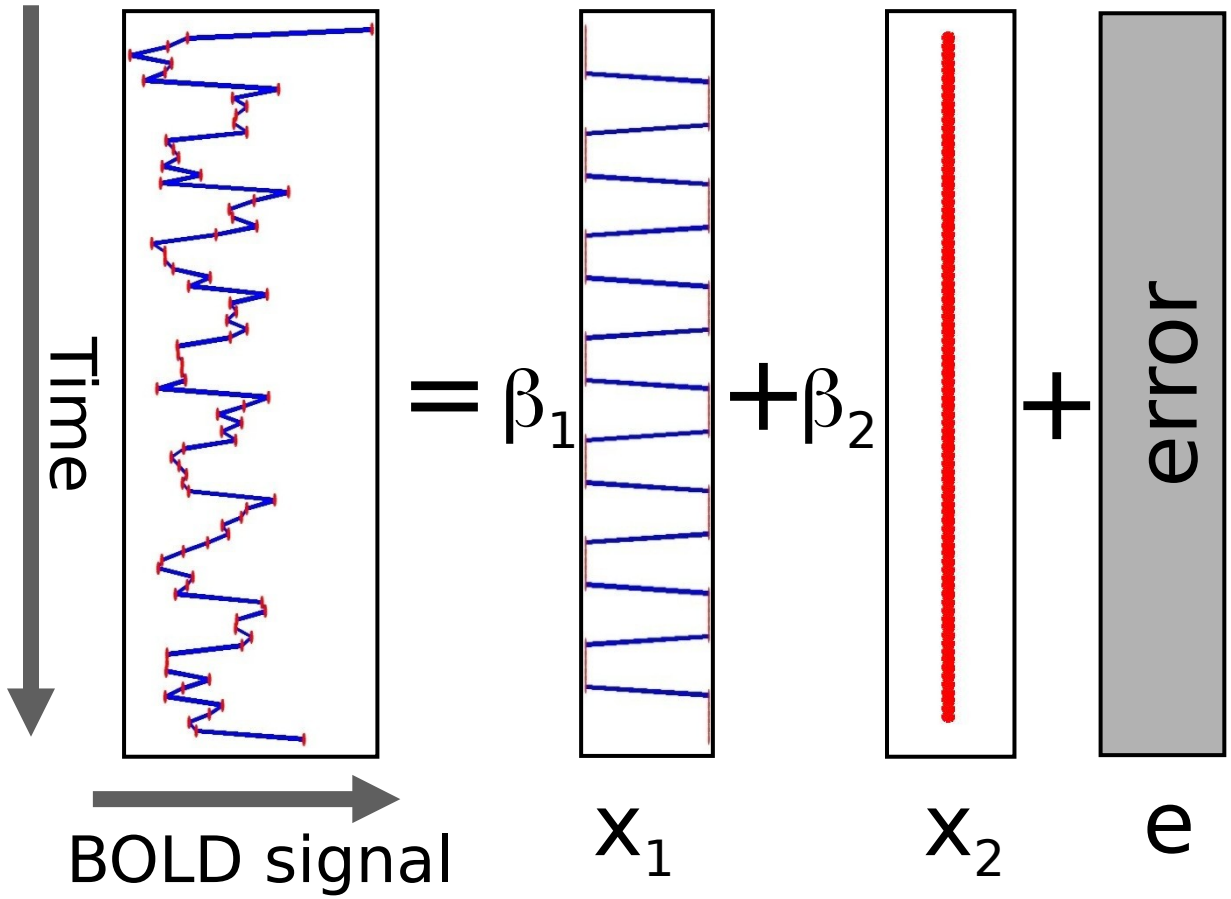
7 cycles of rest and listening

Blocks of 6 scans with 7 sec TR



Question: Is there a change in the BOLD response between listening and rest?

Single voxel regression model



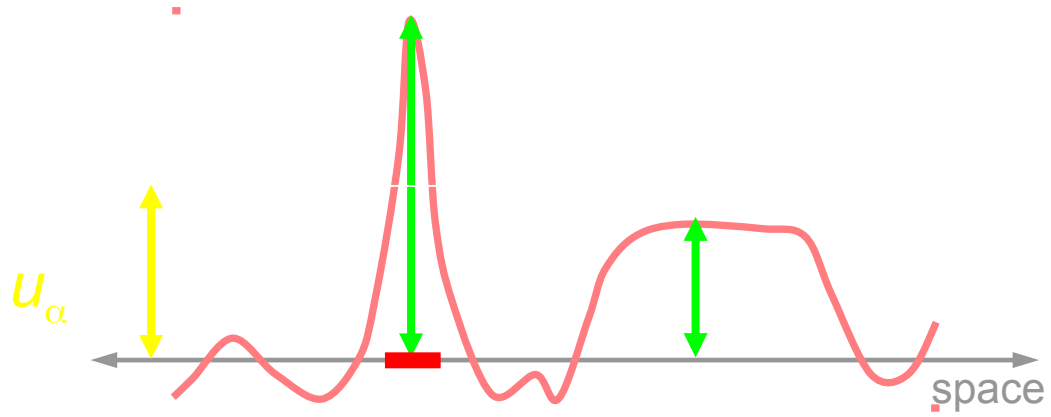
$$y = x_1\beta_1 + x_2\beta_2 + e$$

Voxel-level Inference

Retain voxels above α-level threshold u_α

Gives best spatial specificity

- The null hyp. at a single voxel can be rejected



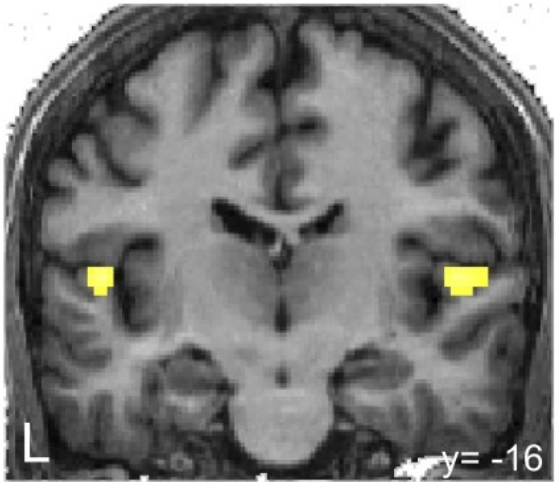
Significant Voxels

No significant Voxels

Statistical Parameter Maps

At each voxel, a T-value is computed and used to test if the effect of interest (β_1 coefficient) is significantly larger than 0:

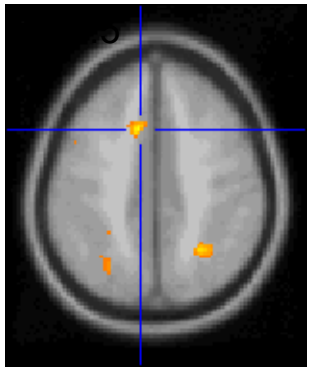
A brain map of p-values is obtained, which is thresholded at a given alpha level (e.g. $p < 0.0001$) and displayed overlaying an anatomical scan.



Assessing Statistic Images

Where's the signal?

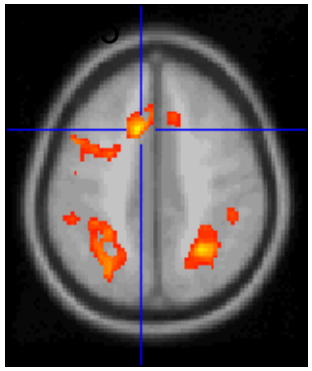
High Threshold



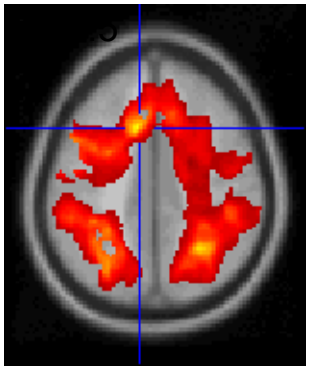
Good Specificity

Poor Power
(risk of false
negatives)

Med. Threshold



Low Threshold



Poor Specificity
(risk of false
positives)

Good Power

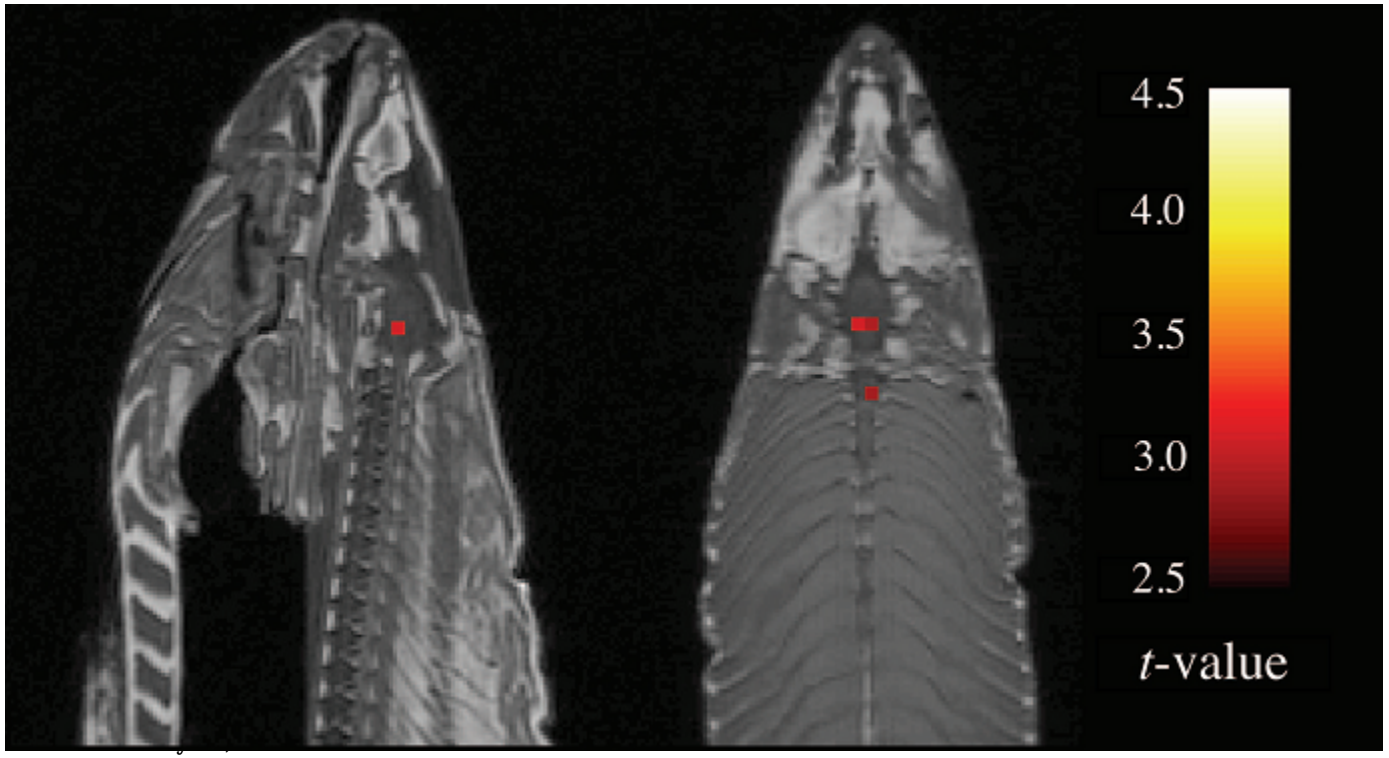
...but why threshold?!

The problem of multiple comparisons

Example: probability that a “5%” event is observed at least once in 'n' trials

Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon (Bennett et al. 2009)

In this study, the salmon was shown images of people in social situations, either socially inclusive situations or socially exclusive situations.



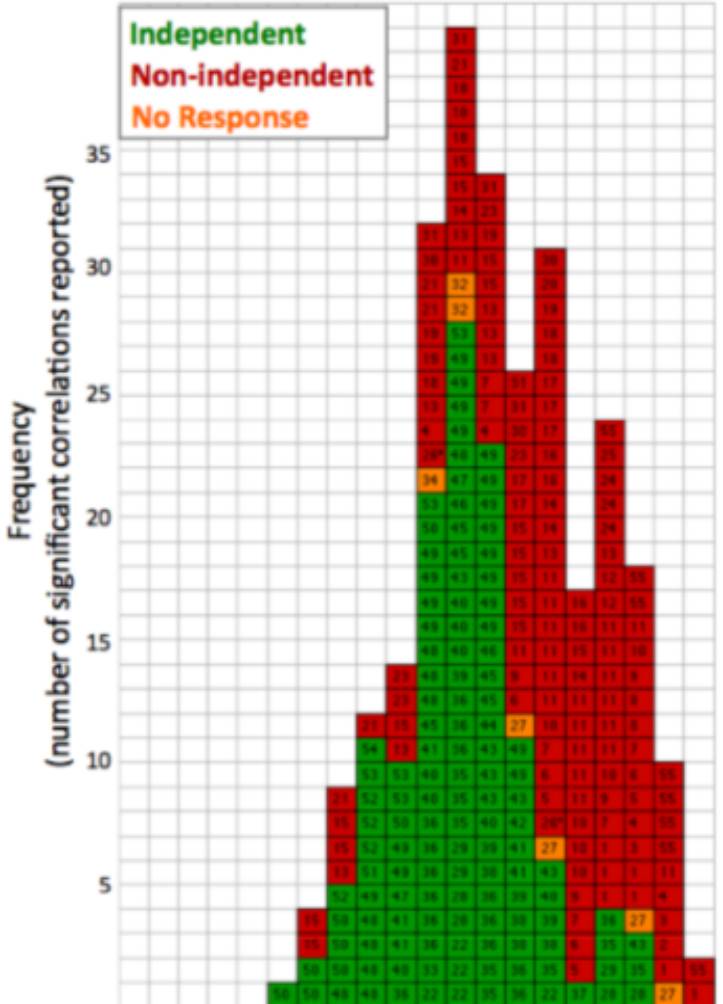
“Voodoo correlations”

(Vul et al (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Psych. Sci.)*

In the Social neuroscience, many studies have reported very high correlations (r>.8) between self-reported measure (of presonality...) and BOLD signal.

These values are implausible given that reliability of the measures.

It is due to the non-independent selection of voxels where the correlations were computed.



Addressing the Multiple Comparison Issue

Family-Wise-Error threshold

Given the variance in the image and its spatial correlation, this procedure estimates a threshold that enforces a error rate at the whole brain level. PFWE<.05 means ther 0.05 % probability that one voxel at least is detected under the null hypothesis.

False Discovery Rate threshold

Controls the relative number of false alarm voxels in an image (I.e PFDR < 0.05 means of the voxels in the mans are Falese alarms)

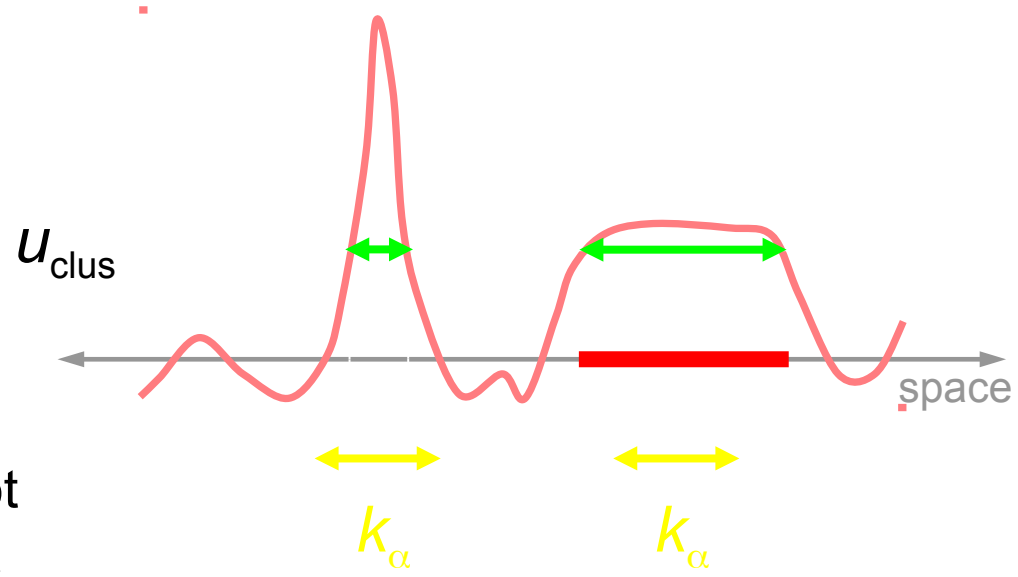
Cluster-based thresholding

A cluster is a set of contiguous voxels. Given a voxel threshold and the smoothness of th data, one can compute the distribution of Cluster sizes under the null hypothesis of activation.

A priori Regions of interest/ Parcellation

Cluster-level Inference

- Two step-process
  - Define clusters by arbitrary threshold u_{clus}
  - Retain clusters larger than α-level threshold k_α



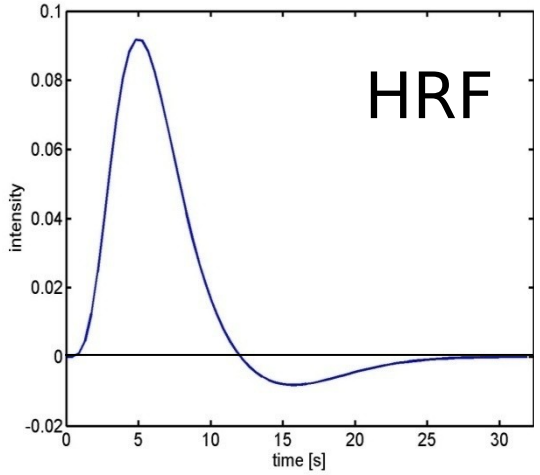
Cluster not significant

Cluster significant

IMPROVING THE MODEL

What are the problems of this model?

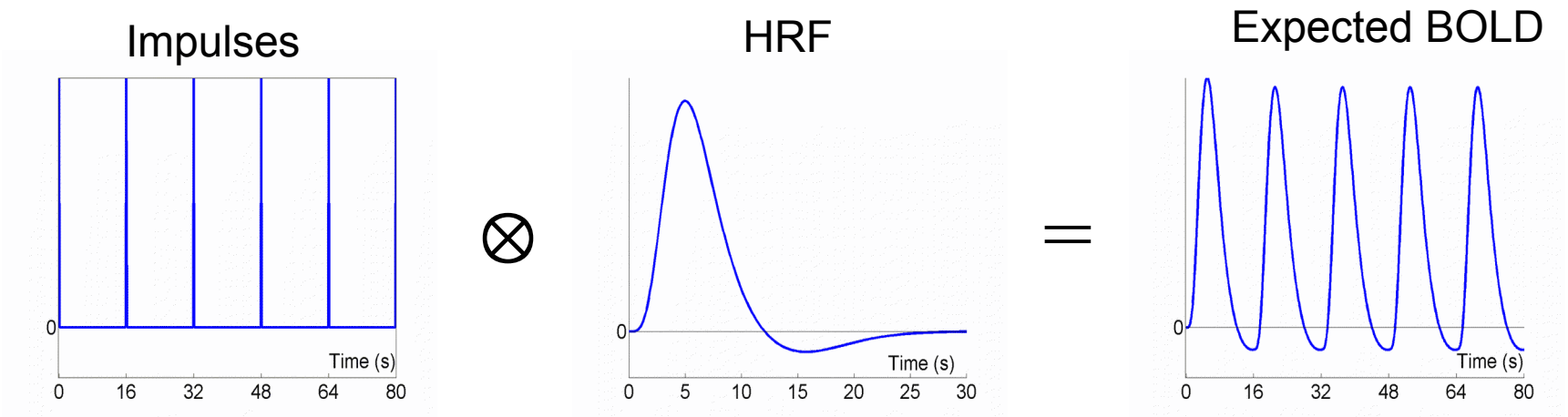
1. BOLD responses have a delayed and dispersed form.



2. The BOLD signal includes substantial amounts of low-frequency noise (eg due to scanner drift).

3. Due to breathing, heartbeat & unmodeled neuronal activity, the errors are serially correlated. This violates the assumptions of the noise model in the GLM

Problem 1: Shape of BOLD response
Solution: Convolution model



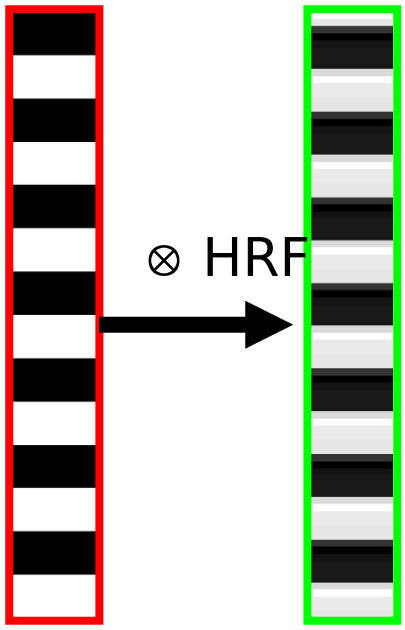
$$f \otimes g(t) = \int_0^t f(\tau)\, g(t-\tau)\, d\tau$$

expected BOLD response

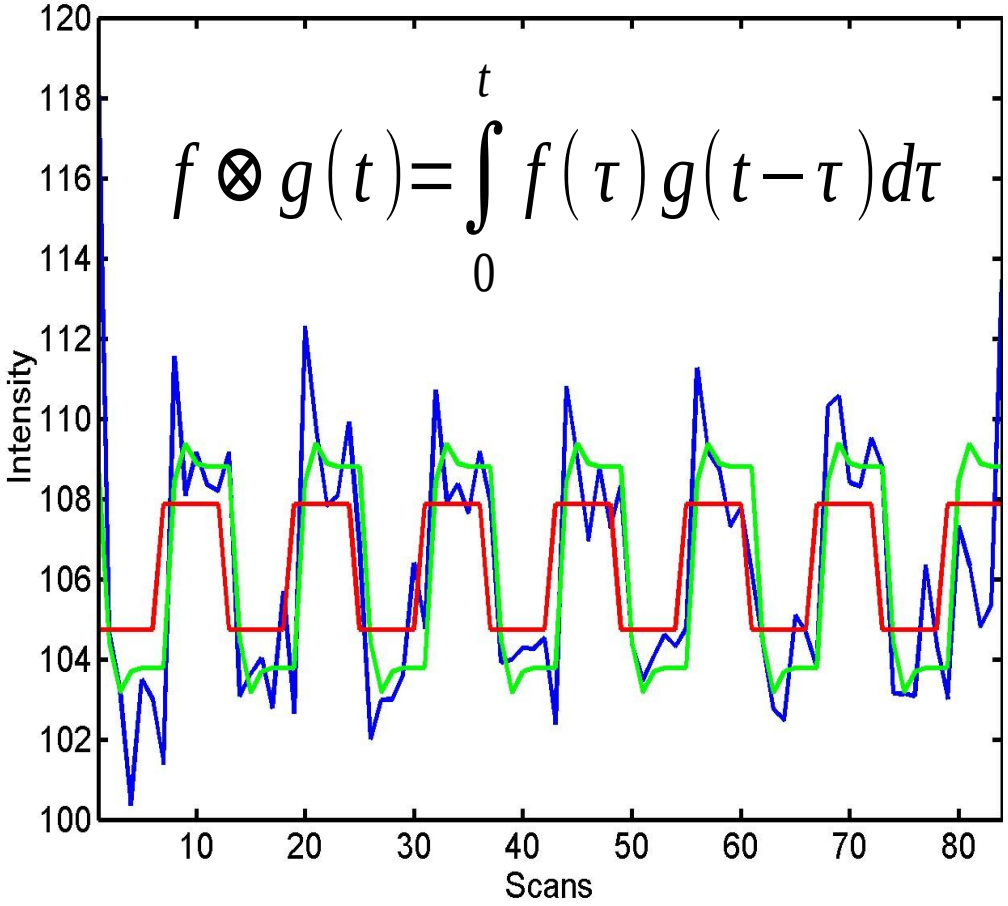
= input function ⊗ impulse response function (HRF)

Convolution model of the BOLD response

Convolve stimulus function with a canonical hemodynamic response function (HRF):



$$f \otimes g(t) = \int_0^t f(\tau)\, g(t-\tau)\, d\tau$$



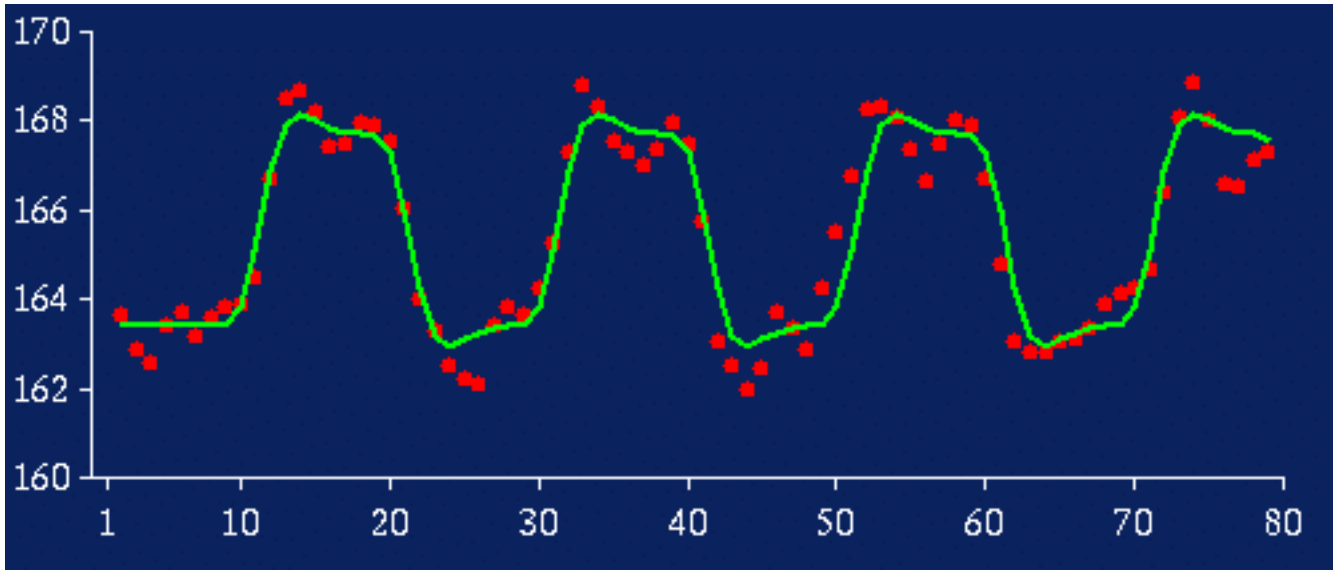
At this stage, you have obtained, for a single participants, maps of parameters that assess *the effect size for each condition* in your experiment.

You can then perform statistical tests to detect which conditions produce **statistically significant effects**, or compare them, ***for this particular participant****.*

Yet “science” requires reliability between subjects and you must therefore conduct a **group analysis** (or between participants) to make inference at the human population level.

Subject 1

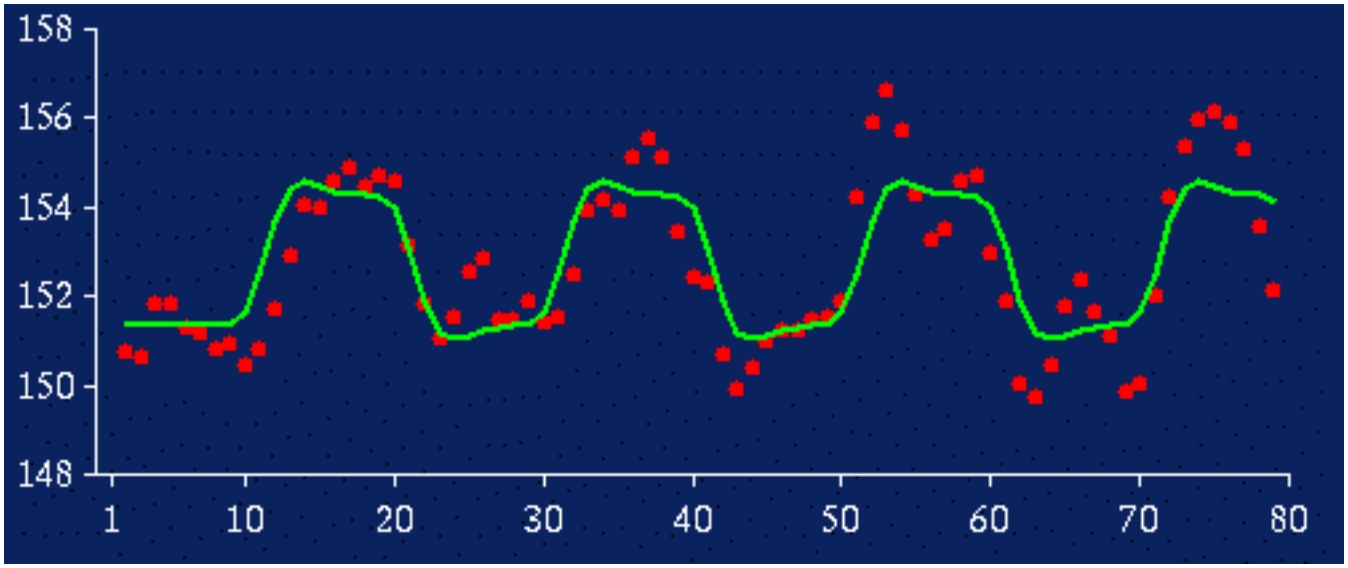
For voxel v in the brain



Effect size, c ~ 4

Subject 3

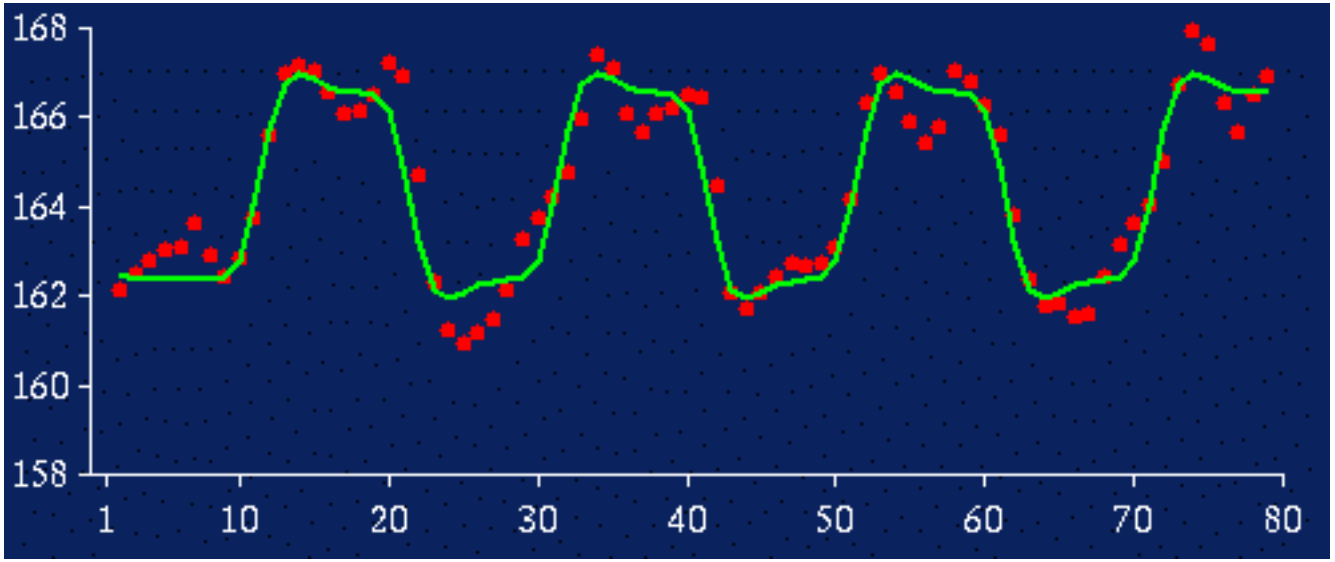
For voxel v in the brain



Effect size, c ~ 2

Subject 12

For voxel v in the brain

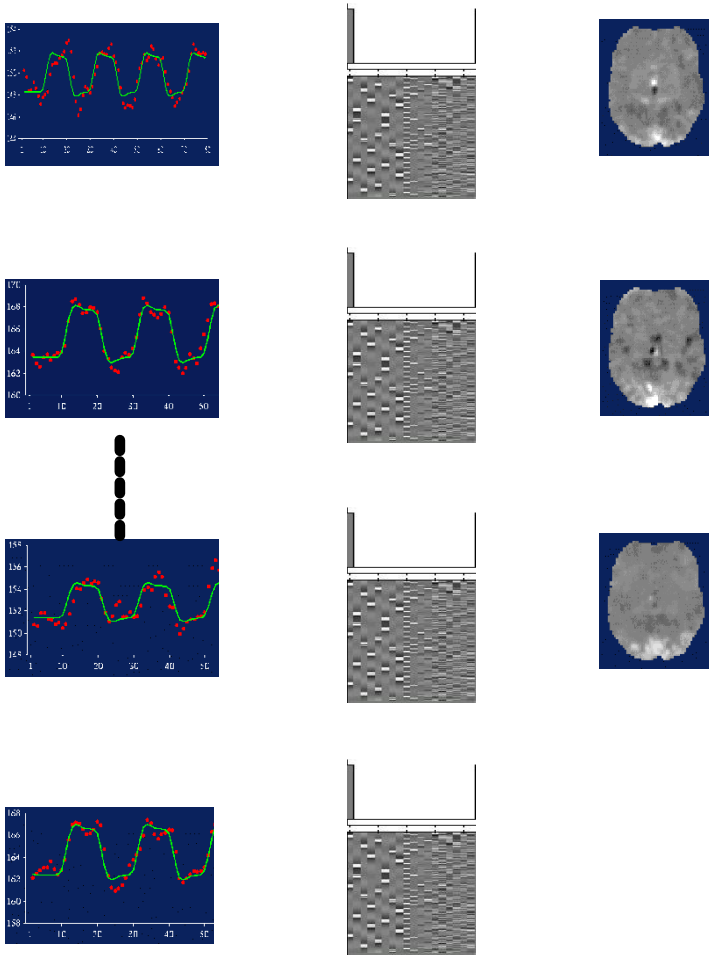


Effect size, c ~ 4

RFX: Summary Statistic

First level

Data | Design Matrix | Contrast Images



RFX: Summary Statistic

First level

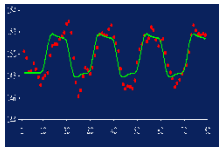
Second level

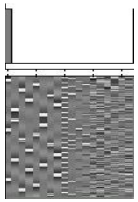
$$t = \frac{c^T \hat{\alpha}}{\sqrt{V\hat{a}r(c^T \hat{\alpha})}}$$

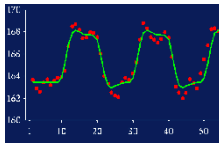
Data

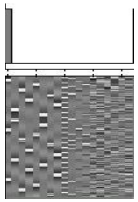
Design Matrix

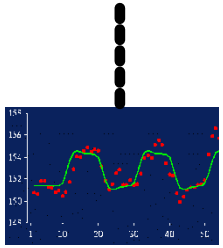
Contrast Images

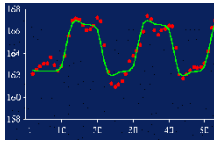


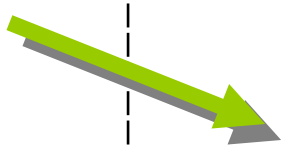


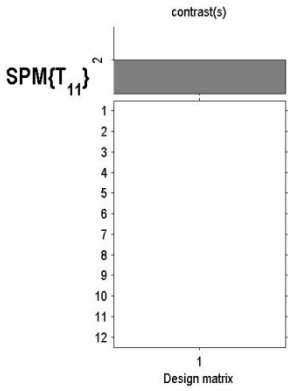


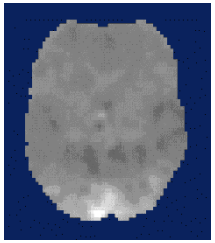




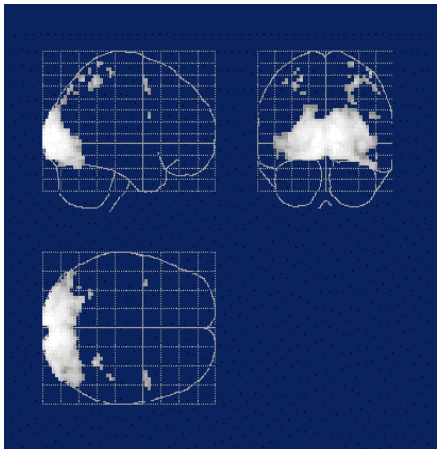








SPM(t)



One-sample t-test @ 2nd level

Summary

1. We specify a general linear model of the data

2. The model is combined with the HRF, high-pass filtered and serial correlations corrected

3. The model is applied to every voxel, producing beta images.

4. Next we'll compare betas to make inferences

http://www.fil.ion.ucl.ac.uk/~pzeidman/teaching/GLM.ppt

Multivariate analyses

The massive **univariate** approach to analyse fMRI data described earlier compare activity independently at each voxel.

Multivariate analyses rely on activity patterns from several voxels (the whole brain or, more typically, Regions of Interest).

.

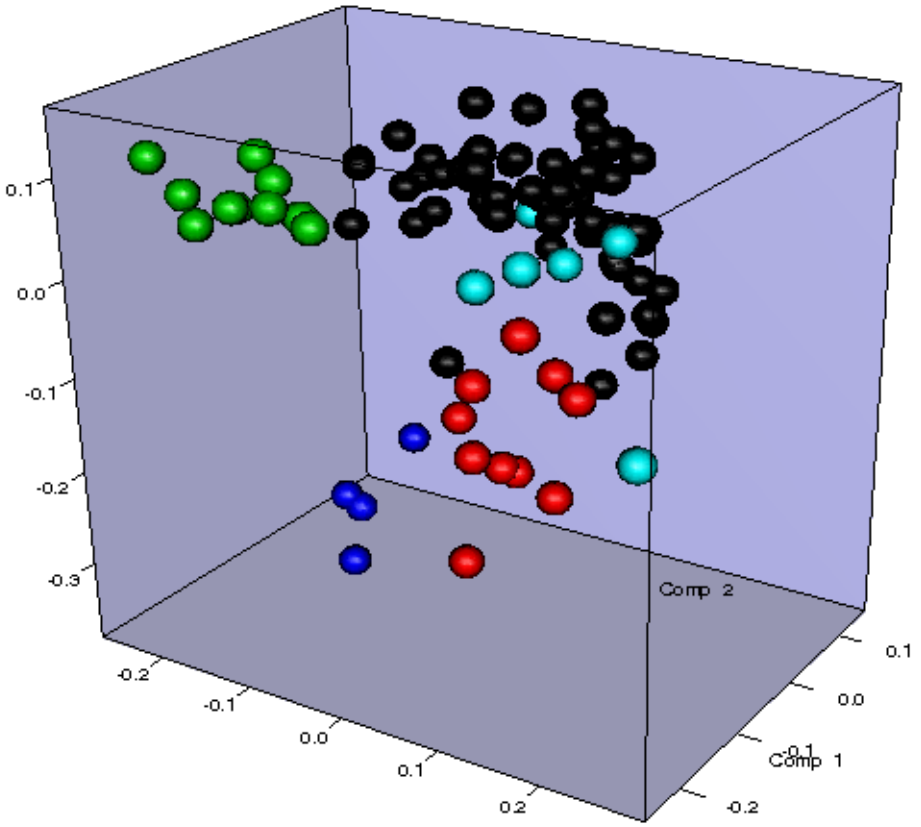
Classification algorithms:

Linear Discriminant Analysis (LDA),

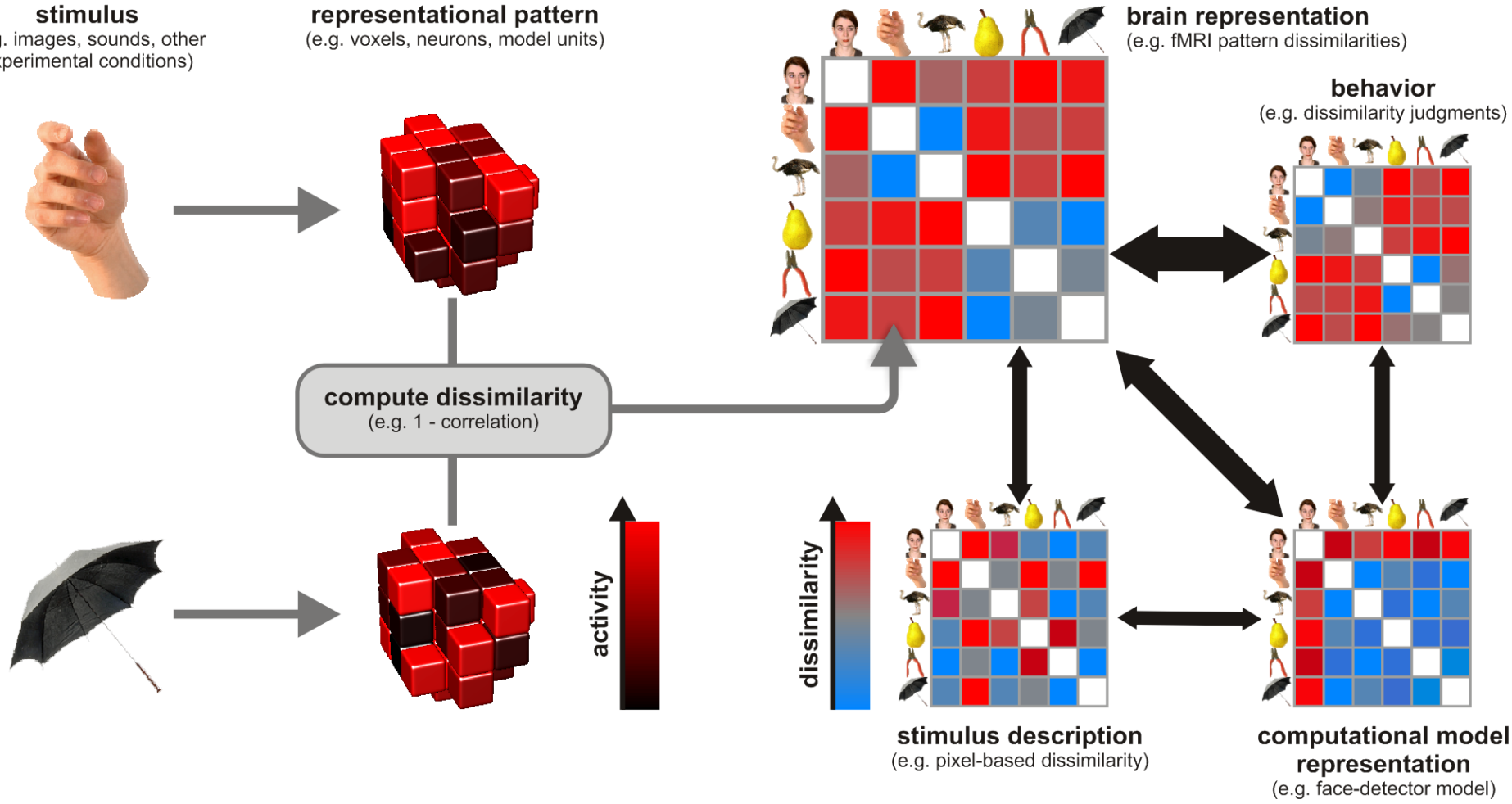
neural networks

support vector machine (SVM)

Penalized Logistic Regression



Representational Similarity Analysis



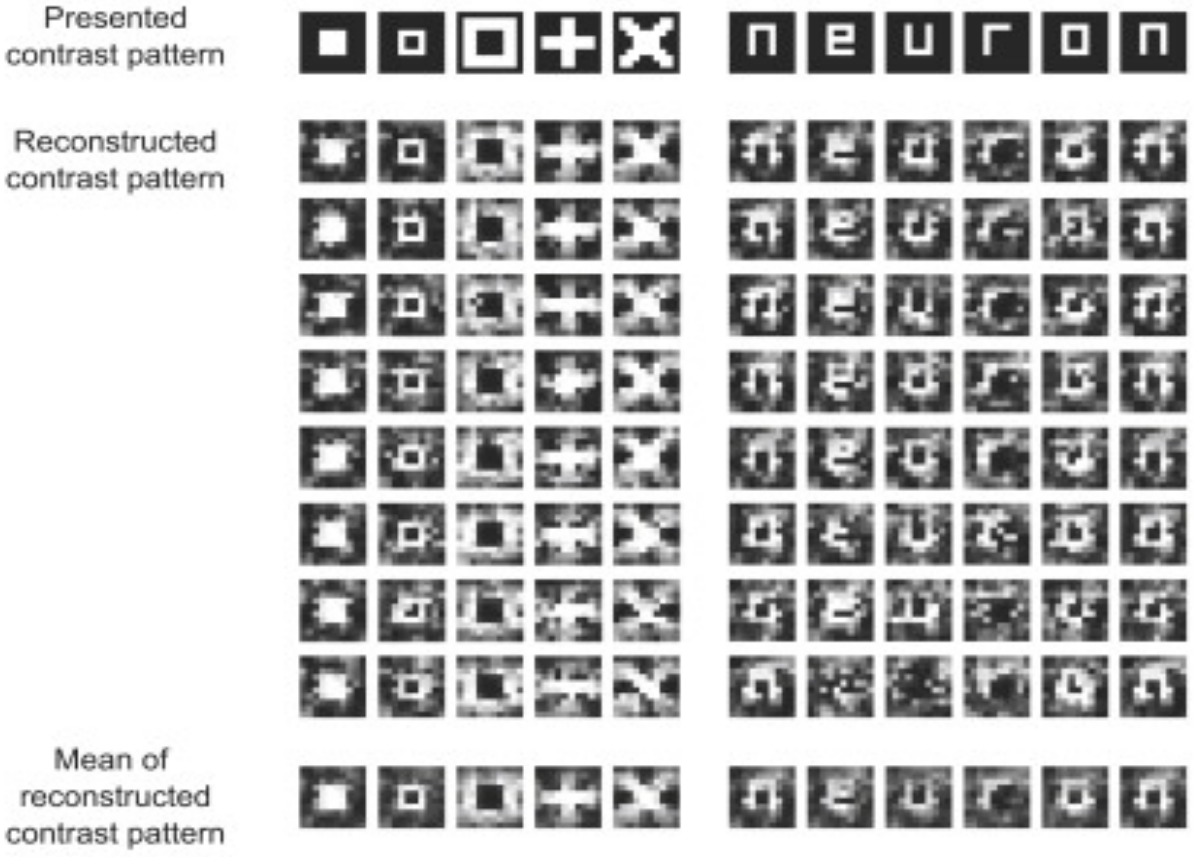
See Kriegeskorte and Kievit. (2013) “Representational Geometry: Integrating Cognition, Computation, and the Brain.” *Trends in Cognitive Sciences*

Visual stimulus reconstruction from fMRI activation in the visual cortex

(Miyawaki et al. 2008. *Neuron*)

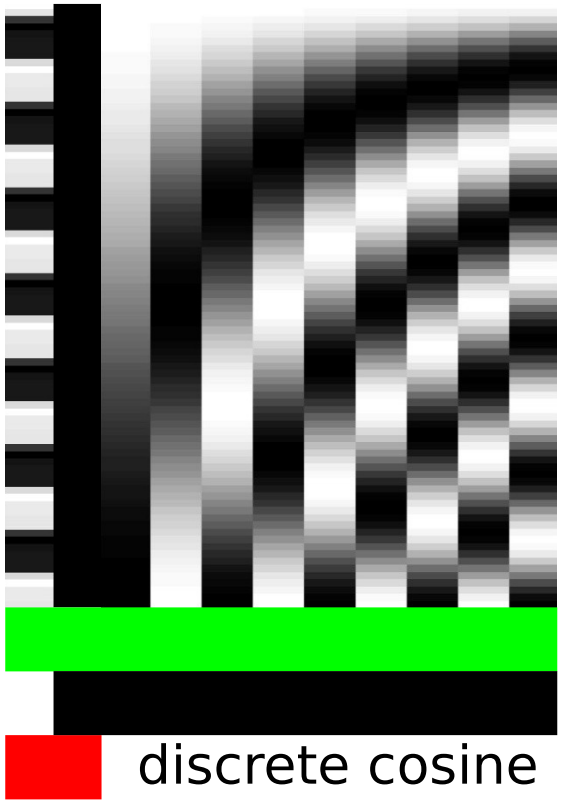
Activity was measured in early visual cortex (V1-V4) for numerous 10×10 binary patterns of visual stimuli. The authors looked at correlations in bins of voxel activity to hundreds of visual test patterns.

Then they displayed novel visual input and used a linear combination of the local image element responses to predict the visual input from the brain activity alone.

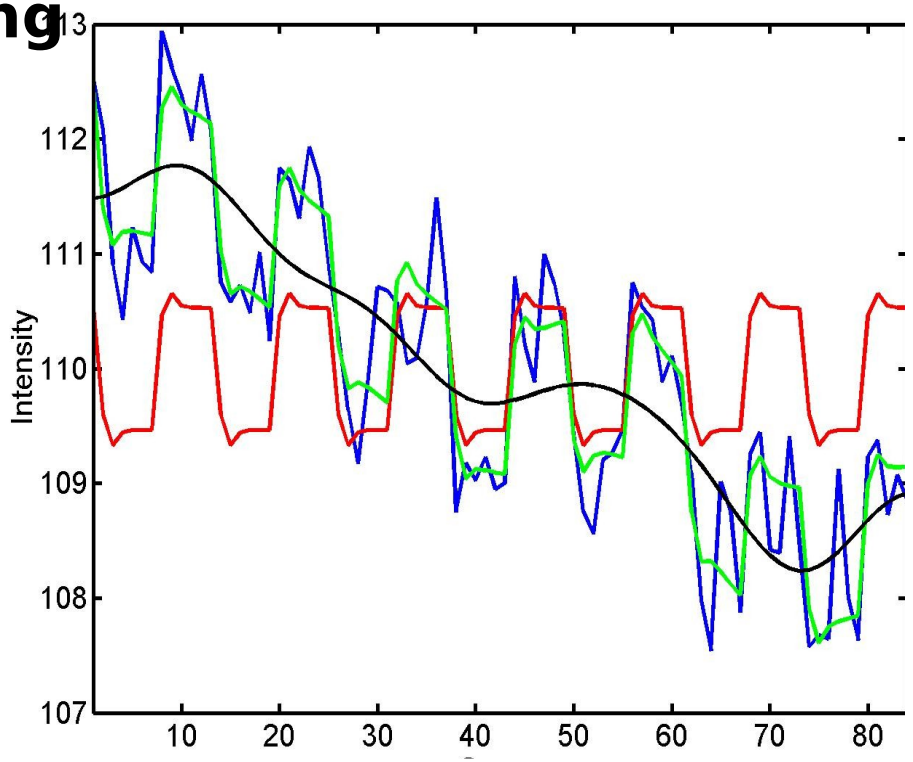


Problem 2: Low-frequency noise

Solution: High pass filtering



discrete cosine transform (DCT) set



blue = data

black = mean + low-frequency drift

green = predicted response, taking into account low-frequency drift

red = predicted response, NOT taking into account low-frequency drift

Problem 3: Serial correlations

$$e_t = a e_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim N(0, \sigma^2)$$

1st order autoregressive process: AR(1)