

# General linear model: basic

- Introducing General Linear Model (GLM): Start with an example
- Properties of the BOLD signal
  - Linear Time-Invariant (LTI) system
  - The hemodynamic response function
- (Briefly) Evaluating efficiency of a design
- (Briefly) Going back to the example

# Statistical analysis of BOLD data

Let's start with an example:  
Tom et al. (2007, Science)

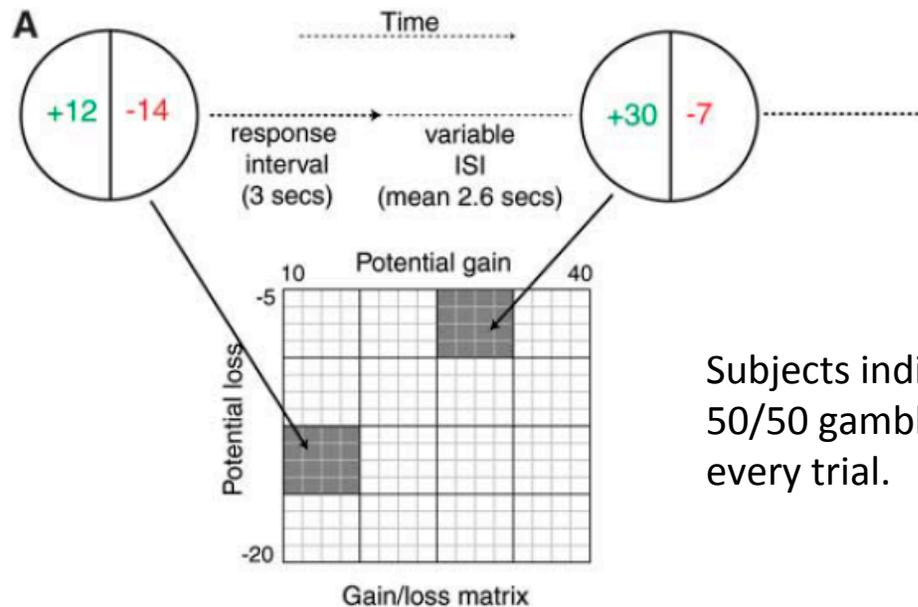
# Tom et al. (2007, Science)

- Question: Where and how are monetary gains and losses represented in the brain when people are making risky decisions?



Russ Poldrack

- Experimental design:



Subjects indicate whether s/he wanted to play a 50/50 gamble on either winning \$x or losing \$y in every trial.

## Tom et al. (2007)

- Construct a **General Linear Model** (GLM) to analyze data

$$Y(t) = \beta_0(t) + \beta_1 x_I(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

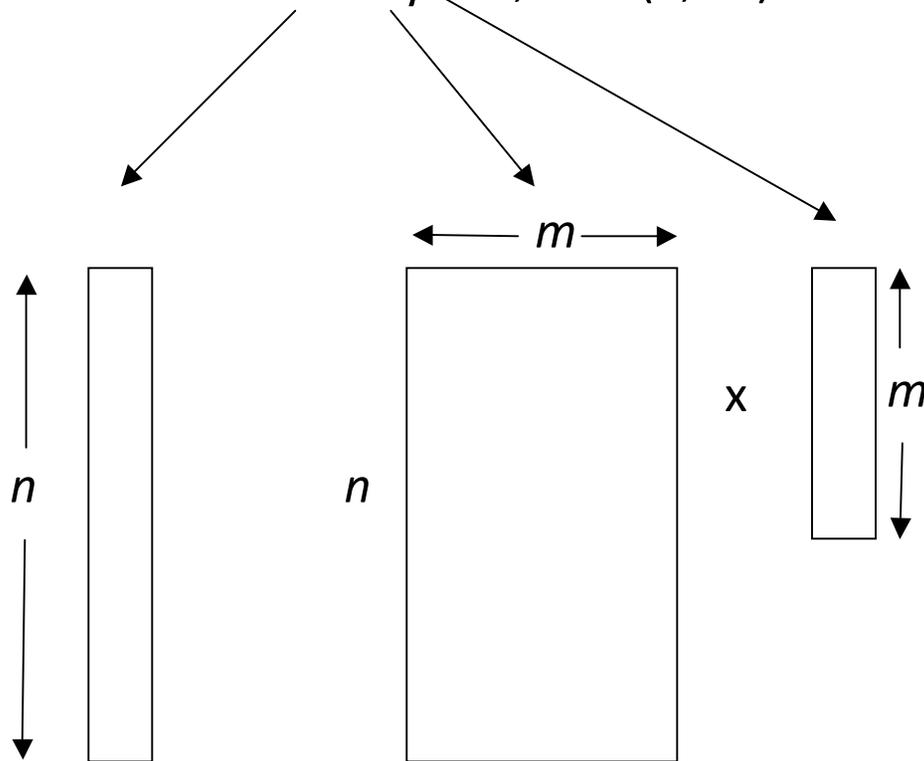
Indicator variable      The amount of gain      The amount of loss      Euclidian distance

By having this model, we can look for brain areas correlated with gains, losses, or both.

We will return to this example at the end of the lecture ...

# General linear model

$$Y = X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2)$$



BOLD times series

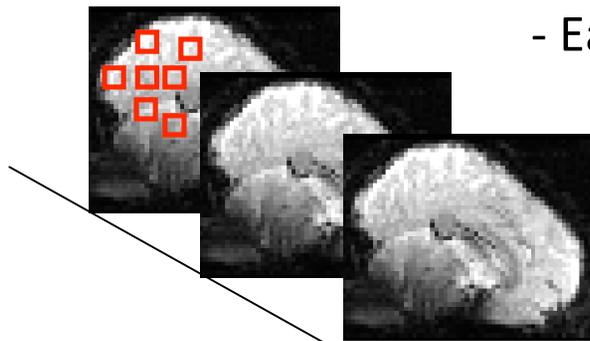
Design matrix

Parameter vector

(BOLD: Blood Oxygenation Level Dependent)

# Statistical analysis of BOLD data

Univariate approach



- Each voxel in the brain is analyzed *separately*
- Each voxel presents a **time-series** data

...

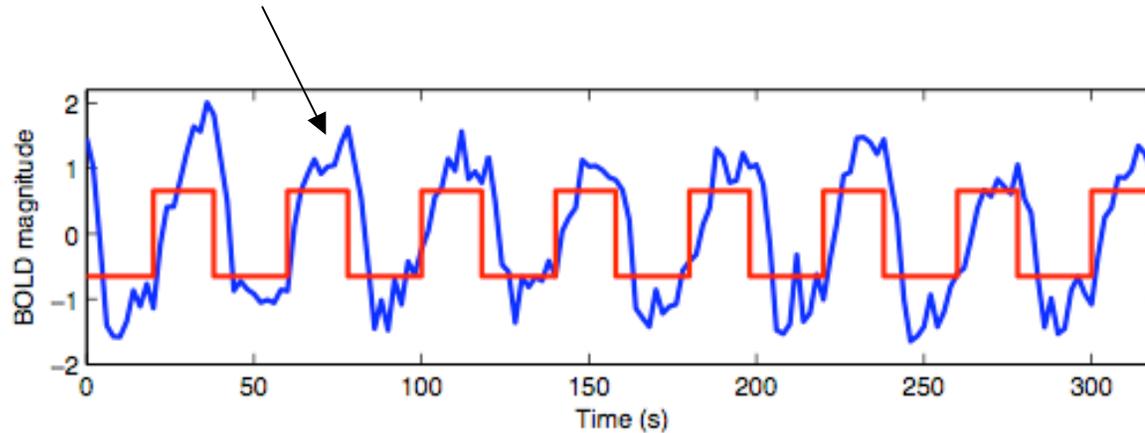
Time

## Key

- Understanding the relation between BOLD signal and neural activity
- Understanding the properties of the BOLD signal
- Understanding the characteristics of noise in the BOLD signal

# The BOLD signal

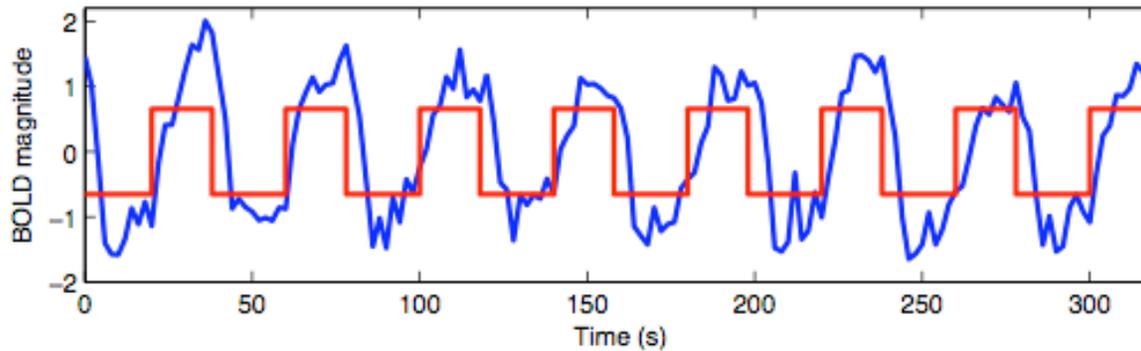
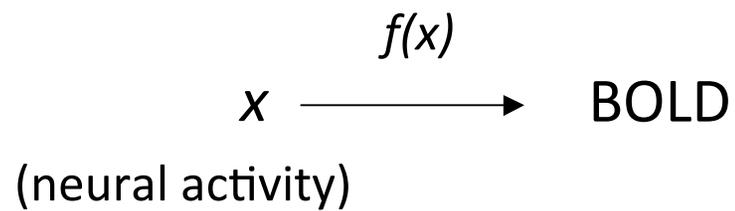
- Observed BOLD signal (example)



- BOLD signal did not look exactly like the predicted neural activity (in red)

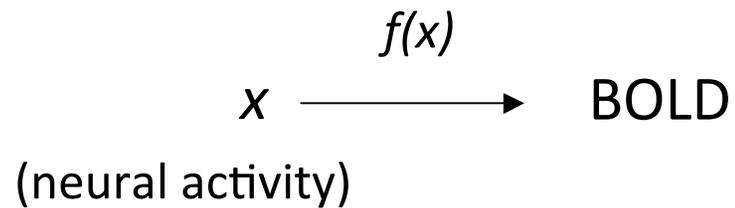
# The BOLD signal

- BOLD signal as a transformation of neural activity



# The BOLD signal

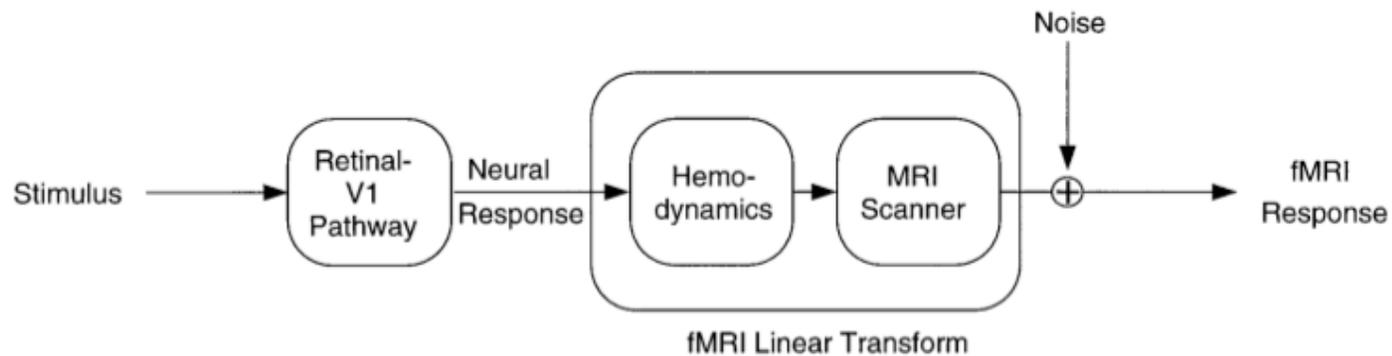
- BOLD signal as a transformation of neural activity



- Identifying the properties of the transformation function is critical
- The first thing to check is if BOLD is a **linear transform** of neural activity

## Around the mid 1990s ...

- Boynton et al. (1996, J Neurosci.) tested the linear transform model in primary visual cortex (V1)

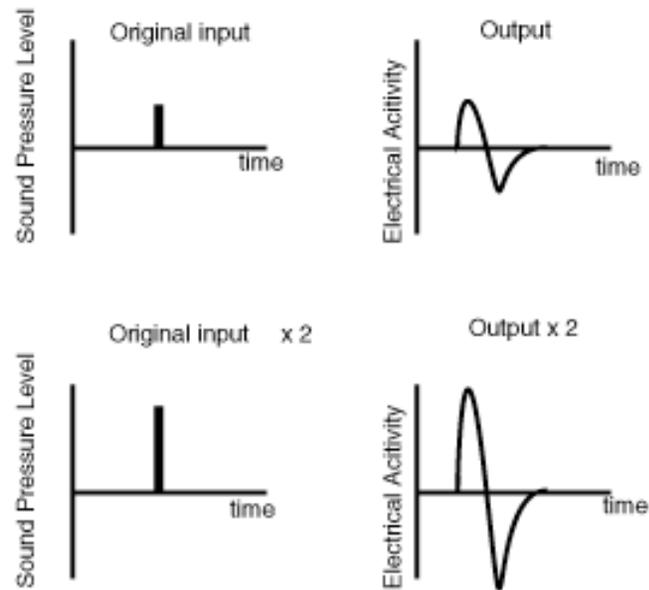


**Important:** It was only assumed that the transformation from neural response to fMRI response is linear.

# Properties of a Linear Time-Invariant (LTI) system

## - Homogeneity

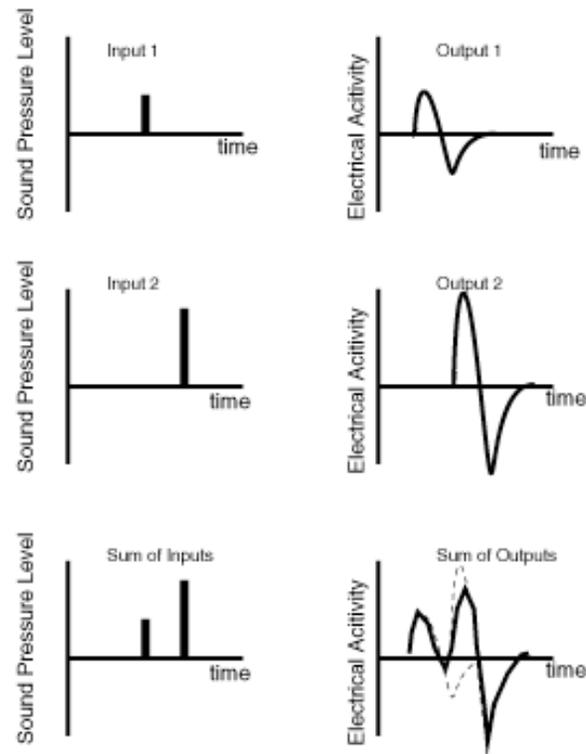
### Scalar Rule



When the input magnitude is doubled, the output response is also doubled

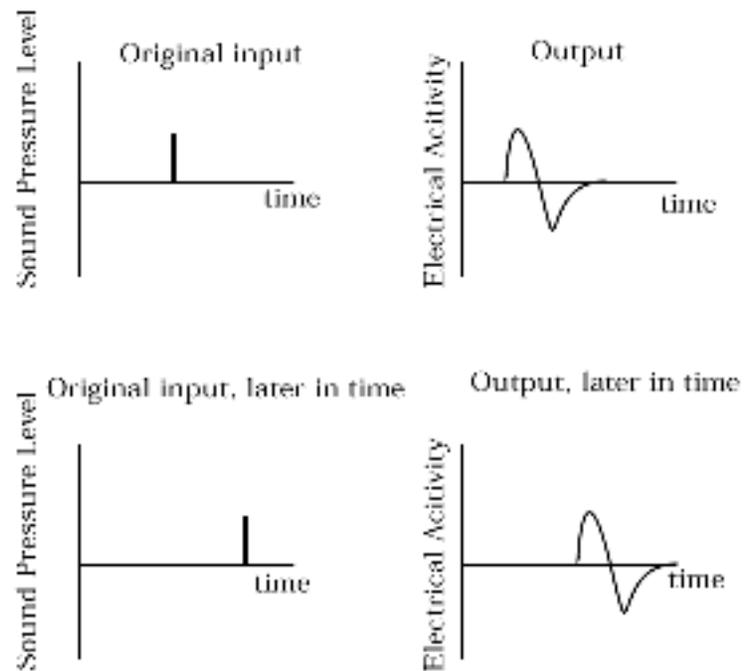
# Properties of a Linear Time-Invariant (LTI) system

## - Additivity



# Properties of a Linear Time-Invariant (LTI) system

- Shift invariance



Shift invariance with respect to time (time invariant)

Why is LTI system important?

If LTI holds,

Convolution

$$(h * f)(t) = \int_0^{\infty} h(\tau) f(t - \tau) d\tau.$$



The output of a LTI system is simply the convolution of the input and the **impulse response function,  $h(t)$**

If LTI holds,

Convolution

$$(h * f)(t) = \int_0^{\infty} h(\tau) f(t - \tau) d\tau.$$

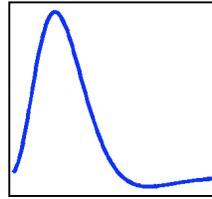


The output of a LTI system is simply the convolution of the input and the **impulse response function,  $h(t)$**

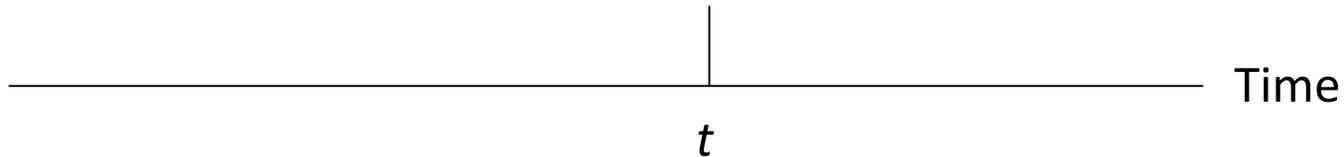
# Visualizing convolution

$$(h * f)(t) = \int_0^{\infty} h(\tau) f(t - \tau) d\tau.$$

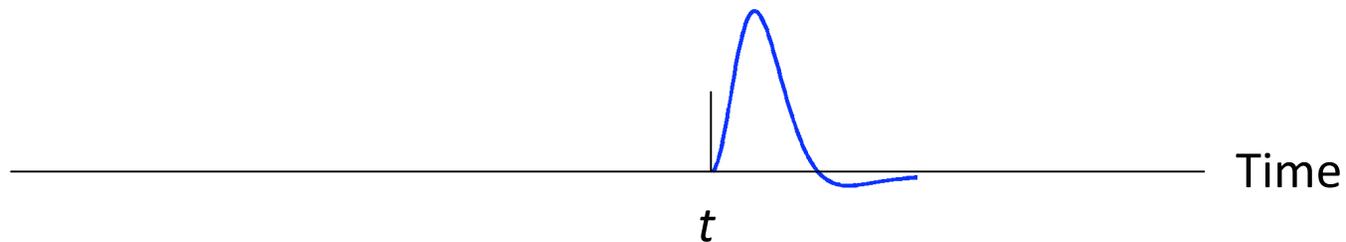
- Suppose  $h(t)$  looks like



- There is an impulse (an input) at time  $t$



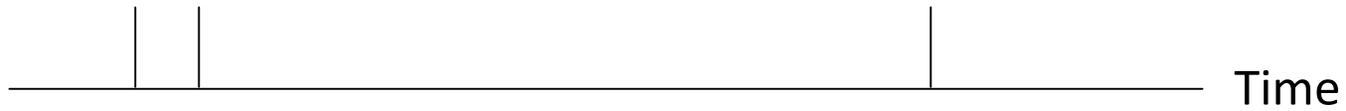
- The impulse response to this input is obtained by convolving the impulse with the impulse response function  $h(t)$



# Visualizing convolution

$$(h * f)(t) = \int_0^{\infty} h(\tau) f(t - \tau) d\tau.$$

- Input time series



- Output (response) time series (the sum of the 3 time series)

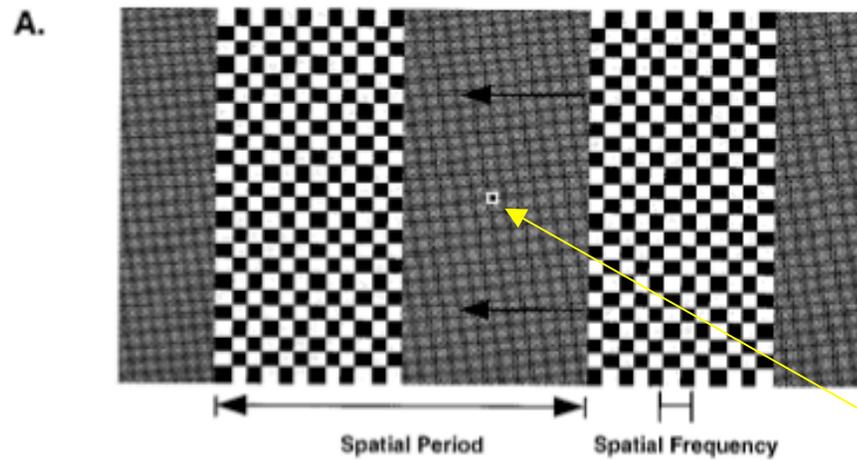


## Why is LTI system important?

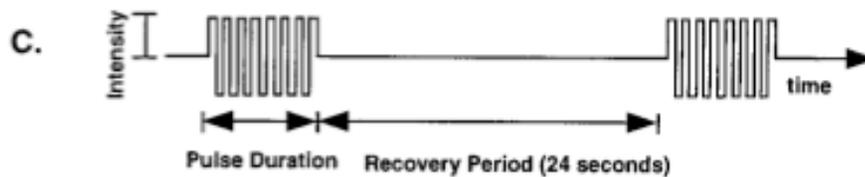
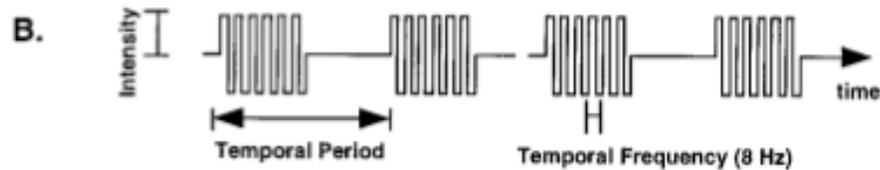
- The output of a LTI system is simply the convolution of the input and the **impulse response function,  $h(t)$**
- All we need to know is  **$h(t)$**

# Boynton et al. (1996)

- flickering checkerboard task



- 4 pulse durations (3, 6, 12, 24 s)
- 4 stimulus contrasts (0, 0.25, 0.5, 1)

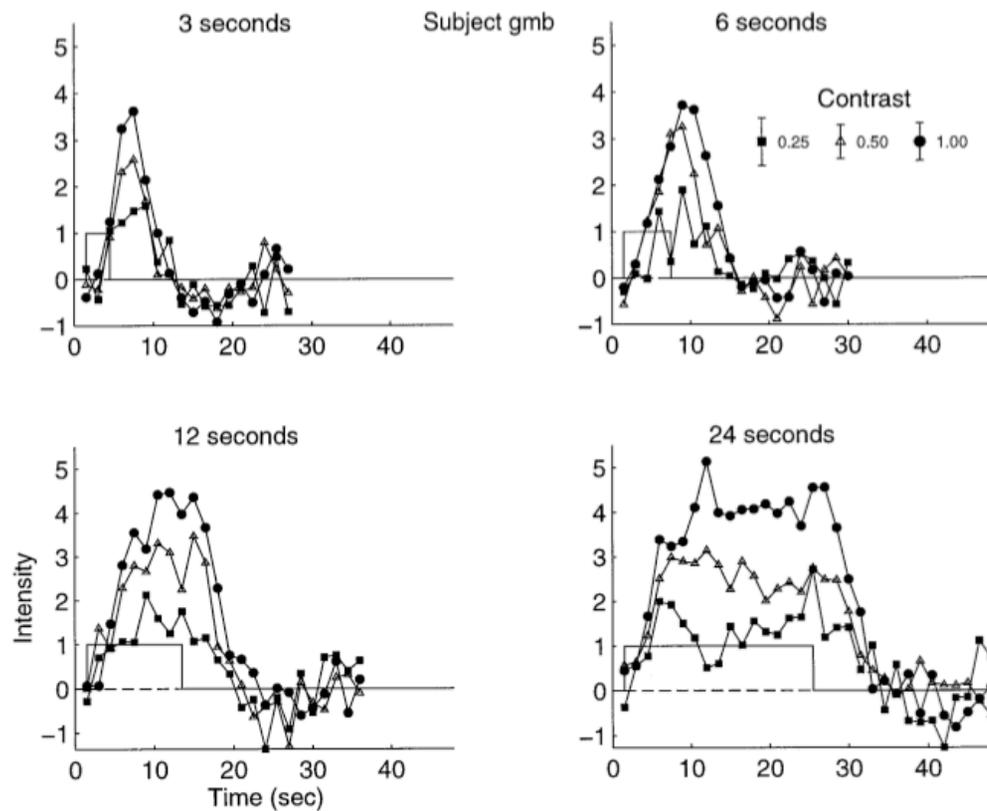


Eye fixation

# Boynton et al. (1996)

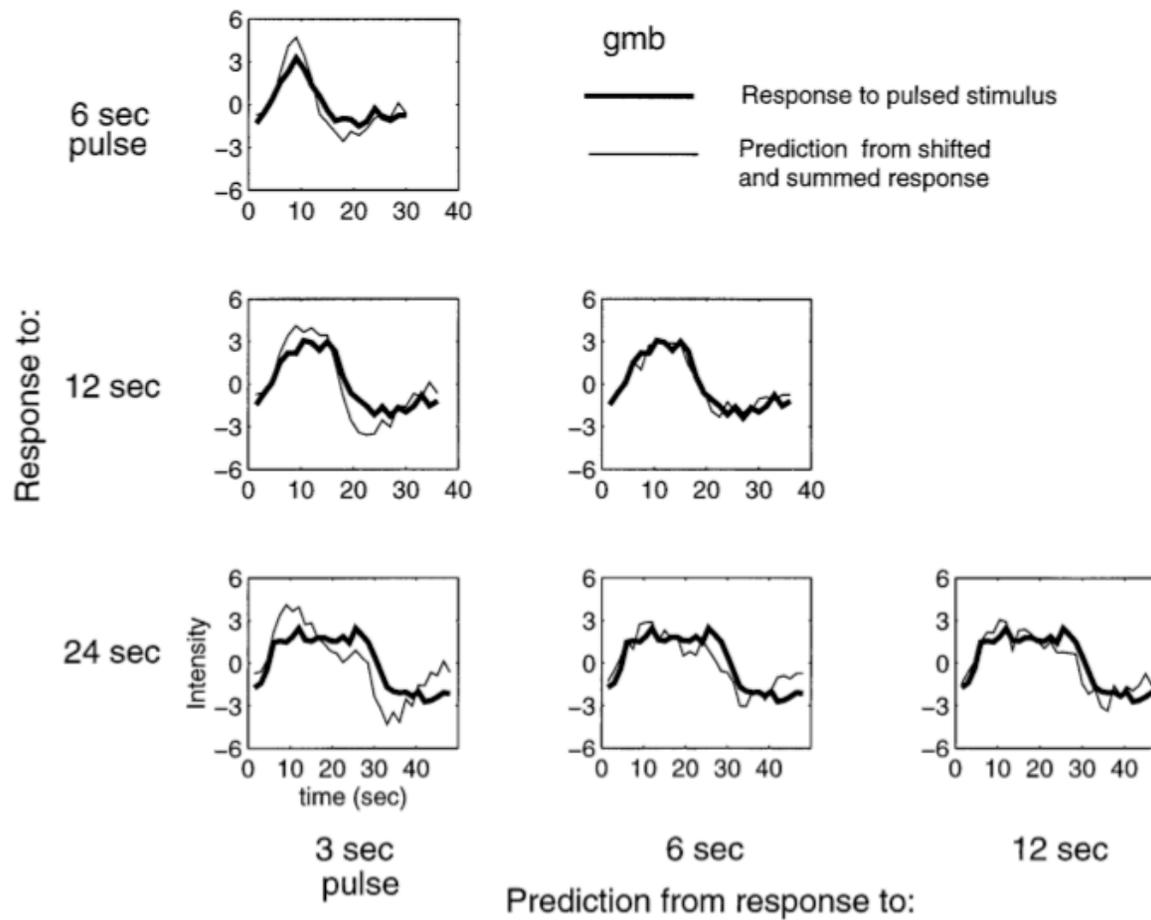
## - Results

Look how response varies as a function of contrast and duration; Could you tell if LTI is hold?



# Boynton et al. (1996)

- LTI assumption holds in most cases



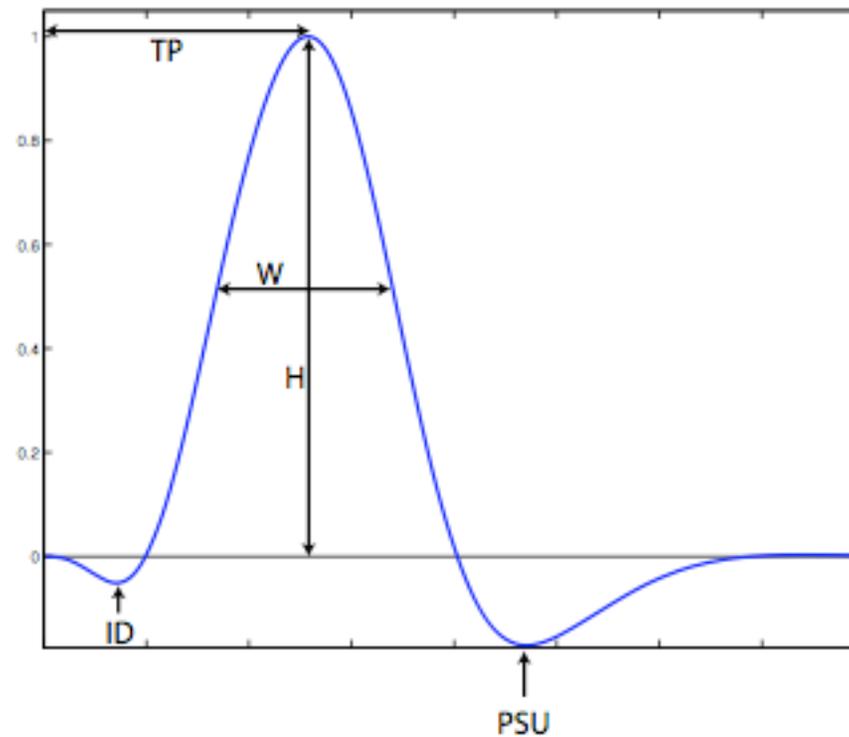
## Modeling $h(t)$

We need to have a good estimate of  $h(t)$

# Modeling $h(t)$

We need to have a good estimate of  $h(t)$

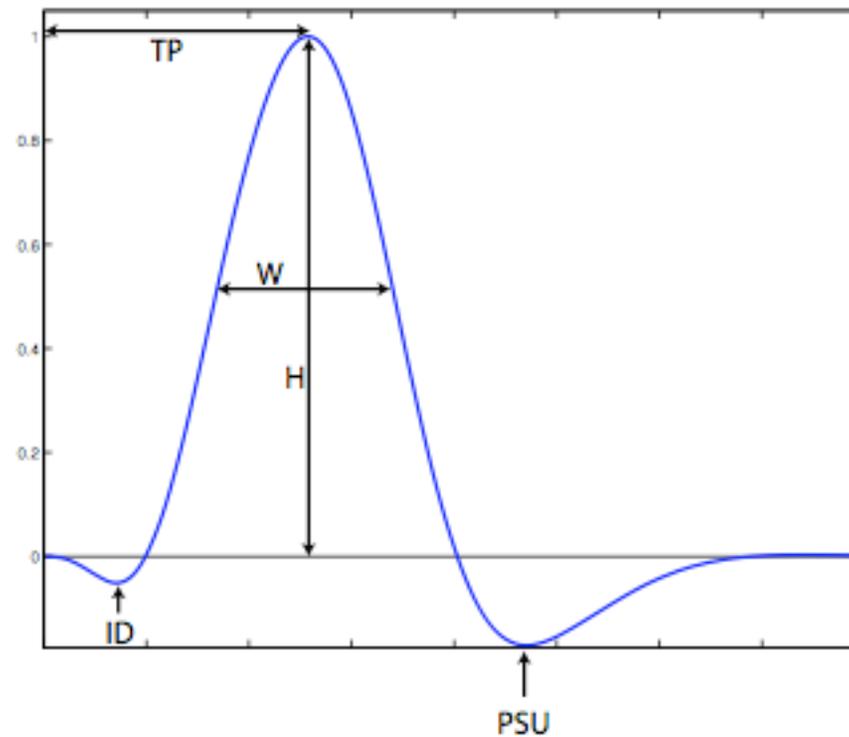
Friston et al. (1994), Lange & Zeger (1997)



# Modeling $h(t)$

We need to have a good estimate of  $h(t)$

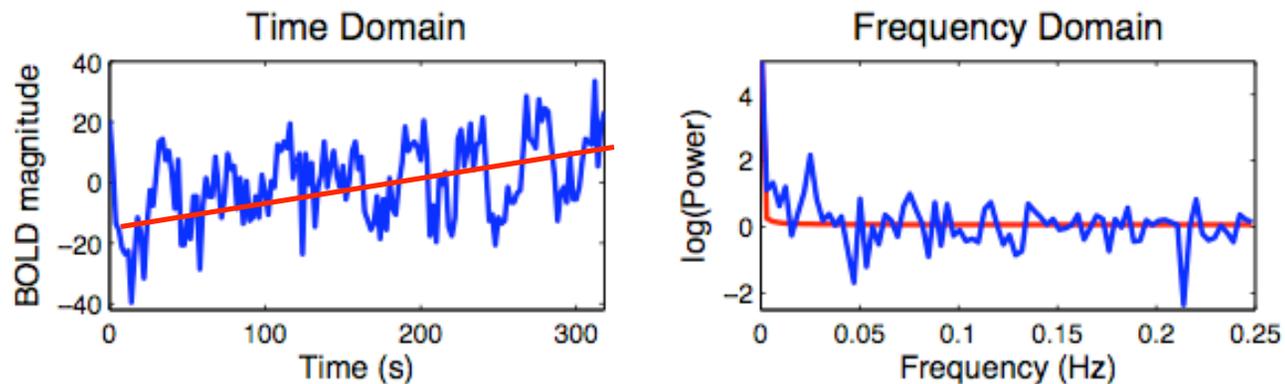
Canonical HRF (double-gamma function)



The BOLD noise

# Characterizing the noise

- low frequency drift



Source: scanner as an additional source of structured noise

**Important:** avoid designing experiment between 0-0.015 Hz  
(block length no more than 35 s for on-off block design)

# Characterizing the noise

- low frequency drift

Source: scanner as an additional source of structured noise

**Important:** avoid designing experiment between 0-0.015 Hz  
(block length no more than 35 s for on-off block design)

- How to remove

Step 1: high-pass filtering

Step 2: pre-whitening to remove **temporal autocorrelation**

## Brief summary

- BOLD signal is the output of an approximate LTI system
- Sources of noise in the BOLD signal can be approximately identified and removed during the preprocessing stage of the data

The question, then, goes back to

- Is my design good?

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

- Covariance

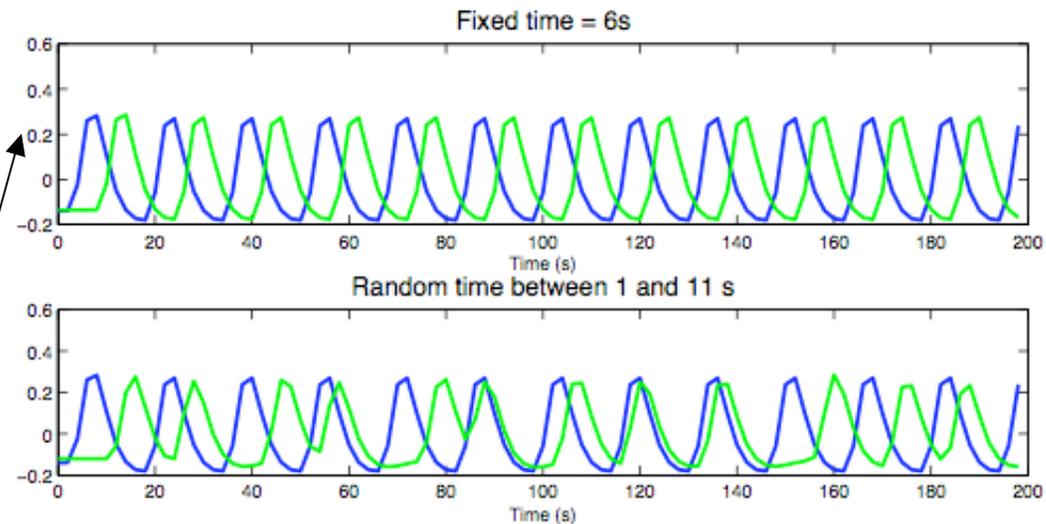
$$\text{cov}(\hat{\beta}) = (X'X)^{-1} \sigma^2$$

- **Design efficiency** usually refers to the variance due to the design,  
So for a contrast  $c$ , the efficiency for that contrast is

$$\text{eff}(c\hat{\beta}) = \frac{1}{c(X'X)^{-1}c'}$$

# Thinking about timing ...

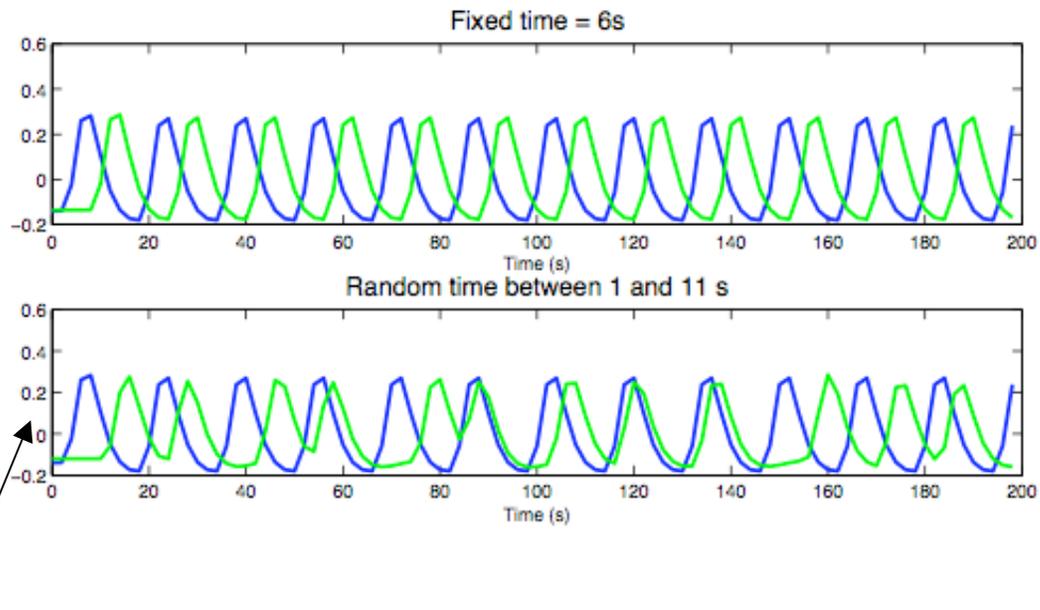
Suppose you have 2 trial types and you want to model them separately



When you have a fixed ISI (inter-stimulus interval), the correlation between your trial-type regressors is going to be high (poor efficiency).

# Thinking about timing ...

Suppose you have 2 trial types and you want to model them separately



You can *reduce* that correlation (improving efficiency) by introducing random timing jitters between trials

## Thinking about timing ...

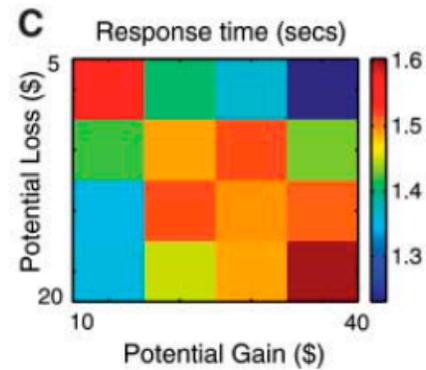
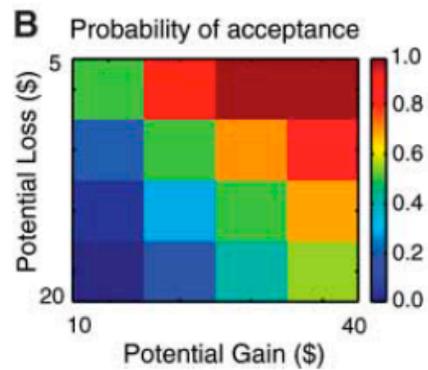
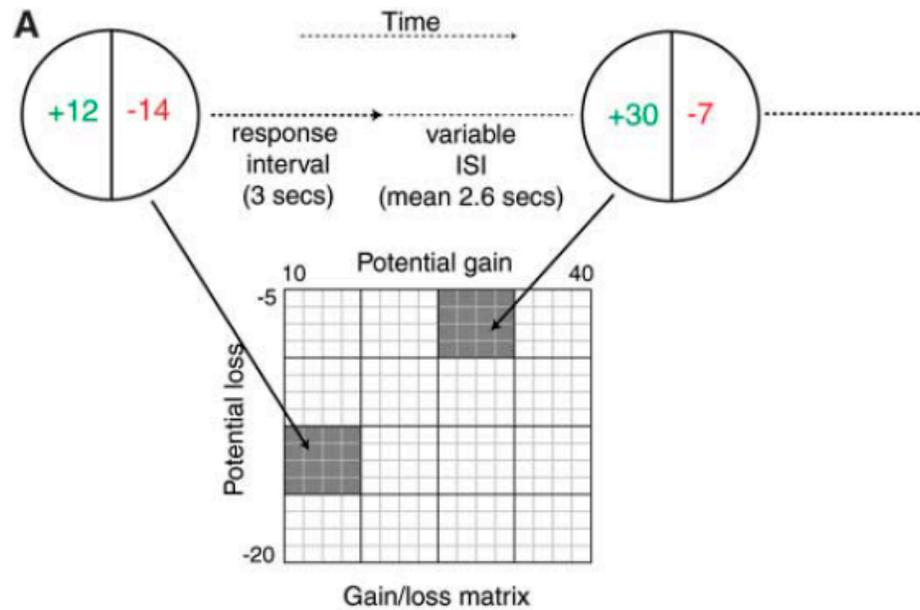
Final note: it's all about how you would like to analyze your data.

- The correlation is introduced because you wish to model the different trial types separately (a legitimate model), and the fact that one trial type always follows the other (in this example)
- It is always a good exercise to come up with a design matrix (or multiple design matrices) for analyzing your data before you even start collecting data!

Going back to the example:  
Tom et al. (2007, Science)

# Tom et al. (2007)

- Design



## Tom et al. (2007)

- Construct General Linear Model (GLM) to analyze data

$$Y(t) = \beta_0(t) + \beta_1 x_I(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

Indicator variable      The amount of gain      The amount of loss      Euclidian distance

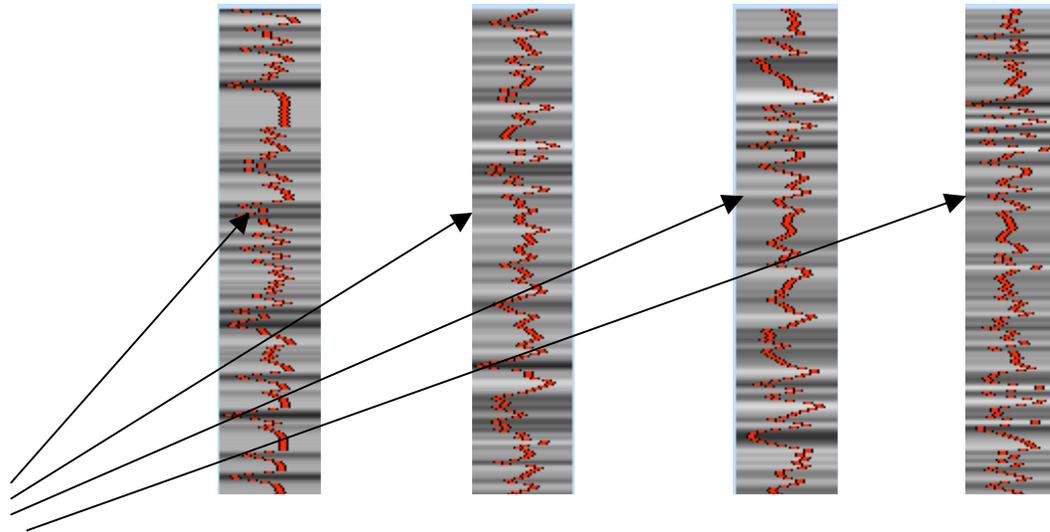
**Important:**

This is the GLM that analyzes data of a single run of a single subject

# Tom et al. (2007)

- Construct General Linear Model (GLM) to analyze data

$$Y(t) = \beta_0(t) + \beta_1 x_I(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

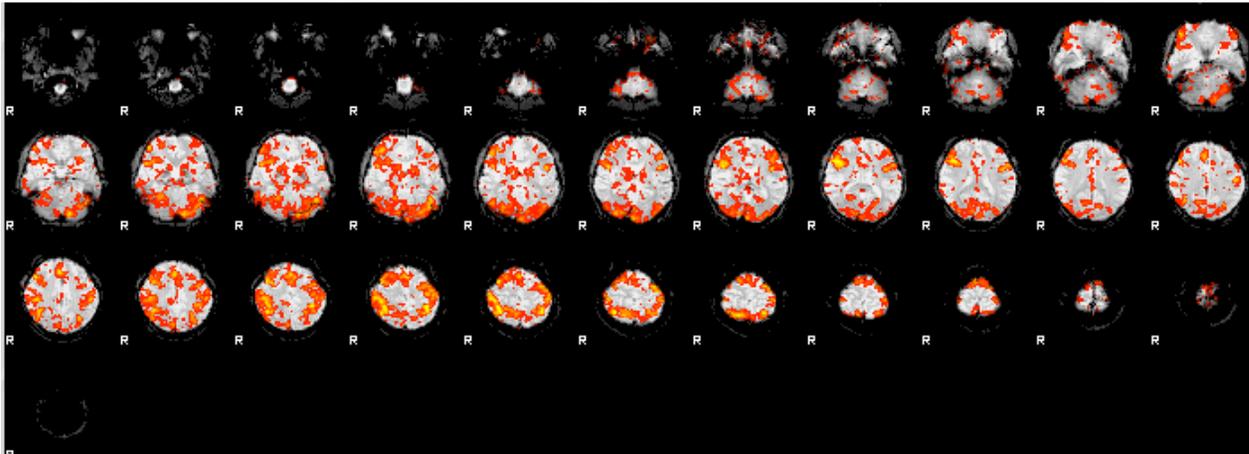


HRF-convolved time course

Example from a single subject's run

# Single-subject results (Tom et al. 2007)

Not the pretty pictures you often see in fMRI papers ...



For fMRI experiments, we typically need ~20 subjects' data to obtain meaningful results

# Multi-level GLM

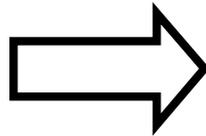
## Level 1: Single-subject GLM

$$\hat{\beta}_{\text{subject 1}}$$

$$\hat{\beta}_{\text{subject 2}}$$

⋮

$$\hat{\beta}_{\text{subject K}}$$



## Level 2: Group-level GLM

$$\hat{\beta}_{\text{subject 1}}$$

$$\hat{\beta}_{\text{subject 2}}$$

$$= X_G \beta_G + \eta$$

⋮

$$\hat{\beta}_{\text{subject K}}$$

The goal is to estimate  $\beta_G$

Multi-level GLM adopts a “summary statistics” approach  
(Beckmann et al. 2003)

# Multi-level GLM

- Multi-level GLM adopts a “summary statistics” approach (Beckmann et al. 2003); more computationally efficient
- $\beta_G$  is a weighted average of the beta from individual subjects (weight is inversely proportional to subject-specific variance)
- Variance at the group level is the sum of individual subjects' variance and the variance at the group level (between subjects)

$$\text{Var}(\hat{\beta}_G) = \frac{\sigma_W^2 + \sigma_S^2}{N}$$

$\sigma_W^2$  Within-subject variance

$\sigma_S^2$  between-subject variance

$$t = \frac{\hat{\beta}_G}{\sqrt{\text{Var}(\hat{\beta}_G)}}$$

## Group results (Tom et al., 2007)

- Network of regions positively correlated with gains (orange) and negatively correlated with losses (blue)

