Basics of fMRI Analysis: Preprocessing, First Level Analysis, and Group Analysis

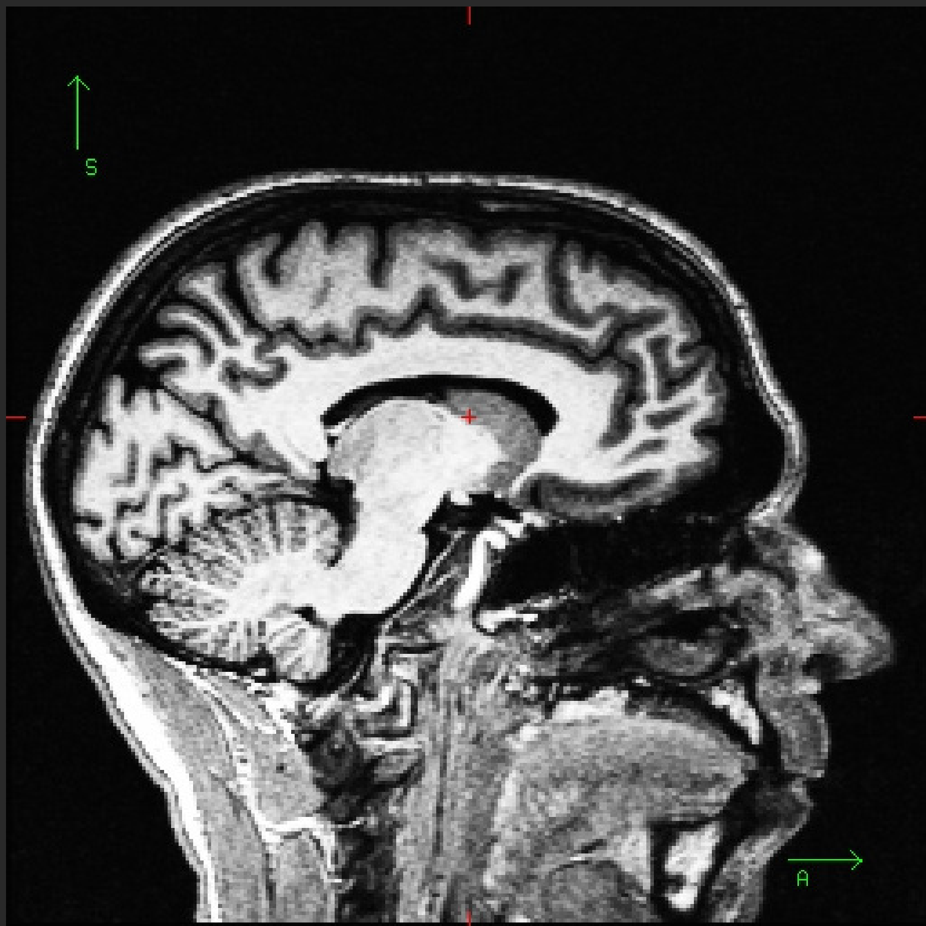
MASSACHUSETTS GENERAL HOSPITAL

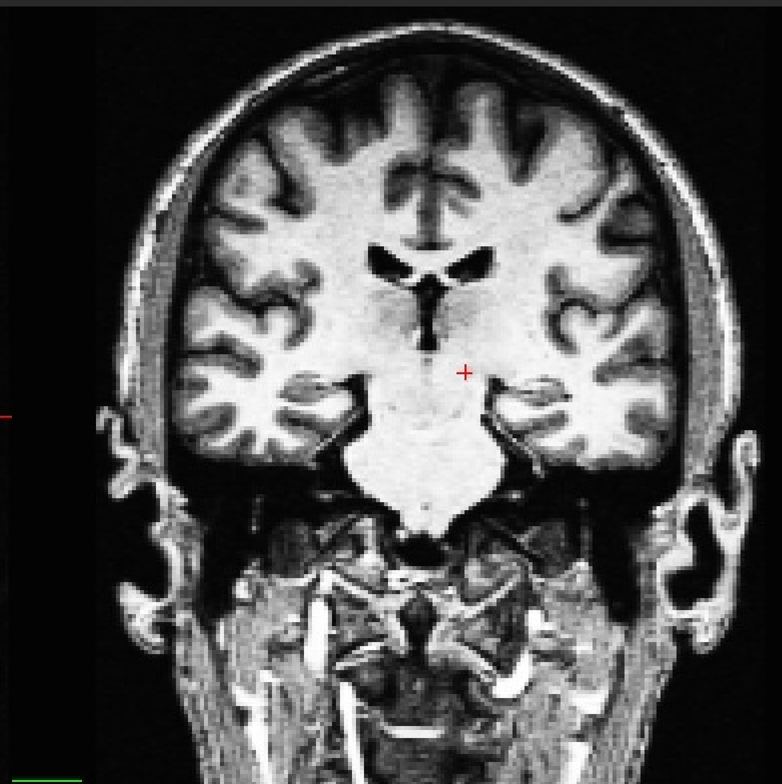
HARVARD MEDICAL SCHOOL

Overview

- Neuroanatomy 101 and fMRI Contrast Mechanism
- Preprocessing
- Hemodynamic Response
- “Univariate” GLM Analysis
- Hypothesis Testing
- Group Analysis (Random, Mixed, Fixed)

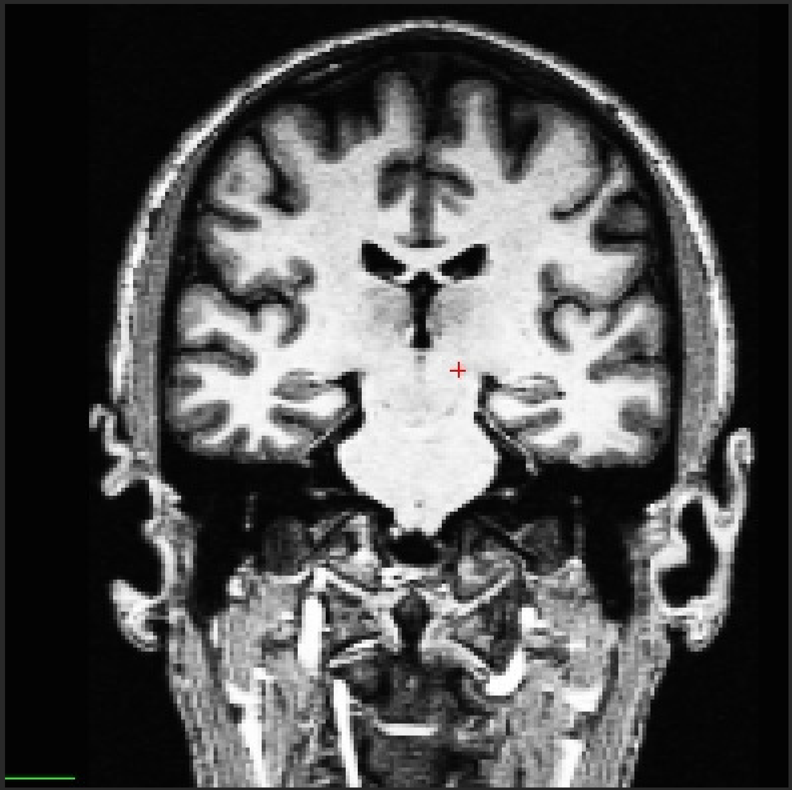
Neuroantomy

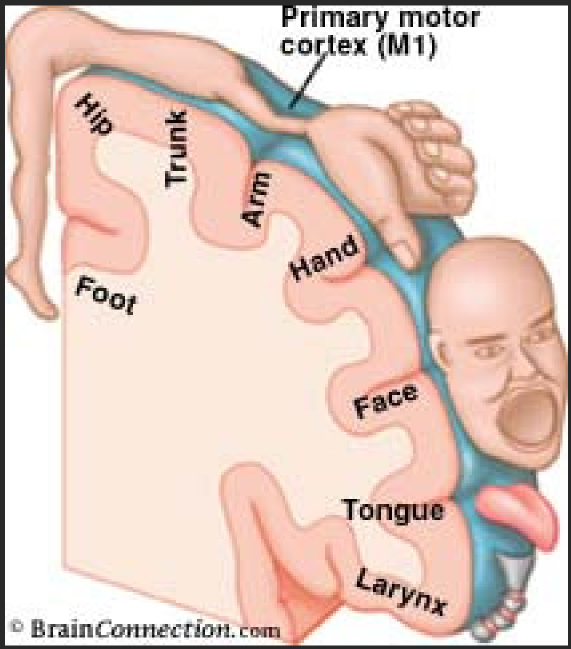




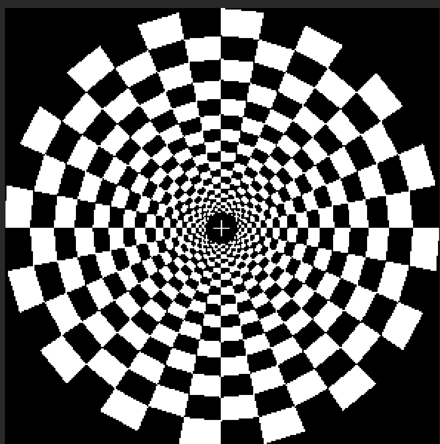
- Gray matter
- White matter
- Cerebrospinal Fluid

Functional Anatomy/Brain Mapping

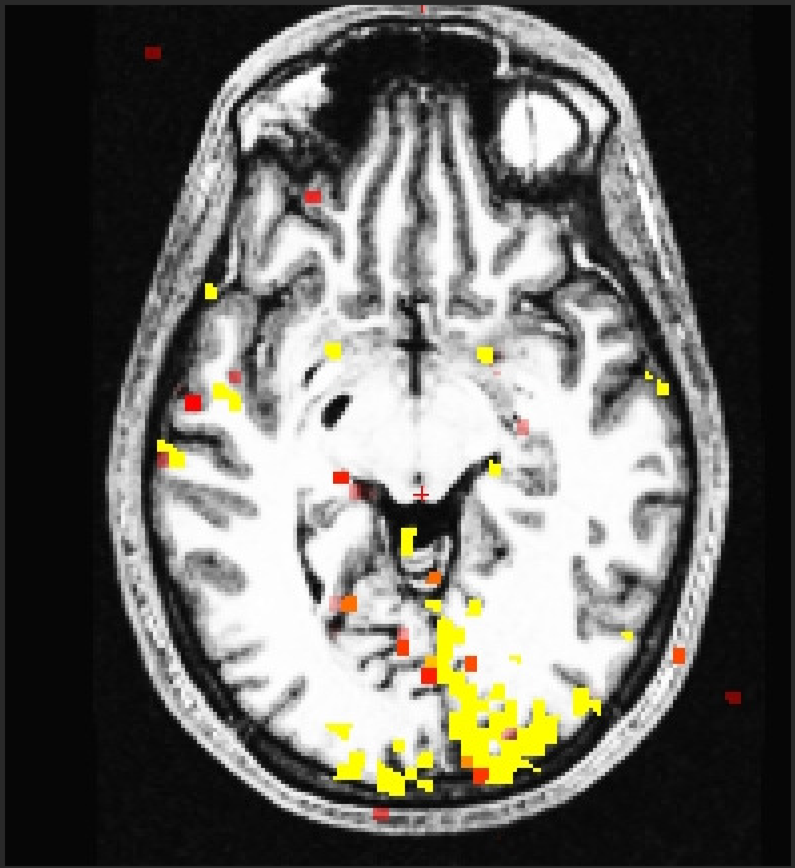




Visual Activation Paradigm



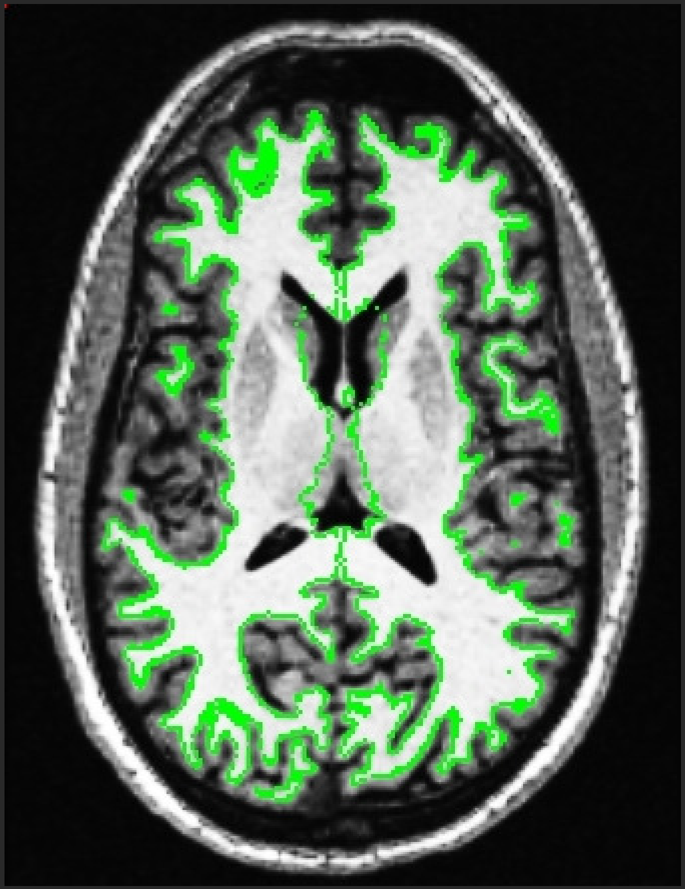
Flickering Checkerboard



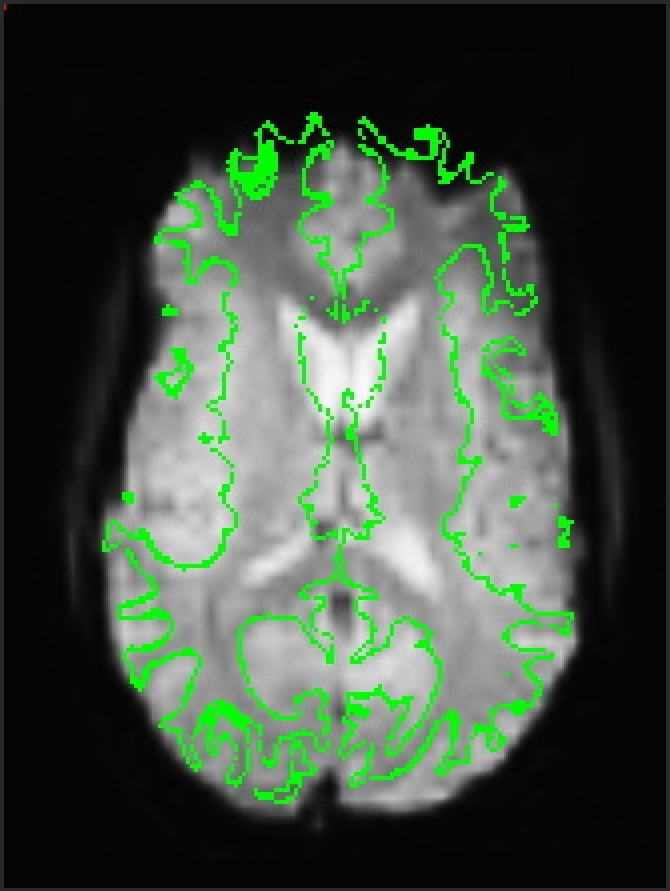
Visual, Auditory, Motor, Tactile, Pain, Perceptual, Recognition, Memory, Emotion, Reward/Punishment, Olfactory, Taste, Gastral, Gambling, Economic, Acupuncture, Meditation, The Pepsi Challenge, …

- Scientific
- Clinical
- Pharmaceutical

Magnetic Resonance Imaging

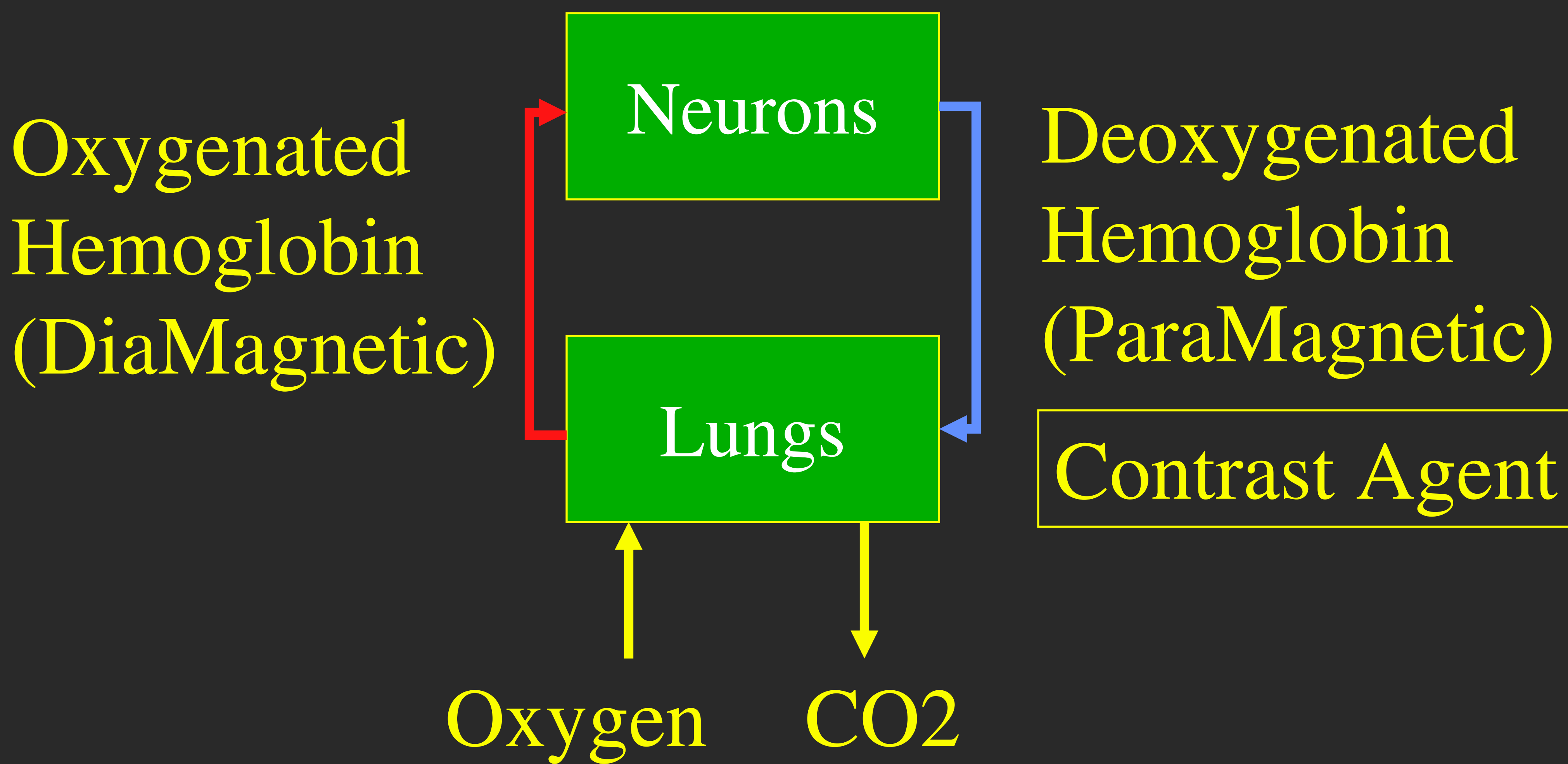


T1-weighted Contrast

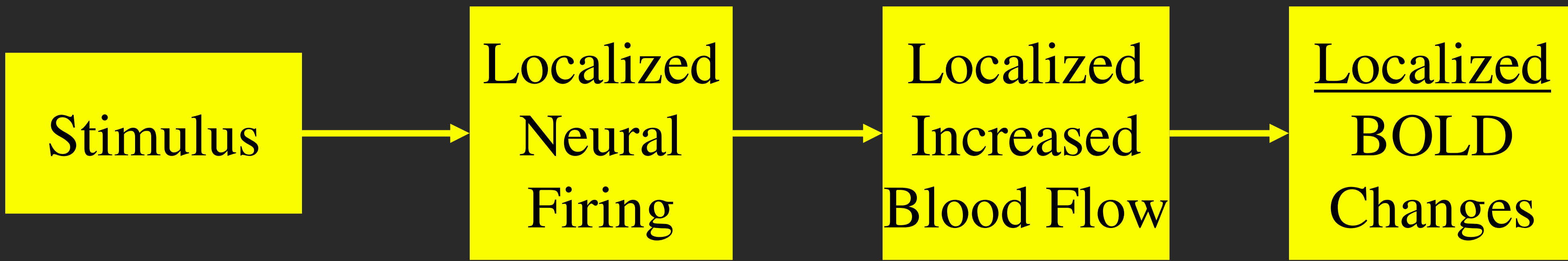


BOLD-weighted Contrast

Blood Oxygen Level Dependence (BOLD)



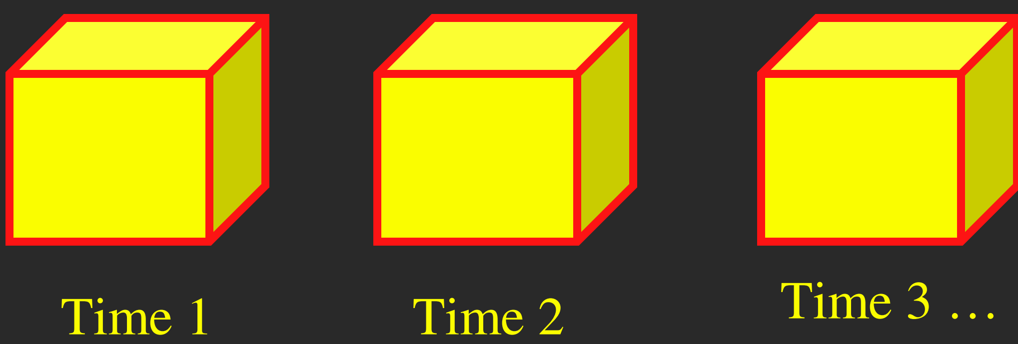
Functional MRI (fMRI)



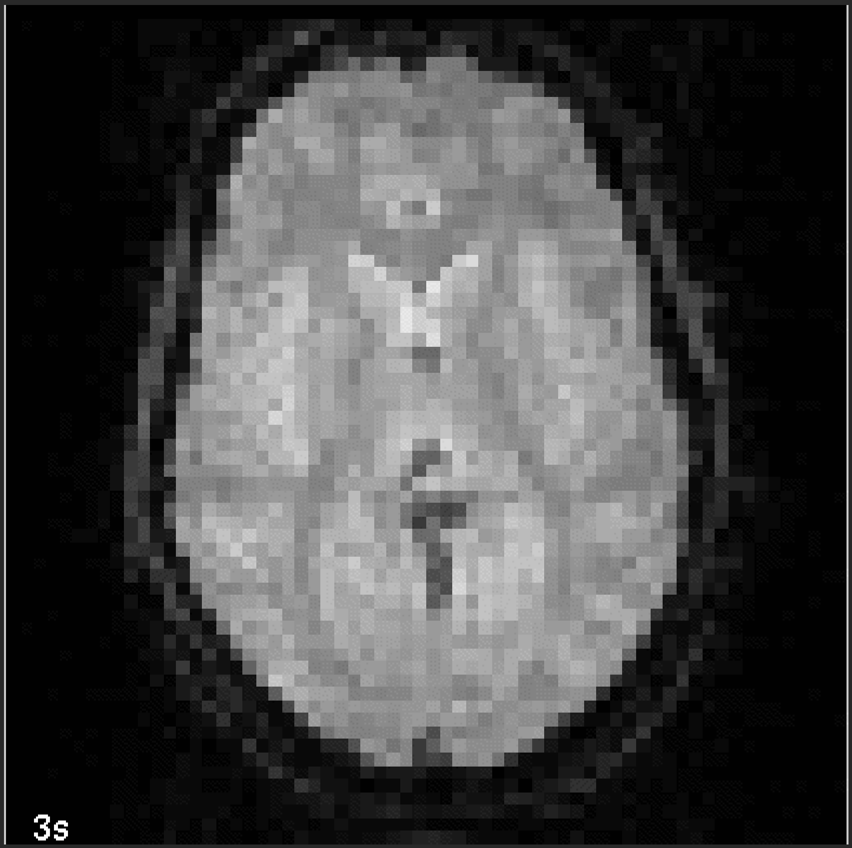
Sample BOLD response in 4D

Space (3D) – voxels (64x64x35, 3x3x5mm^3, ~50,000)

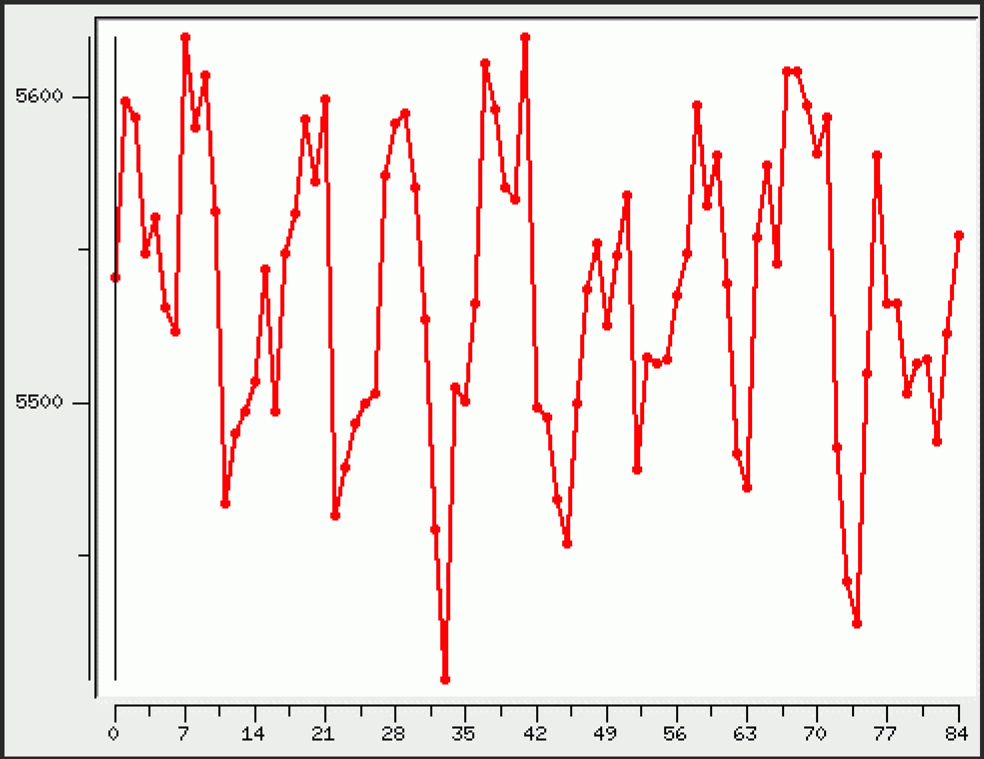
Time (1D) – time points (100, 2 sec) – Movie



4D Volume

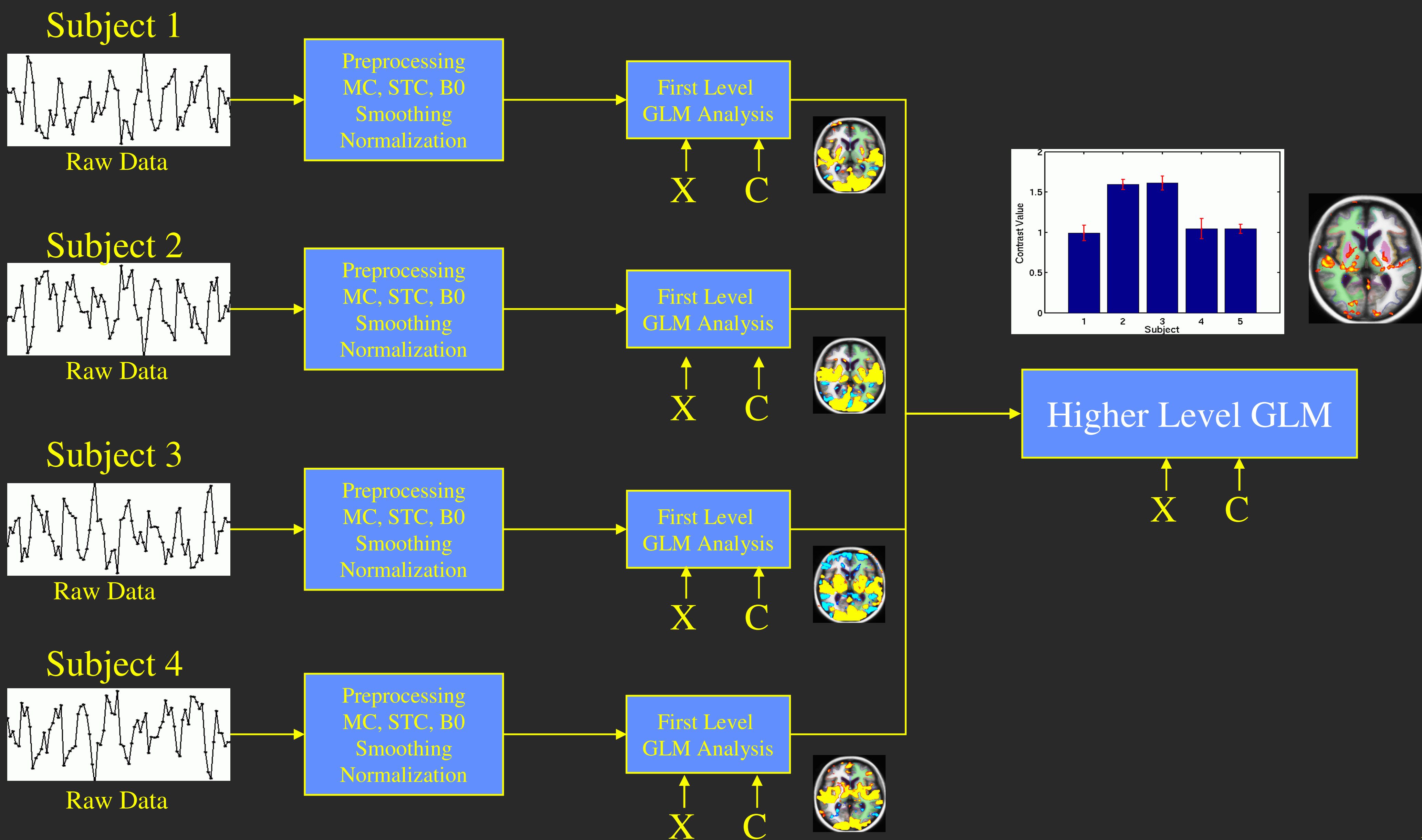


64x64x35



85x1

fMRI Analysis Overview



Preprocessing

- Assures that assumptions of the analysis are met
  - Time course comes from a single location
  - Uniformly spaced in time
  - Spatial “smoothness”
- Analysis – separating signal from noise

Preprocessing

- Start with a 4D data set
1. Motion Correction
2. Slice-Timing Correction
3. B0 Distortion Correction
4. Spatial Normalization
5. Spatial Smoothing
- End with a 4D data set

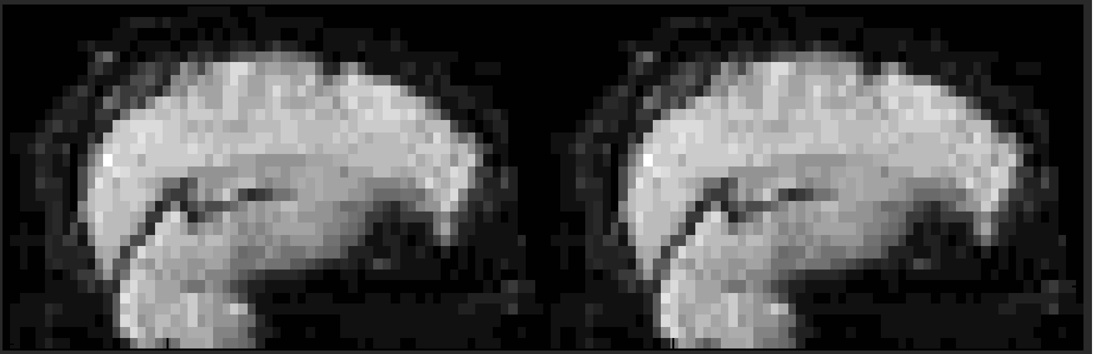
- Can be done in other orders
- Not everything is always done

Motion

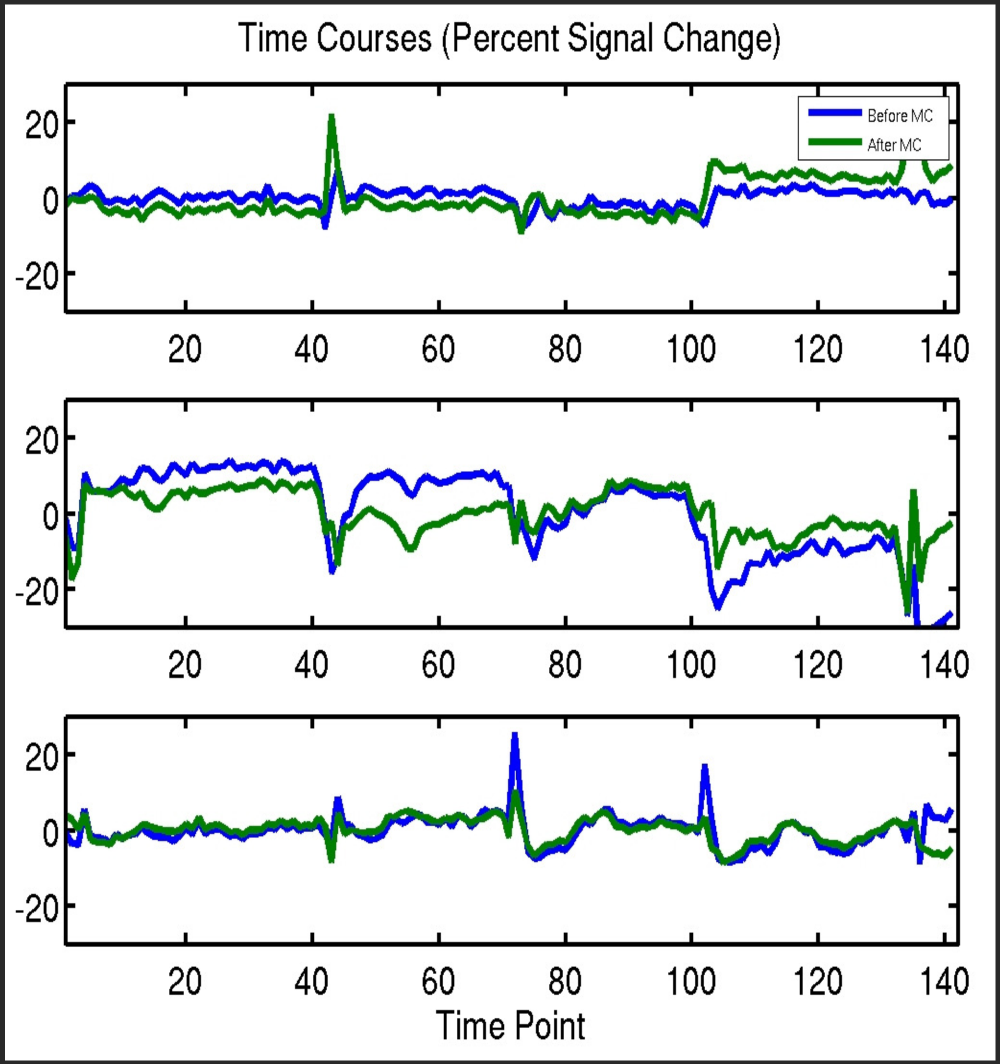
- Analysis assumes that time course represents a value from a single location
- Subjects move
- Shifts can cause noise, uncertainty
  - Edge of the brain and tissue boundaries

Motion and Motion Correction

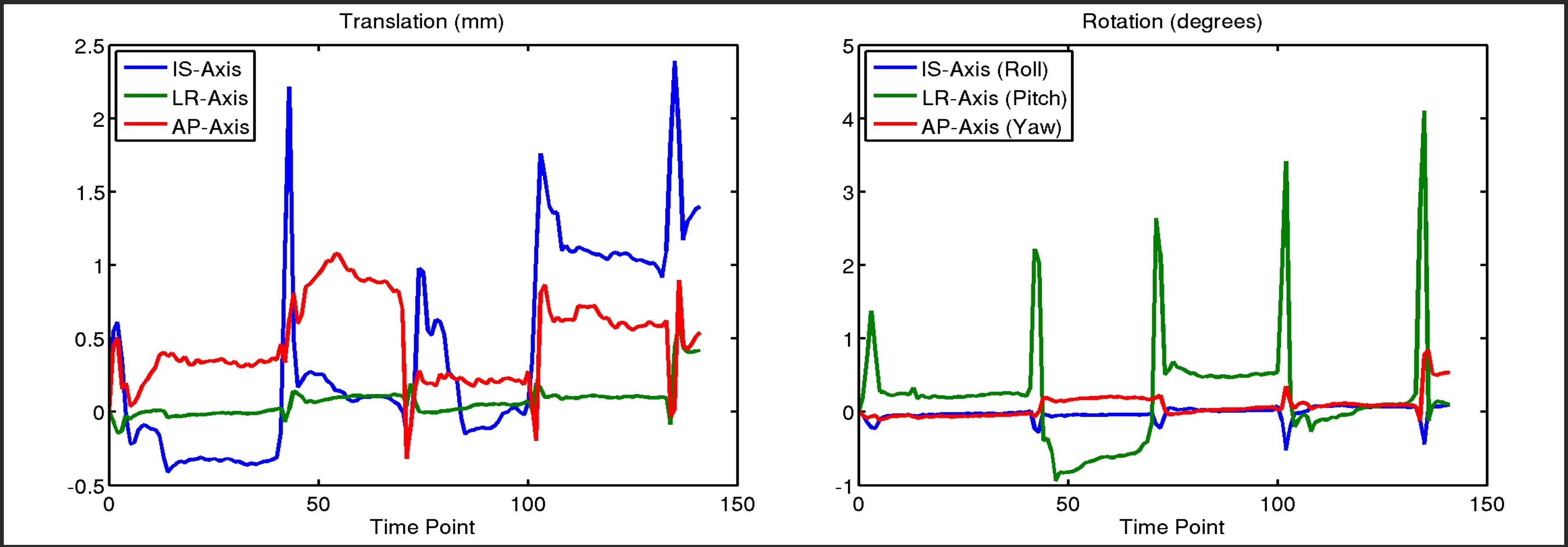
Raw Corrected



- Motion correction reduces motion
- Not perfect



Motion Correction



- Motion correction parameters
- Six for each time point
- Sometimes used as nuisance regressors
- How much motion is too much?

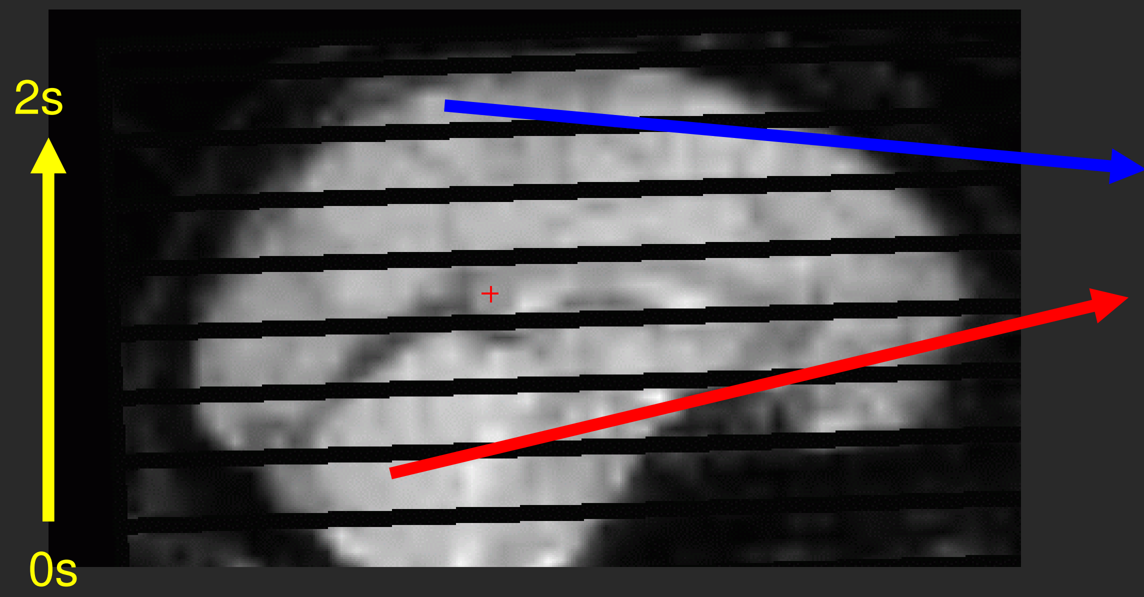
Slice Timing

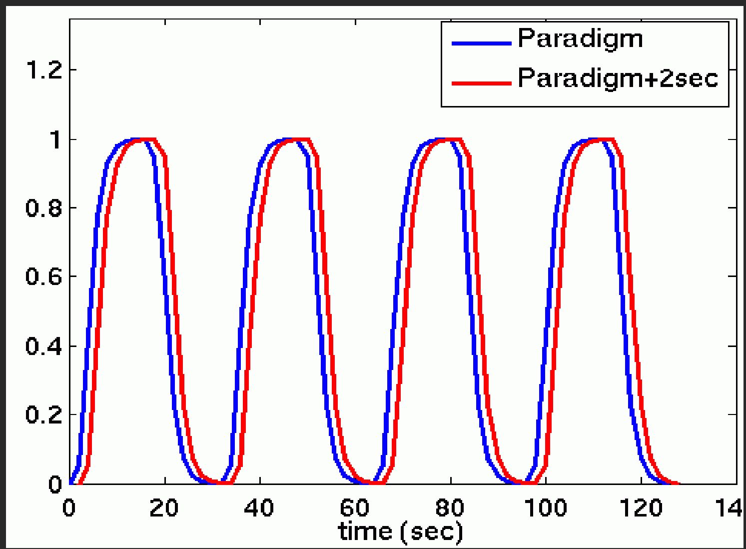
Ascending Interleaved

- Volume not acquired all at one time
- Acquired slice-by-slice
- Each slice has a different delay

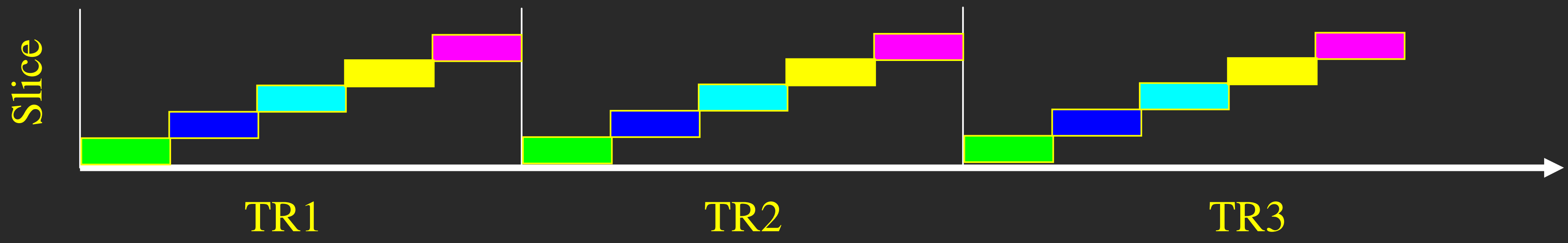
Effect of Slice Delay on Time Course

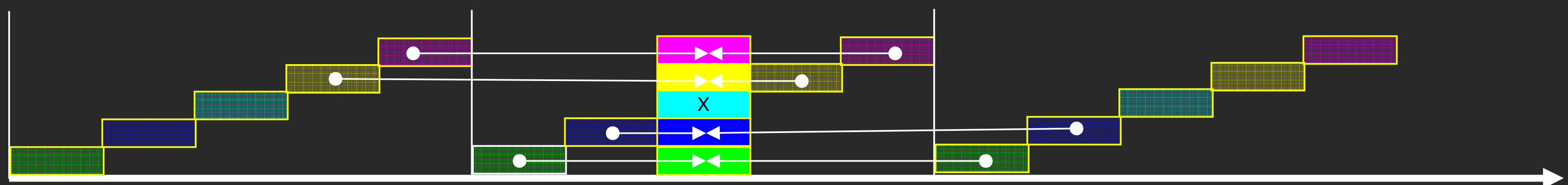
- Volume = 30 slices
- TR = 2 sec
- Time for each slice = 2/30 = 66.7 ms





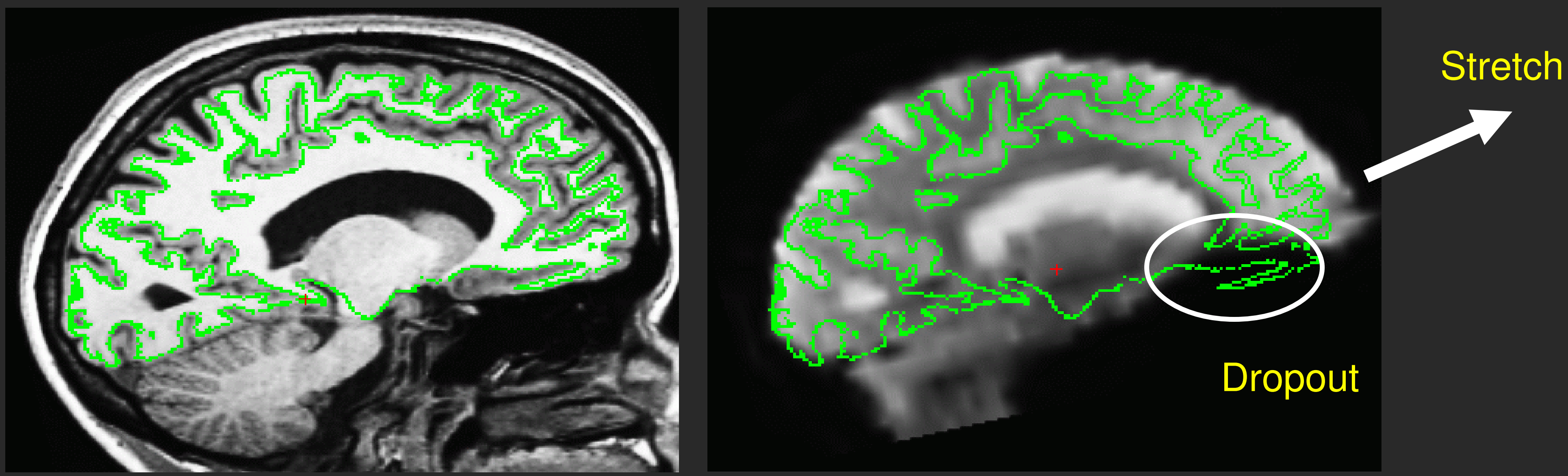
Slice Timing Correction





- Temporal interpolation of adjacent time points
- Usually sinc interpolation
- Each slice gets a different interpolation
- Some slices might not have any interpolation
- Can also be done in the GLM
- You must know the slice order!

B0 Distortion



- Metric (stretching or compressing)
- Intensity Dropout
- A result of a long readout needed to get an entire slice in a single shot.
- Caused by B0 Inhomogeneity

B0 Map

Magnitude

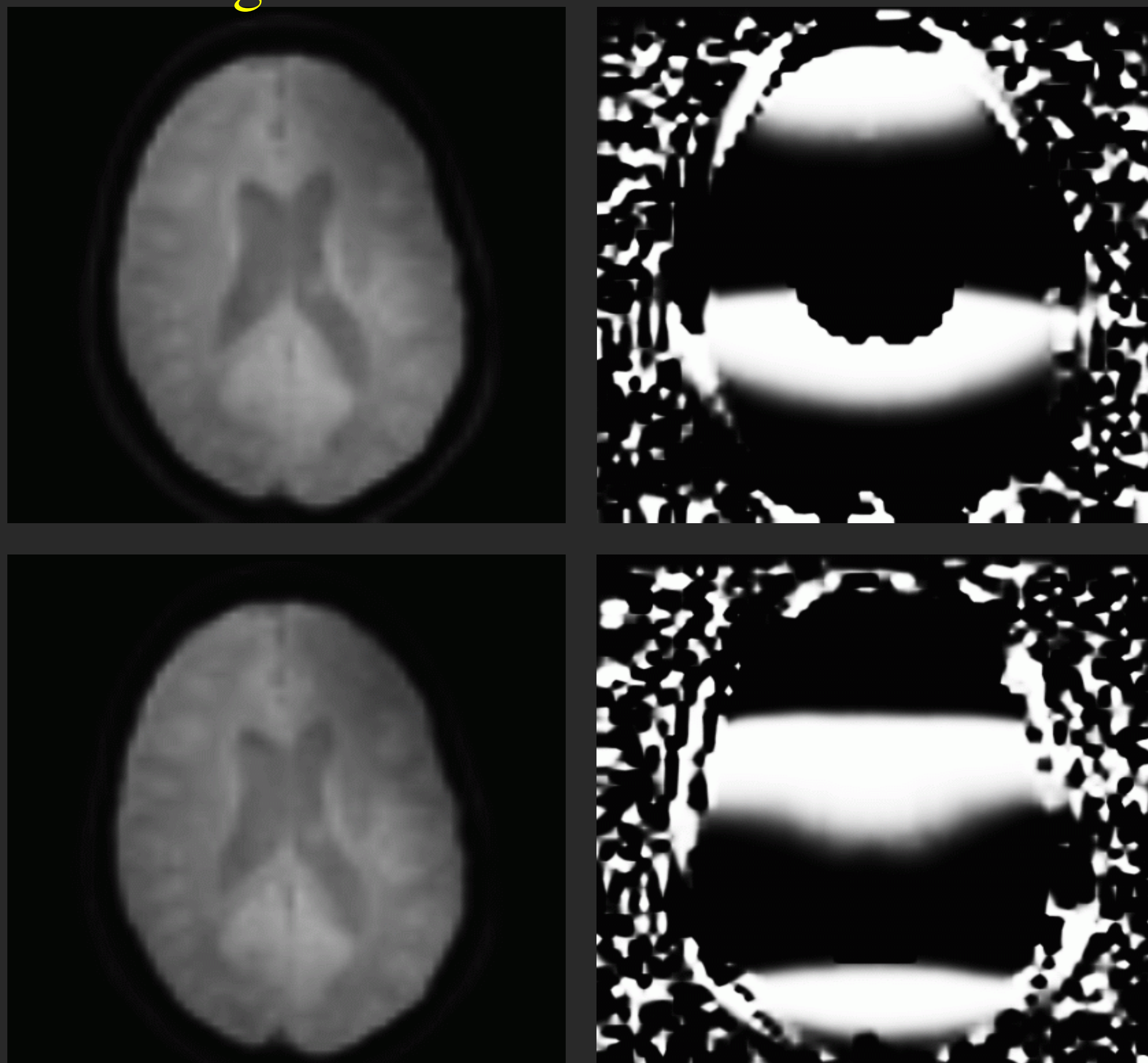
Phase

Echo 1

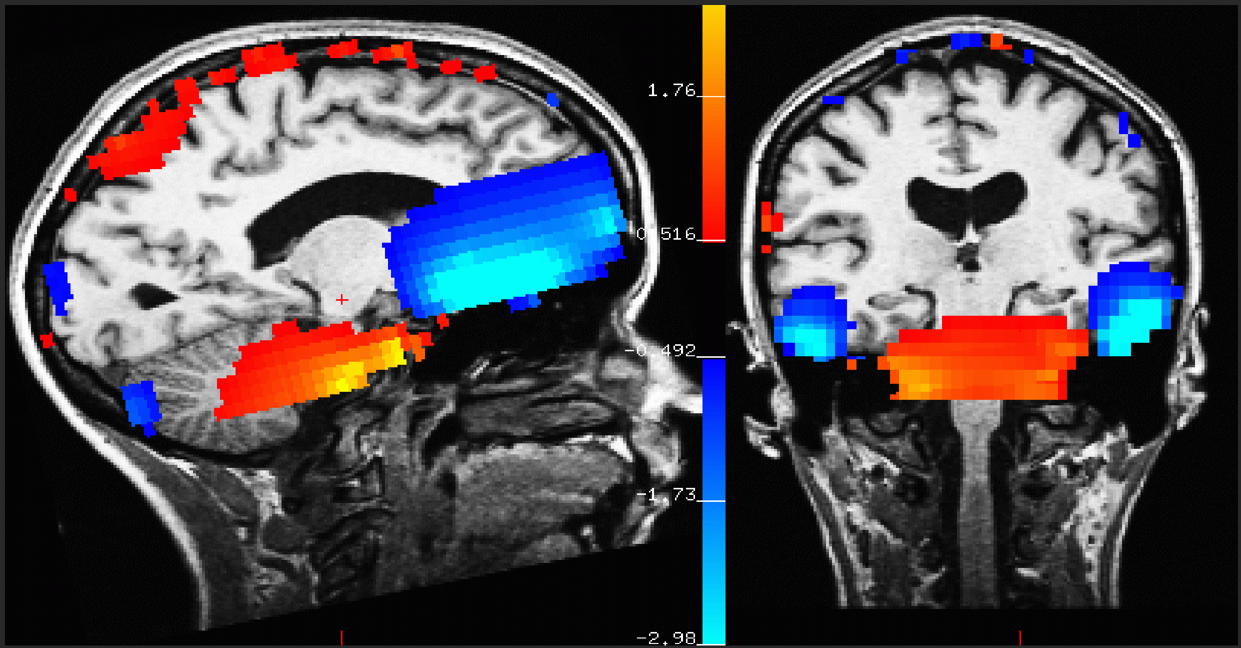
TE1

Echo 2

TE2

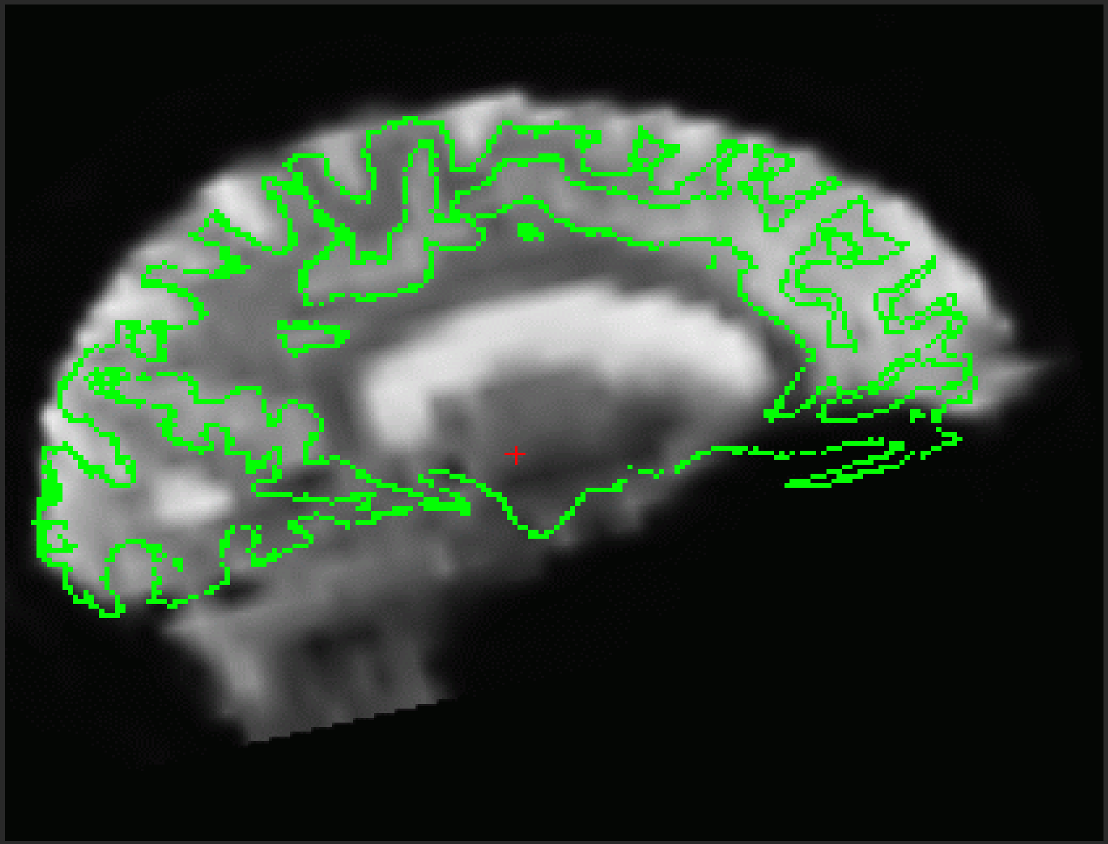


Voxel Shift Map



- Units are voxels (3.5mm)
- Shift is in-plane
- Blue = P→A, Red A→P
- Regions affected near air/tissue boundaries
    - sinuses

B0 Distortion Correction



- Can only fix metric distortion
- Dropout is lost forever

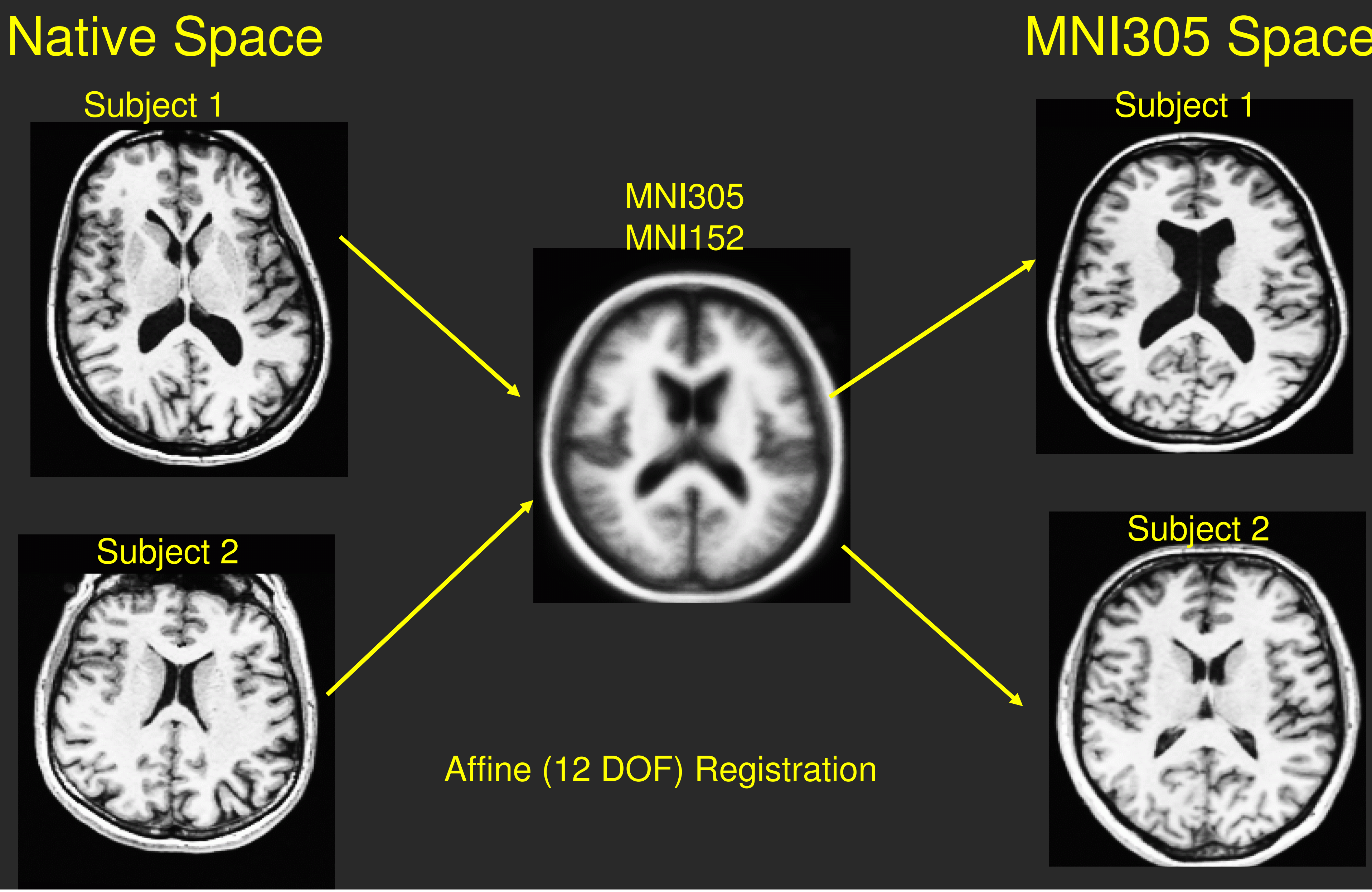
B0 Distortion Correction

- Can only fix metric distortion
- Dropout is lost forever
- Interpolation
- Need:
  - "Echo spacing" – readout time
  - Phase encode direction
- More important for surface than for volume
- Important when combining from different scanners

Spatial Normalization

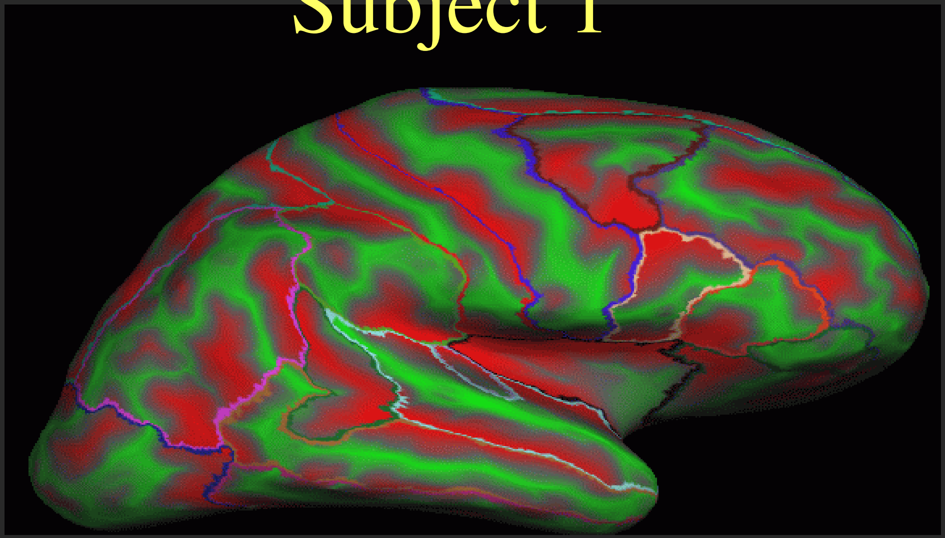
- Transform volume into another volume
  - Re-slicing, re-gridding
- New volume is an “atlas” space
- Align brains of different subjects so that a given voxel represents the “same” location.
- Similar to motion correction
- Preparation for comparing across subjects
- Volume-based
- Surface-based
- Combined Volume-surface-based (CVS)

Spatial Normalization: Volume

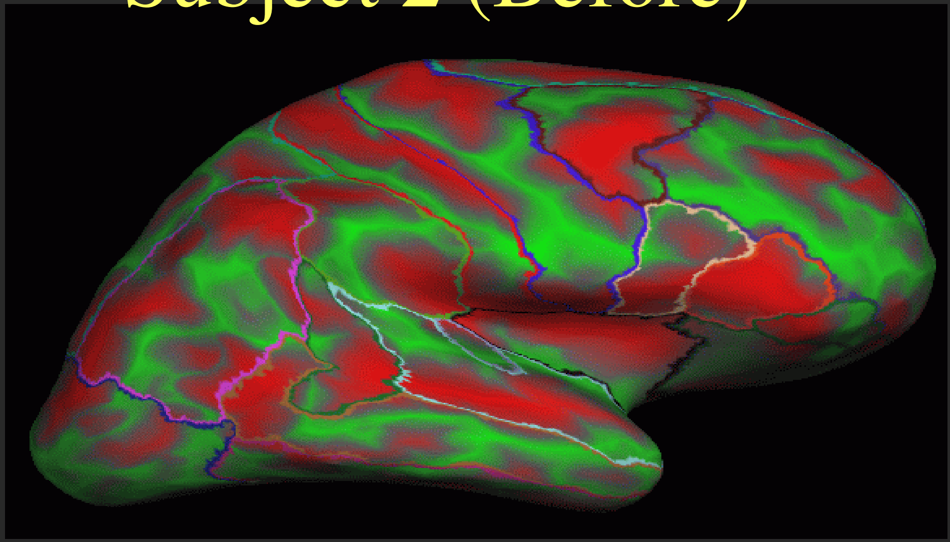


Spatial Normalization: Surface

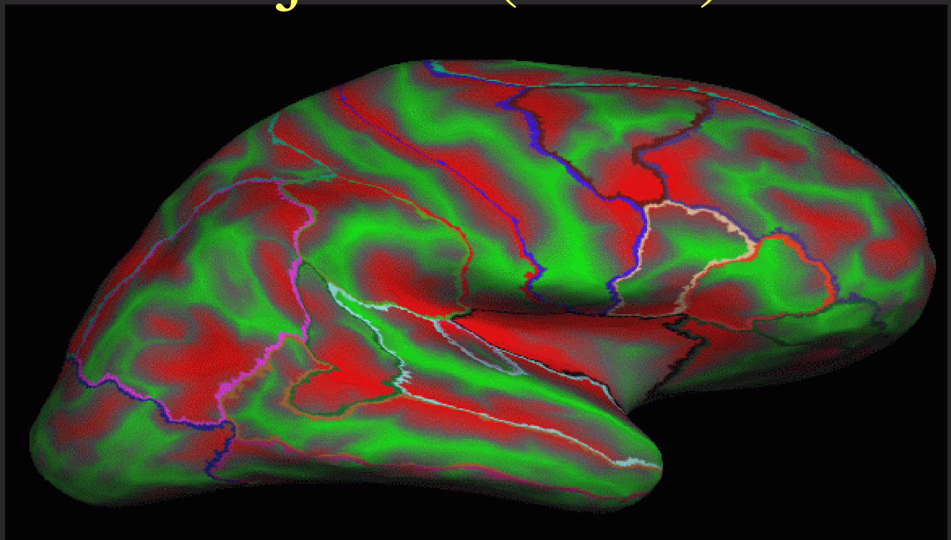
Subject 1



Subject 2 (Before)



Subject 2 (After)



- Shift, Rotate, Stretch
- High dimensional (~500k)
- Preserve metric properties
- Take variance into account
- Common space for group analysis (like Talairach)

Spatial Smoothing

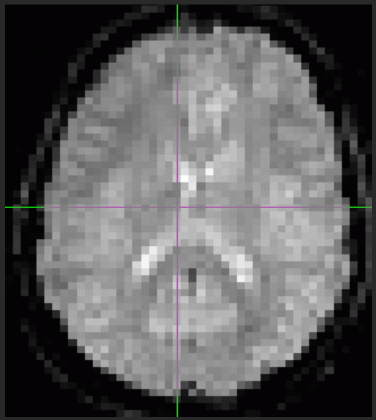
- Replace voxel value with a weighted average of nearby voxels (spatial convolution)
- Weighting is usually Gaussian
- 3D (volume)
- 2D (surface)
- Do after all interpolation, before computing a standard deviation
- Similarity to interpolation
- Improve SNR
- Improve Intersubject registration
- Can have a dramatic effect on your results

Spatial Smoothing

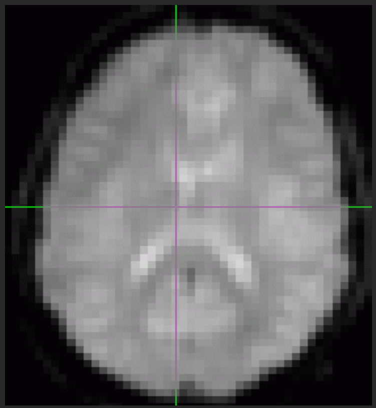
- Spatially convolve image with Gaussian kernel.
- Kernel sums to 1
- Full-Width/Half-max: FWHM = σ/sqrt(log(256))

σ = standard deviation of the Gaussian

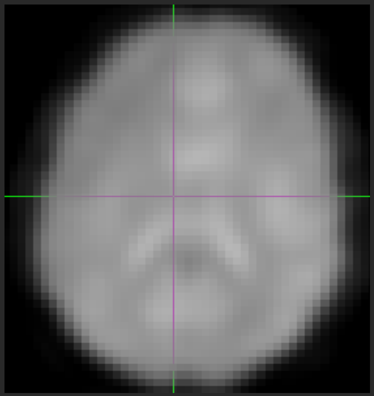
0 FWHM



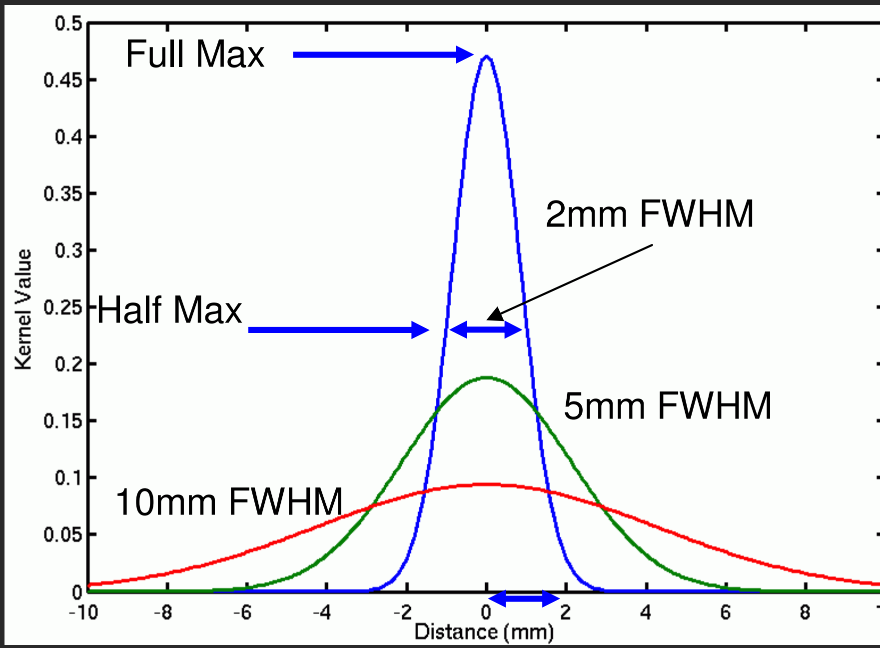
5 FWHM



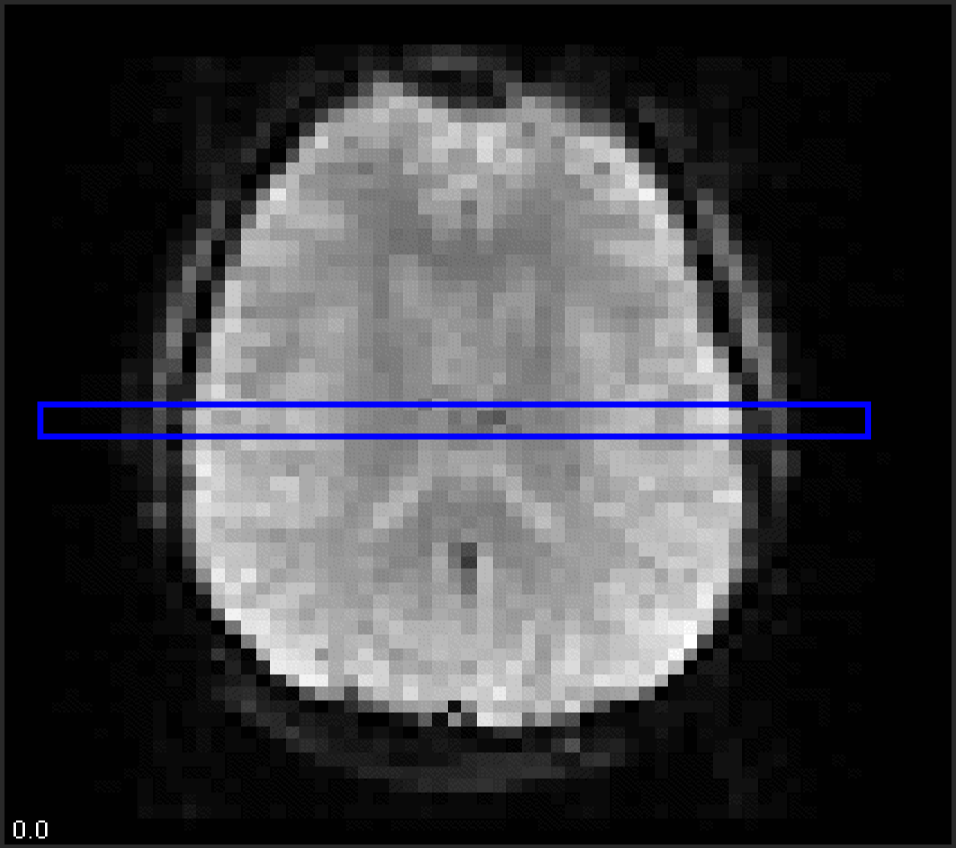
10 FWHM

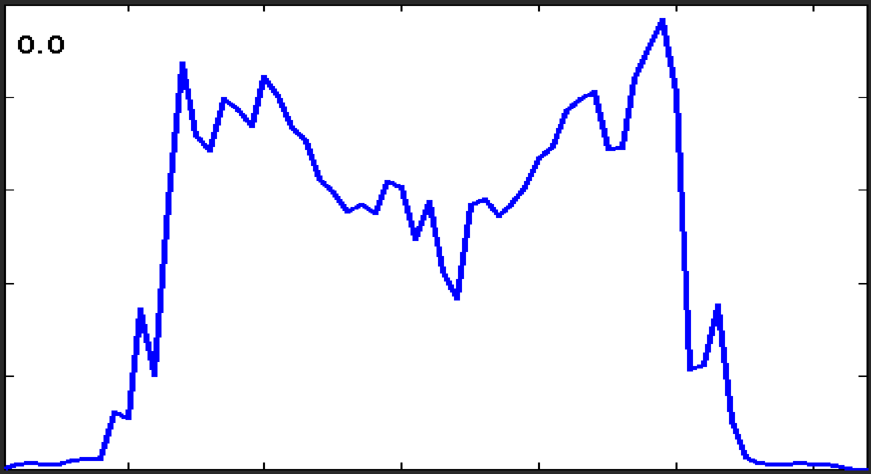


Full-Width/Half-max

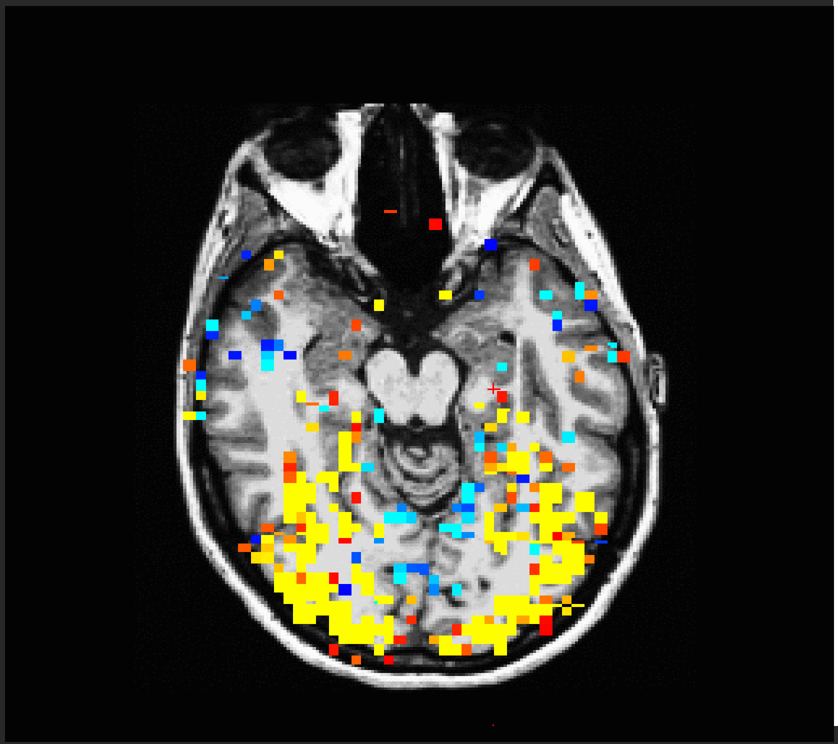


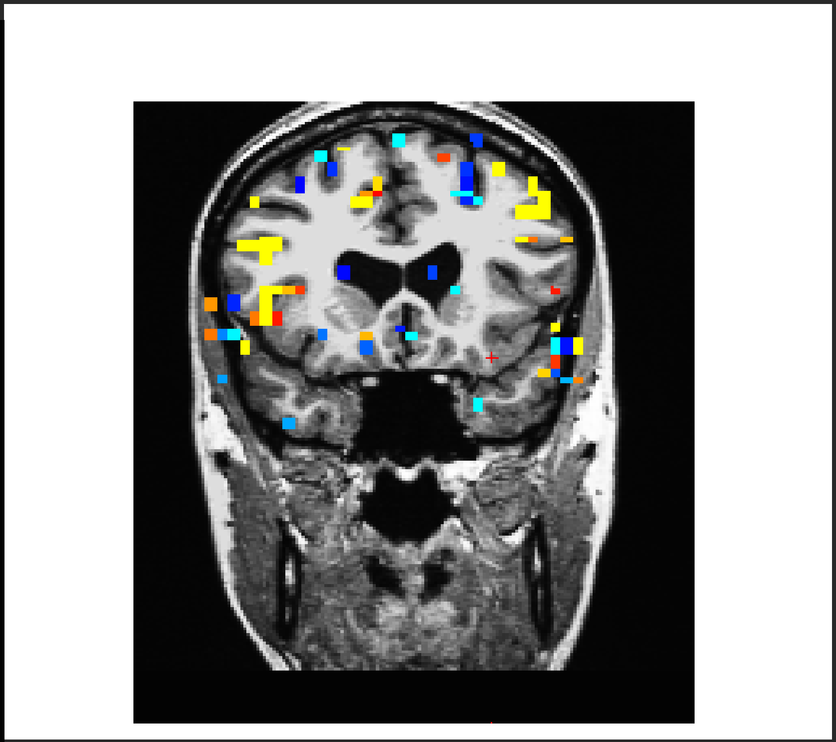
Spatial Smoothing





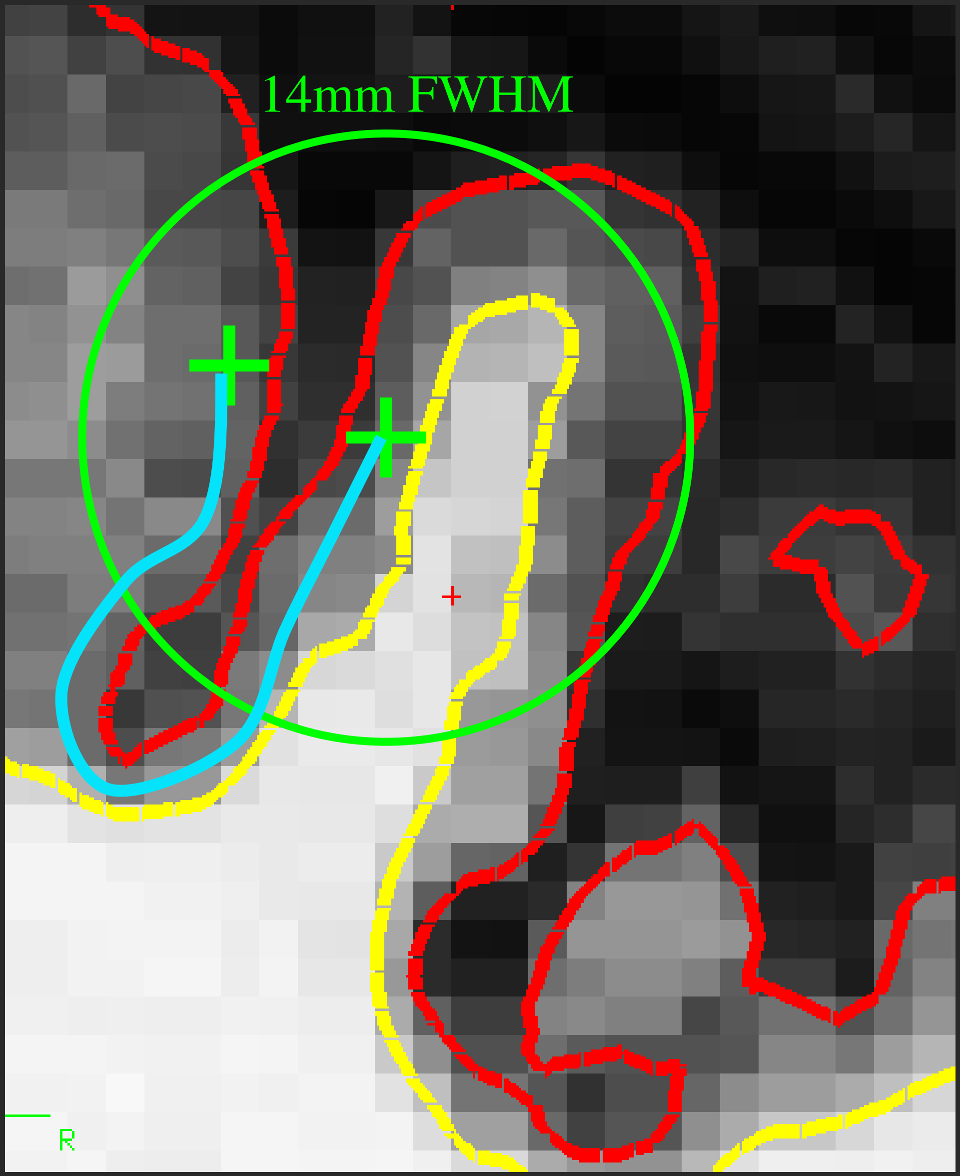
Effect of Smoothing on Activation





- Working memory paradigm
- FWHM: 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20

Volume- vs Surface-based Smoothing



- 5 mm apart in 3D
- 25 mm apart on surface
- Averaging with other tissue types (WM, CSF)
- Averaging with other functional areas

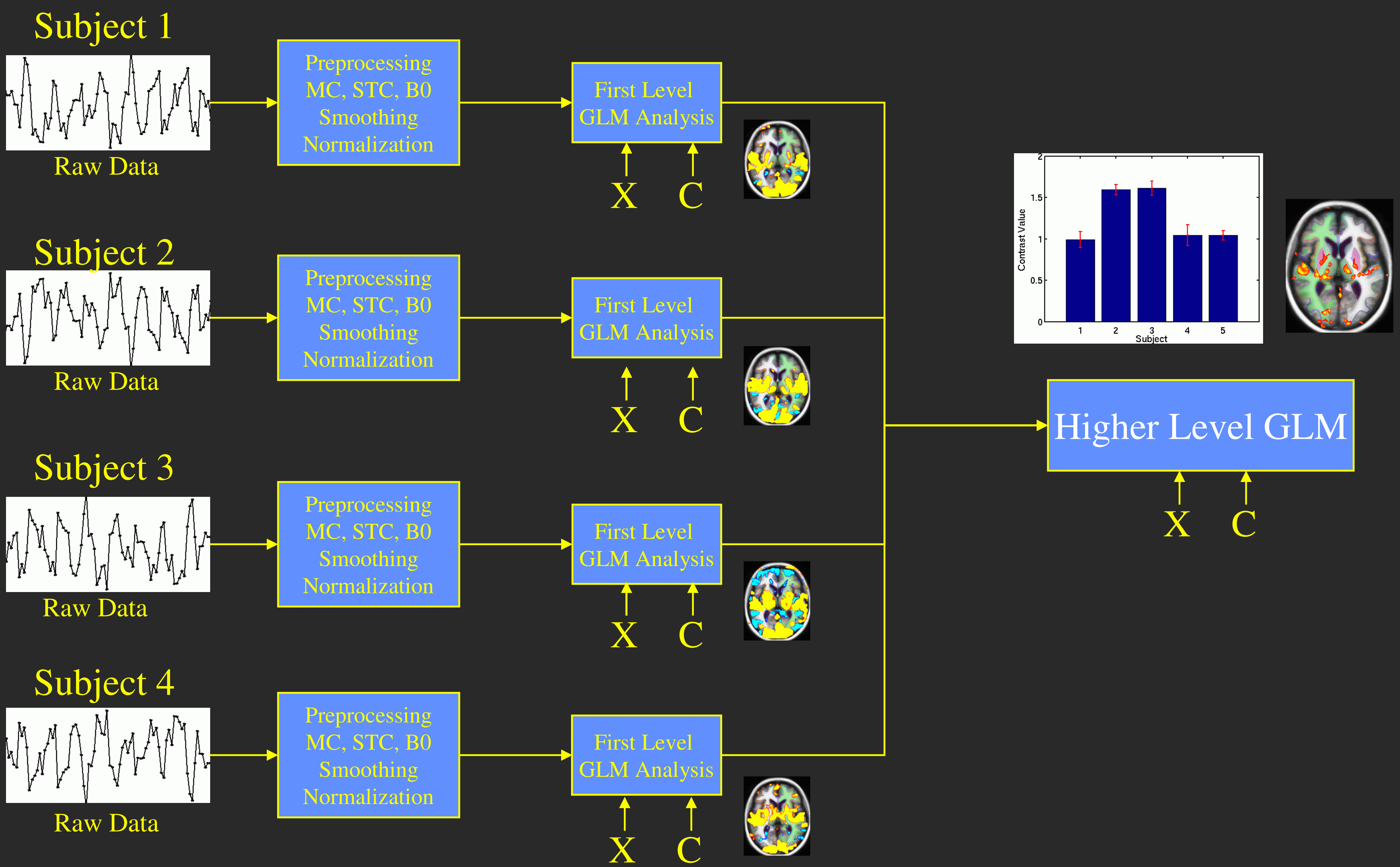
Preprocessing

- Start with a 4D data set
1. Motion Correction - Interpolation
2. Slice-Timing Correction
3. B0 Distortion Correction - Interpolation
4. Spatial Normalization - Interpolation
5. Spatial Smoothing – Interpolation-like
- End with a 4D data set

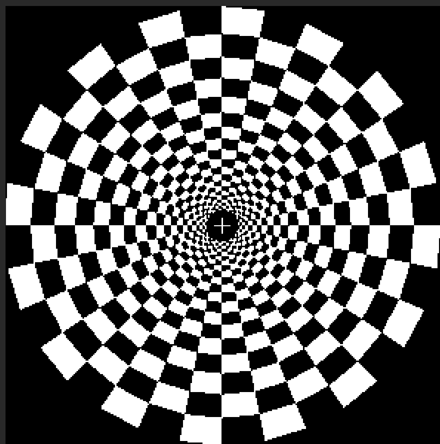
- Can be done in other orders
- Not all are done

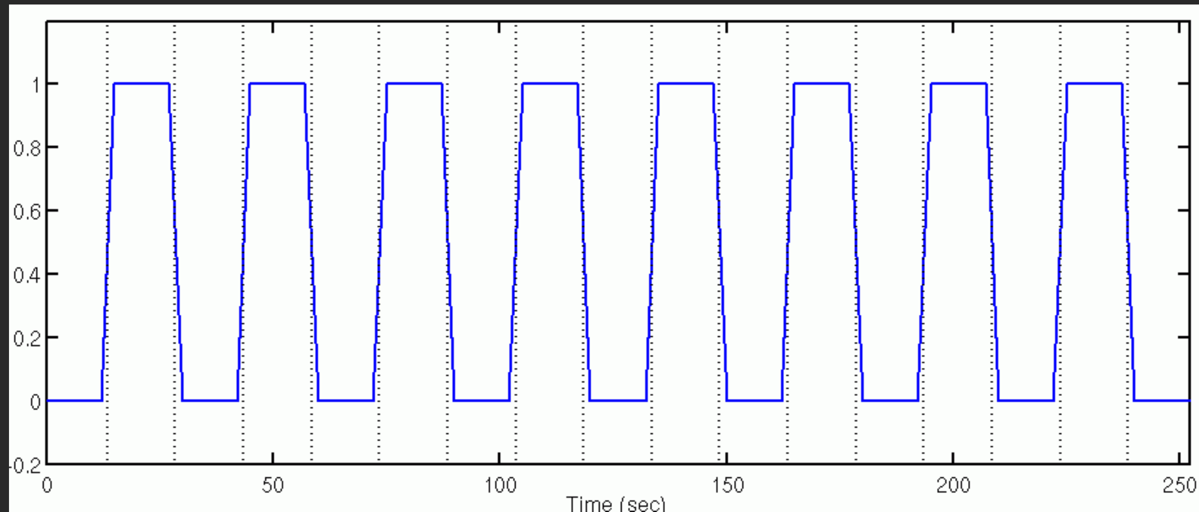
fMRI Time-Series Analysis

fMRI Analysis Overview



Visual/Auditory/Motor Activation Paradigm

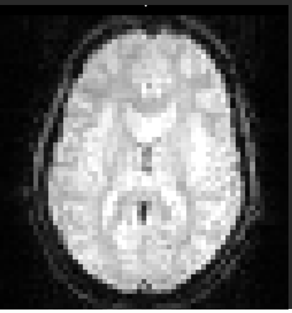


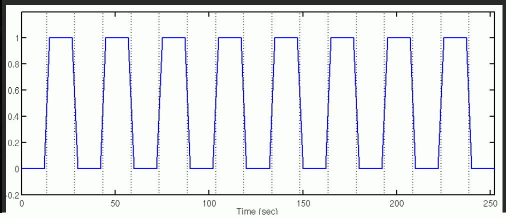


15 sec 'ON', 15 sec 'OFF'

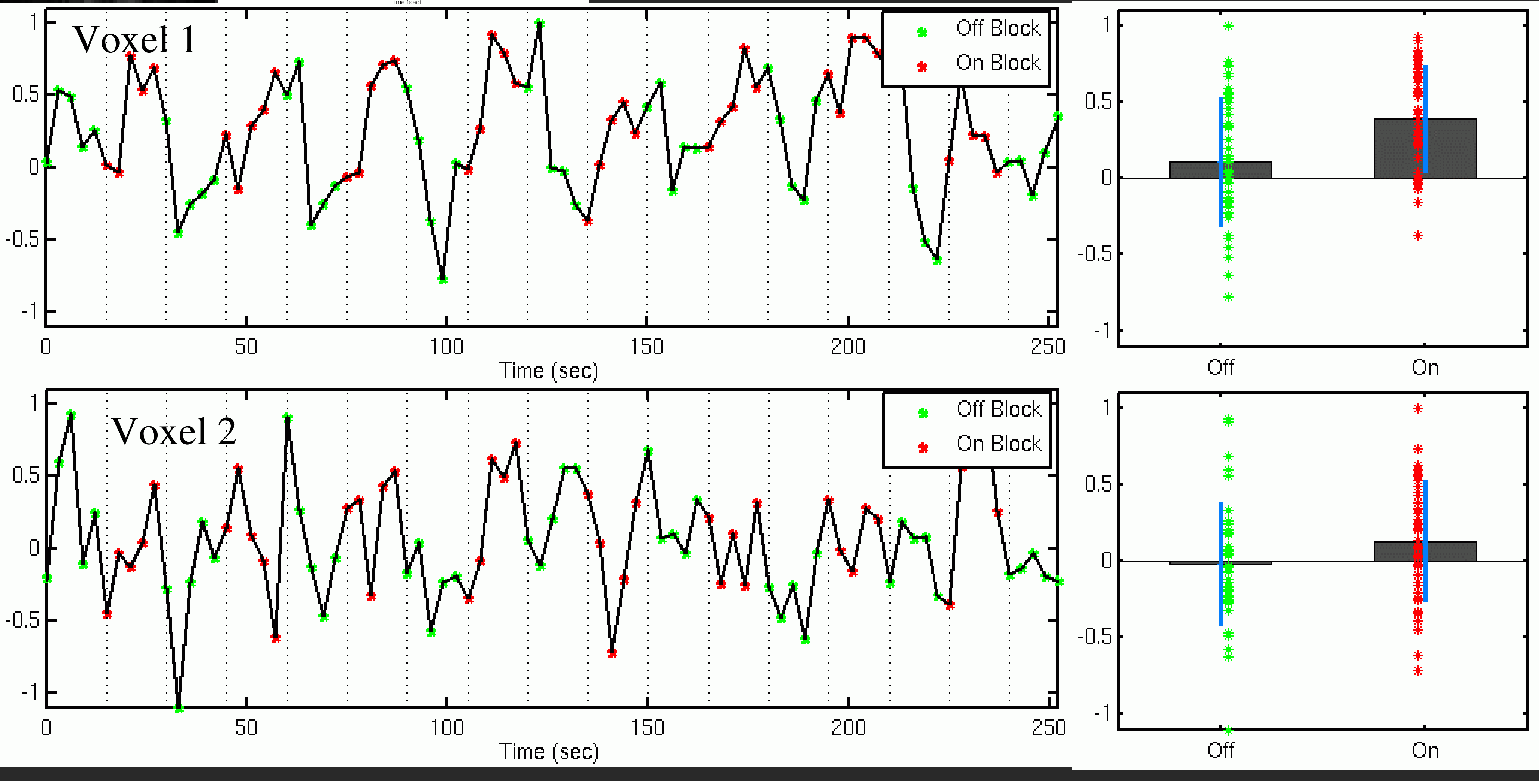
- **Flickering Checkerboard**
- **Auditory Tone**
- **Finger Tapping**

Block Design: 15s Off, 15s On

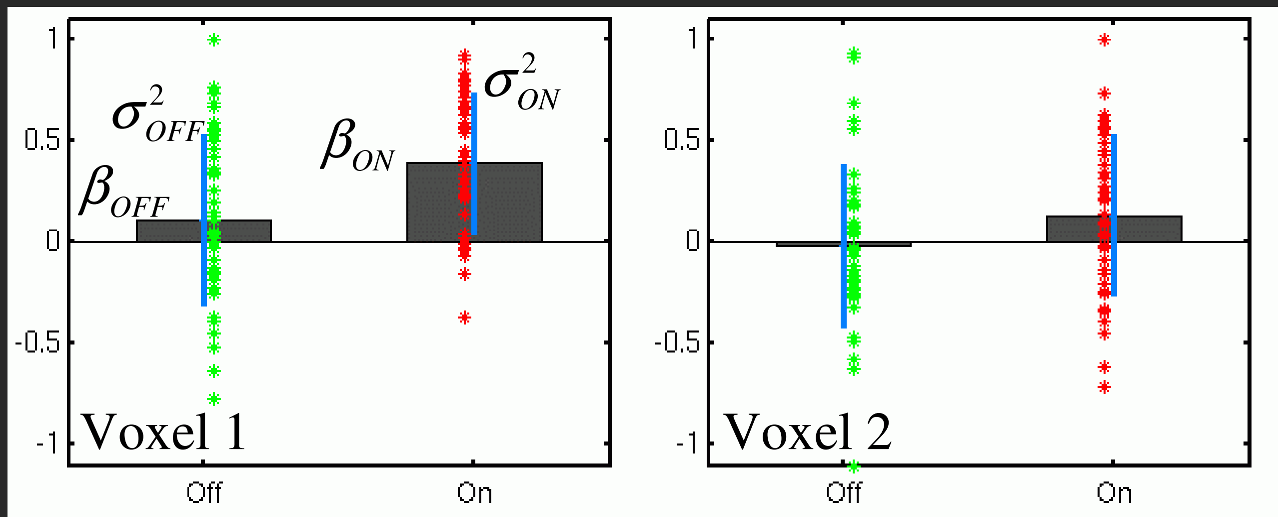




Stimulus Schedule
Paradigm File



Contrasts and Inference



$$= \frac{\beta_{ON} - \beta_{OFF}}{\sqrt{\frac{(N_{ON}-1)\sigma^2_{ON} + (N_{OFF}-1)\sigma^2_{OFF}}{(N_{ON}+N_{OFF}-2)^2}}}$$

$\beta_{ON}, \sigma^2_{ON}, N_{ON}$ – Mean, Var, N in ON

$\beta_{OFF}, \sigma^2_{OFF}, N_{OFF}$ – Mean, Var, N in OFF

$$\text{Contrast} = \beta_{ON} - \beta_{OFF}$$

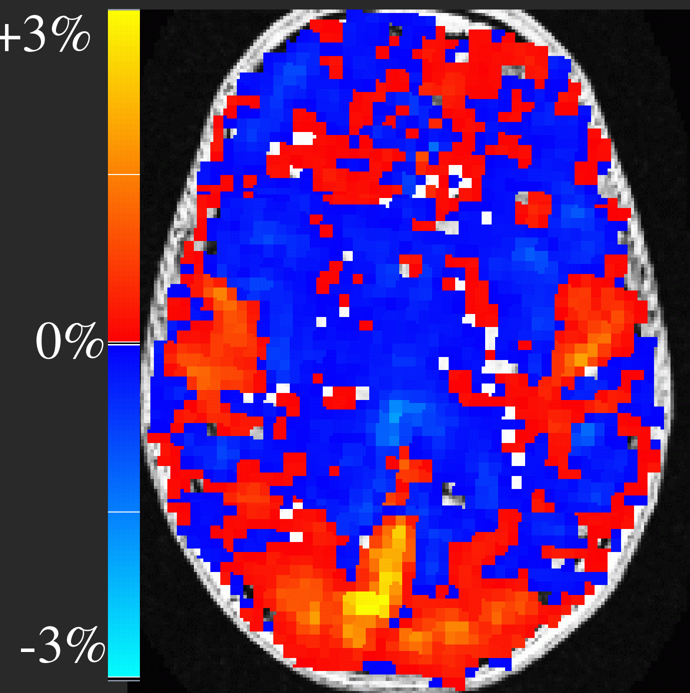
$$\text{Var(Contrast)} = \frac{(N_{ON}-1)\sigma^2_{ON} + (N_{OFF}-1)\sigma^2_{OFF}}{(N_{ON}+N_{OFF}-2)^2}$$

Note: z, t, F monotonic with p

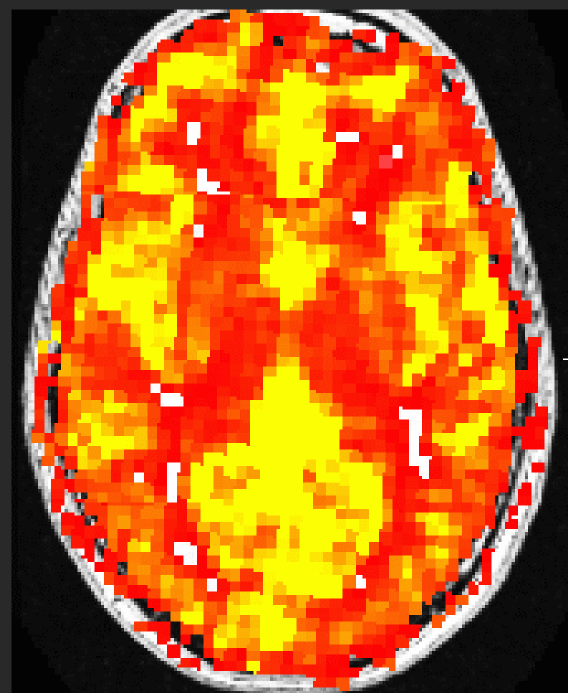
$p = 10^{-11}$, sig=-log10(p) =11

$p = .10$, sig=-log10(p) =1

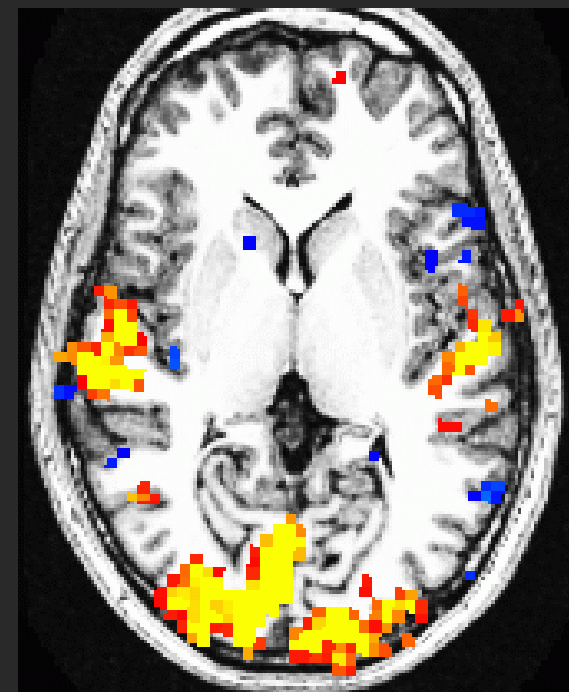
Statistical Parametric Map (SPM)



Contrast Amplitude
CON, COPE, CES



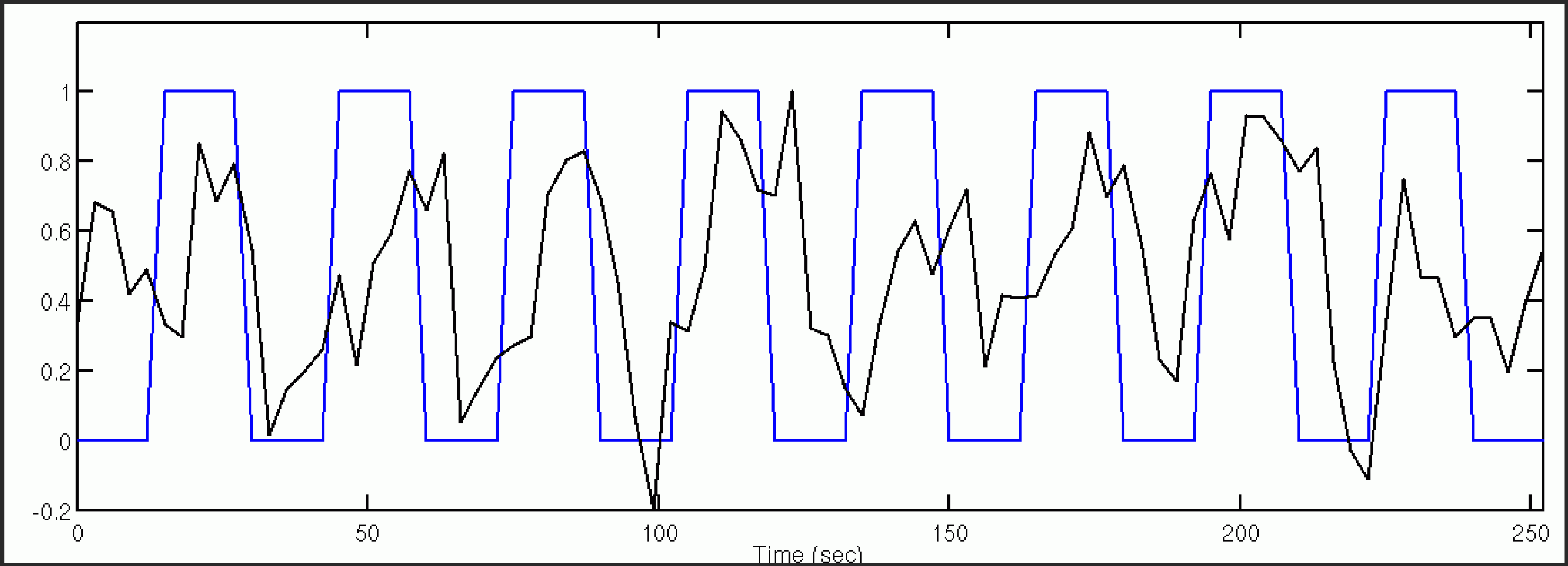
Contrast Amplitude
Variance
(Error Bars)
VARCOPE, CESVAR



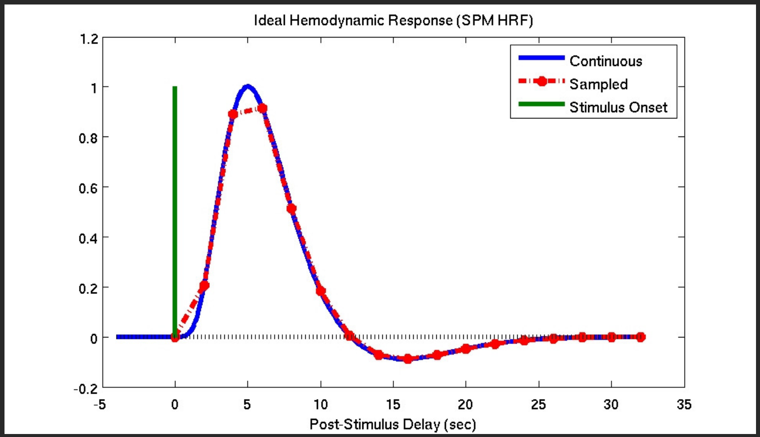
Significance
t-Map (p,z,F)
(Thresholded p<.01)
sig=-log10(p)

"Massive Univariate Analysis"
-- Analyze each voxel separately

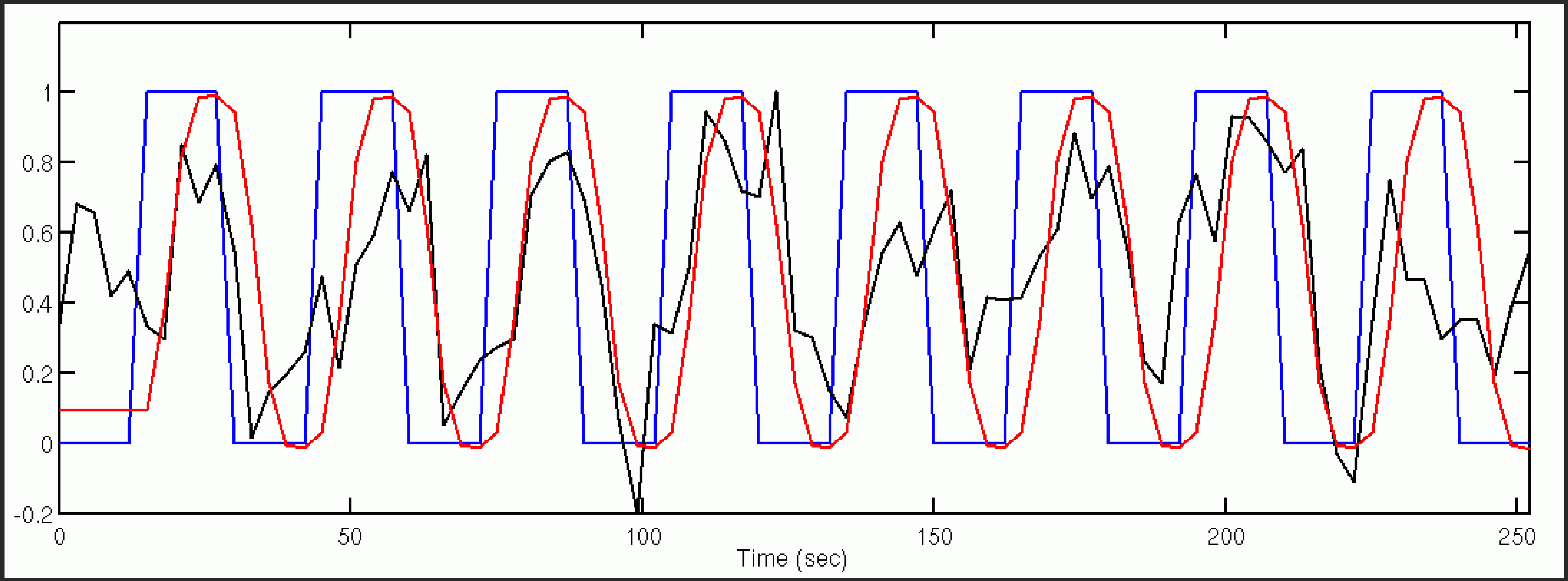
Hemodynamics



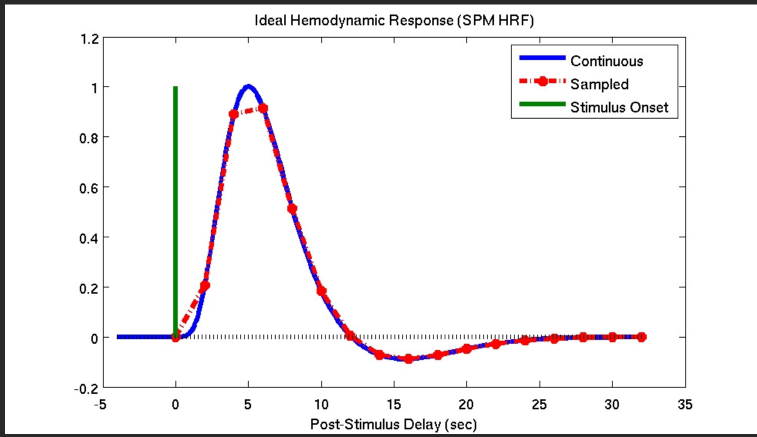
- Delay
- Dispersion
- Grouping by simple time point inaccurate



Convolution with HRF

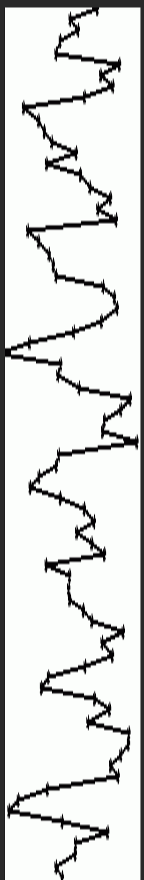


- Shifts, rolls off; more accurate
- Loose ability to simply group time points
- More complicated analysis
- General Linear Model (GLM)



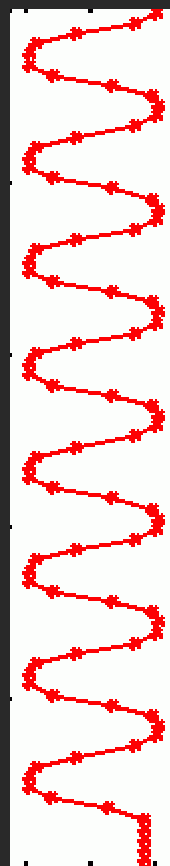
GLM

Data fom one voxel



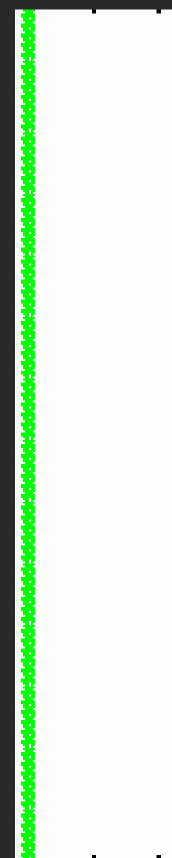
$=\ \beta_{Task}$

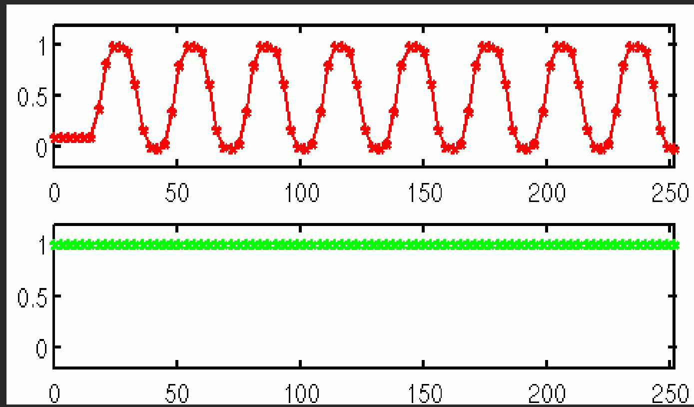
Task



$+\ \beta_{base}$

Baseline Offset (Nuisance)

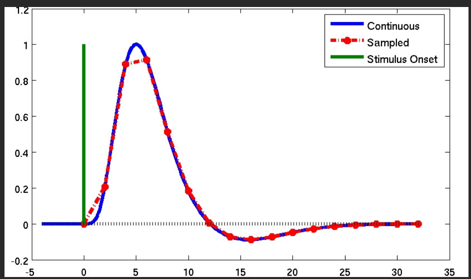




$\beta_{base} = \beta_{off}$

$\beta_{Task} = \beta_{on} - \beta_{off}$

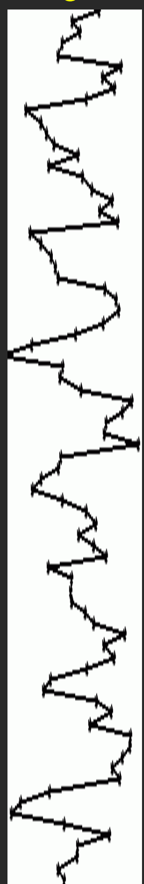
- Implicit Contrast
- HRF Amplitude



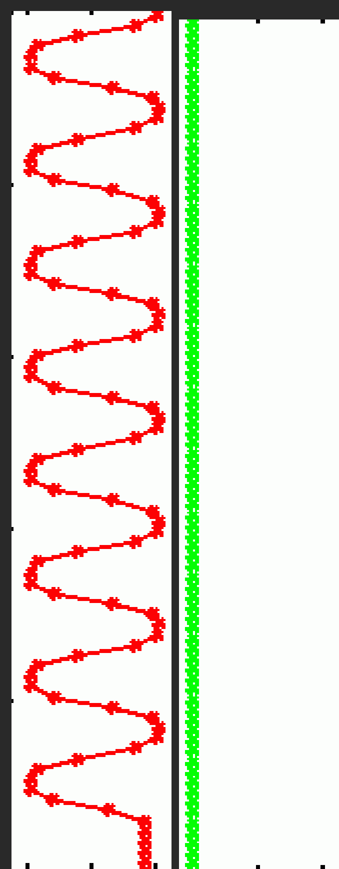
Matrix Model

$$y = X * \beta$$

Observations



$$= \qquad \begin{bmatrix} \beta_{Task} \\ \beta_{base} \end{bmatrix}$$



Data from one voxel

Design Matrix Regressors

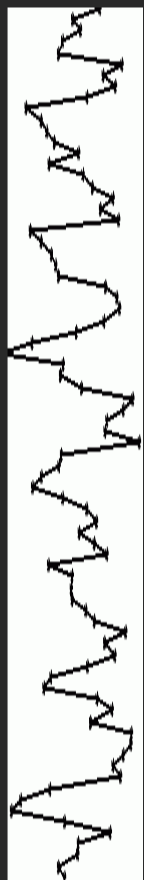
Vector of Regression Coefficients ("Betas")

Design Matrix

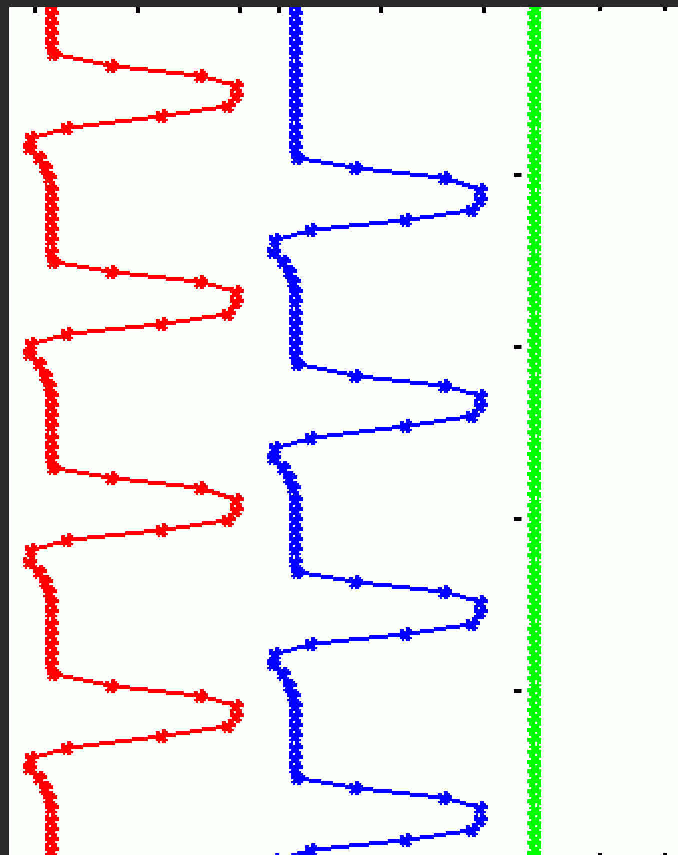
Two Task Conditions

$$y = X * \beta$$

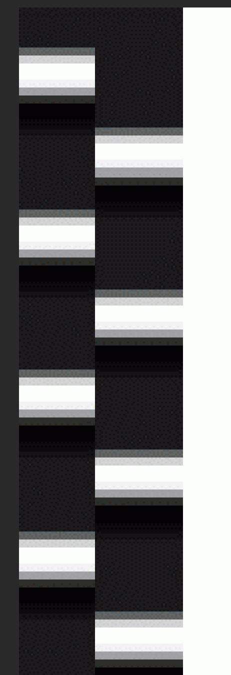
Observations



=



$$\begin{bmatrix} \beta_{Odd} \\ \beta_{Even} \\ \beta_{base} \end{bmatrix}$$

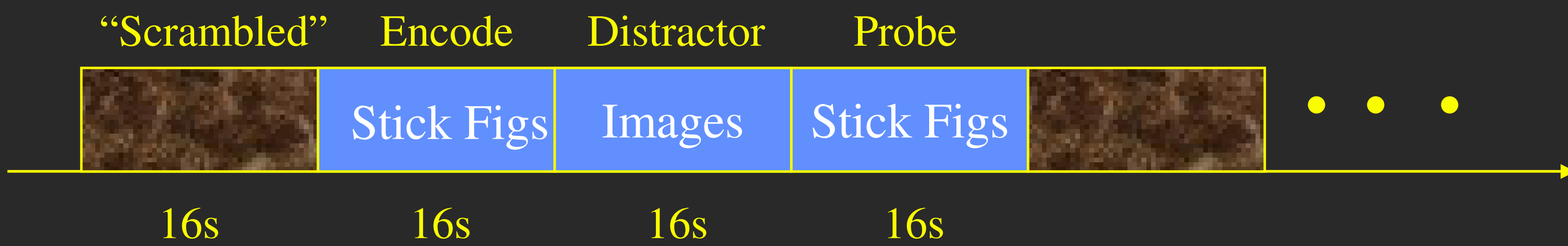


Design Matrix

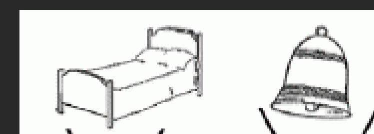
Data from
one voxel

Design Matrix
Regressors

Working Memory Task (fBIRN)



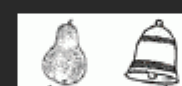
0. "Scrambled" – low-level baseline, no response
1. Encode – series of passively viewed stick figures

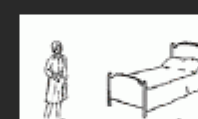


Distractor – respond if there is a face

   2. Emotional
   3. Neutral

Probe – series of two stick figures (forced choice)



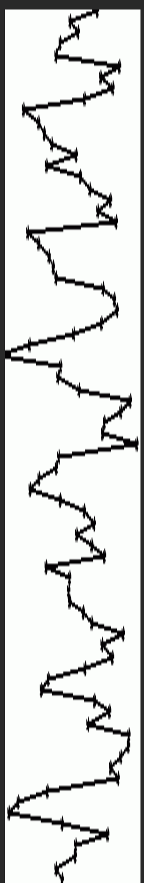
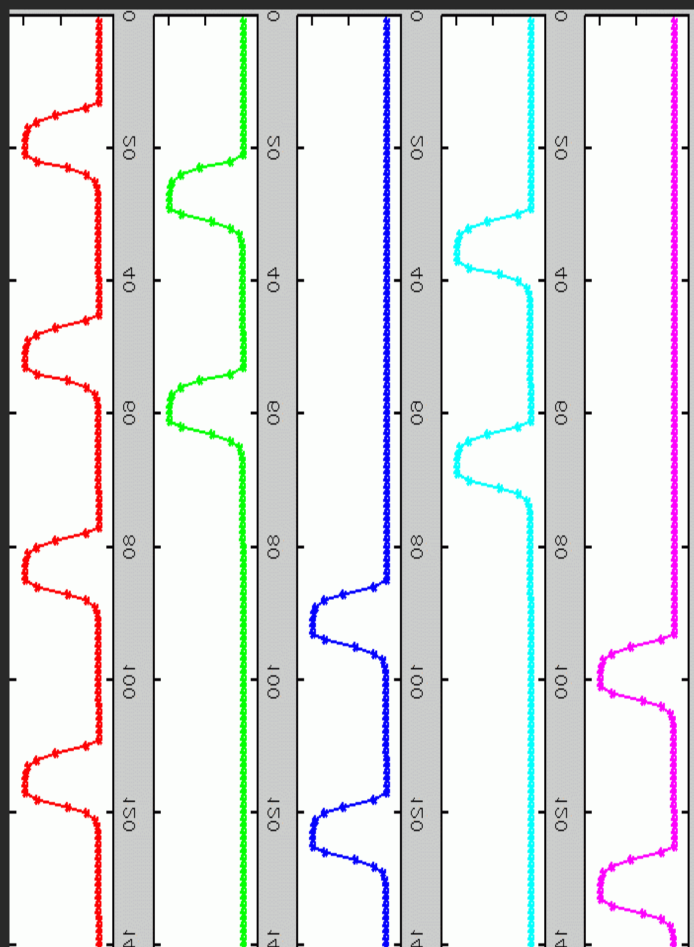


   4. Following Emotional Distractor
   5. Following Neutral Distractor

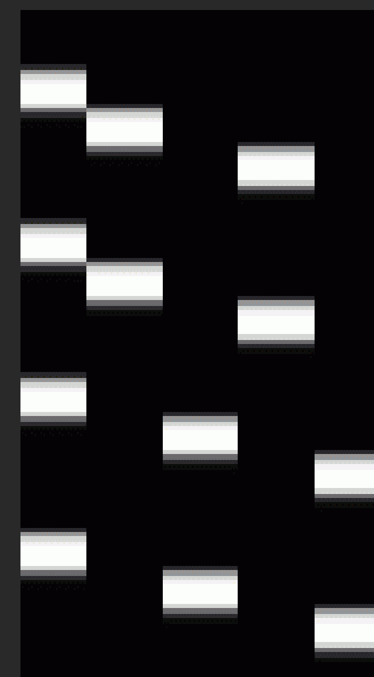
fBIRN: Functional Biomedical Research Network (www.nbirn.net)

Five Task Conditions

$$y = X * \beta$$

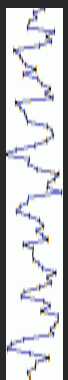
 = 

$$\begin{pmatrix} \beta_{Encode} \\ \beta_{EmotDist} \\ \beta_{NeutDist} \\ \beta_{EmotProbe} \\ \beta_{NeutProbe} \end{pmatrix}$$

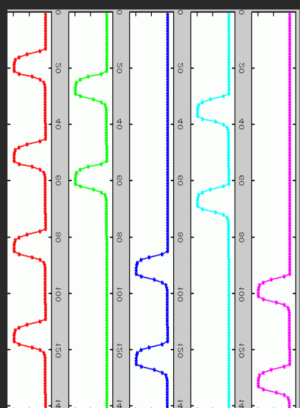


GLM Solution

$y = X * \beta$



$=$

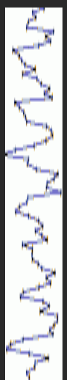


$$\begin{bmatrix} \beta_{Encode} \\ \beta_{EmotDist} \\ \beta_{NeutDist} \\ \beta_{EmotProbe} \\ \beta_{NeutProbe} \end{bmatrix}$$

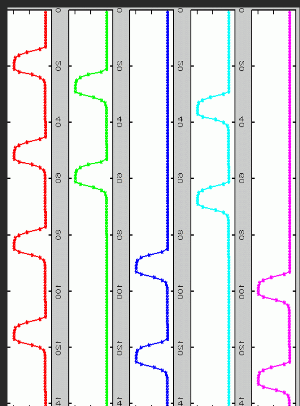
- Set of simultaneous equations
- Each row of X is an equation
- Each column of X is an unknown
- βs are unknown
- 142 Time Points (Equations)
- 5 unknowns

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

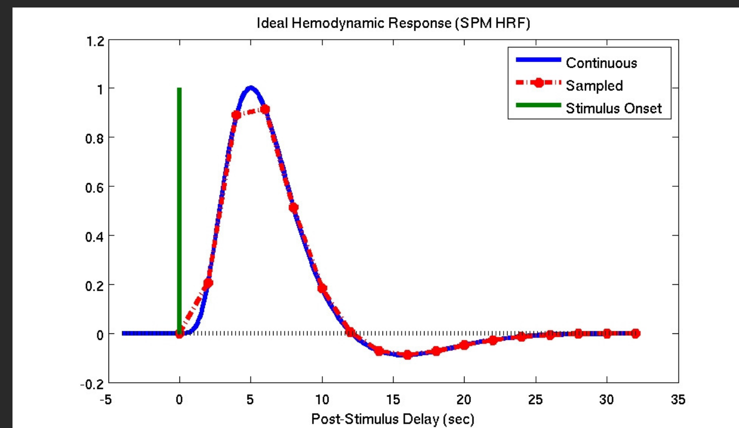
Estimates of the HRF Amplitude



$=$



$$\begin{pmatrix} \beta_{Encode} \\ \beta_{EmotDist} \\ \beta_{NeutDist} \\ \beta_{EmotProbe} \\ \beta_{NeutProbe} \end{pmatrix}$$



$$y = X\beta + n, \quad \hat{\beta} = (X^T X)^{-1} X^T y$$

$\hat{\beta}_{Encode}$ = Hemodynamic amplitude in response to Encode

$\hat{\beta}_{EmotDist}$ = Hemodynamic amplitude in response to Emotional Distractor

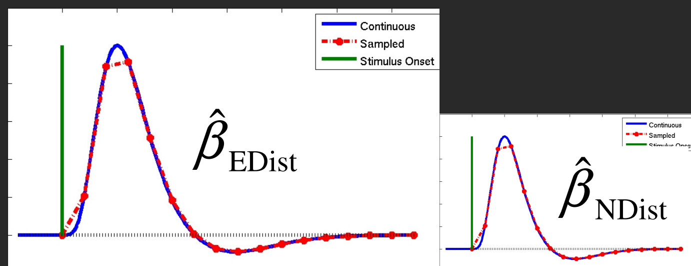
$\hat{\beta}_{NeutDist}$ = Hemodynamic amplitude in response to Neutral Distractor

$\hat{\beta}_{EmotProbe}$ = Hemodynamic amplitude in response to Probe following Emotional Distractor

$\hat{\beta}_{NeuttProbe}$ = Hemodynamic amplitude in response to Probe following Neutral Distractor

Hypotheses and Contrasts

Which voxels respond more/less/differently to the Emotional Distractor than to the Neutral Distractor?



$$\hat{\beta}_{\text{EDist}} > \hat{\beta}_{\text{NDist}}$$

$$\hat{\beta}_{\text{EDist}} < \hat{\beta}_{\text{NDist}}$$

$$\hat{\beta}_{\text{EDist}} \neq \hat{\beta}_{\text{NDist}}$$

$$\gamma = \hat{\beta}_{\text{EDist}} - \hat{\beta}_{\text{NDist}} > 0, < 0, \neq 0$$

Contrast: Assign Weights to each Beta

$$\gamma = c_{\text{Encode}}\,\hat{\beta}_{\text{Encode}} + c_{\text{EDist}}\,\hat{\beta}_{\text{EDist}} + c_{\text{NDist}}\,\hat{\beta}_{\text{NDist}} + c_{\text{EProbe}}\,\hat{\beta}_{\text{EProbe}} + c_{\text{NProbe}}\,\hat{\beta}_{\text{NProbe}}$$

$$c_{\text{Encode}} = 0$$

$$c_{\text{EDist}} = +1$$

$$c_{\text{NDist}} = -1$$

$$c_{\text{EProbe}} = 0$$

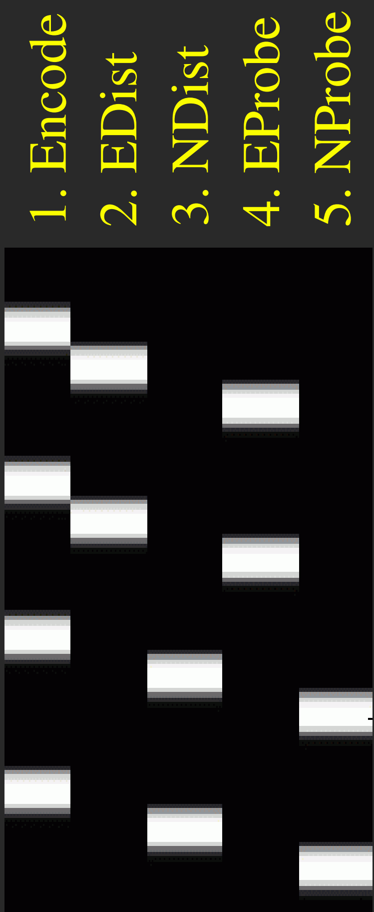
$$c_{\text{NProbe}} = 0$$

$$C = [0 \quad +1 \quad -1 \quad 0 \quad 0]\ \text{Contrast Matirx}$$

Hypotheses

- Which voxels respond more to the Emotional Distractor than to the Neutral Distractor?
- Which voxels respond to Encode (relative to baseline)?
- Which voxels respond to the Emotional Distractor?
- Which voxels respond to either Distractor?
- Which voxels respond more to the Probe following the Emotional Distractor than to the Probe following the Neutral Distractor?

Which voxels respond more to the Emotional Distractor than to the Neutral Distractor?



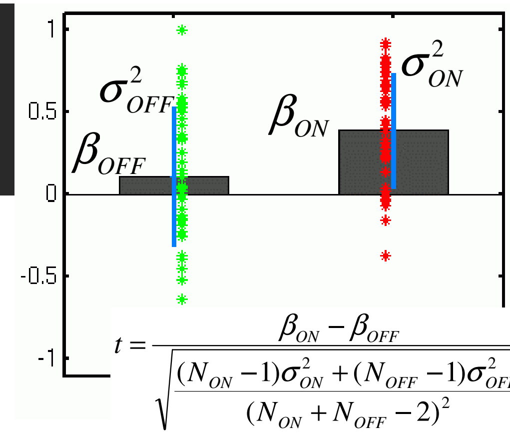
- Only interested in Emotional and Neutral Distractors
- No statement about other conditions

Condition:	1	2	3	4	5
Weight	0	+1	-1	0	0

Contrast Matrix

C = [0 +1 -1 0 0]

Contrasts and the Full Model



$$y = X\beta + n, \quad y = s + n, \quad n \sim N(0, \sigma_n^2)$$

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad \text{Parameter Estimates}$$

$$\hat{\sigma}_n^2 = \frac{\hat{n}^T \hat{n}}{DOF} \quad \text{Residual Variance,} \quad \hat{n} = y - X\hat{\beta}$$

$$\hat{\gamma} = C\hat{\beta} \quad \text{Contrast}$$

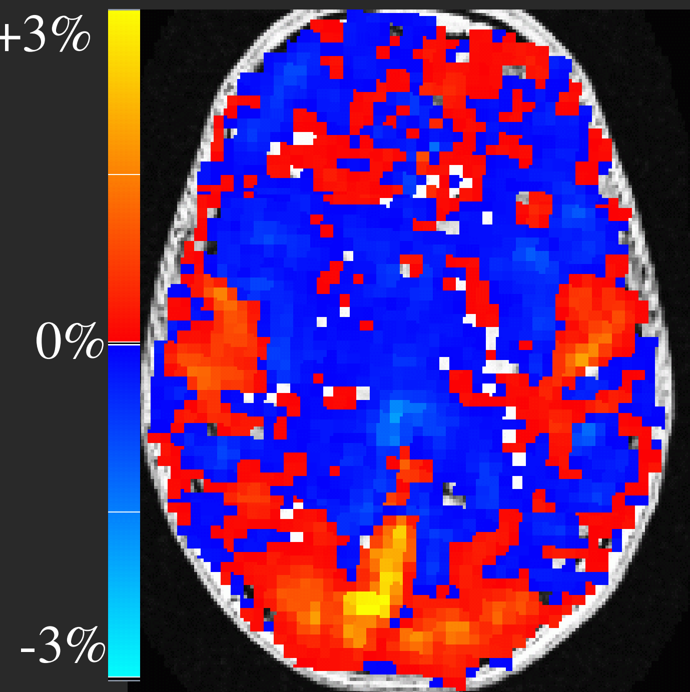
$$\hat{\sigma}_\gamma^2 = \hat{\Sigma}_\gamma = \frac{1}{J}\left(C(X^T X)^{-1} C^T\right)\hat{\sigma}_n^2 \quad \text{Contrast Variance Estimate}$$

J = Rows in C

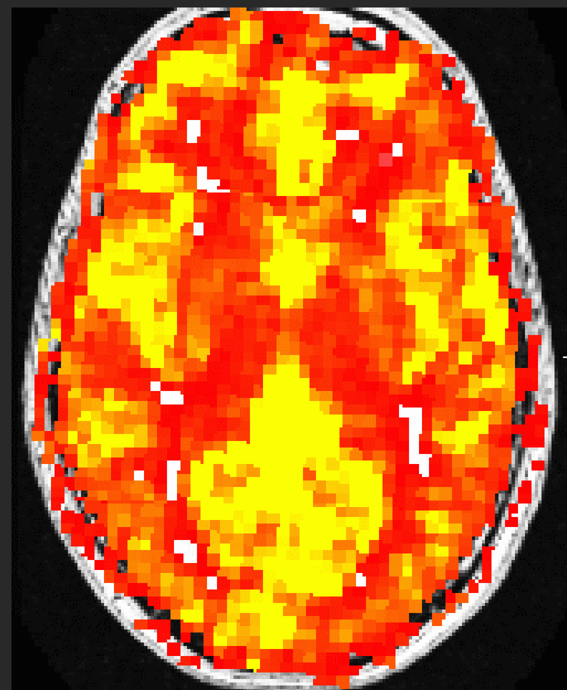
$$\text{t}_{\text{DOF}} = \frac{\hat{\gamma}}{\hat{\sigma}_\gamma} = \frac{C\hat{\beta}}{\sqrt{\left(C(X^T X)^{-1} C^T\right)\hat{\sigma}_n^2}} \quad \text{t - Test (univariate)}$$

$$\text{F}_{\text{DOF, J}} = \hat{\gamma}^T \hat{\Sigma}_\gamma^{-1} \hat{\gamma} \quad \text{F - Test (multivariate)}$$

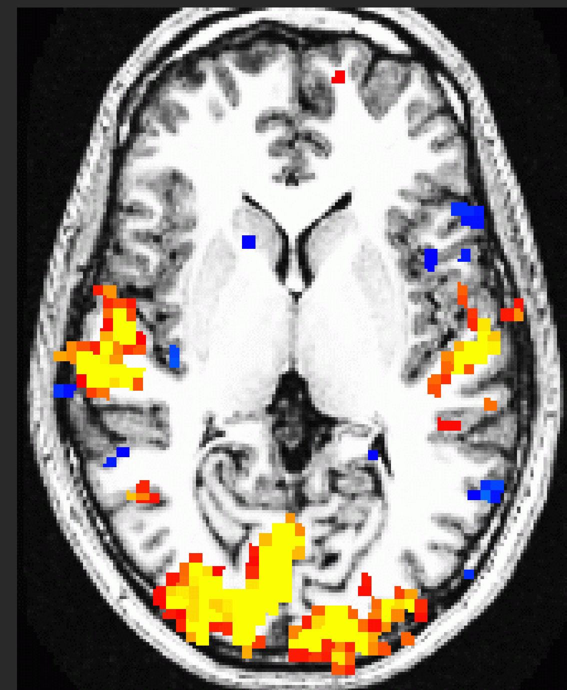
First Level GLM Outputs



Contrast Amplitude
CON, COPE, CES



Contrast Amplitude
Variance
(Error Bars)
VARCOPE, CESVAR



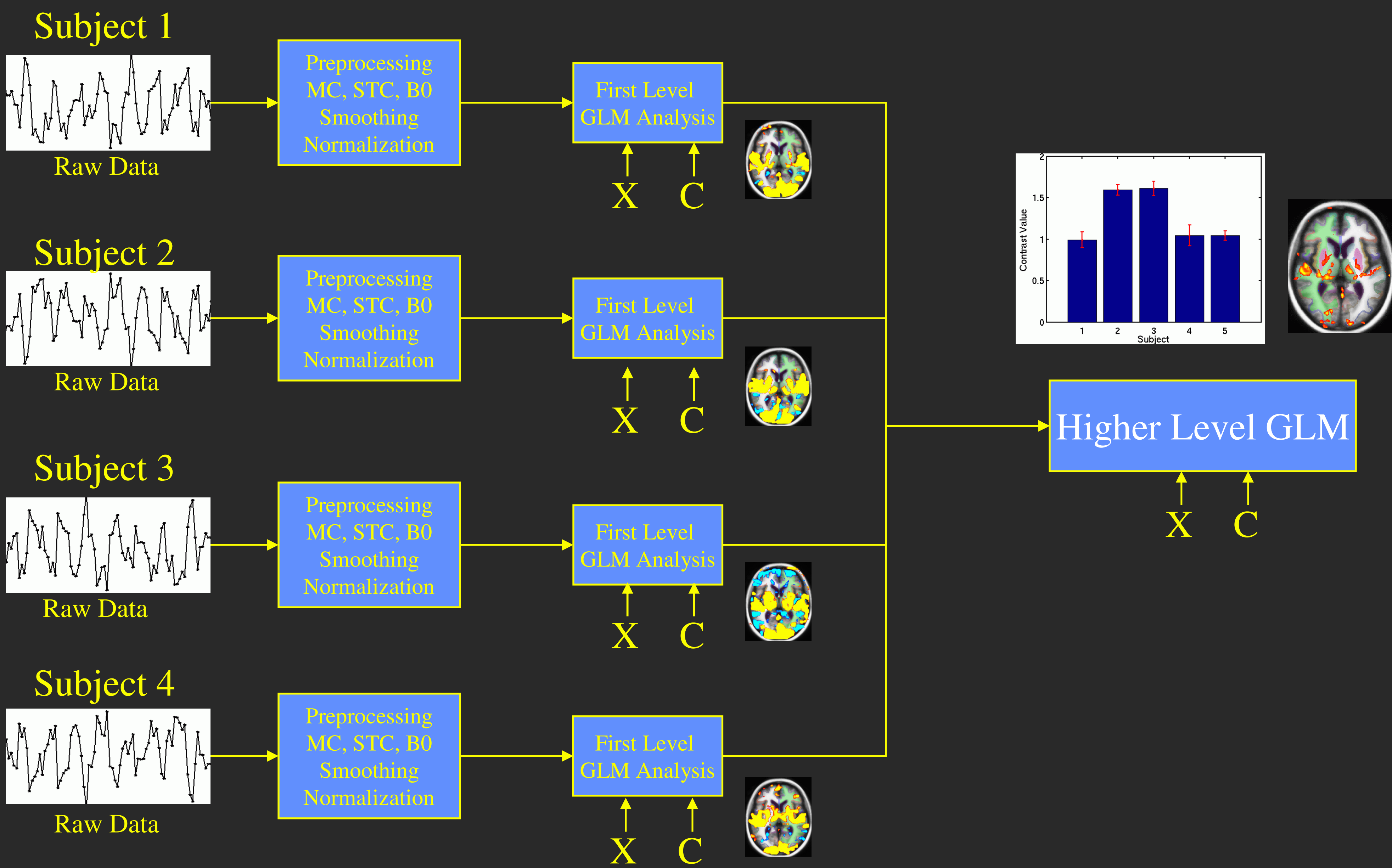
Significance
t-Map (p,z,F)
(Thresholded p<.01)
sig=-log10(p)

Time Series Analysis Summary

- Correlational
- Design Matrix (HRF shape)
- Estimate HRF amplitude (Parameters)
- Contrasts to test hypotheses
- Results at each voxel:
    - Contrast Value
    - Contrast Value Variance
    - p-value (Volume of Activation)
- Pass Contrast Value and Variance up to higher level analyses

fMRI Group Analysis

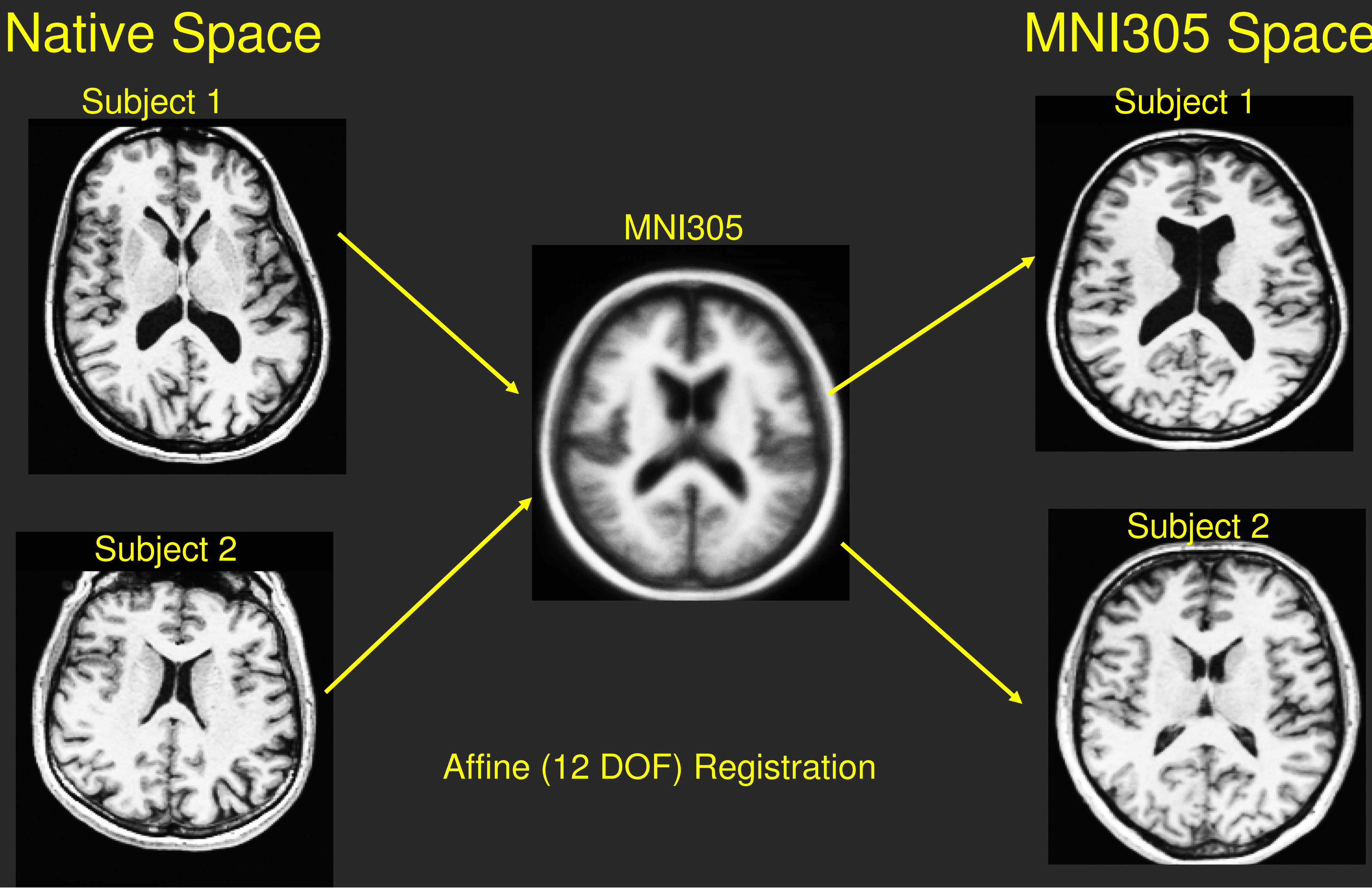
fMRI Analysis Overview



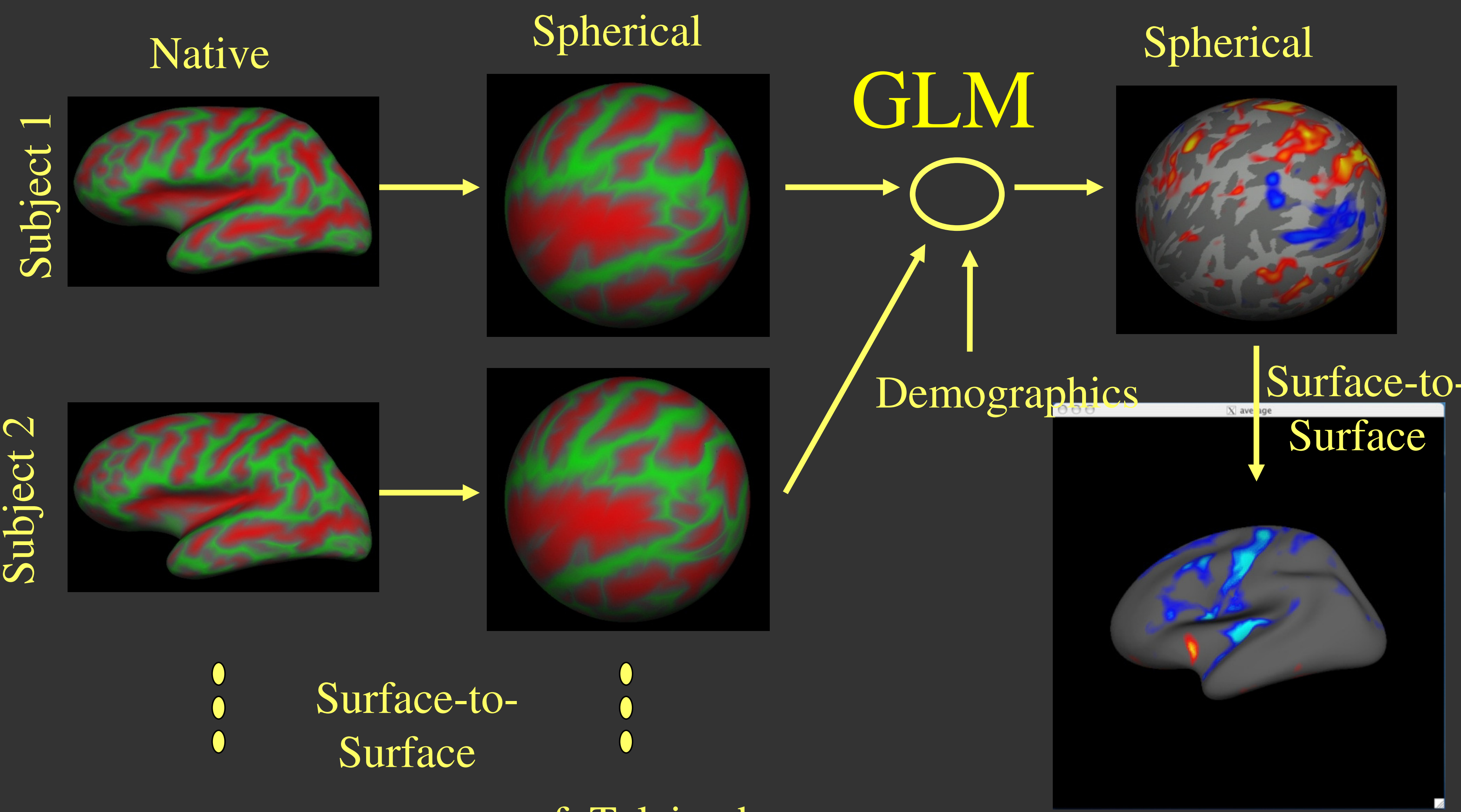
Overview

- Goal of Group Analysis
- Types of Group Analysis
  - Random Effects, Mixed Effects, Fixed Effects
- Multi-Level General Linear Model (GLM)

Spatial Normalization, Atlas Space



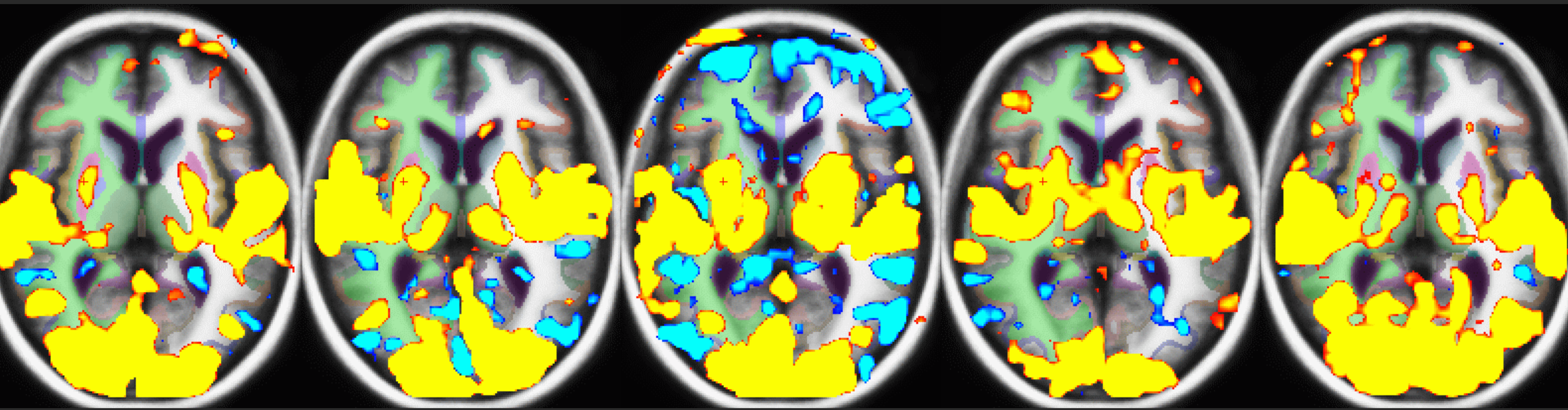
Inter-Subject Averaging



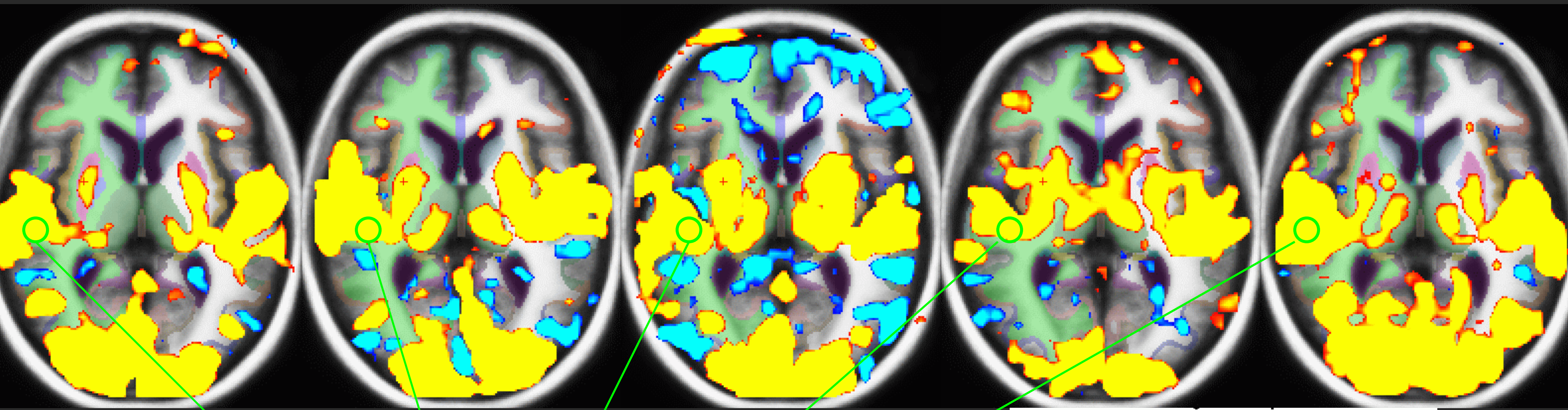
cf. Talairach

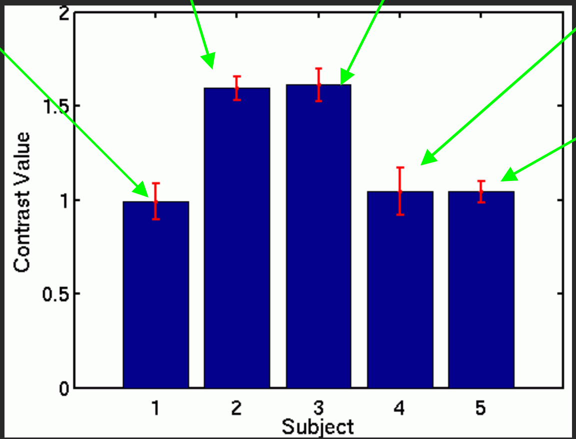
Is Pattern Repeatable Across Subject?

Subject 1 Subject 2 Subject 3 Subject 4 Subject 5

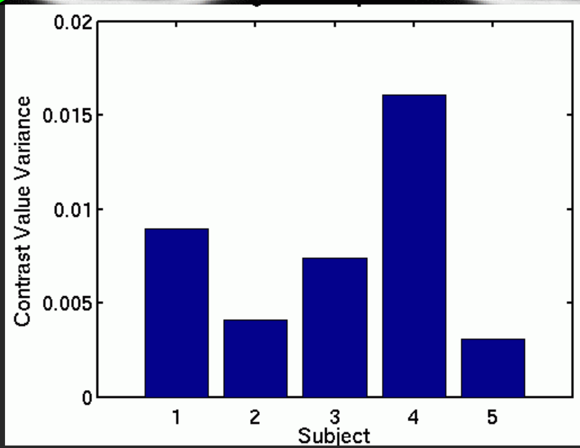


Group Analysis





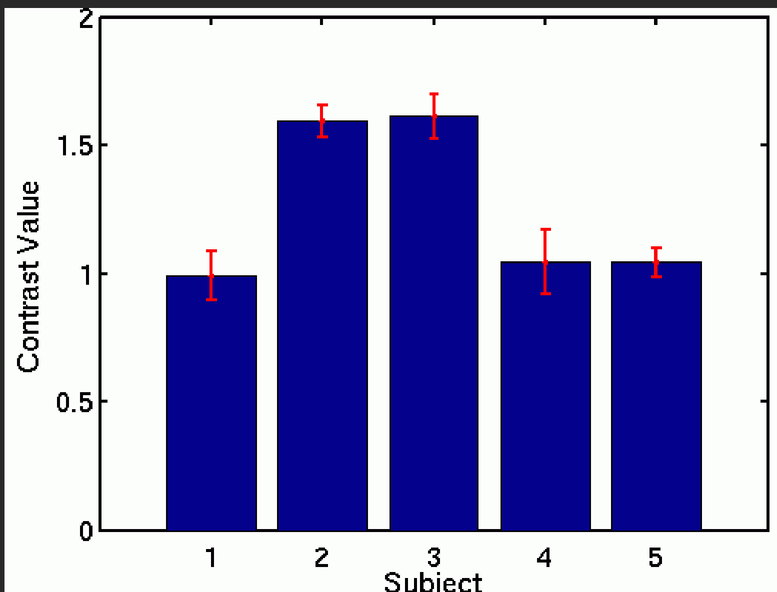
Contrast Amplitudes

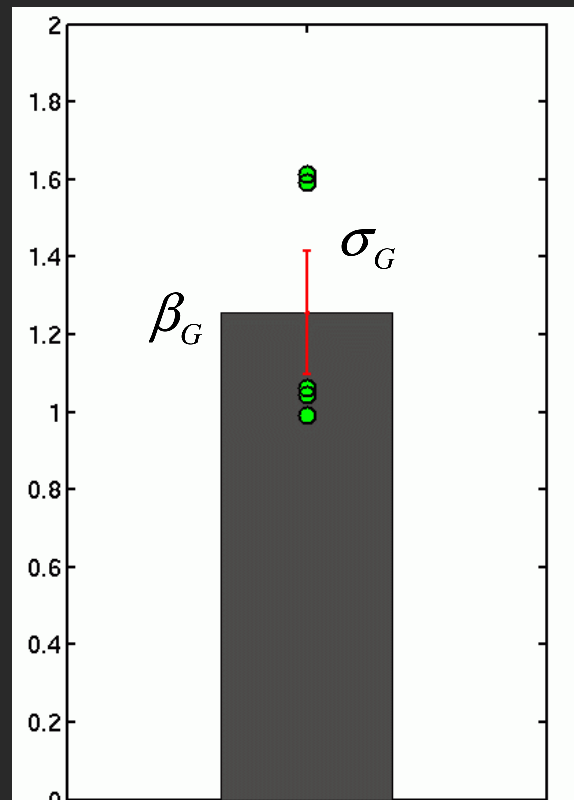


Contrast Amplitudes Variances (Error Bars)

Does not have to be all positive!

“Random Effects (RFx)” Analysis

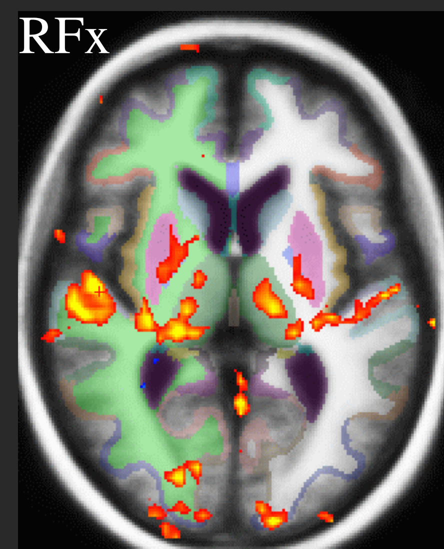


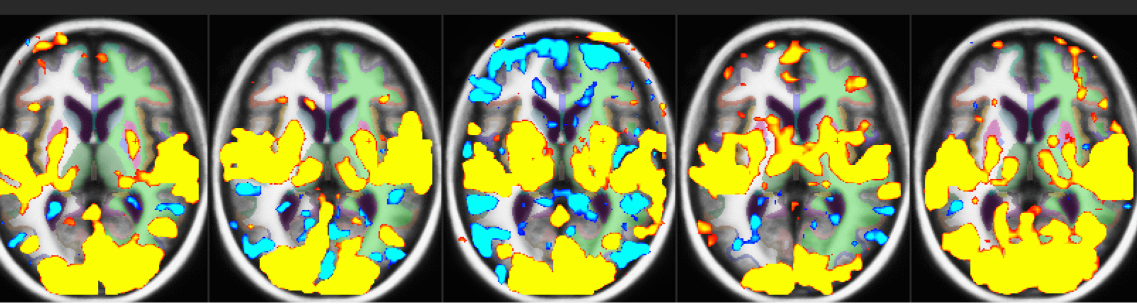


$$t = \frac{\beta_G}{\sigma_{\beta_G}}, \sigma^2_{\beta_G} = \frac{\sigma^2_G}{N_G}$$

$$\sigma^2_G = \frac{\sum (\beta_G - \beta_i)^2}{(N_G - 1)}$$

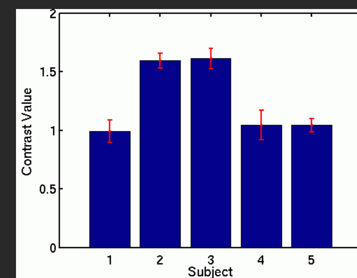
$$DOF = N_G - 1$$

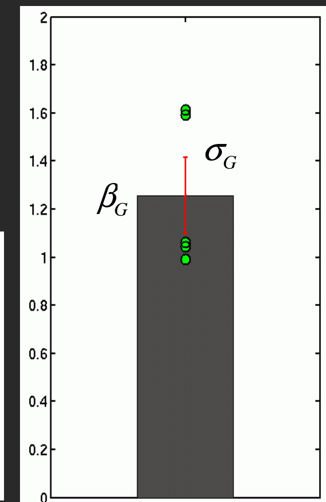


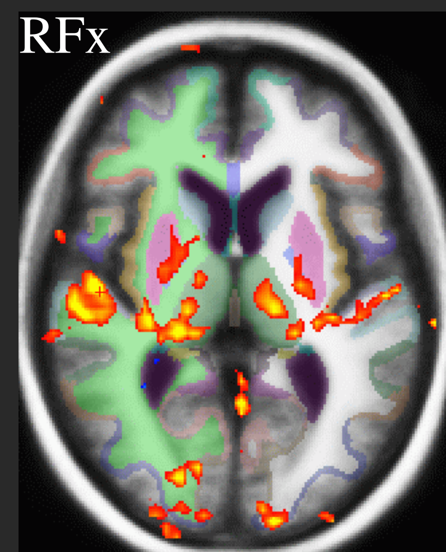


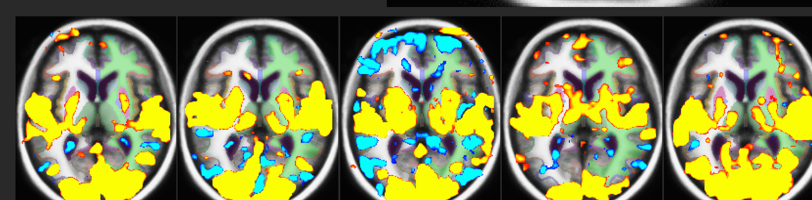
“Random Effects (RFx)” Analysis

- Model Subjects as a Random Effect
- Variance comes from a single source: variance across subjects
  - Mean at the <u>population</u> mean
  - Variance of the <u>population</u> variance
- Does not take first-level noise into account (assumes 0)
- “Ordinary” Least Squares (OLS)
- Usually less activation than individuals

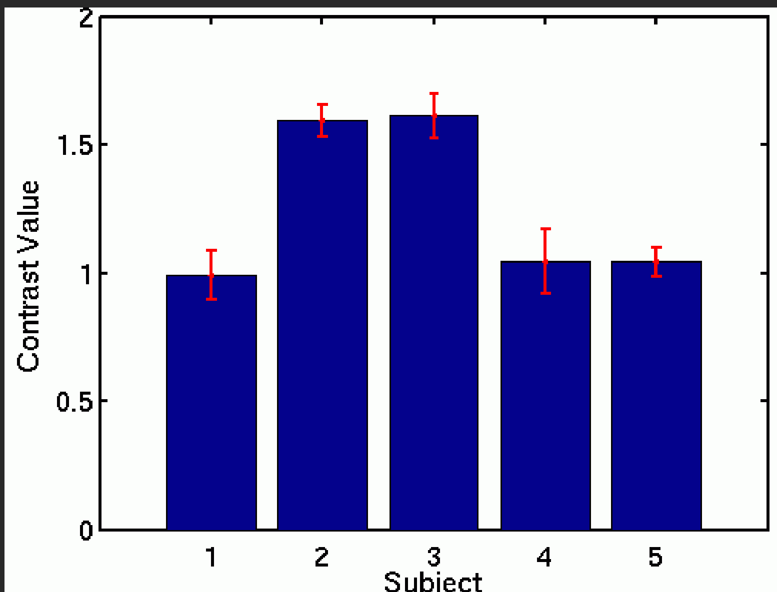


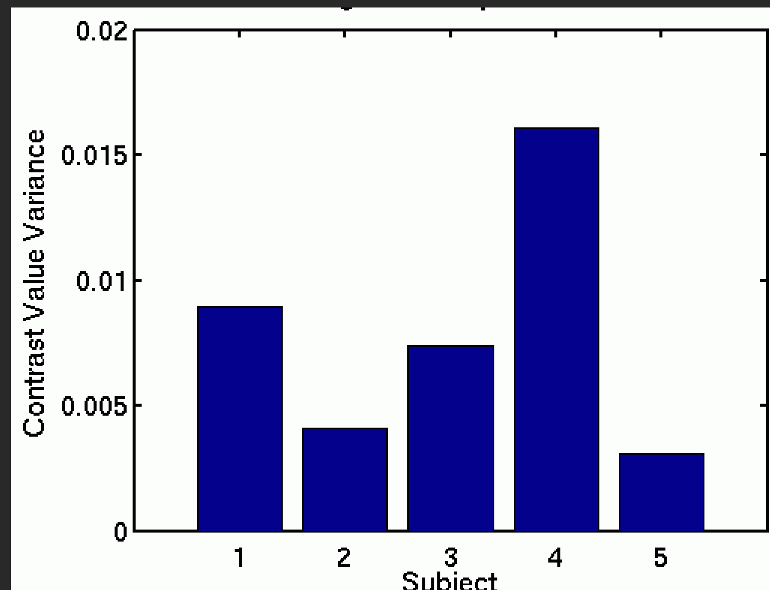


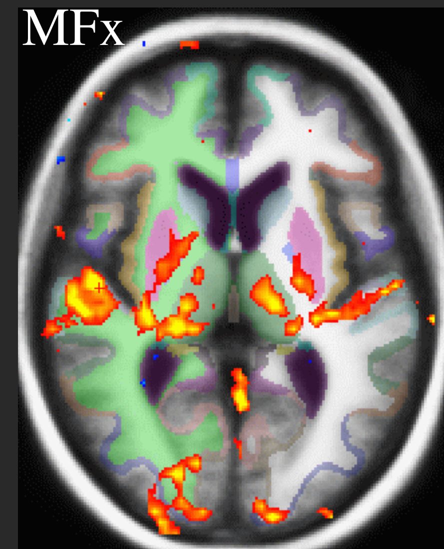


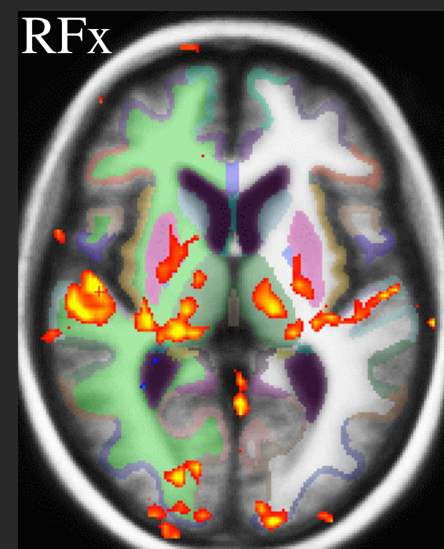


“Mixed Effects (MFx)” Analysis

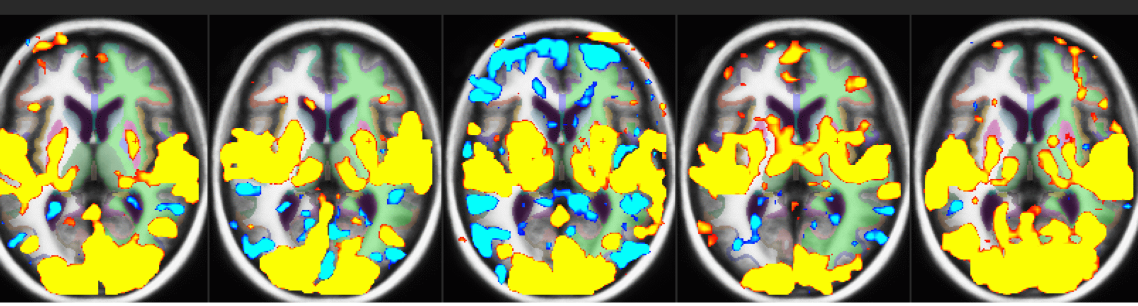






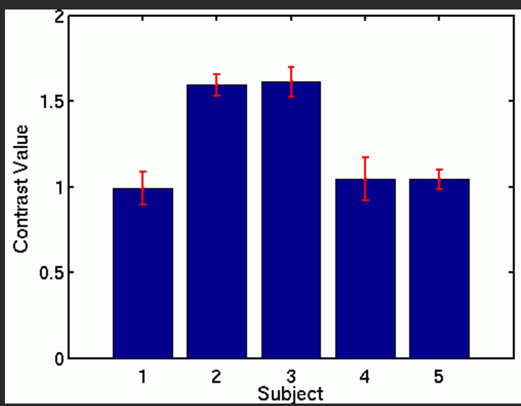


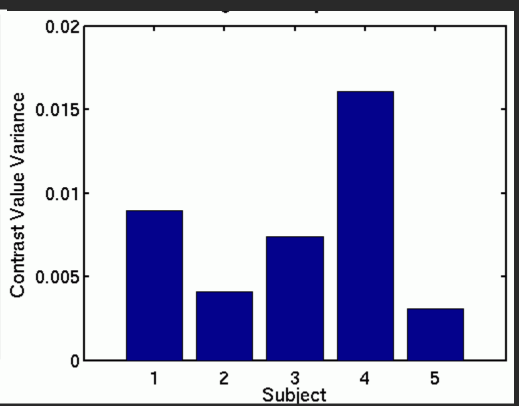
- Down-weight each subject based on variance.
- Weighted Least Squares vs (“Ordinary” LS)

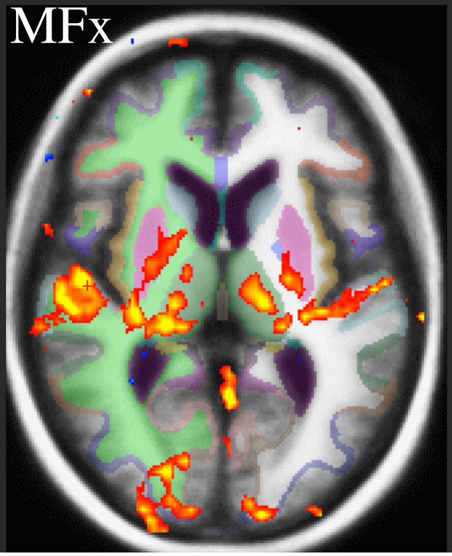


"Mixed Effects (MFx)" Analysis

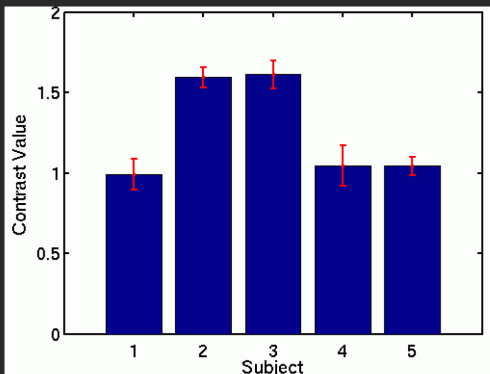
- Down-weight each subject based on variance.
- Weighted Least Squares vs ("Ordinary" LS)
- Protects against unequal variances across group or groups ("heteroskedasticity")
- May increase or decrease significance with respect to simple Random Effects
- More complicated to compute
- "Pseudo-MFx" – simply weight by first-level variance (easier to compute)

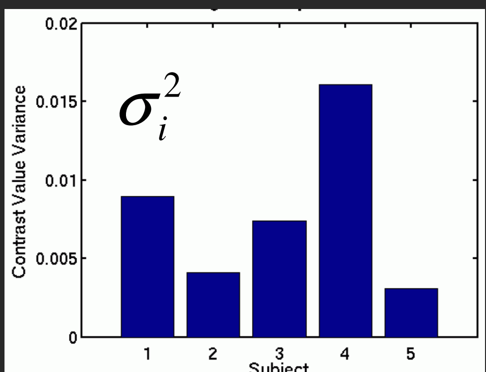






"Fixed Effects (FFx)" Analysis

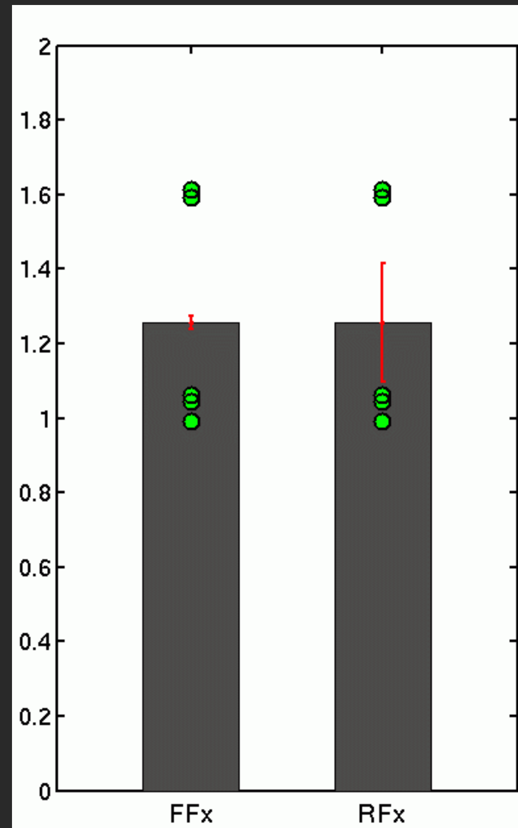


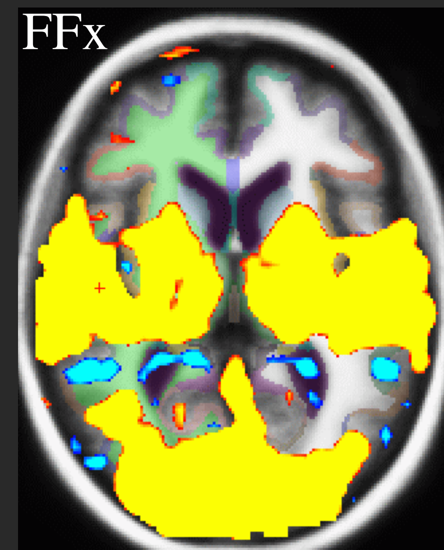


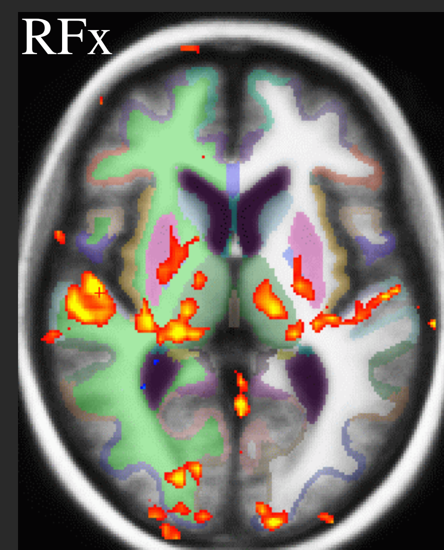
$$t = \frac{\beta_G}{\sqrt{\sigma^2_{\beta_G}}}$$

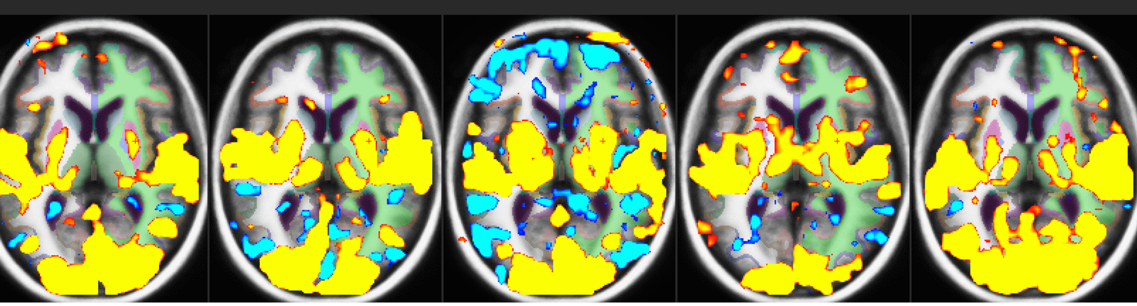
$$\sigma^2_{\beta_G} = \frac{\sum \sigma_i^2}{(N_G)^2}$$

$$DOF = \sum DOF_i$$









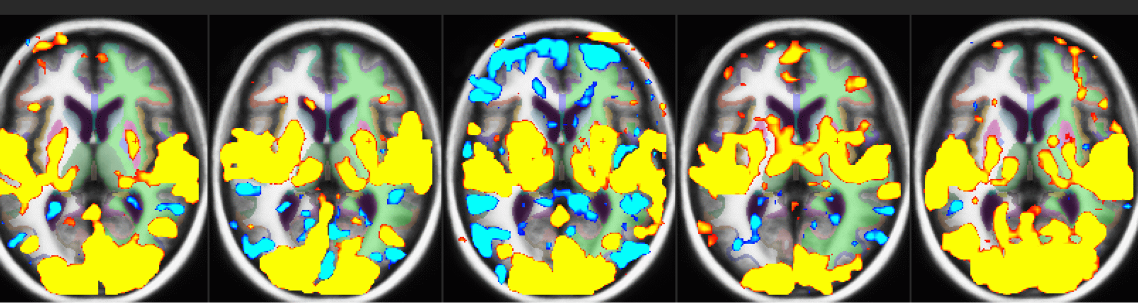
“Fixed Effects (FFx)” Analysis

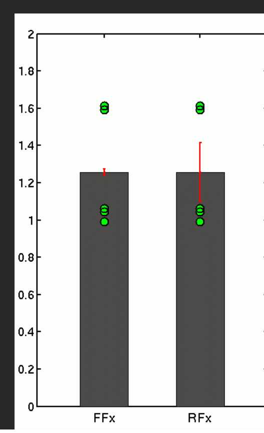
- As if all subjects treated as a single subject (fixed effect)
- Small error bars (with respect to RFx)
- Large DOF
- Same mean as RFx
- Huge areas of activation
- Not generalizable beyond <u>sample</u>.

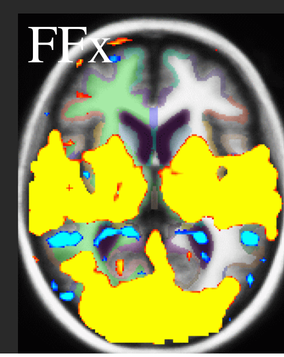
$$t = \frac{\beta_G}{\sqrt{\sigma^2_{\beta_G}}}$$

$$\sigma^2_{\beta_G} = \frac{\sum \sigma_i^2}{(N_G)^2}$$

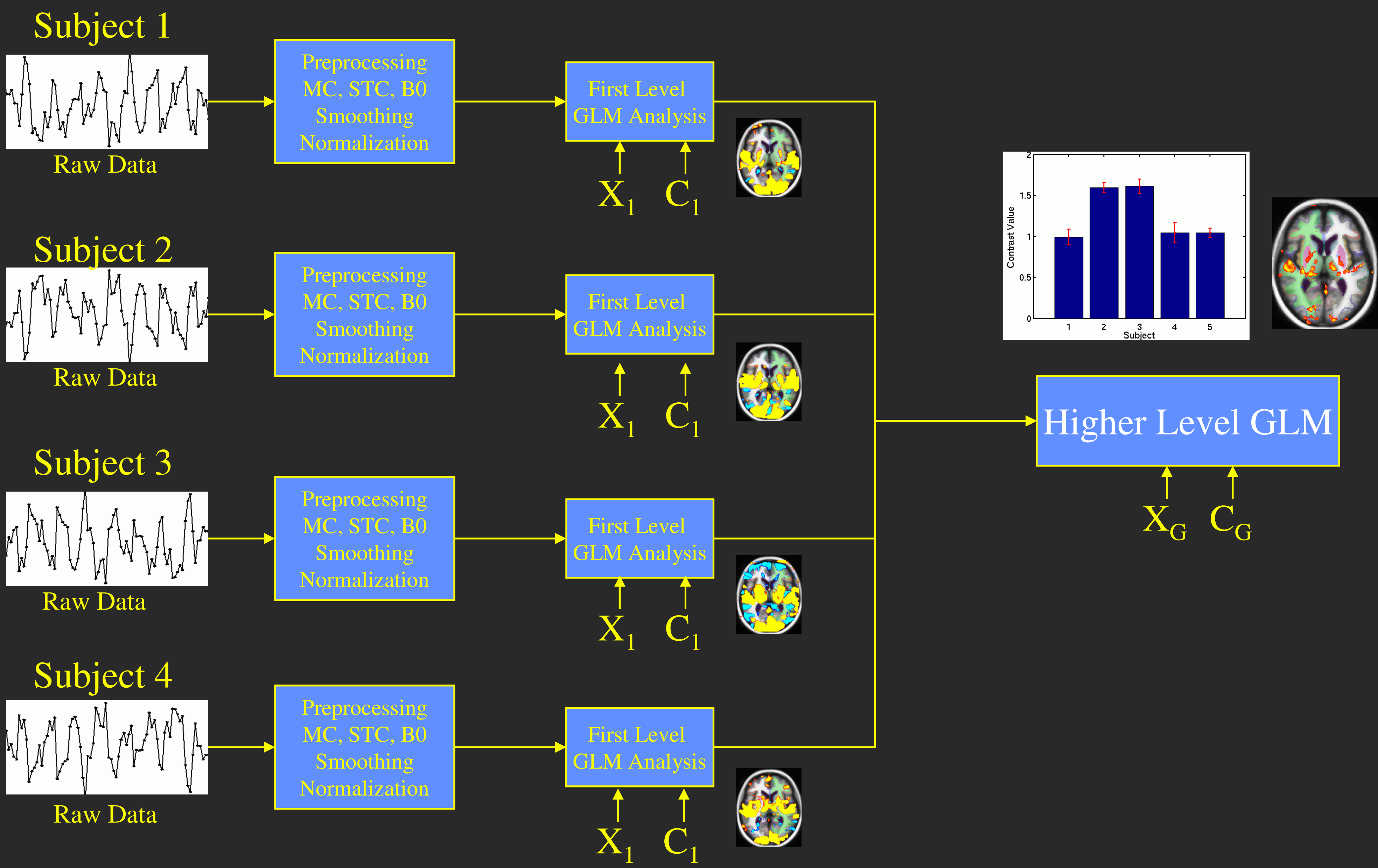
$$DOF = \sum DOF_i$$







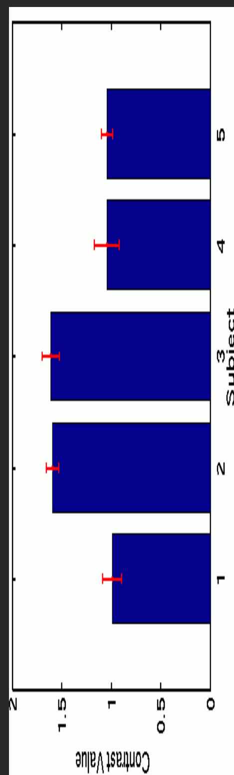
fMRI Analysis Overview



Higher Level GLM Analysis

$$y \quad = \quad X * \beta$$

Observations
(Low-Level Contrasts)

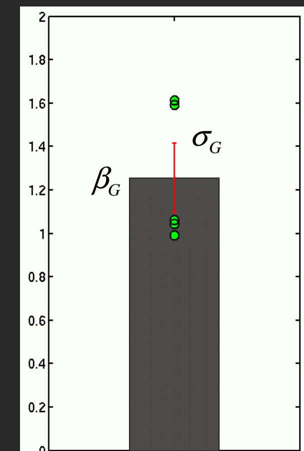


Data from
one voxel

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \beta_G \end{bmatrix}$$

Design Matrix
(Regressors)

Vector of
Regression
Coefficients
(“Betas”)



Contrast Matrix:
C = [1]
Contrast = $C*\beta = \beta_G$

Summary

- Preprocessing – MC, STC, B0, Normalize, Smooth
- First Level GLM Analysis – Design matrix, HRF, Nuisance
- Contrasts, Hypothesis Testing – contrast matrix
- Group Analysis
  - Random, Mixed, Fixed
  - Multi-level GLM (Design and Contrast Matrices)