TMBIC 2017 資料分析助理研習營

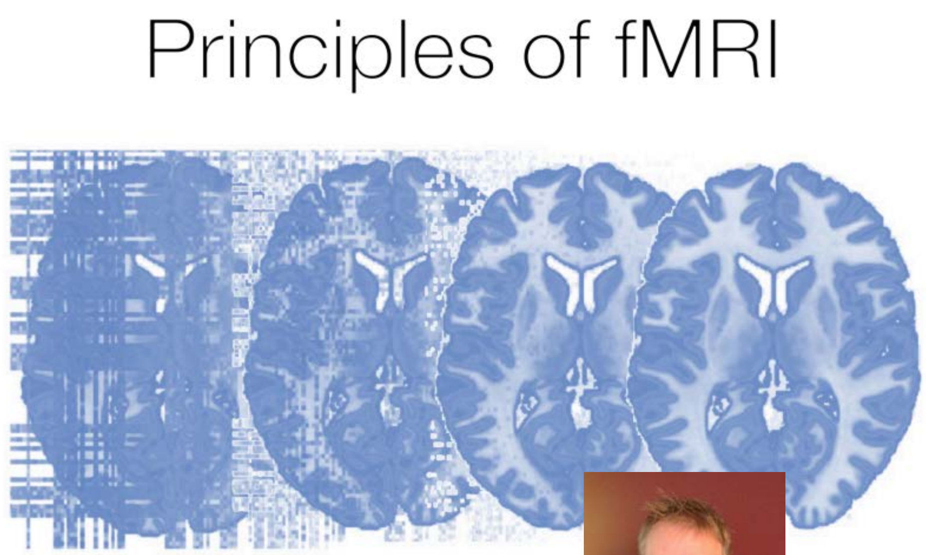
FMRI實驗設計與資料前處理

FMRI Experimental Designs and Data Preprocessing

張智宏副教授　中央大學認知神經科學研究所

認知神經科學研究所
Institute of Cognitive Neuroscience
National Central University, Taiwan

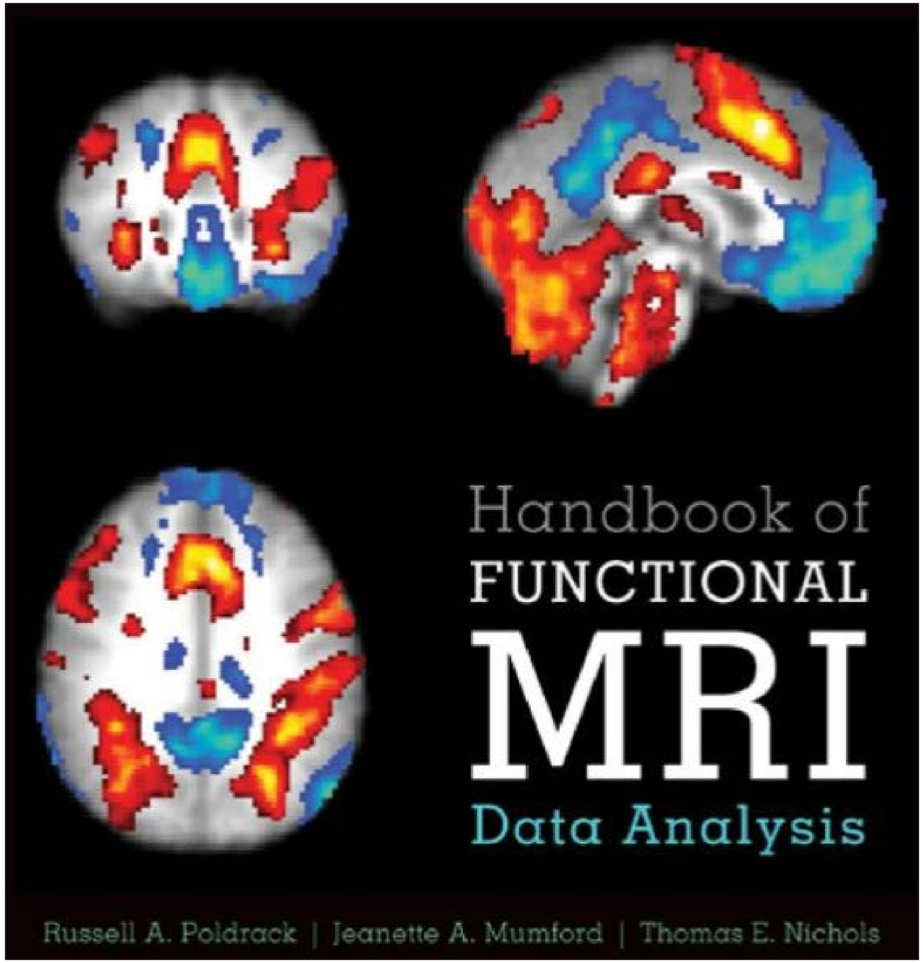
參考書籍



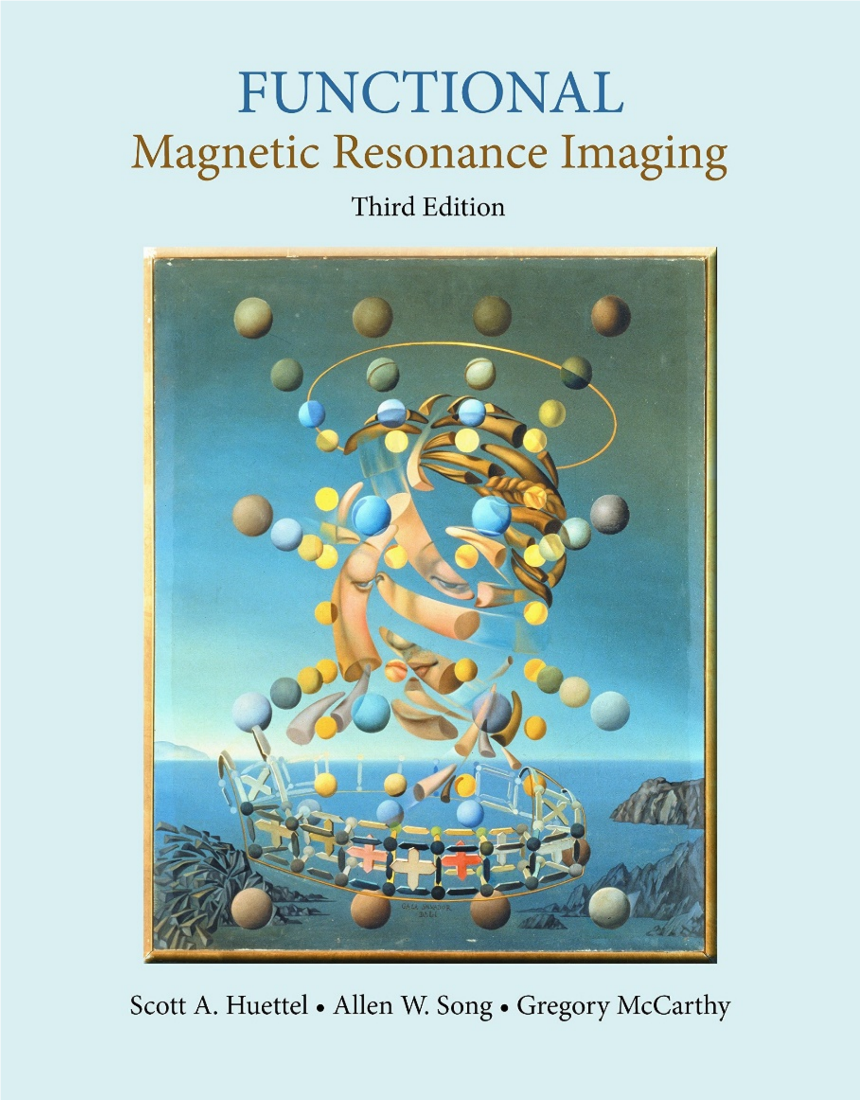
Martin A. Lindquist, Ph.D.

Tor D. Wager, Ph.D.

https://leanpub.com/principlesoffmri

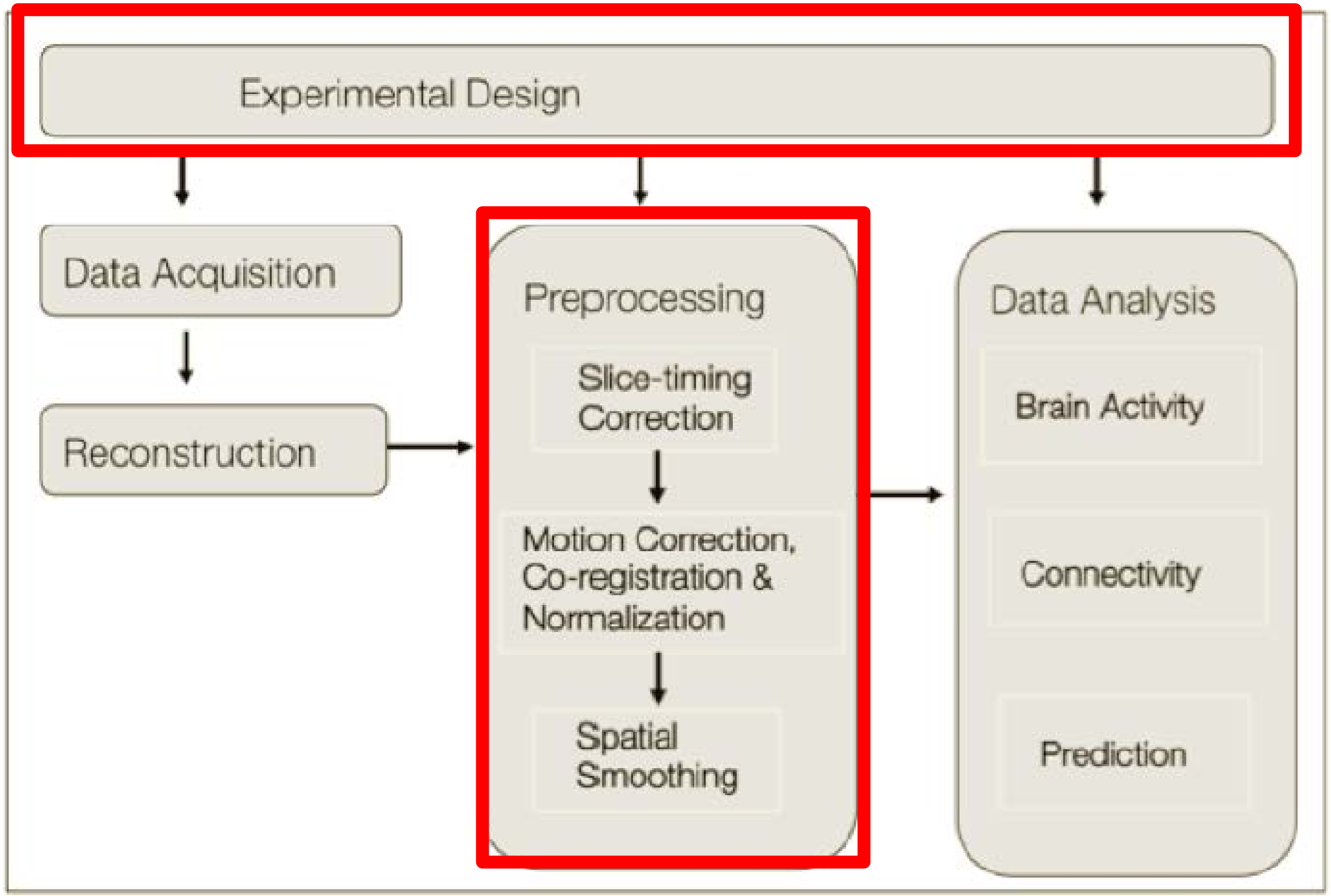


http://www.fmri-data-analysis.org



http://www.sinauer.com/functional-magnetic-resonance-imaging-737.html

The Big Picture

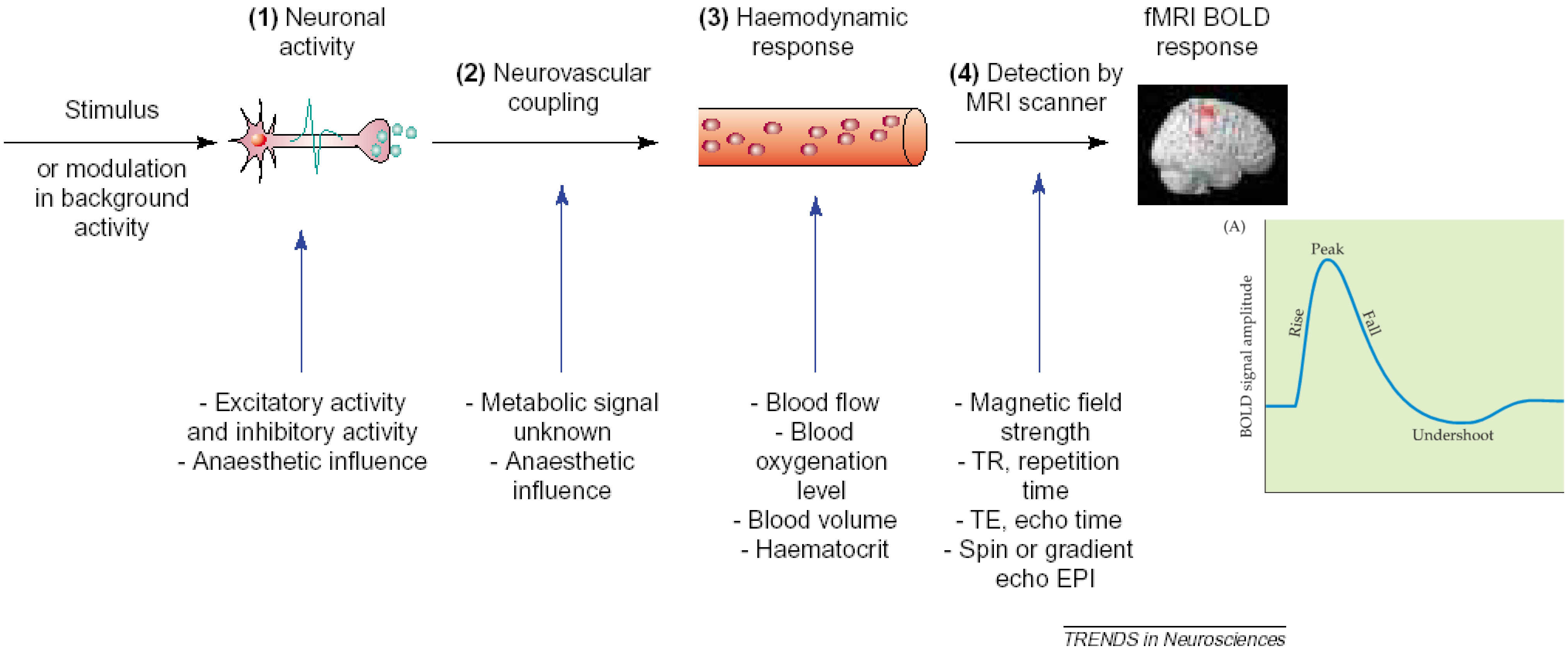


Source: Lindquist & Tor (2015)

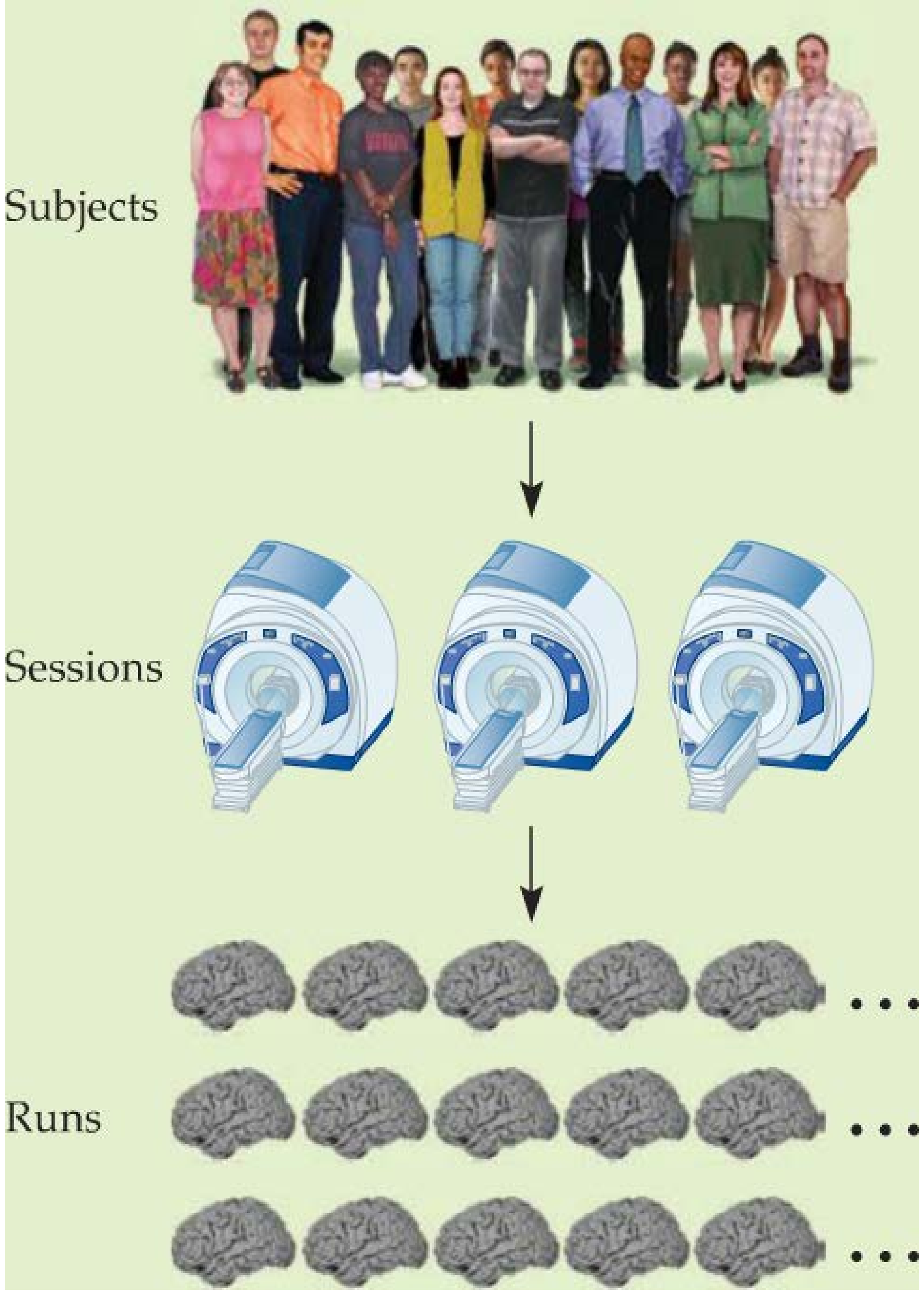
實驗基本概念

- 受控制的觀察
  - 獨變項（Independent Variables; IV）
    - 由研究者操弄
    - 至少兩個水準
    - 例：運動強度、心理壓力程度、圖形顏色、圖形類別
- 依變項(Dependent Variables, DV)
  - 量化指標
  - 可受獨變項影響而改變
  - 例：心跳、壓力荷爾蒙、反應時間、BOLD

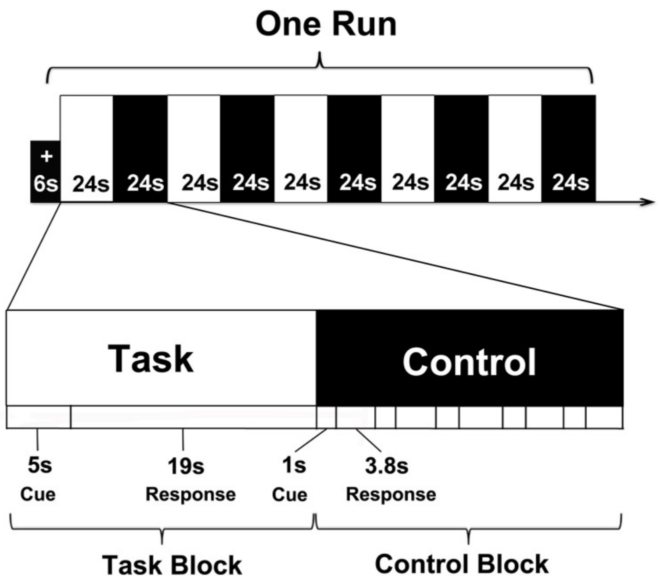
關於FMRI，我們觀察的是...



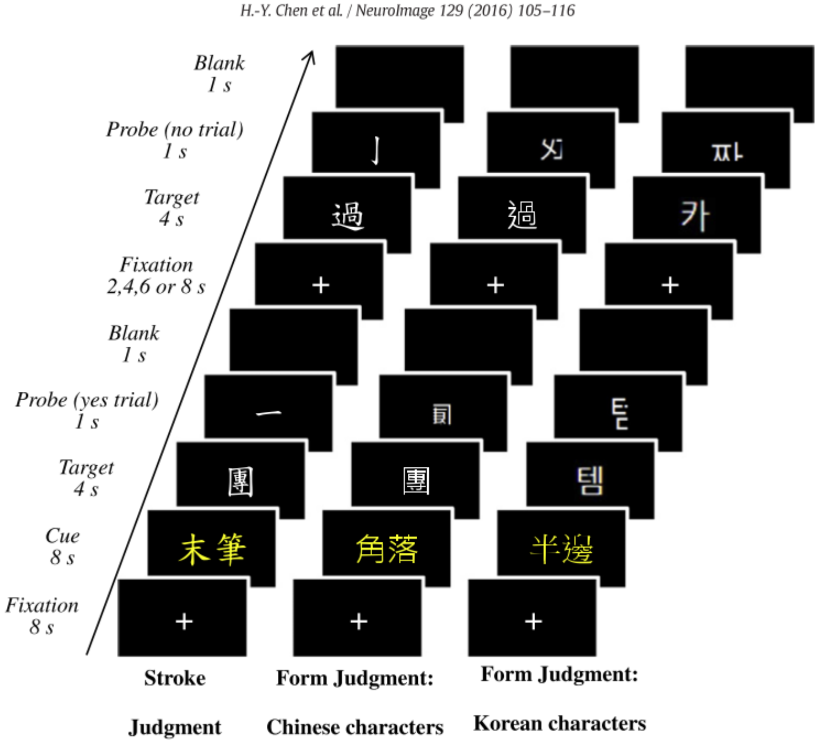
Source: Arthurs & Boniface, 2002, *Trends in Neurosciences*

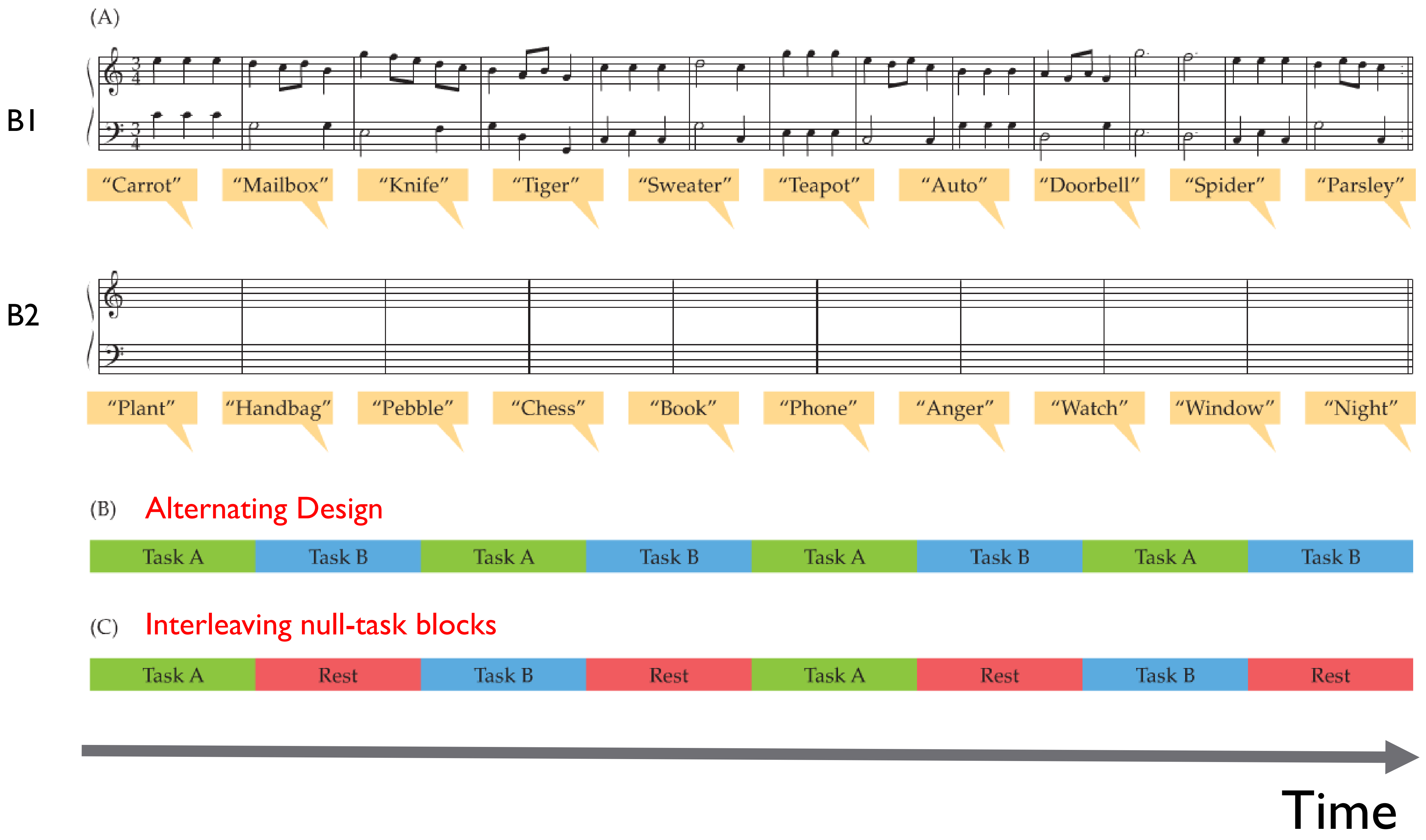


- Blocks
- Conditions

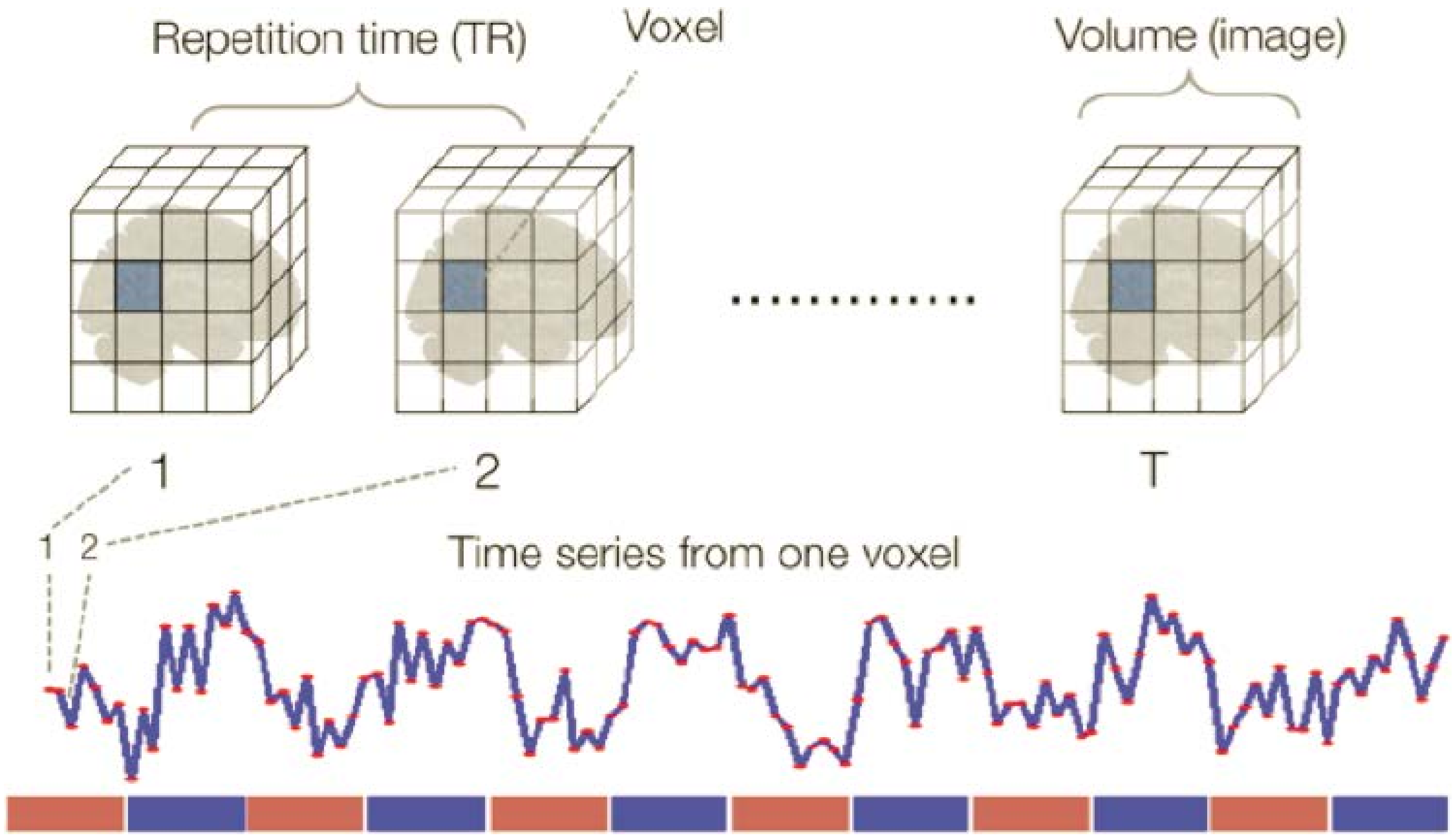


- Trials
- Events




(A)
B1
"Carrot"
"Mailbox"
"Knife"
"Tiger"
"Sweater"
"Teapot"
"Auto"
"Doorbell"
"Spider"
"Parsley"
B2
"Plant"
"Handbag"
"Pebble"
"Chess"
"Book"
"Phone"
"Anger"
"Watch"
"Window"
"Night"
(B) Alternating Design
Task A
Task B
Task A
Task B
Task A
Task B
Task A
Task B
(C) Interleaving null-task blocks
Task A
Rest
Task B
Rest
Task A
Rest
Task B
Rest
Time

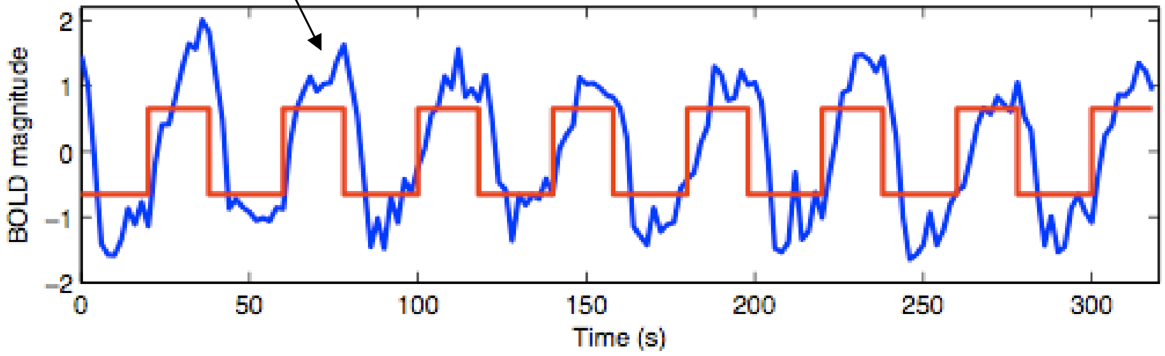
FMRI Time Series



Source: Lindquist & Tor (2015)

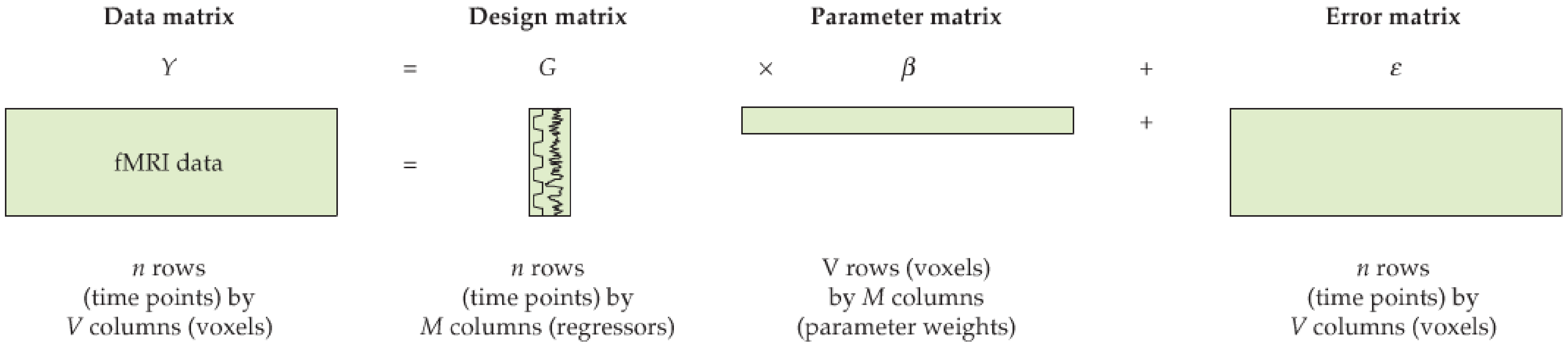
IV and DV from the Perspective of Data Analysis

- Observed BOLD signal (example)



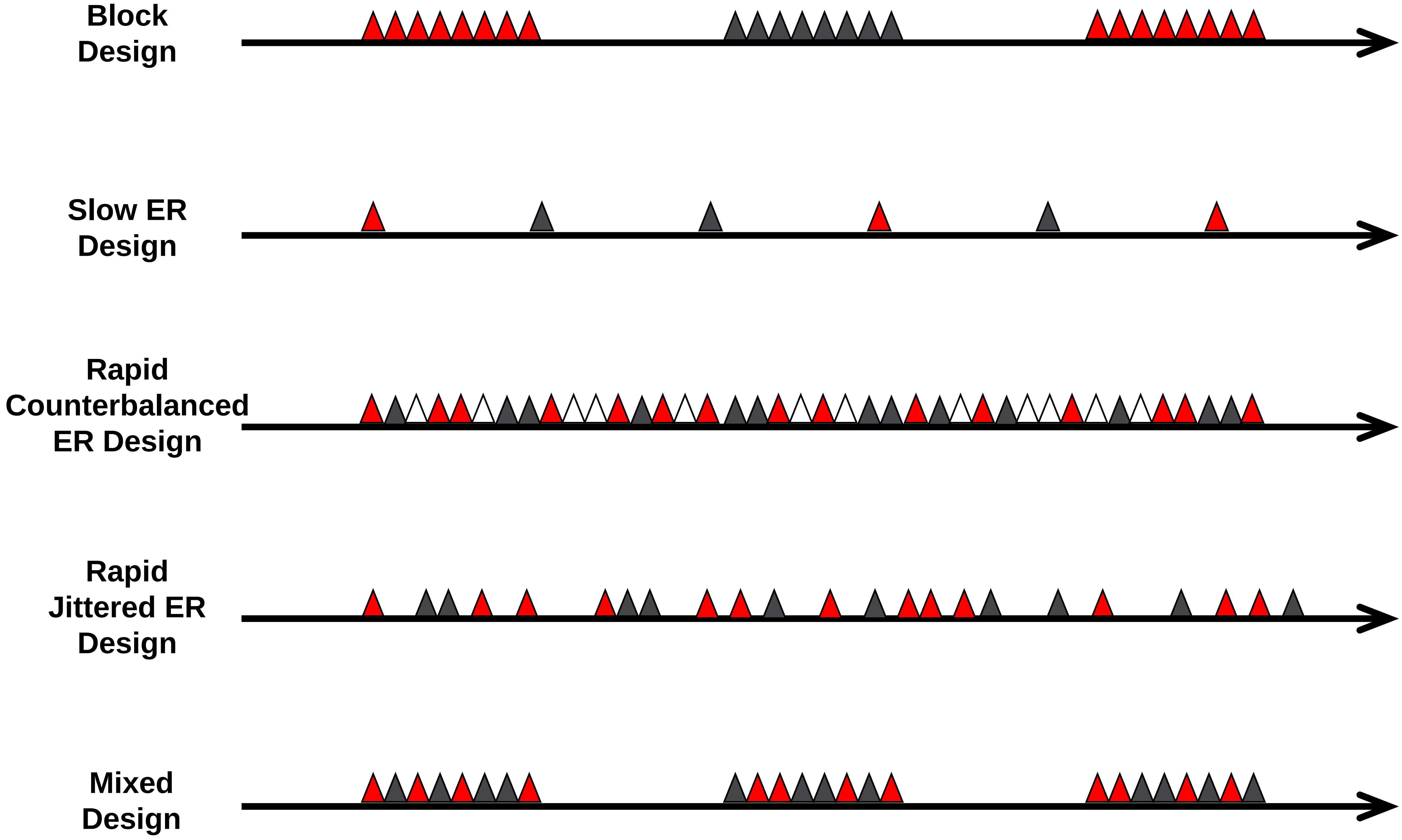
- BOLD signal did not look exactly like the predicted neural activity (in red)

General Linear Model

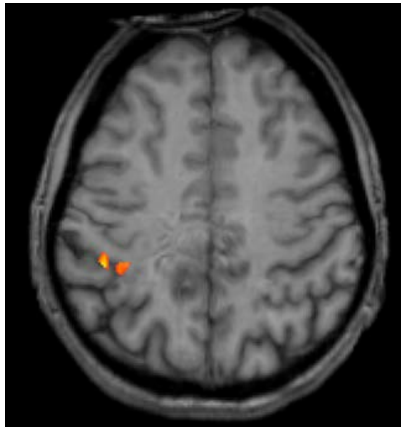


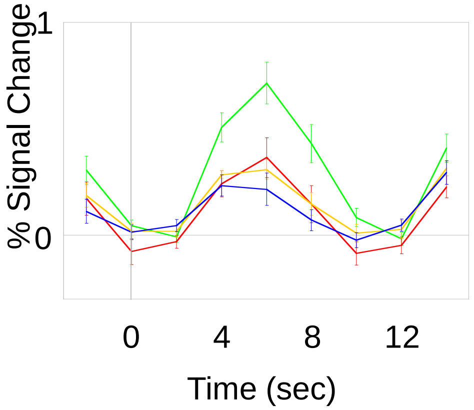
FMRI Experimental Designs

- Blocked designs
- Event-related designs
- Mixed designs


Block Design
Slow ER Design
Rapid Counterbalanced ER Design
Rapid Jittered ER Design
Mixed Design

Detection vs. Estimation

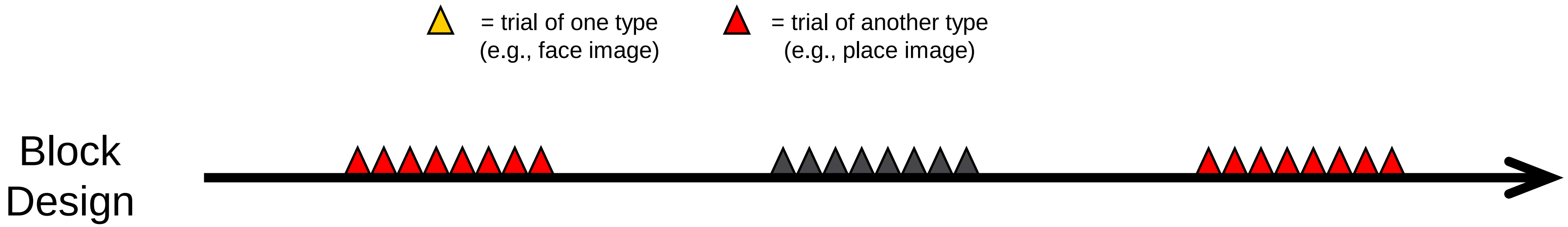




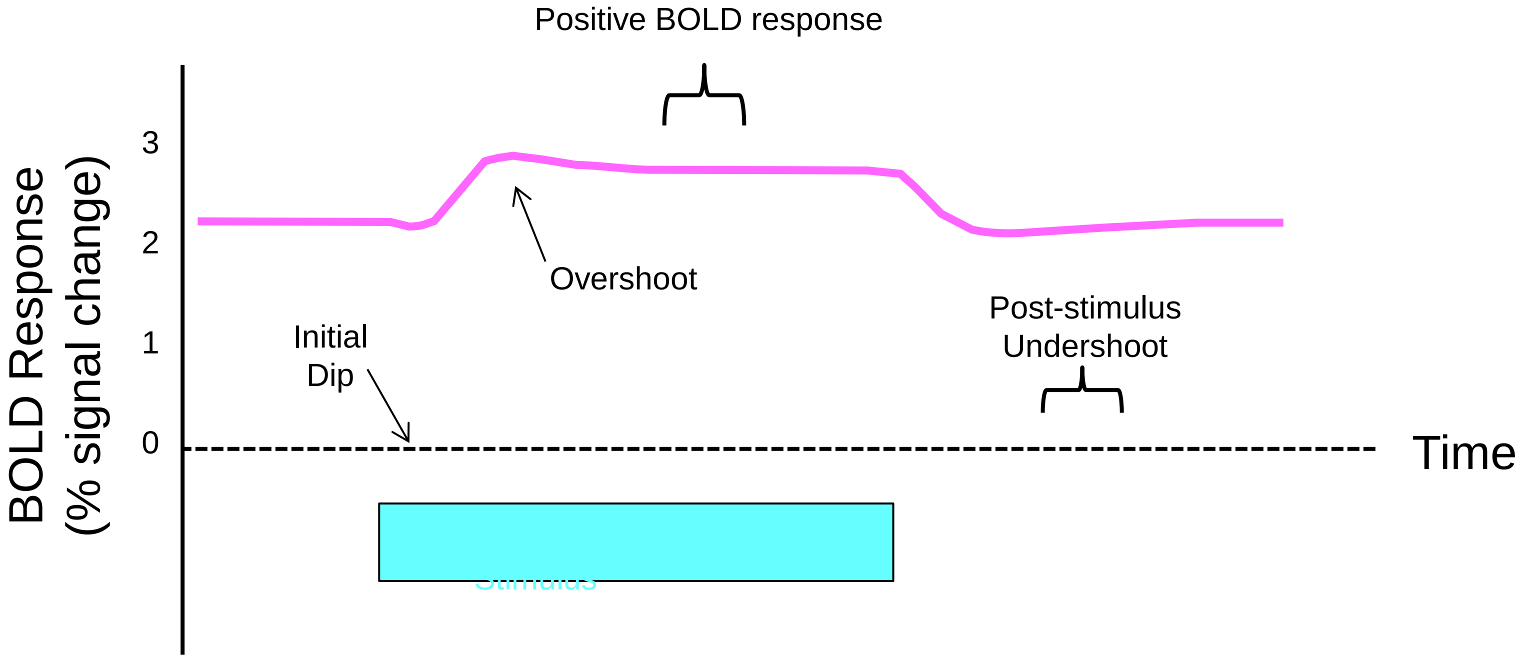
- Detection: determination of whether activity of a given voxel (or region) changes in response to the experimental manipulation
- "which voxel?"
  - Estimation: measurement of the time course within an active voxel in response to the experimental manipulation
  - "How does signal change in a voxel?"

Definitions modified from: Huettel, Song & McCarthy, 2004, Functional Magnetic Resonance Imaging

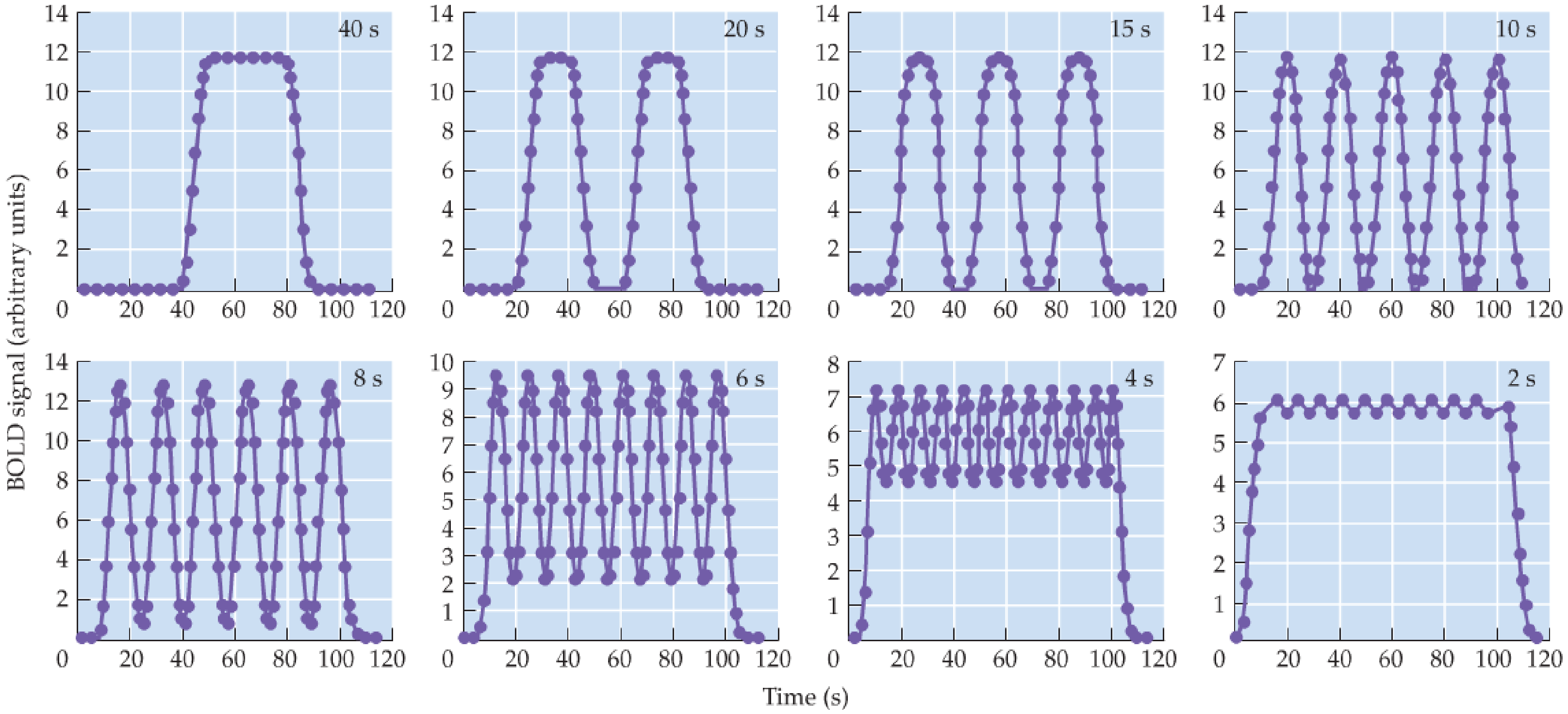
Block Designs



Early Assumption: Because the hemodynamic response delays and blurs the response to activation, the temporal resolution of FMRI is limited.



Effect of Block Interval on FMRI HRF

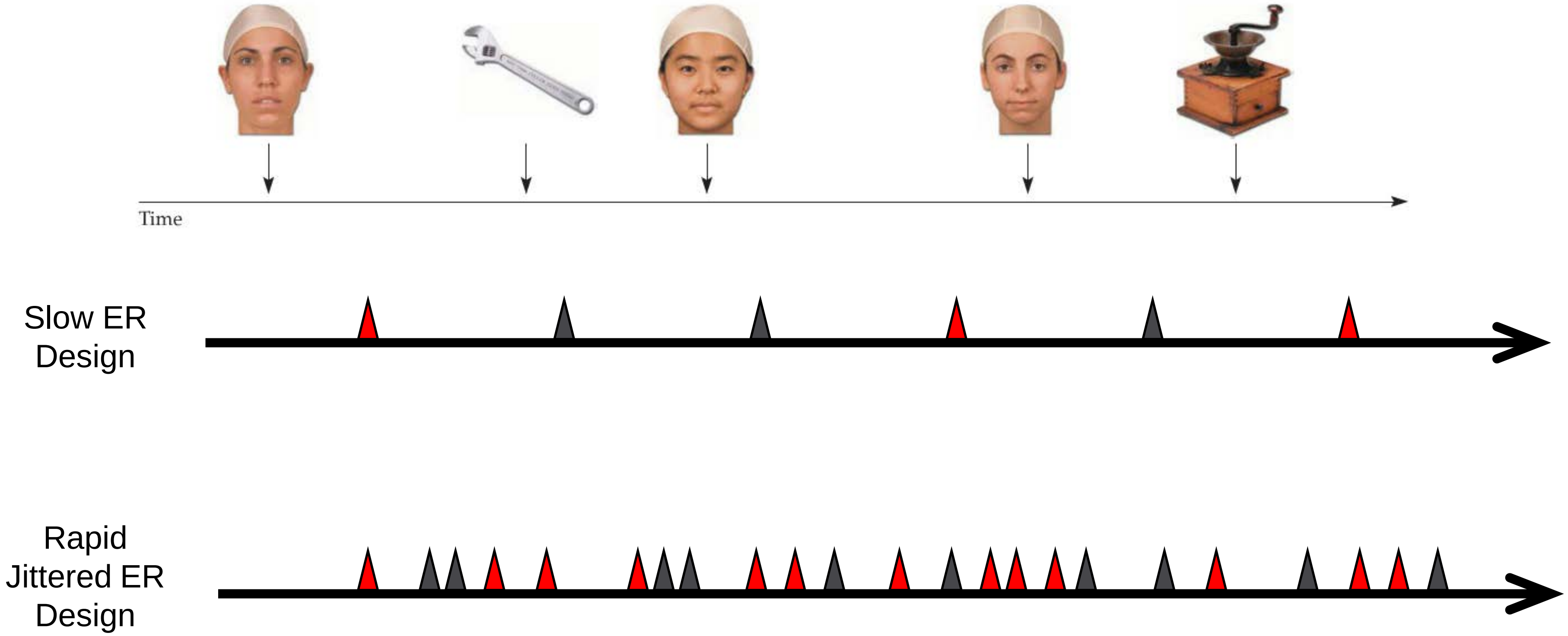


Recommendations for Using Blocked Design

- Length of a block
  - Minimally 10s and optimally 16s (Liu, 2004).
  - Equivalent for conditions or combination of conditions to be compared

- Evoking the same mental process throughout a block

Event-related Designs



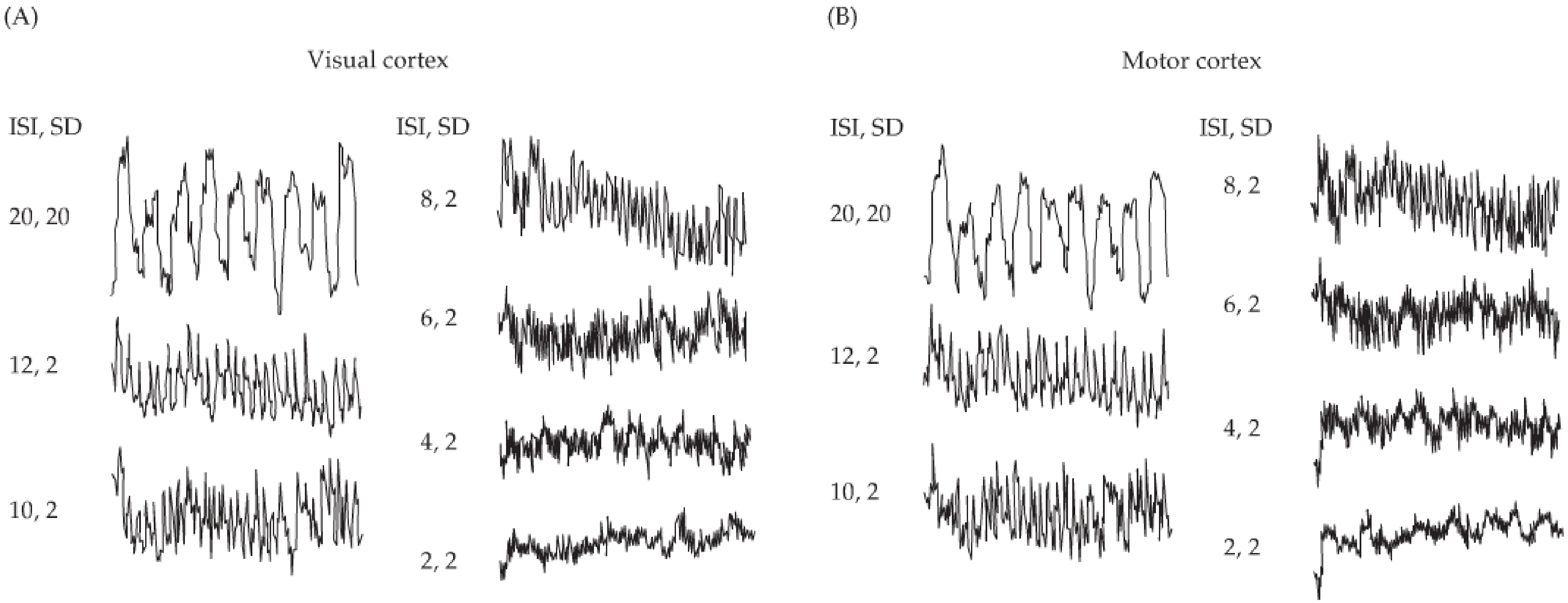
Slow Event-Related Designs

Slow ER Design

Periodic (Slow) ER Design

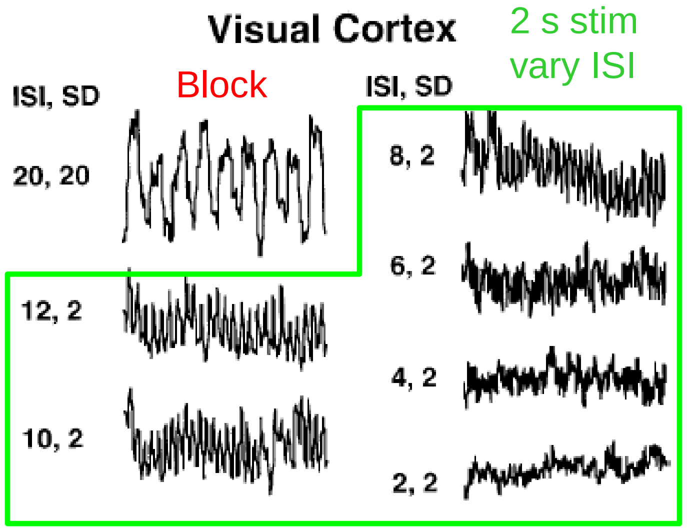
- Fixed and long ISI
  - Usually > 15s
  - Each event evokes a complete HR, and corresponding BOLD are selectively averaged.
  - Inefficient

Effects of ISI on ER-FMRI Activation

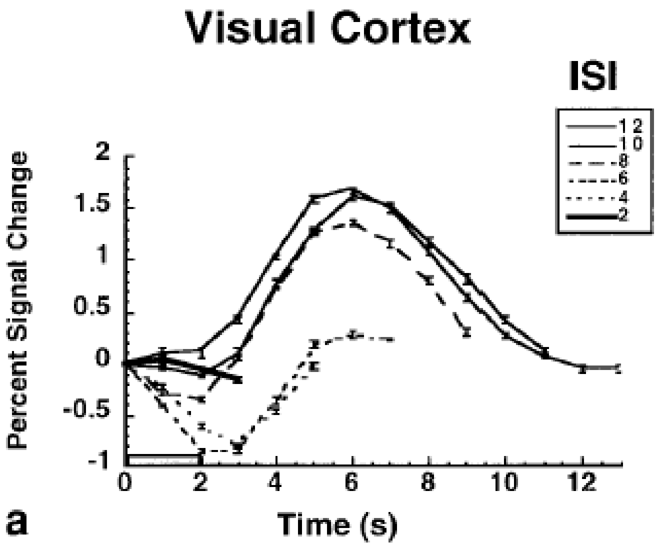


Bandettinni & Cox (2000)

Slow Event-Related Design: Constant ITI



Bandettini et al. (2000)
What is the optimal trial spacing (duration + intertrial interval, ITI) for a Spaced Mixed Trial design with constant stimulus duration?



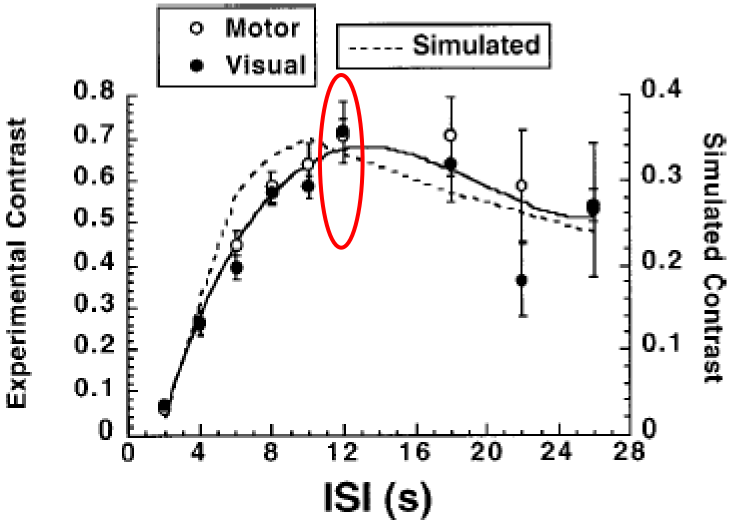
Event-related average

Source: Bandettini et al., 2000

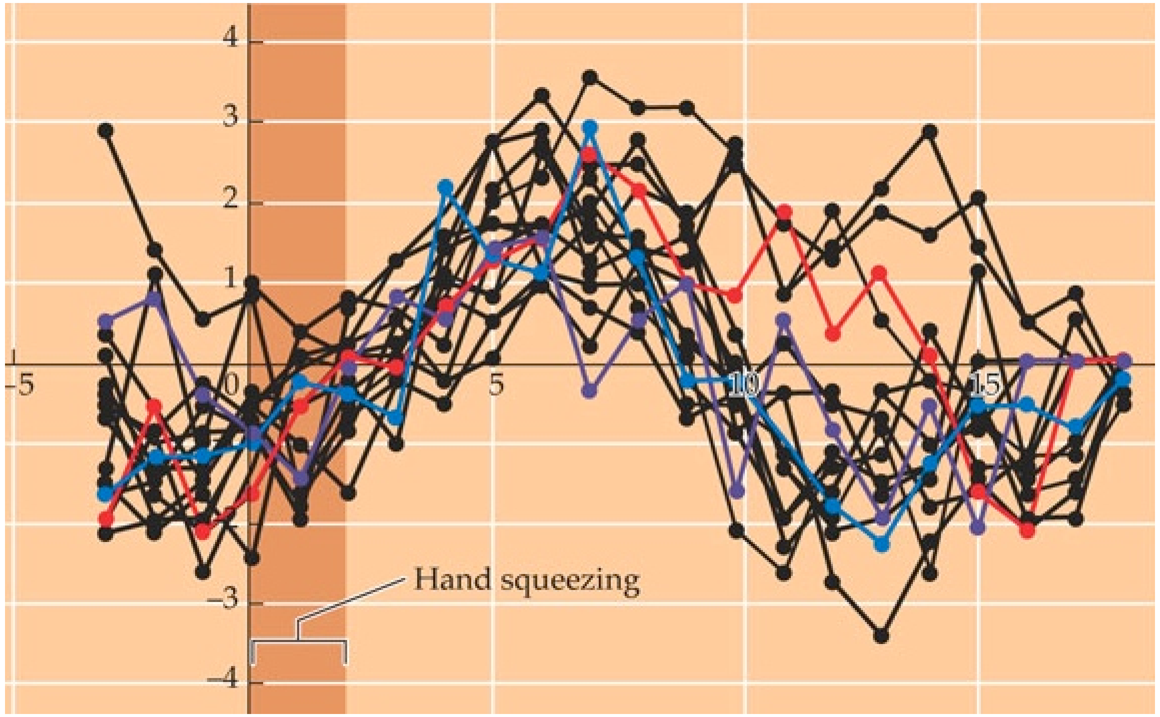
Source: Bandettini et al., 2000

Optimal Constant ITI

- Brief (< 2 sec) stimuli:
  - optimal trial spacing = 12 sec
- For longer stimuli:
  - optimal trial spacing = 8 + 2*stimulus duration
- Effective loss in power of event related design:
  - = -35%
  - i.e., for 6 minutes of block design, run ~9 min ER design

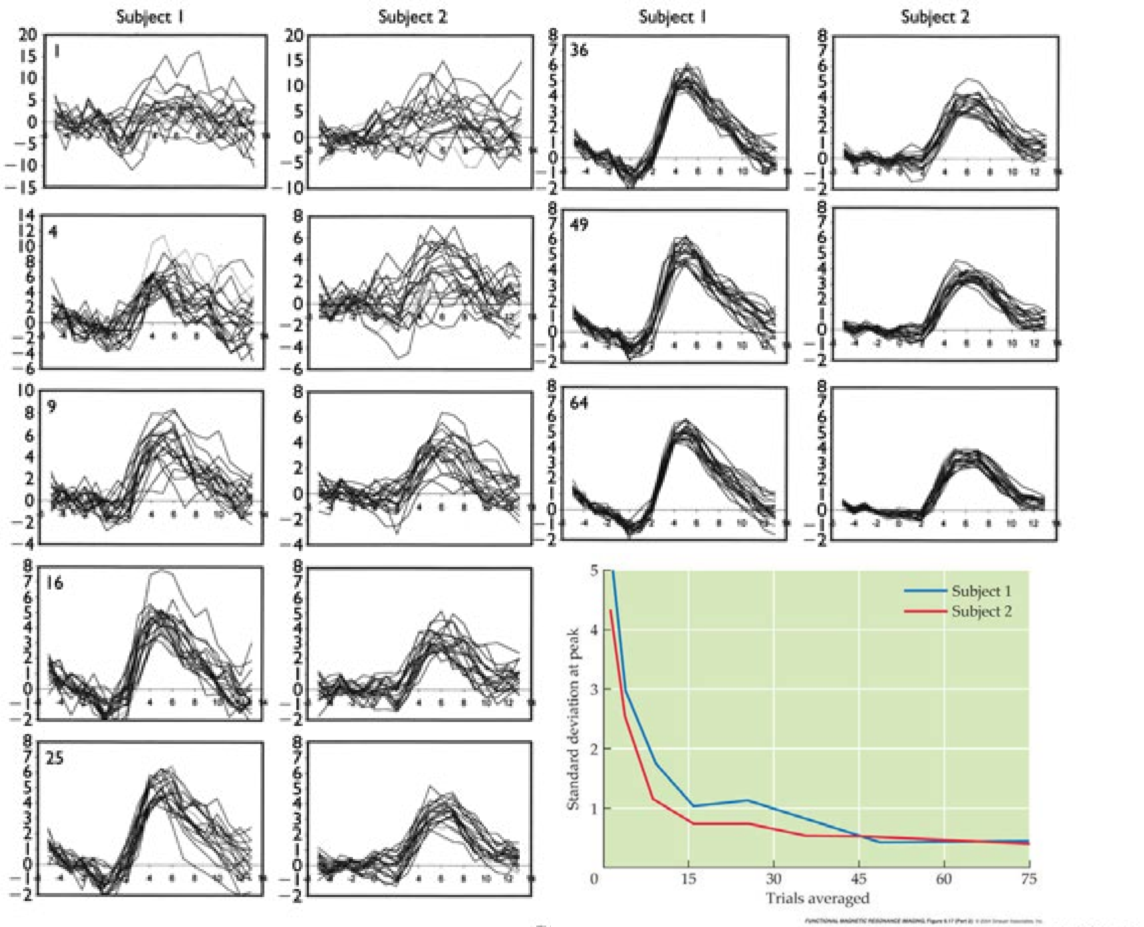


Trial to Trial Variability



Huettel, Song & McCarthy, 2004,
Functional Magnetic Resonance Imaging

How Many Trials Do You Need?

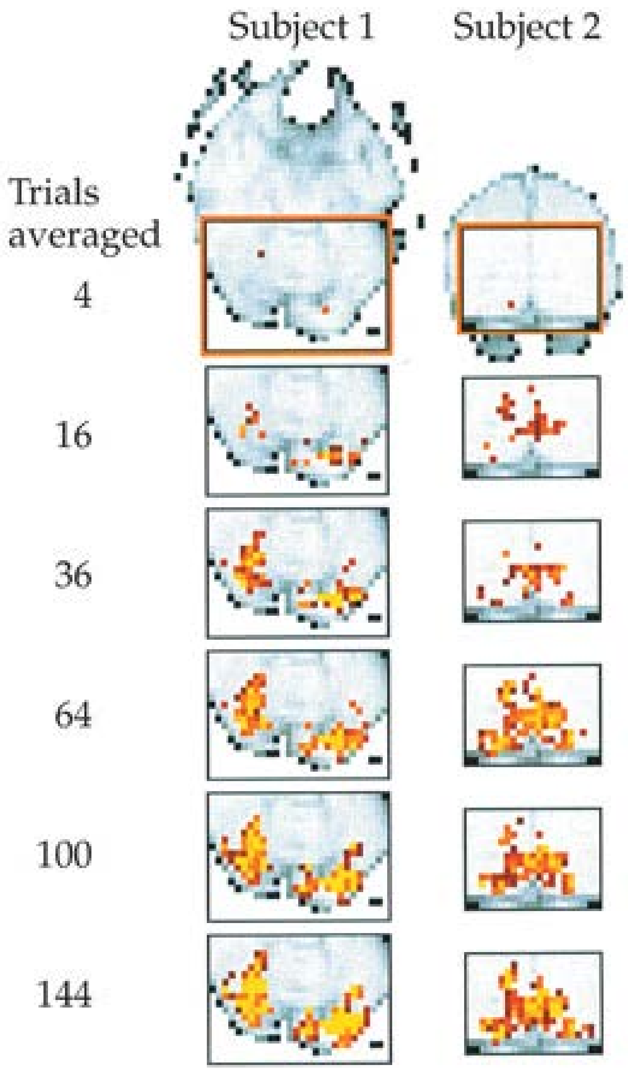


FUNCTIONAL MAGNETIC RESONANCE IMAGING, Figure 9.17 (Part 1) © 2004 Sinauer Associates, Inc.

Huettel, Song & McCarthy, 2004, Functional Magnetic Resonance Imaging

- standard error of the mean varies with square root of number of trials
- Number of trials needed will vary with effect size
- Function begins to asymptote around 15 trials

Effect of Adding Trials



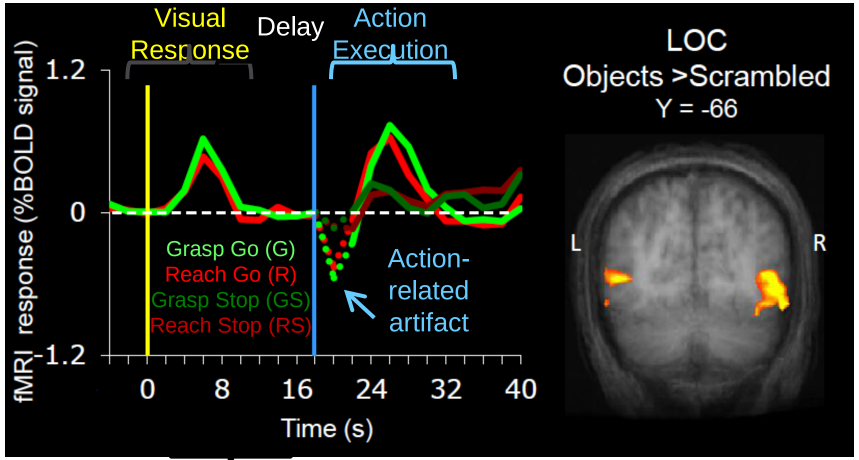
Huettel, Song & McCarthy, 2004, Functional Magnetic Resonance Imaging

Pros & Cons of Slow ER Designs

Pros

- excellent estimation
- useful for studies with delay periods
- very useful for designs with motion artifacts (grasping, swallowing, speech) because you can tease out artifacts
- analysis is straightforward

Example: Delayed Hand Actions
(*Singhal et al., under revision*)



Really long delay:
18 s

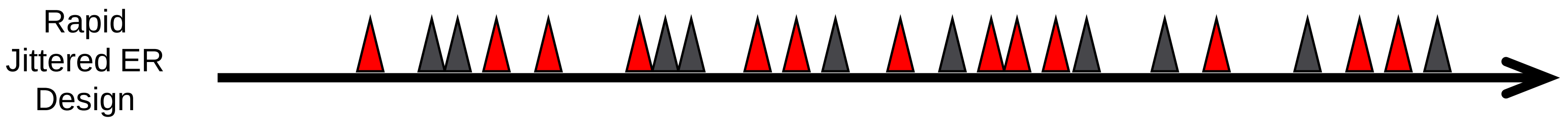
Cons

- poor detection power because you get very few trials per condition by spending most of your sampling power on estimating the baseline
- subjects can get VERY bored and sleepy with long inter-trial intervals

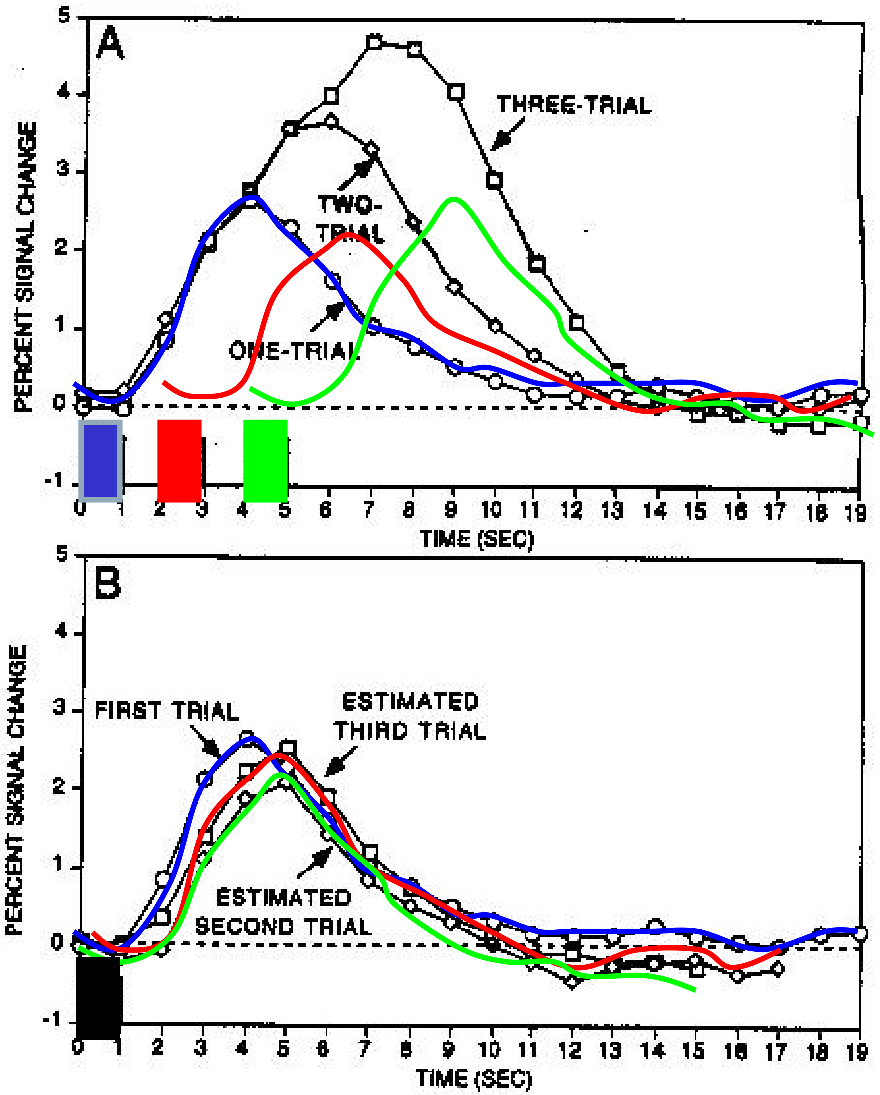
Effect of this design on our subject

要不要再快一點？

- Yes, but we have to test assumptions regarding linearity of BOLD signal first



Linearity of BOLD response



Source: *Dale & Buckner, 1997*

Linearity:
"Do things add up?"

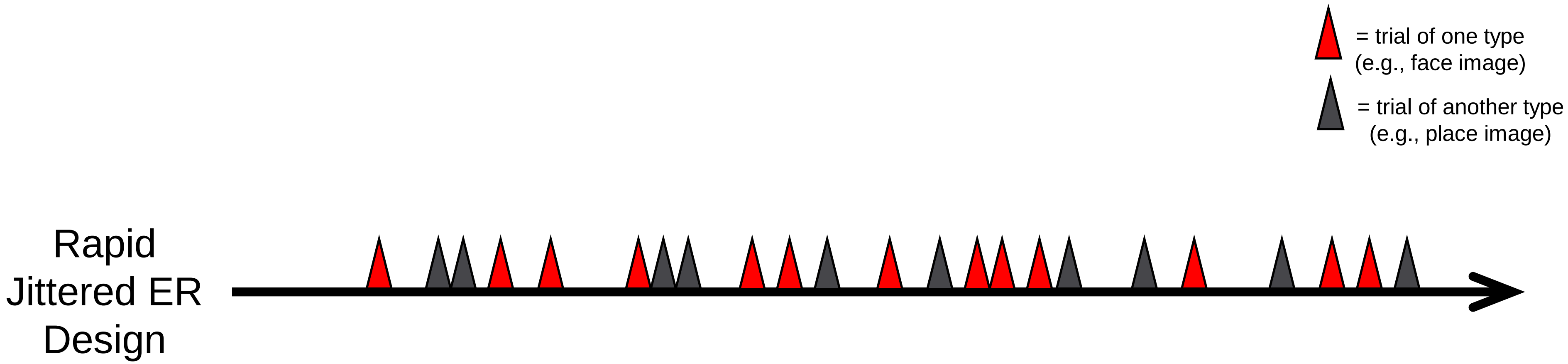
red = 2 - 1

green = 3 - 2

Sync each trial response to start of trial

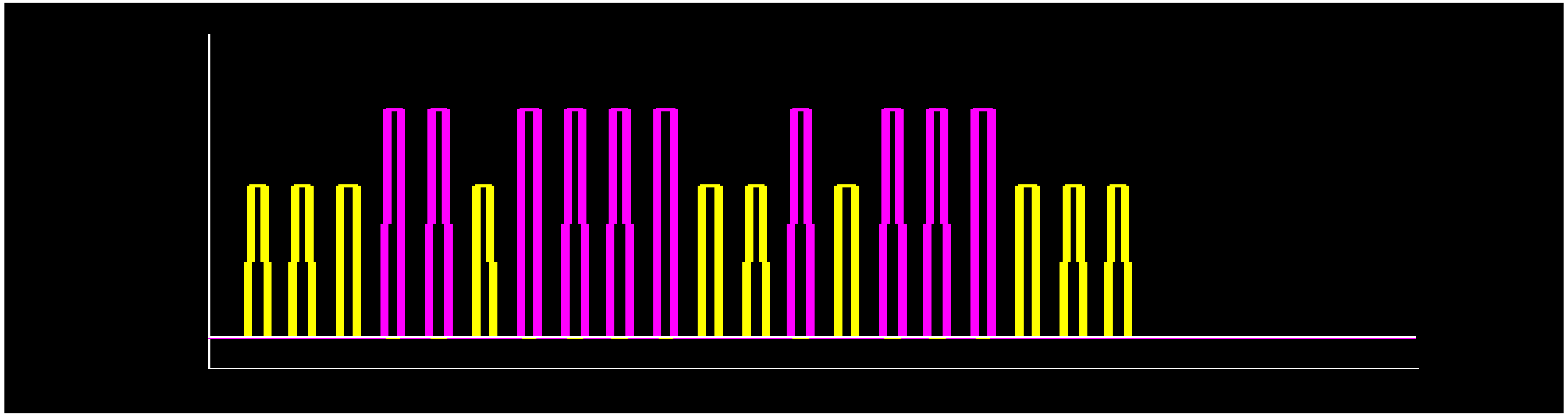
Not quite linear but good enough!

Rapid Jittered ER Design

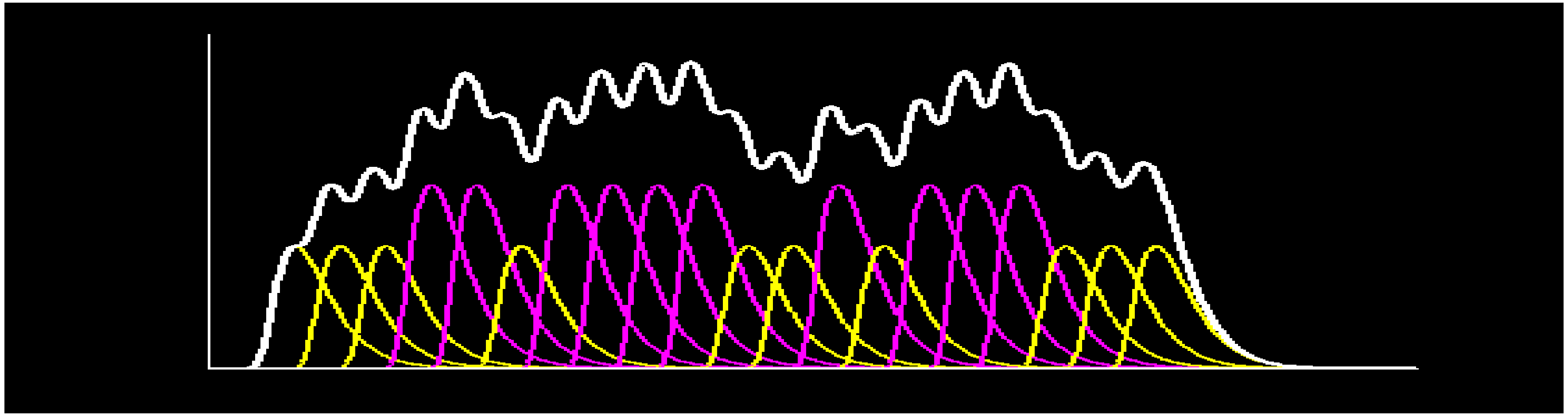


A popular choice is to use 'jittered' designs with inter-stimulus intervals of at least4*s* and with exponentially decreasing delay frequencies up to 16*s*.

BOLD Overlap With Regular Trial Spacing



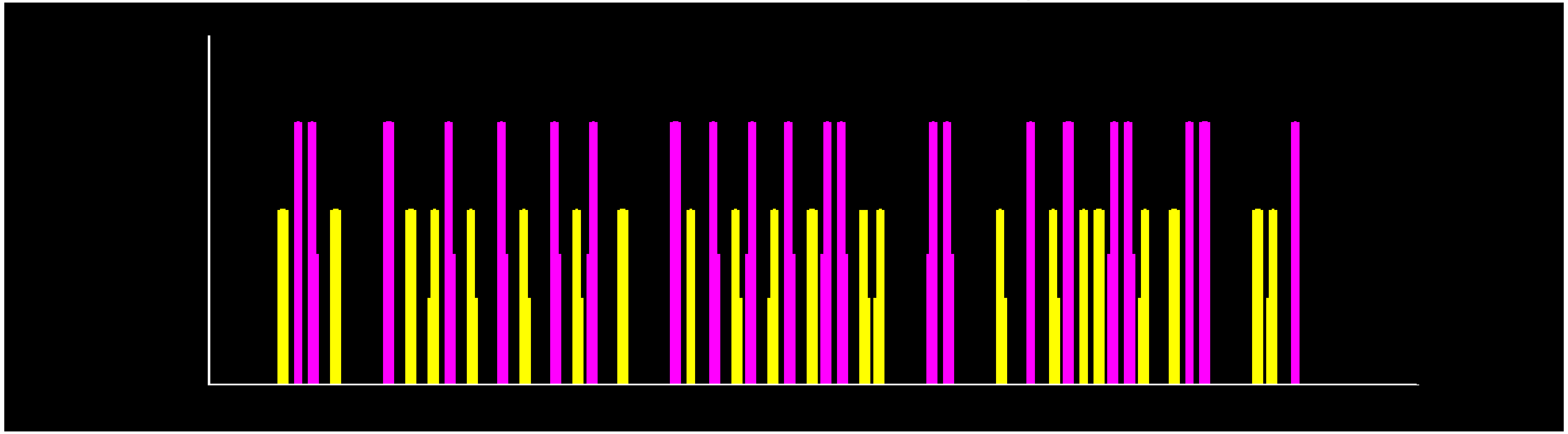
Neuronal activity from **TWO** event types with constant ITI



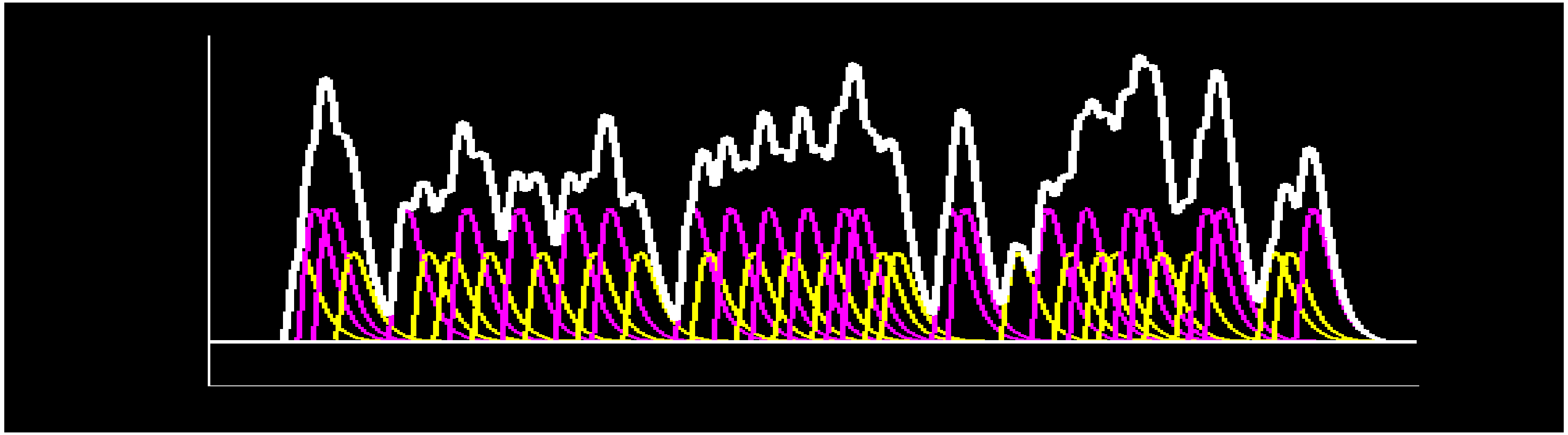
Partial tetanus BOLD activity from two event types

Slide from Matt Brown

BOLD Overlap with Jittering



Neuronal activity from closely-spaced, jittered events

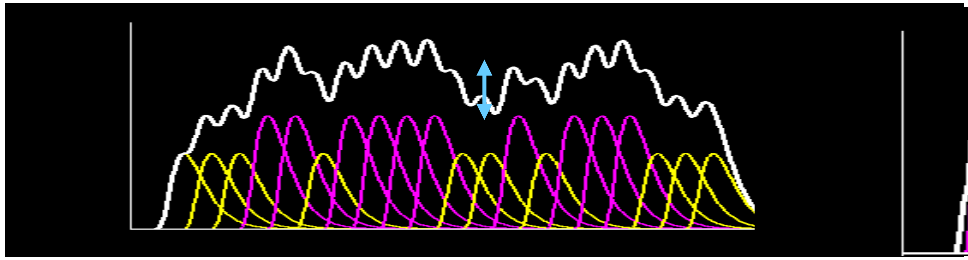


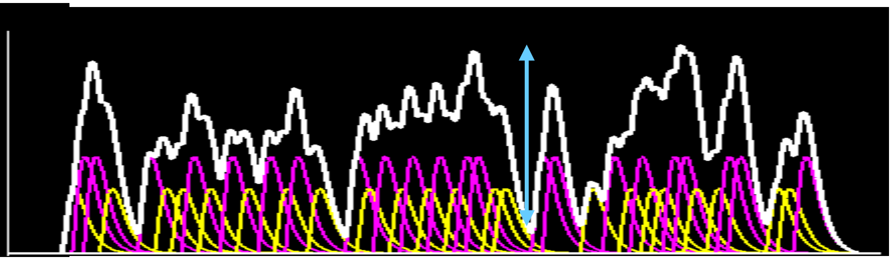
BOLD activity from closely-spaced, jittered events

Slide from Matt Brown

Why jitter?

- Yields larger fluctuations in signal



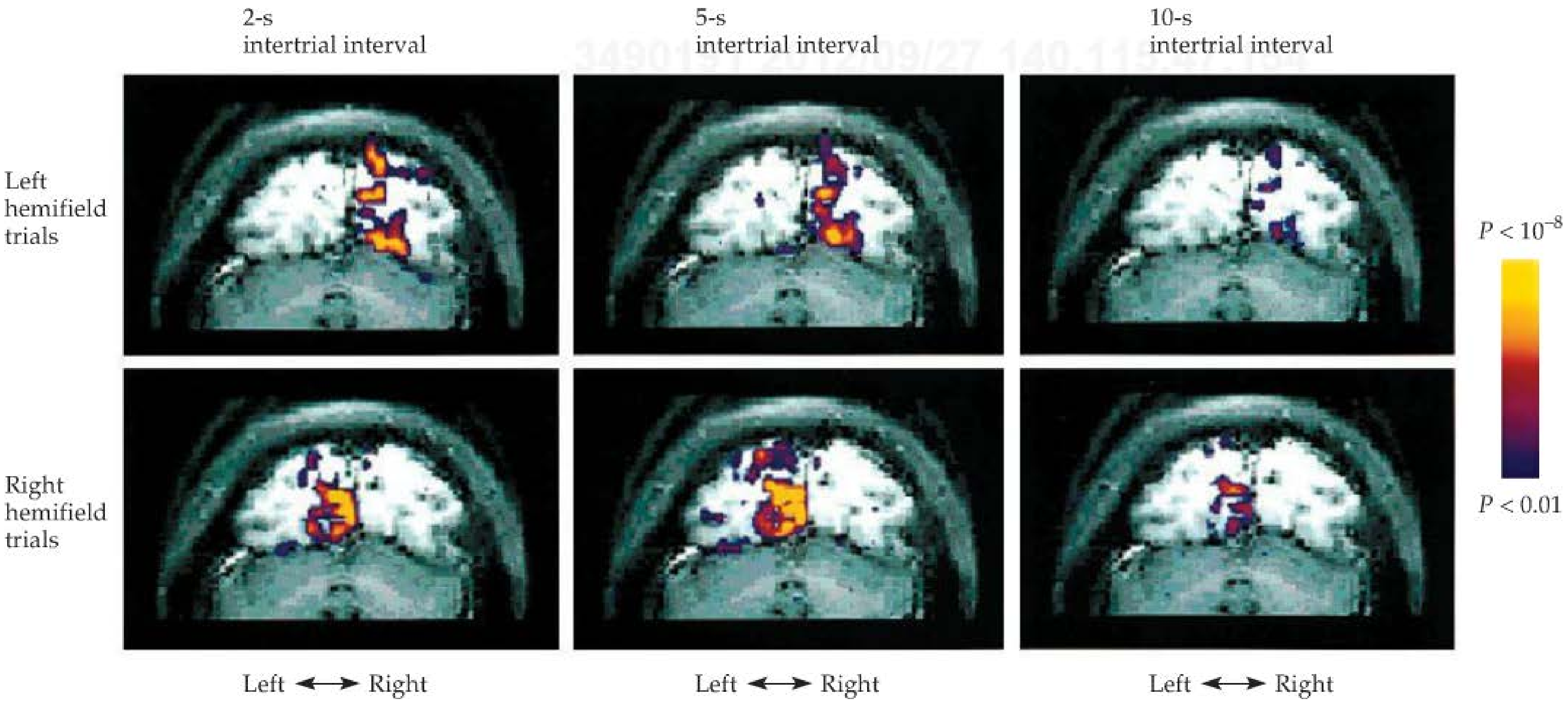


When pink is on, yellow is off
→ pink and yellow are anticorrelated

Includes cases when both pink and yellow are off
→ less anticorrelation

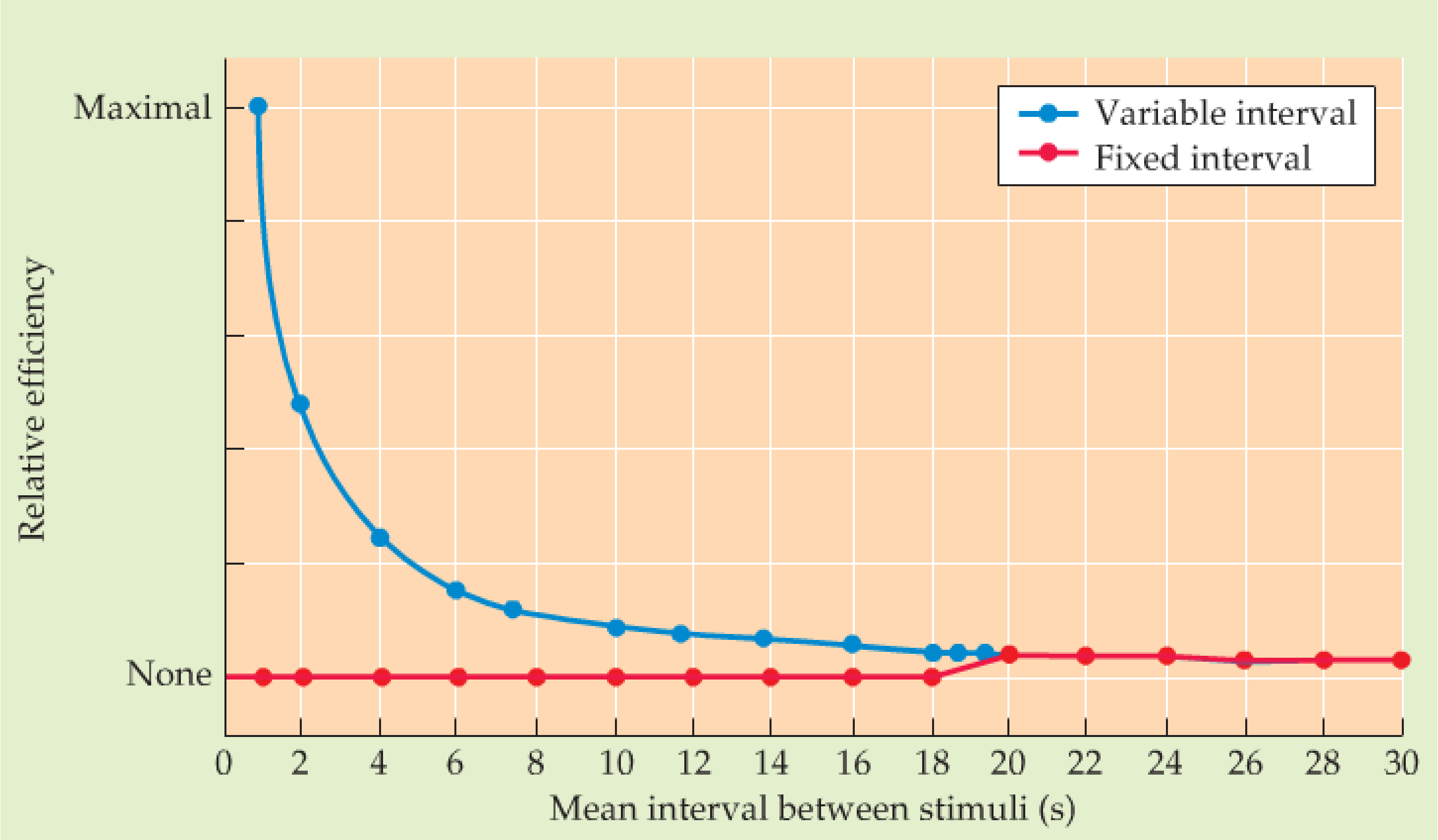
- Without jittering predictors from different trial types are strongly anticorrelated
  - As we know, the GLM doesn't do so well when predictors are correlated (or anticorrelated)

Rapid ER-FMRI with Randomized Stimulus Presentation



Short randomized ITI enhances detection power.

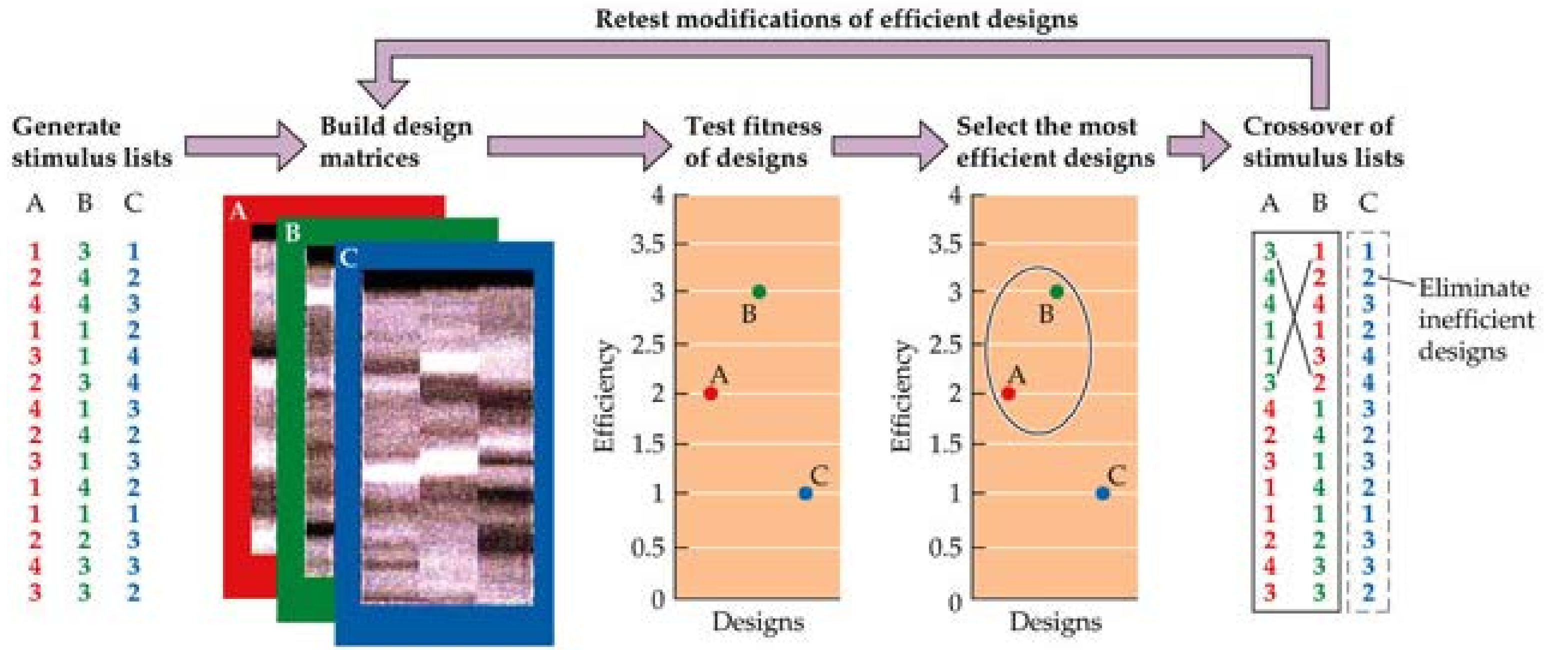
Variable vs. Fixed Intervals



After Dale (1999)

Algorithms for Picking Efficient Designs

Genetic Algorithms



http://wagerlab.colorado.edu/tools

Pros & Cons of Applying Standard GLM to Rapid-ER Designs

Pros

- Acceptable detection power
- trials can be put in unpredictable order
- subjects don't get so bored

Cons and Caveats

- reduced detection compared to block designs
- requires stronger assumptions about linearity
  - BOLD is non-linear with inter-event intervals < 6 sec.
  - Nonlinearity becomes severe under 2 sec.
- errors in HRF model can introduce errors in activation estimates

Good Practices in FMRI

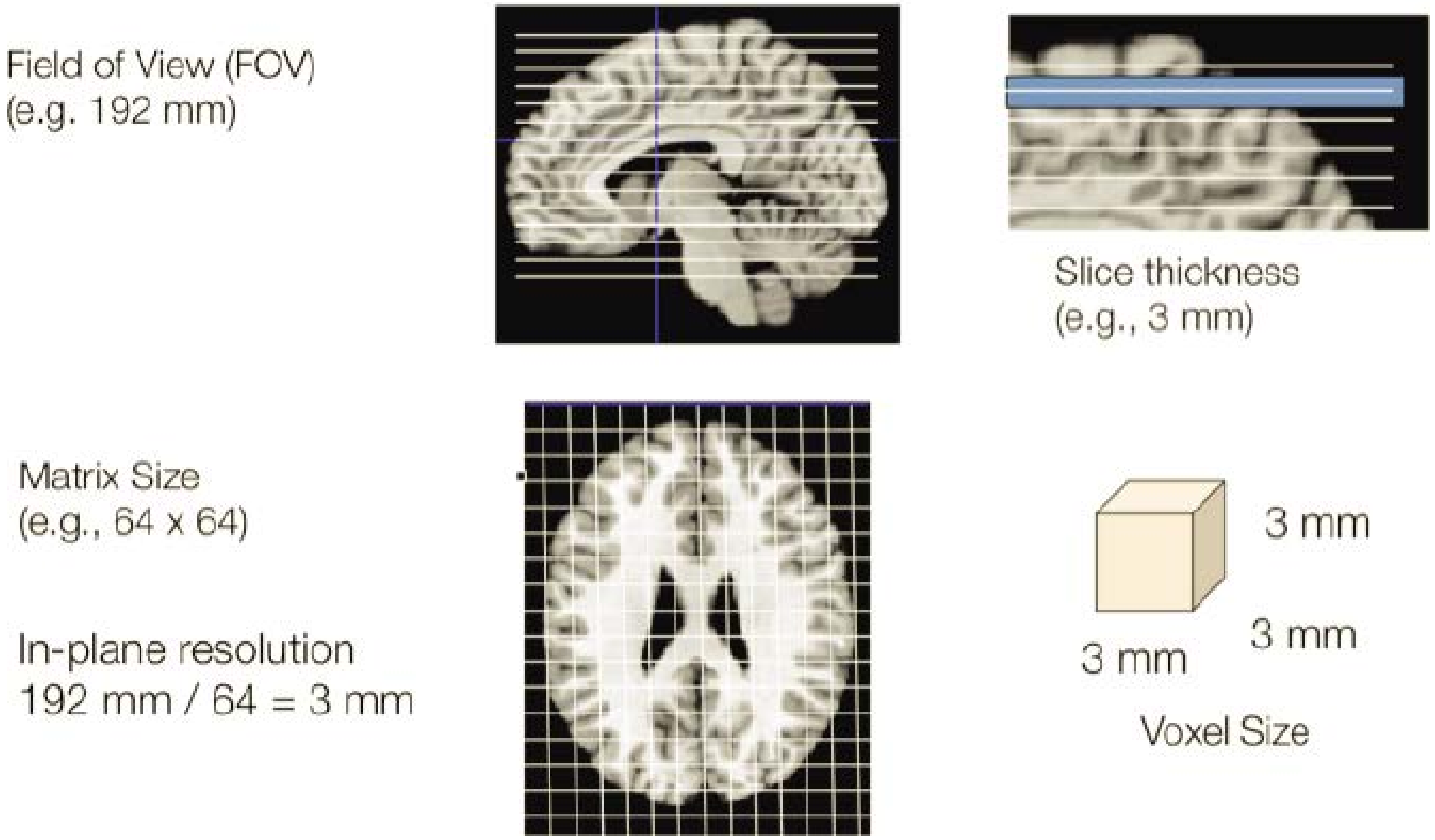
- Evoke the cognitive processes of interest
- Maximize data collection from each subject
- Maximize sample size
- Choose conditions and timings that maximize evoked changes in the process of interests
- Minimize correlation between BOLDs of successive events
- Compute correlation between behavioral performance and activation

腦造影 vs. 腦照影

FMRI資料前處理

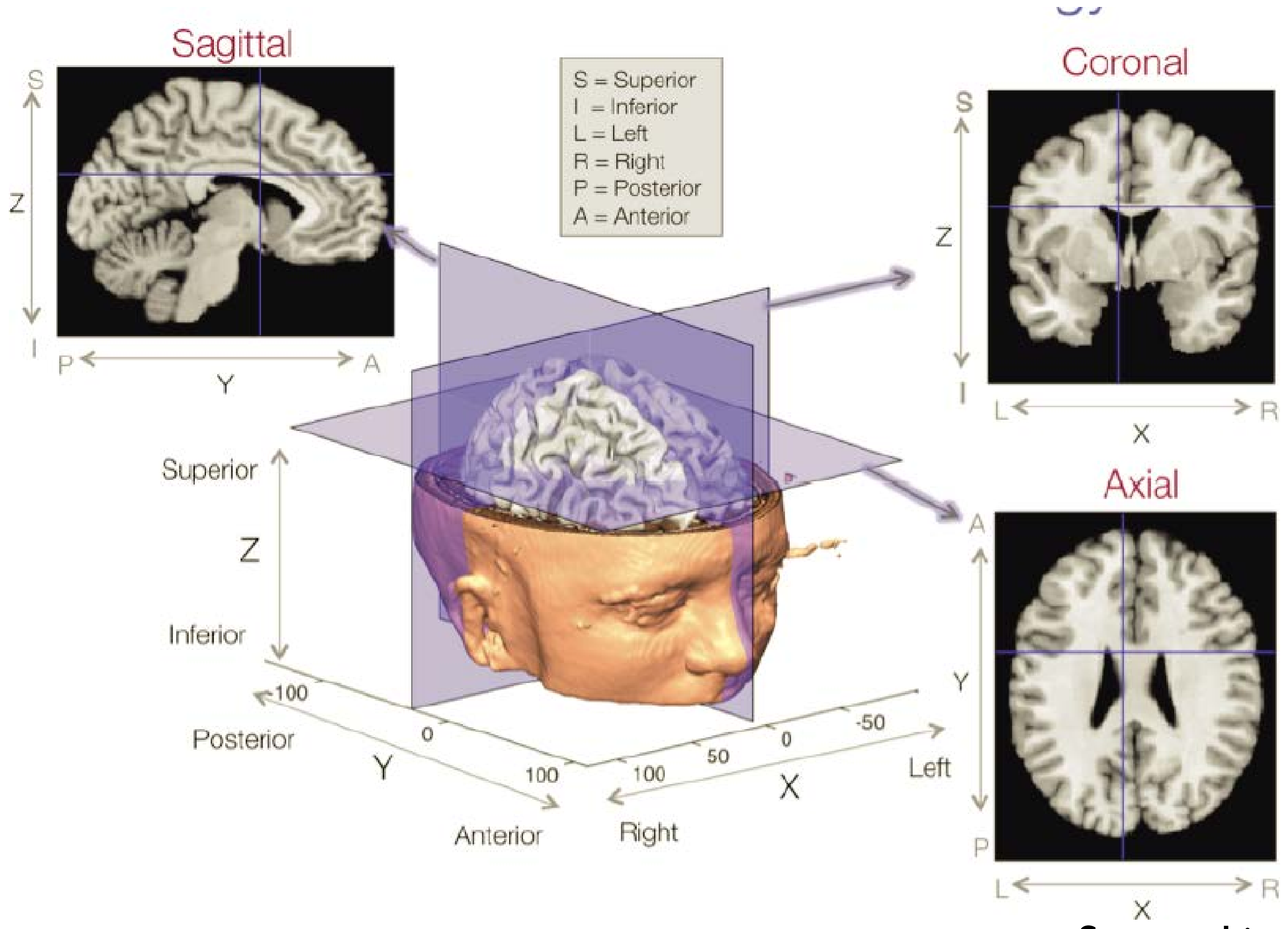
Preprocessing

Images: Basic Terminology

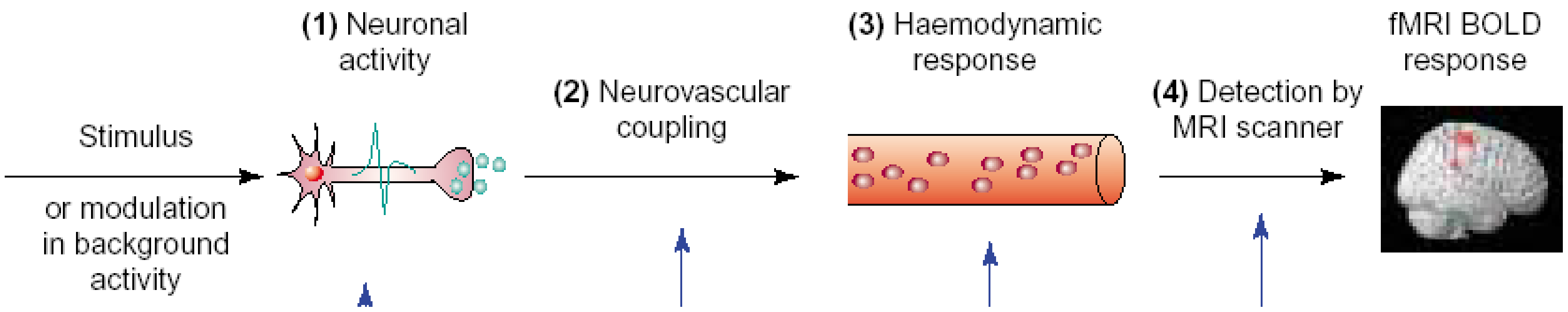


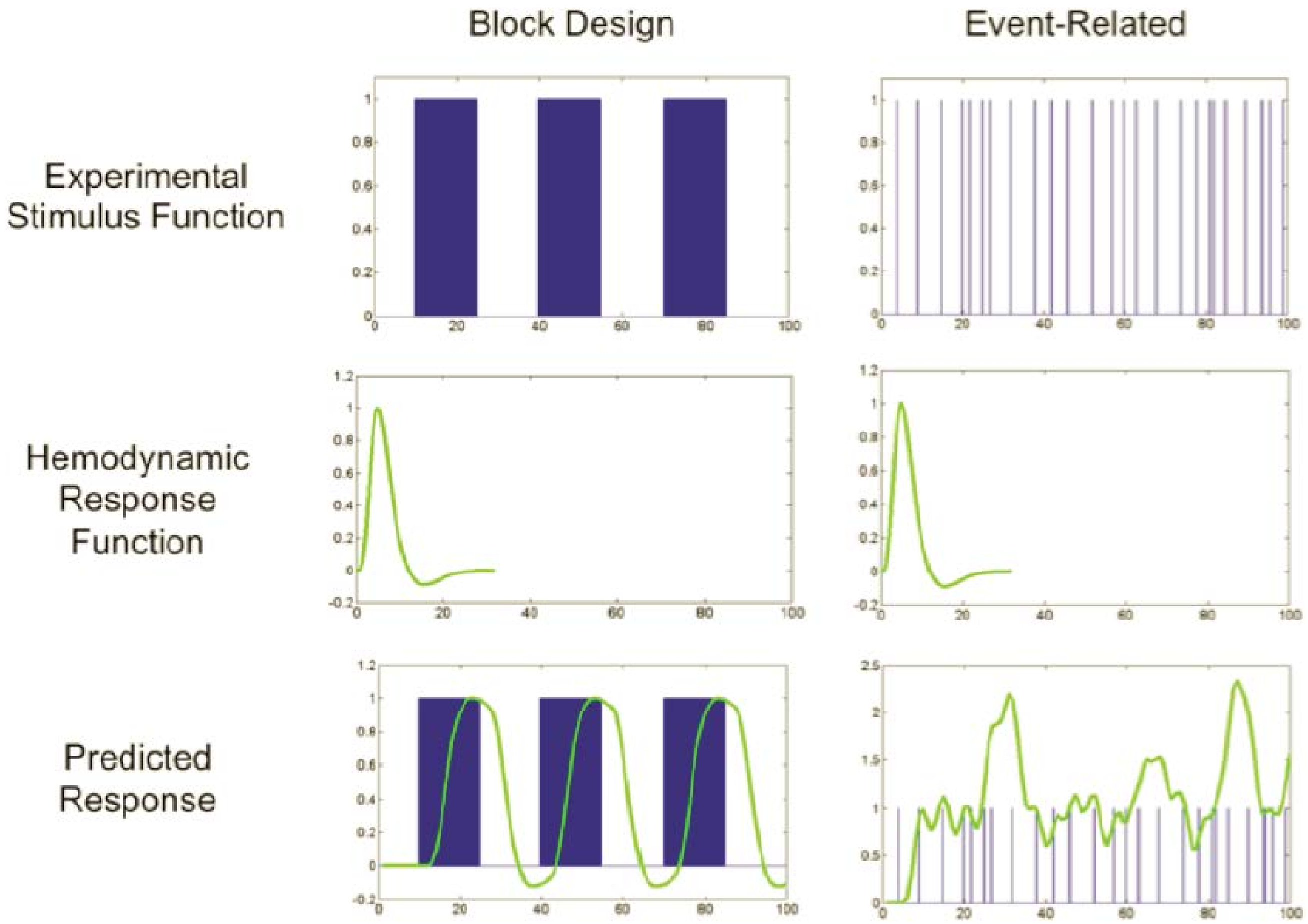
Source: Lindquist & Tor (2015)

Brain Dimensions and Terminology



Source: Lindquist & Tor (2015)

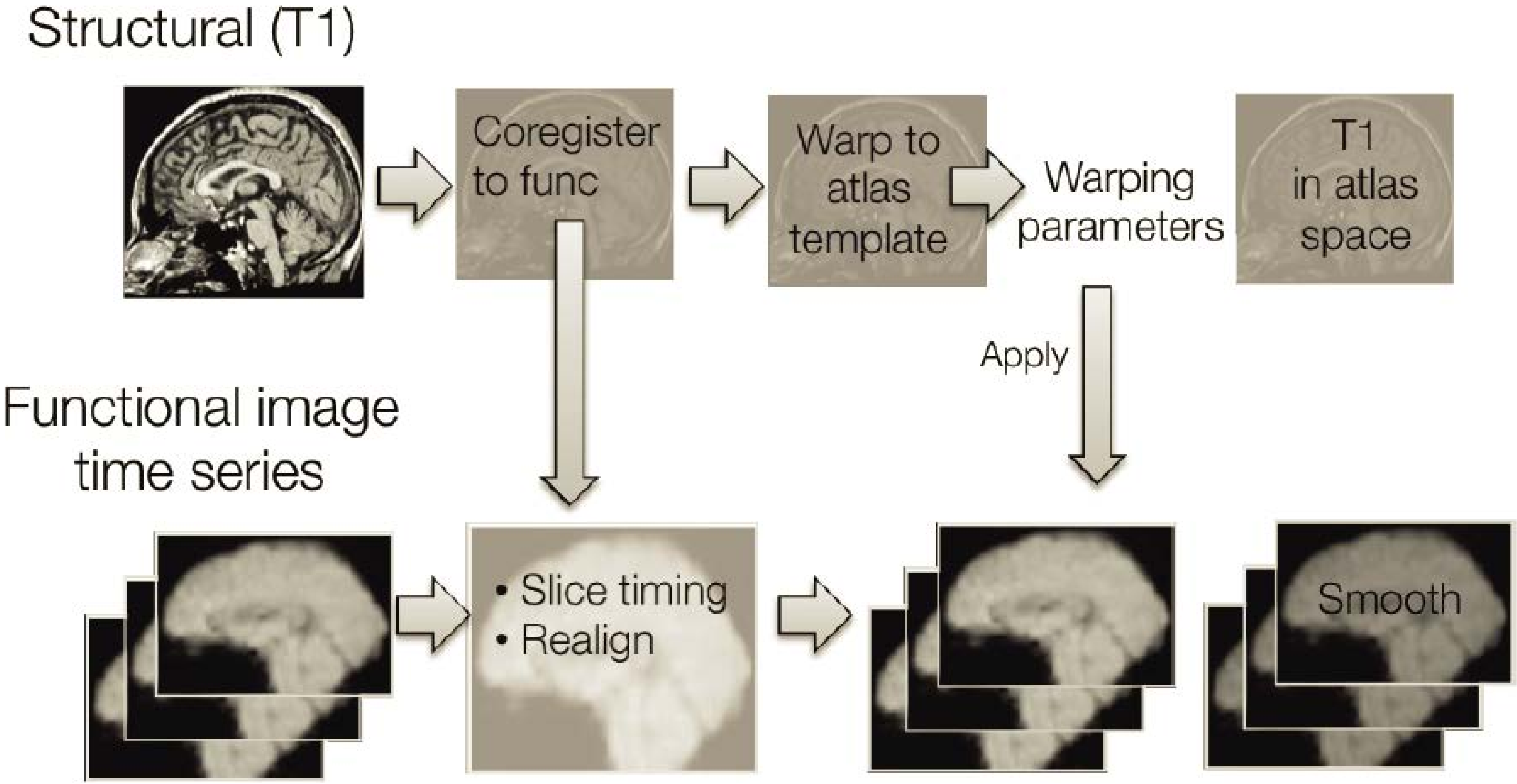




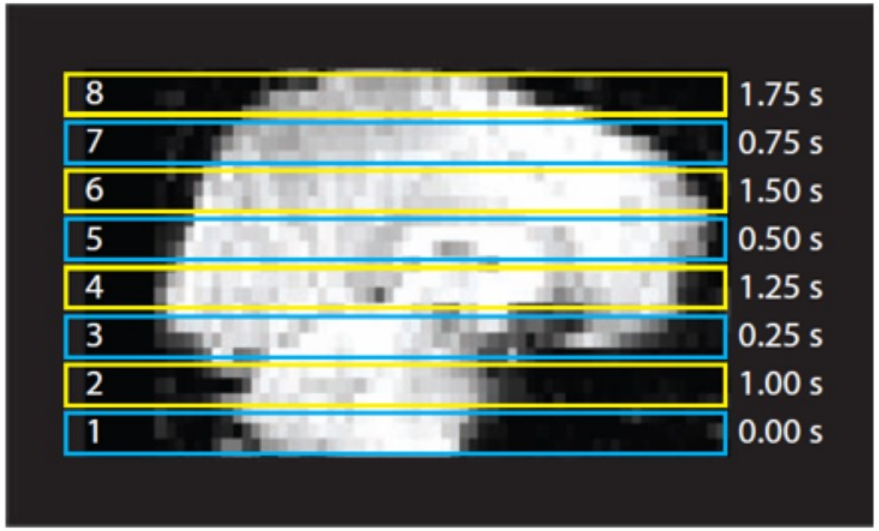
Two-gamma function

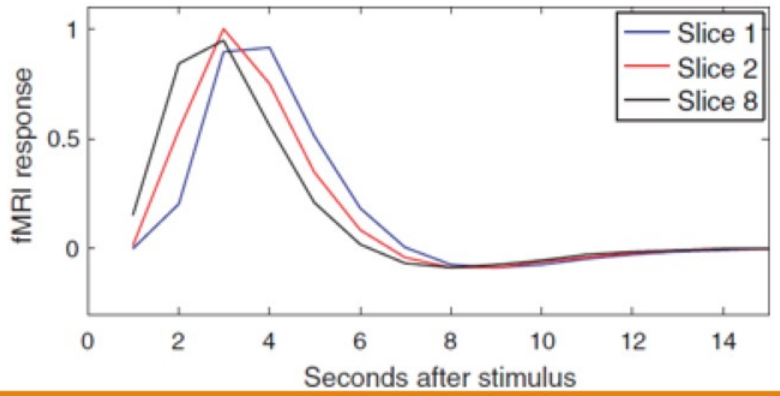
$$h(t) = \frac{t^{\alpha_1 - 1}\beta_1^{\alpha_1} e^{-\beta_1 t}}{\Gamma(\alpha_1)} - c\frac{t^{\alpha_2 - 1}\beta_2^{\alpha_2} e^{-\beta_2 t}}{\Gamma(\alpha_2)}$$

Overview of Preprocessing



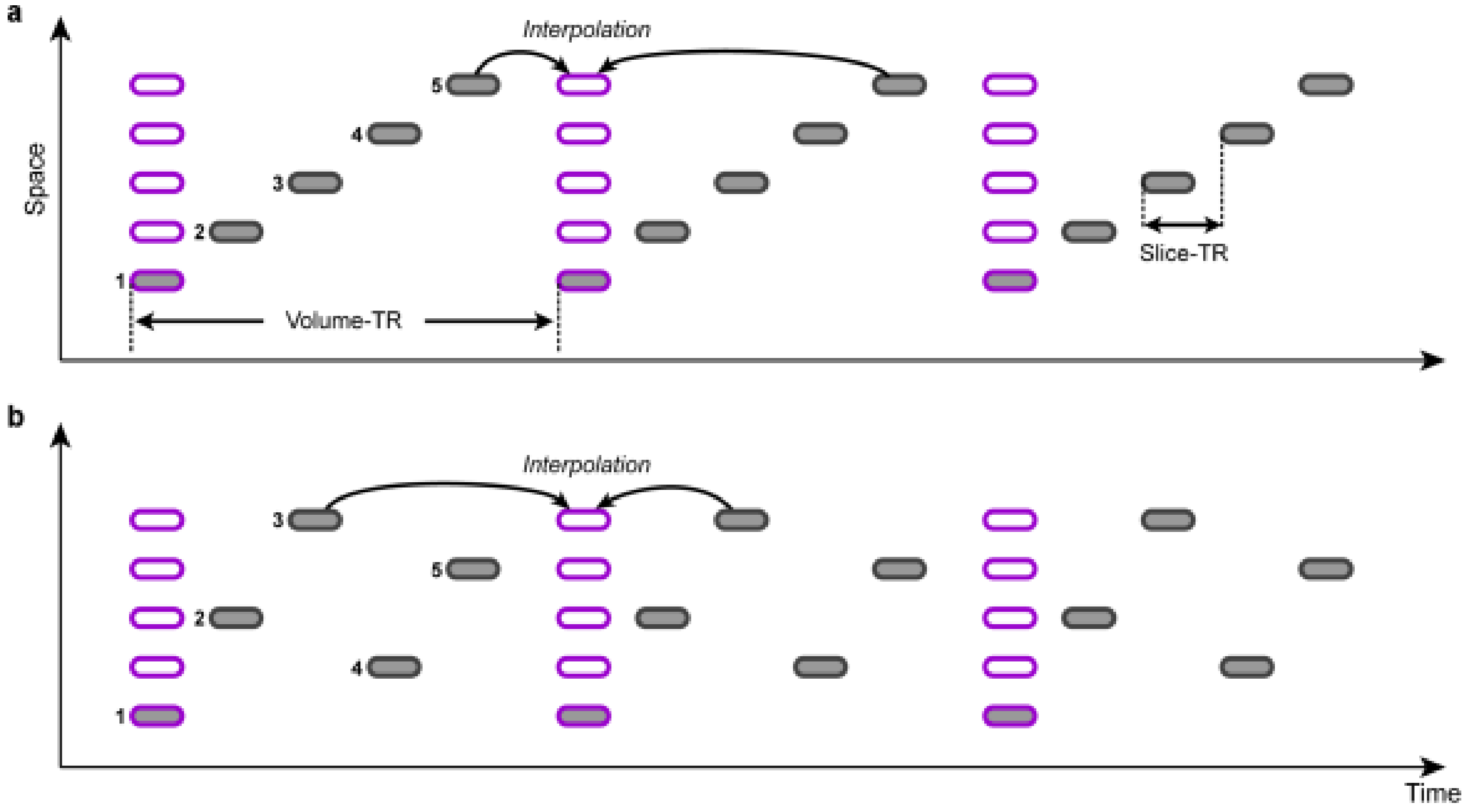
Slice Time Problem



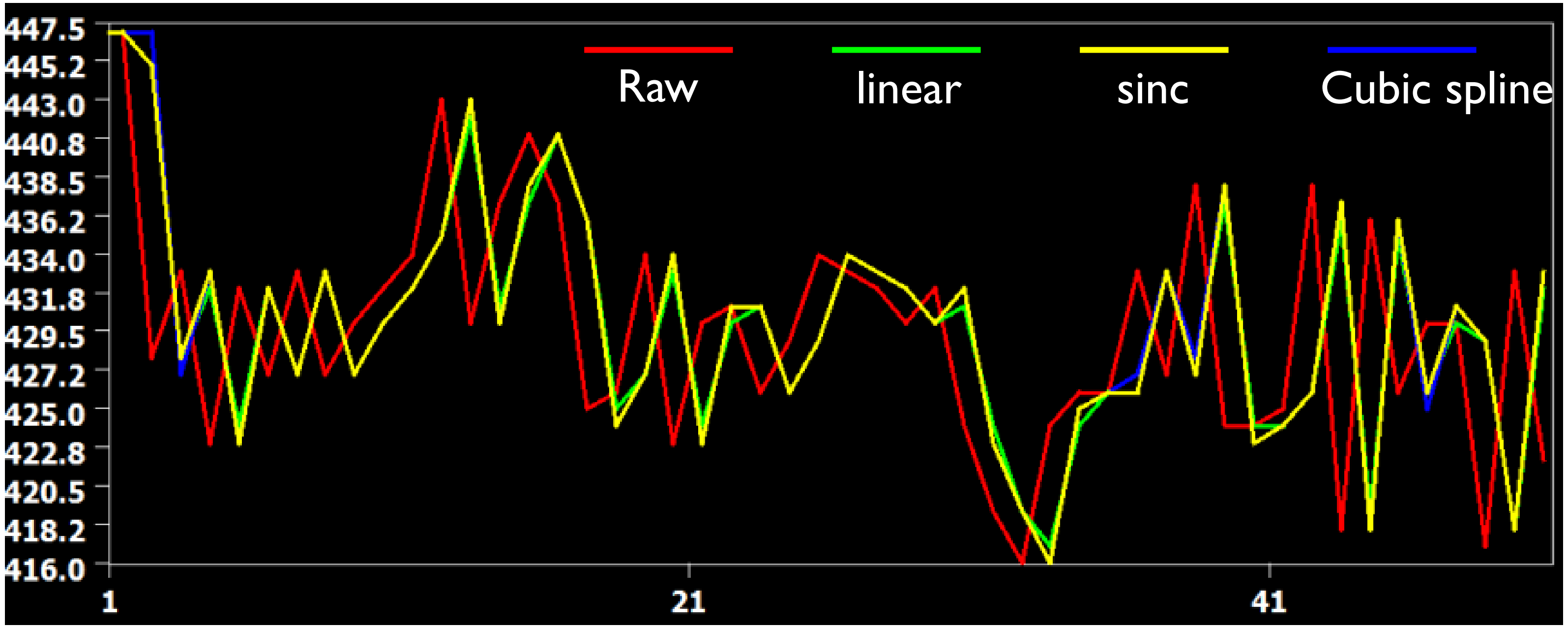


Not accounting for the timing differences between slices may lead to problematic time course differences between voxels on different slices.

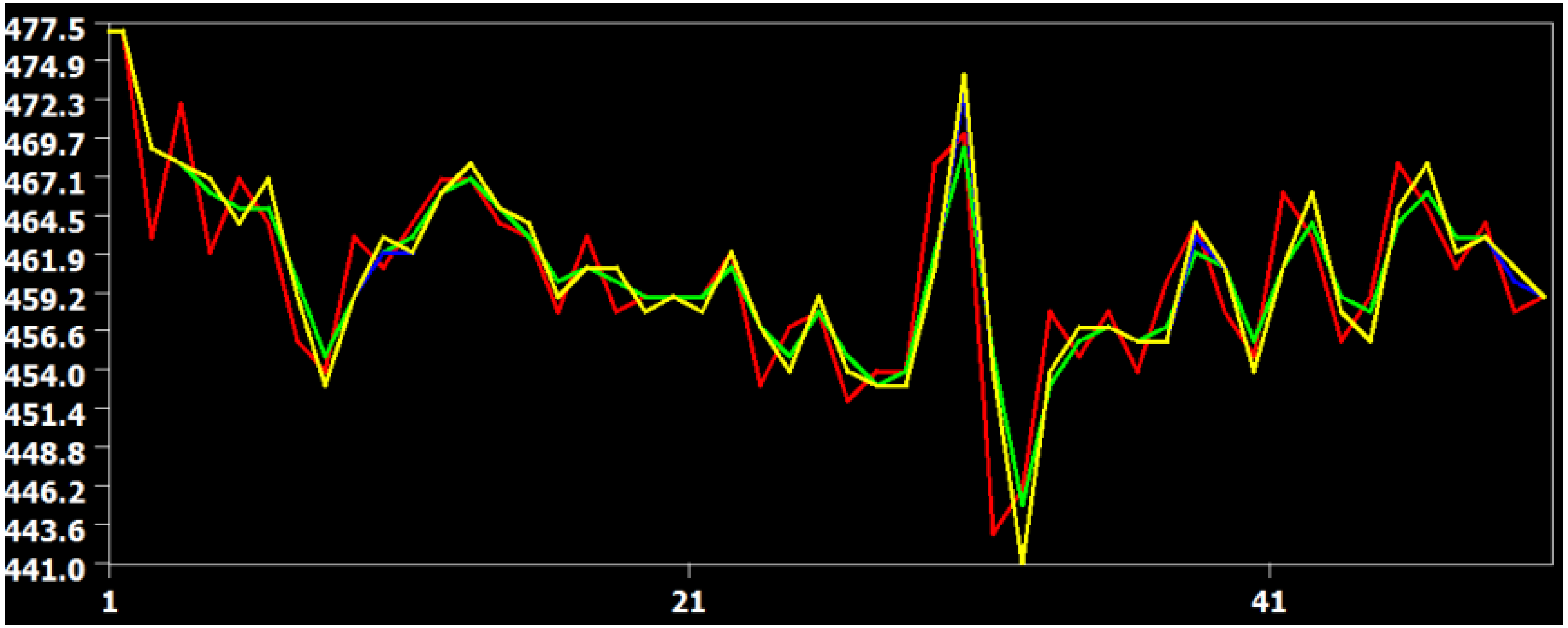
More on Slice Time Correction



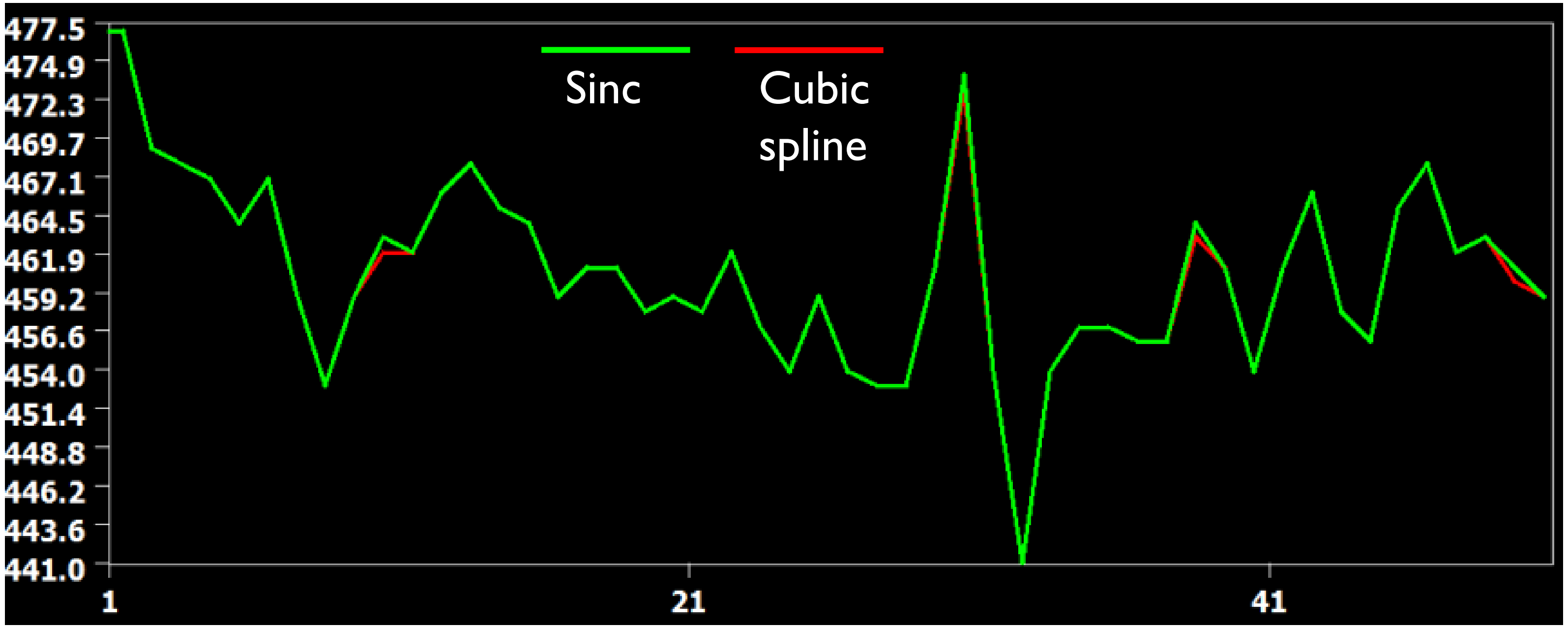
Raw from the Last Slice



Raw Data from Middle Slice

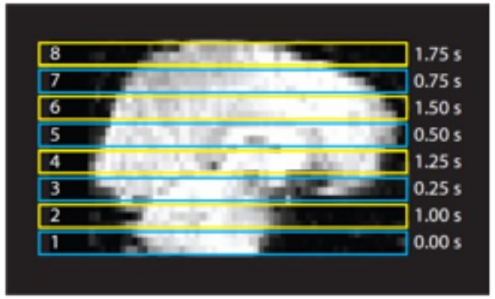


Cubic Spline vs. sinc



Reasons Not to Correct for Slice Time

Propagation of artifacts



With short TR and interleaved acquisition, slice-timing problems is minimal

- Particularly after spatial smoothing

Temporal derivatives absorbs the impacts

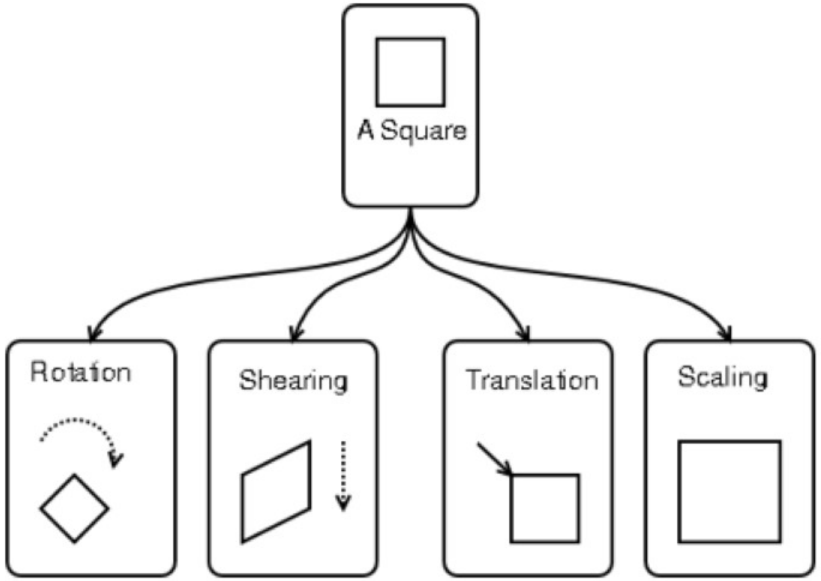
Spatial Transformation

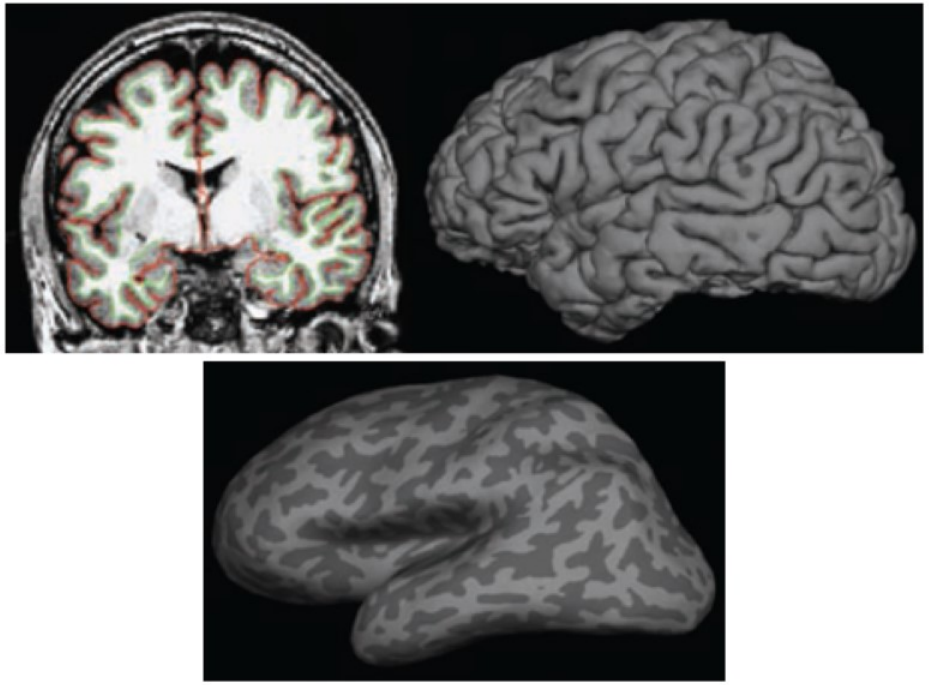
Volume-based transformations

- Changes to 3D volume of data

Surface-based registration

- Changes to surface data





Models for Spatial Transformations

Affine transformation

- Translation, rotation, scaling, shearing
- Rigid-body transformation
  - No scaling and shearing

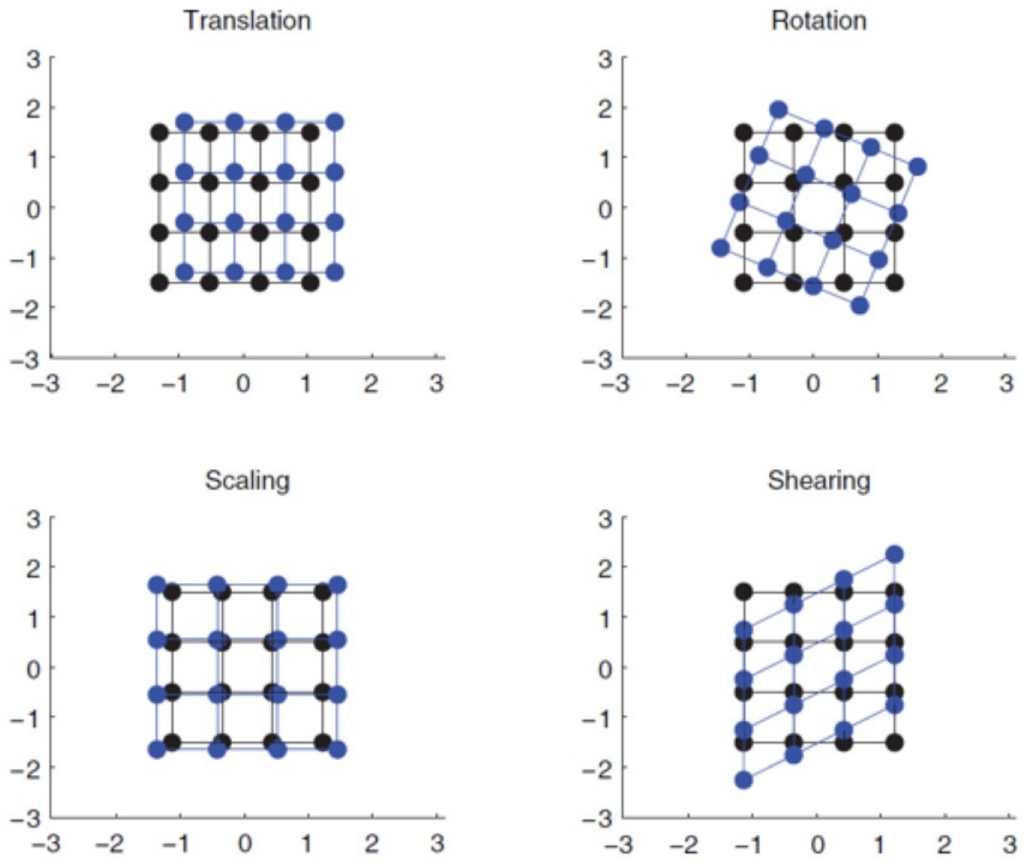
Piecewise linear transformation

- Divide the image into sub-regions and transform each of them respectively

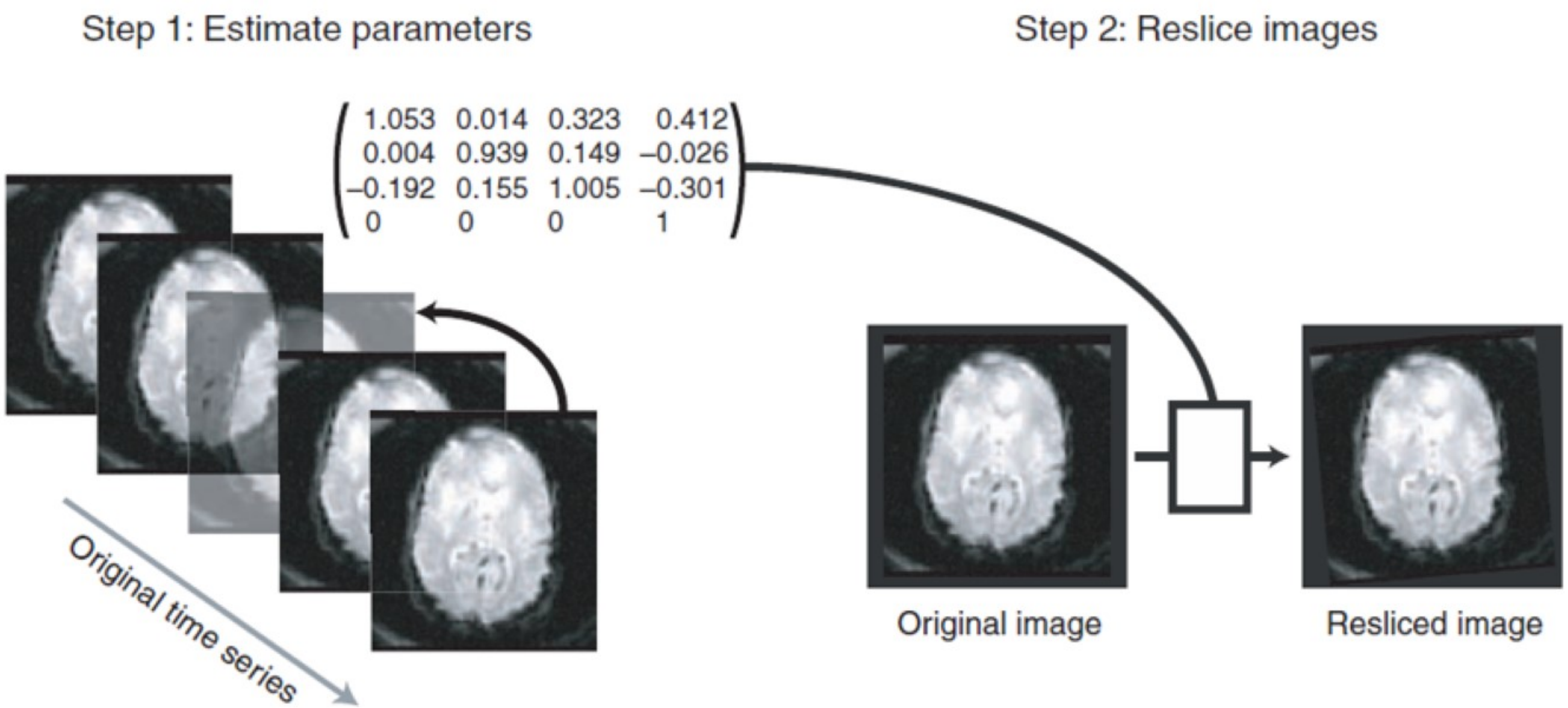
Nonlinear transformation

- Transforming higher-dimensional representations of the image in a non-linear fashion

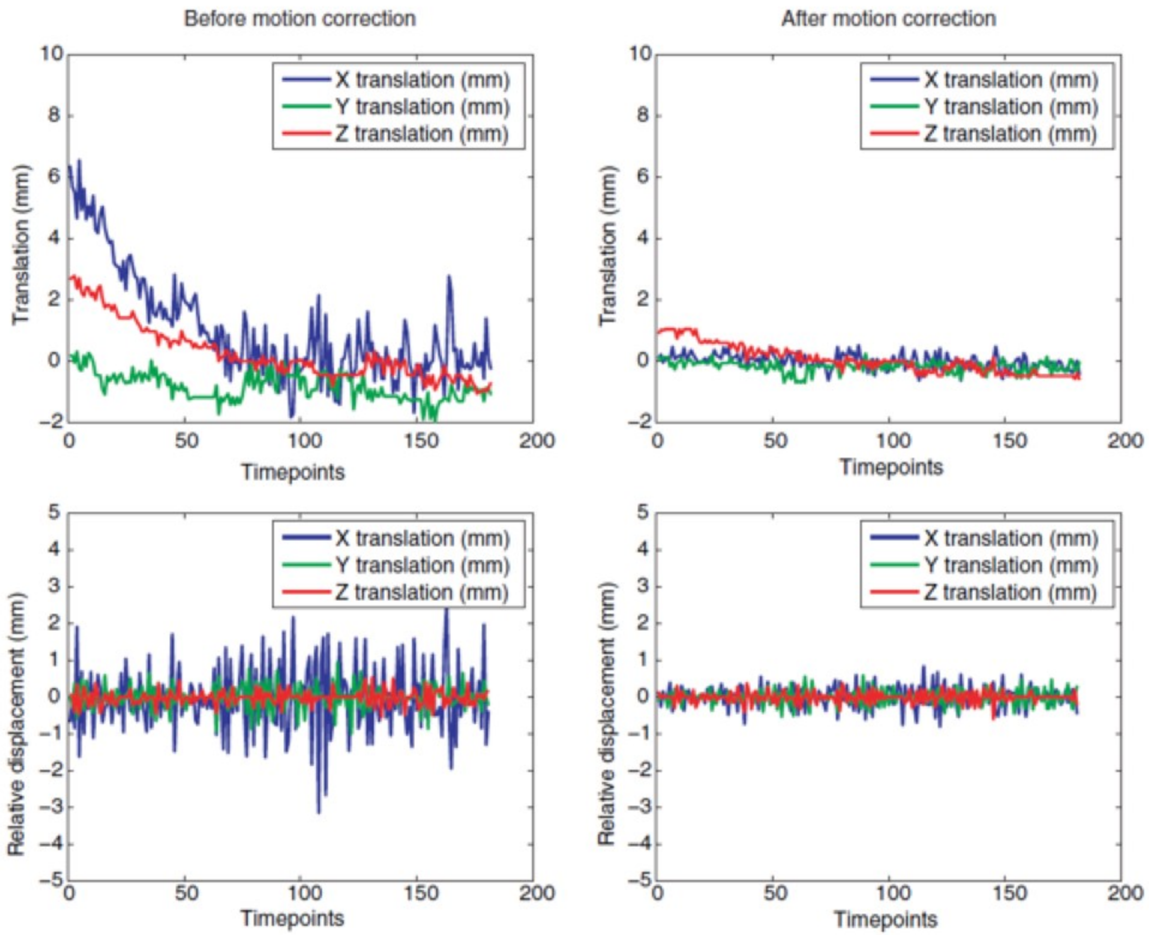
Affine Transformations



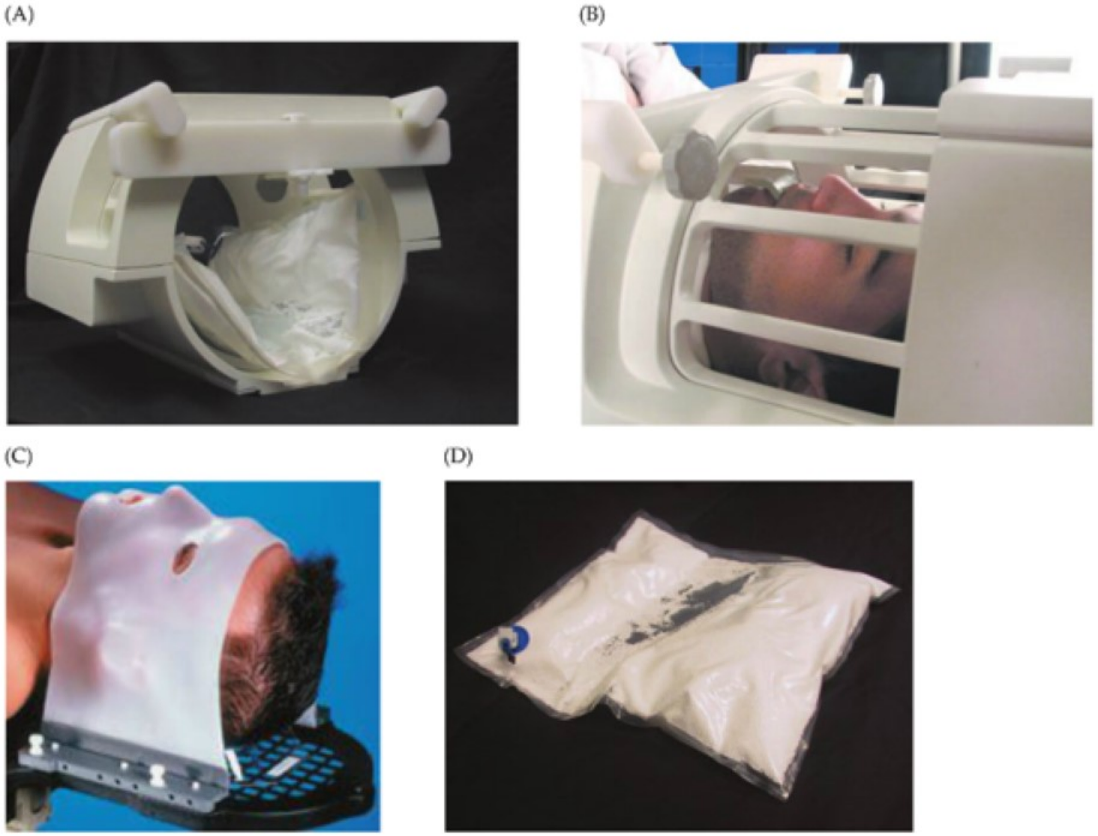
Motion Correction (Realignment)



Plots of Estimated Head Motion



Methods to Prevent Head Motion



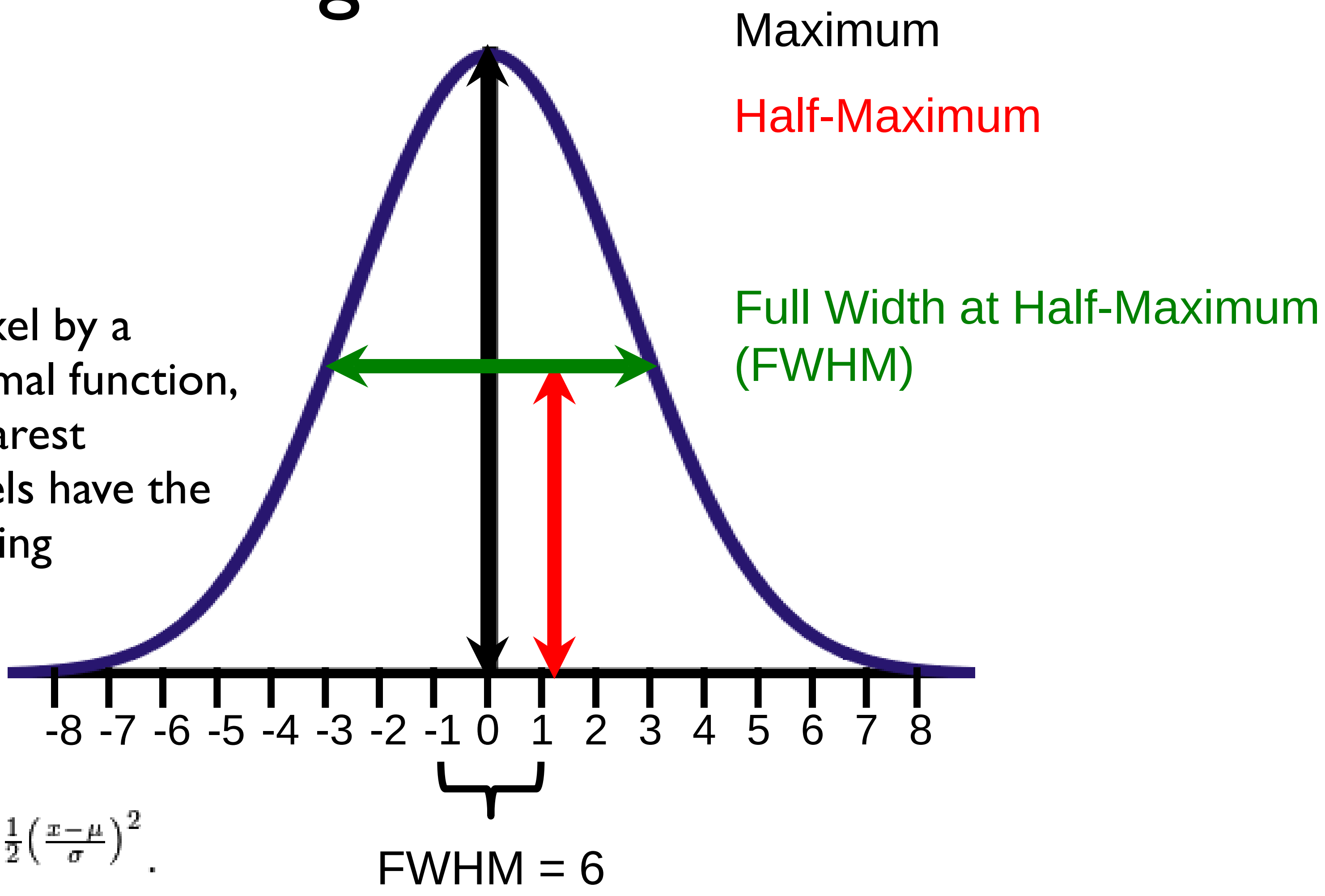
Spatial Smoothing

- Enhancing signal-to-noise ratio
  - By averaged out variation at smaller scale
- Enhancing cross-individual overlap
  - Sacrifice spatial resolution for power
- Fulfilling assumption of data analysis
  - Gaussian random fields

Spatial Smoothing

Gaussian kernel

smooth each voxel by a Gaussian or normal function, such that the nearest neighboring voxels have the strongest weighting

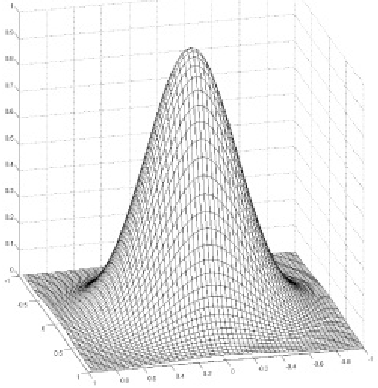


$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

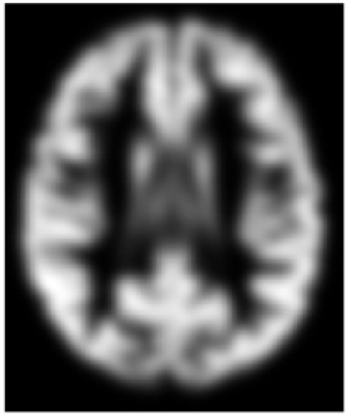
$$FWHM = 2\sigma\sqrt{2\ln(2)}$$

2D Spatial Smoothing

$\otimes$



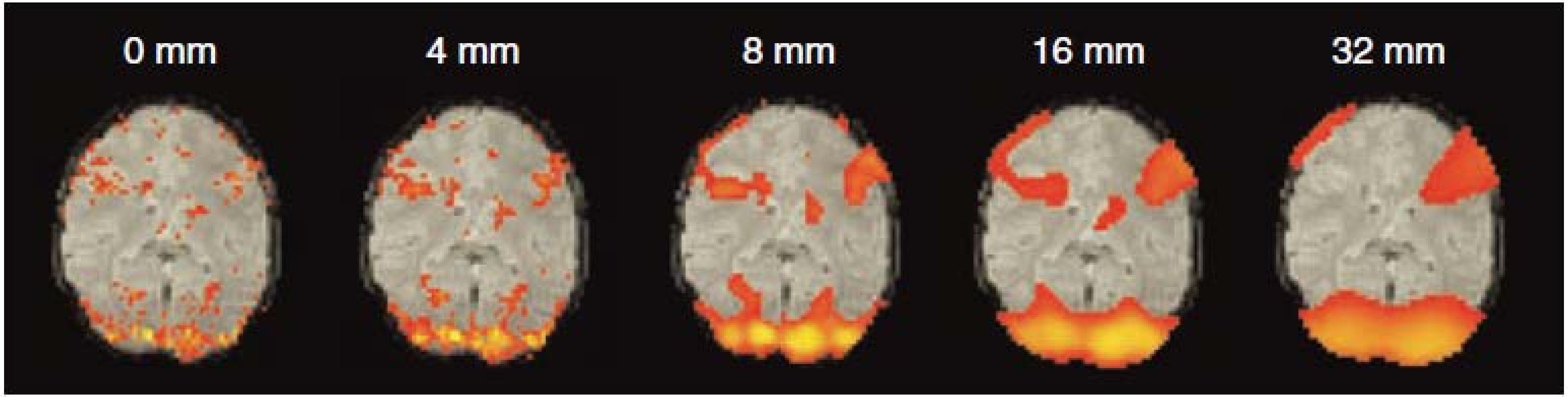
$=$



$$f(x,y) = A \exp\left(-\left(\frac{(x-x_o)^2}{2\sigma_x^2} + \frac{(y-y_o)^2}{2\sigma_y^2}\right)\right).$$

$$FWHM = \sqrt{FWHM_{intrinsic}^2 + FWHM_{applied}^2}$$

Effect of Smoothing on Activation



Should you spatially smooth?

- Advantages
  - Increases Signal to Noise Ratio (SNR)
    - Matched Filter Theorem: Maximum increase in SNR by filter with same shape/size as signal
  - Reduces number of comparisons
    - Allows application of Gaussian Field Theory
  - May improve comparisons across subjects
    - Signal may be spread widely across cortex, due to inter-subject variability

"Why would you spend $4 million to buy an MRI scanner and then blur the data till it looked like PET?"
-- Ravi Menon

- Disadvantages
  - Reduces spatial resolution
  - Challenging to smooth accurately if size/shape of signal is not known

Slide from Duke course

Recommendation for Smoothing

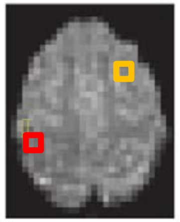
- Noise reduction
  - Filter smaller than expected extent of activation
- Reducing structural variability
  - Variability in the population
  - Efficiency of normalization
- Gaussian random fields assumption
  - FWHM twice the voxel size

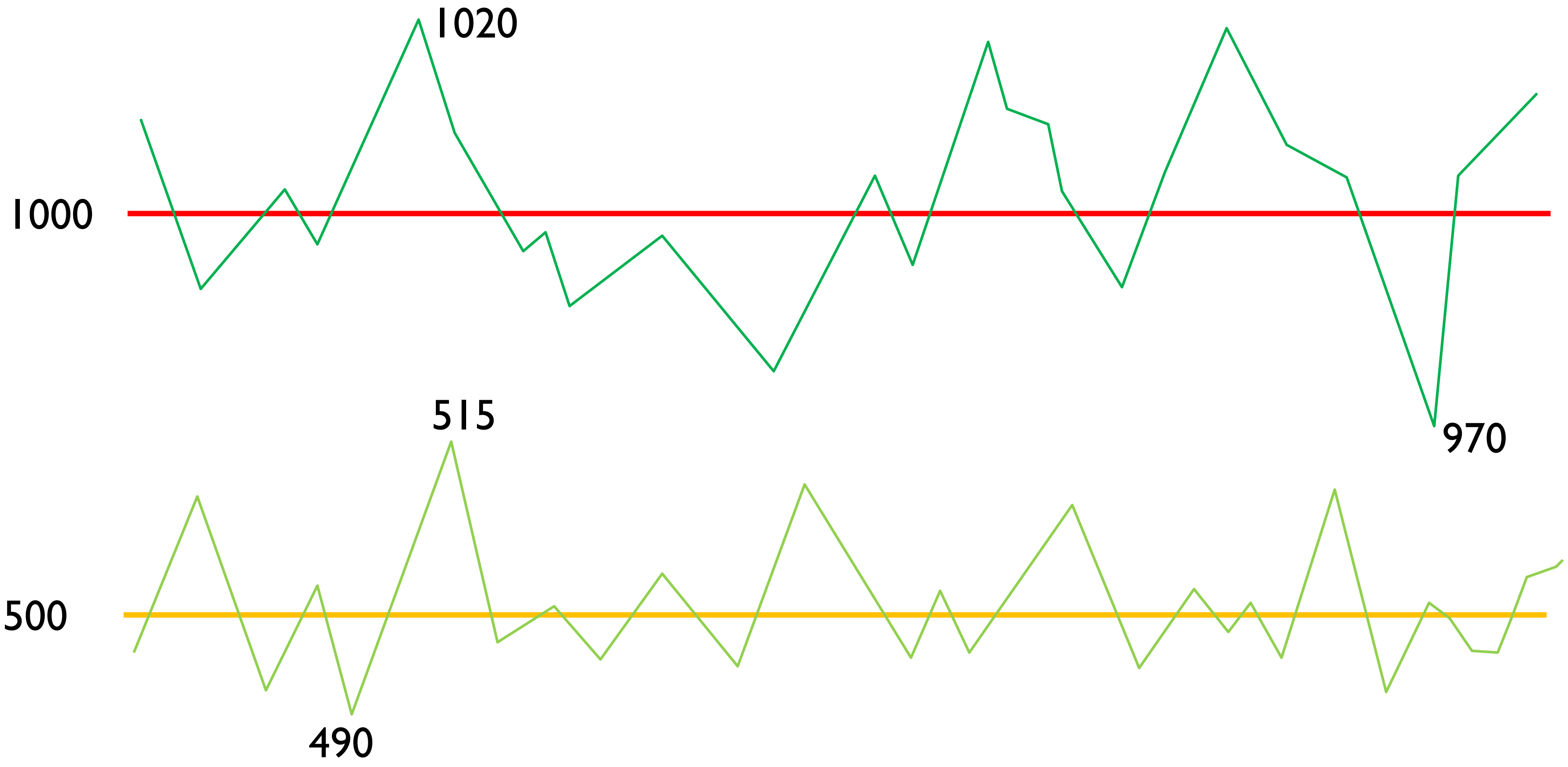
Scaling

- Magnitude of BOLD fluctuation ∝ baseline level

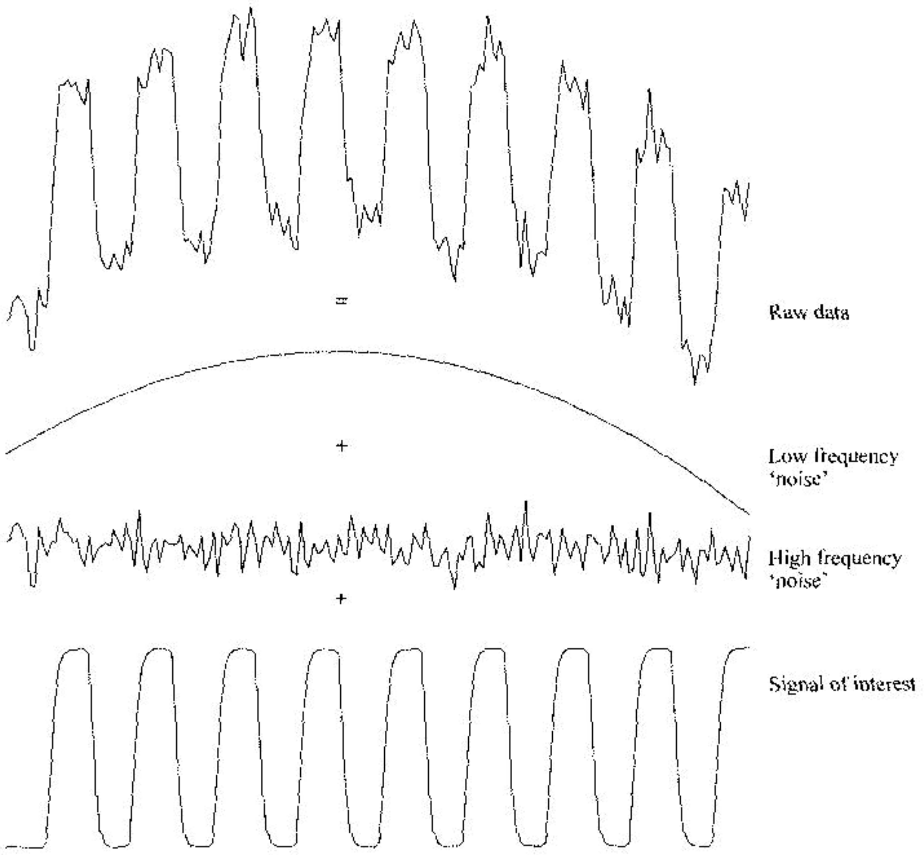
- Scaling time series of each voxel and for each run by the mean of all TRs of that run
  - Percentage
  - Z-score
- For proper comparison of statistics between runs

Voxelwise Scaling



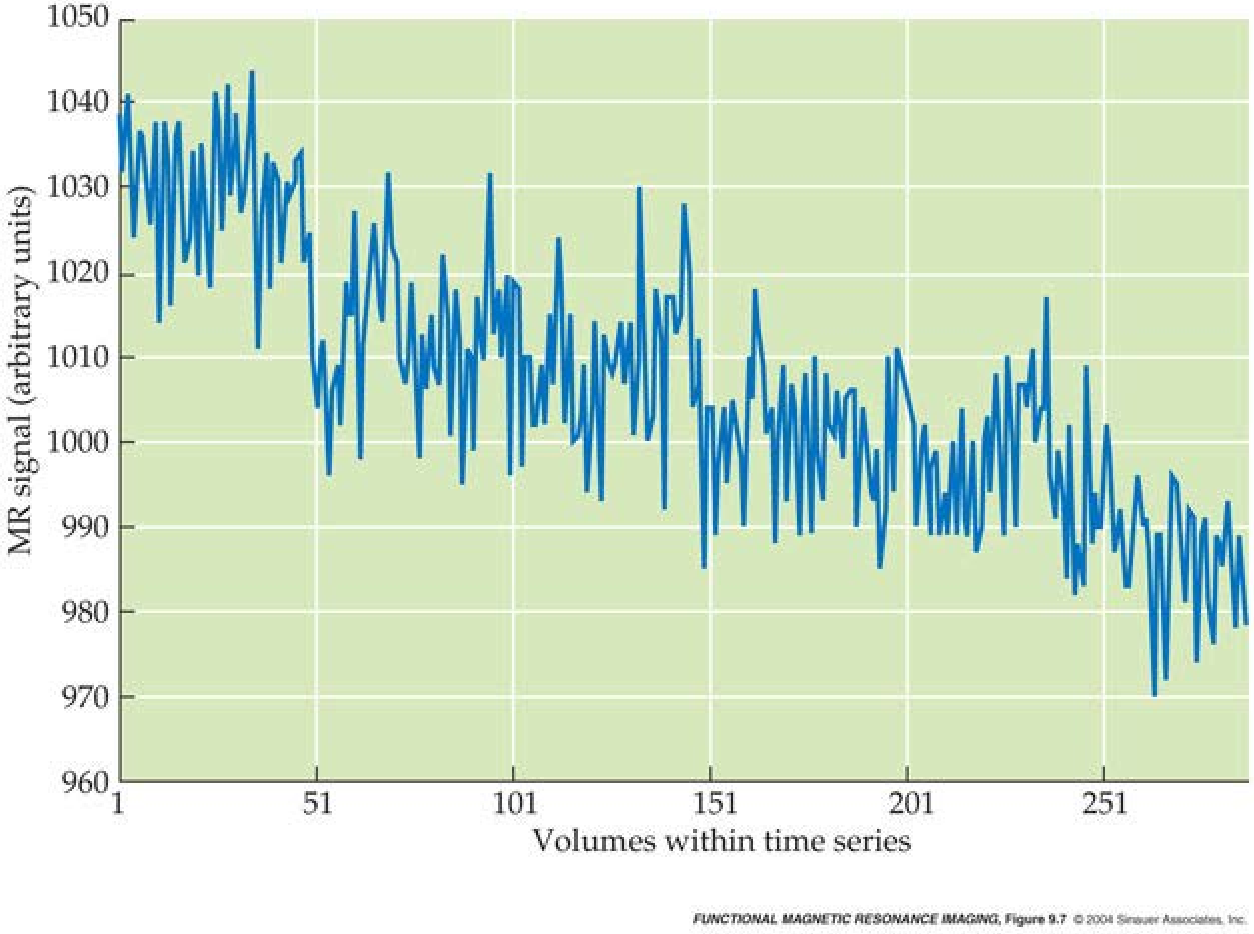


Components of Time Course Data



Source: Smith chapter in Functional MRI: An Introduction to Methods

Linear Drift



Huettel, Song & McCarthy, 2004, Functional Magnetic Resonance Imaging

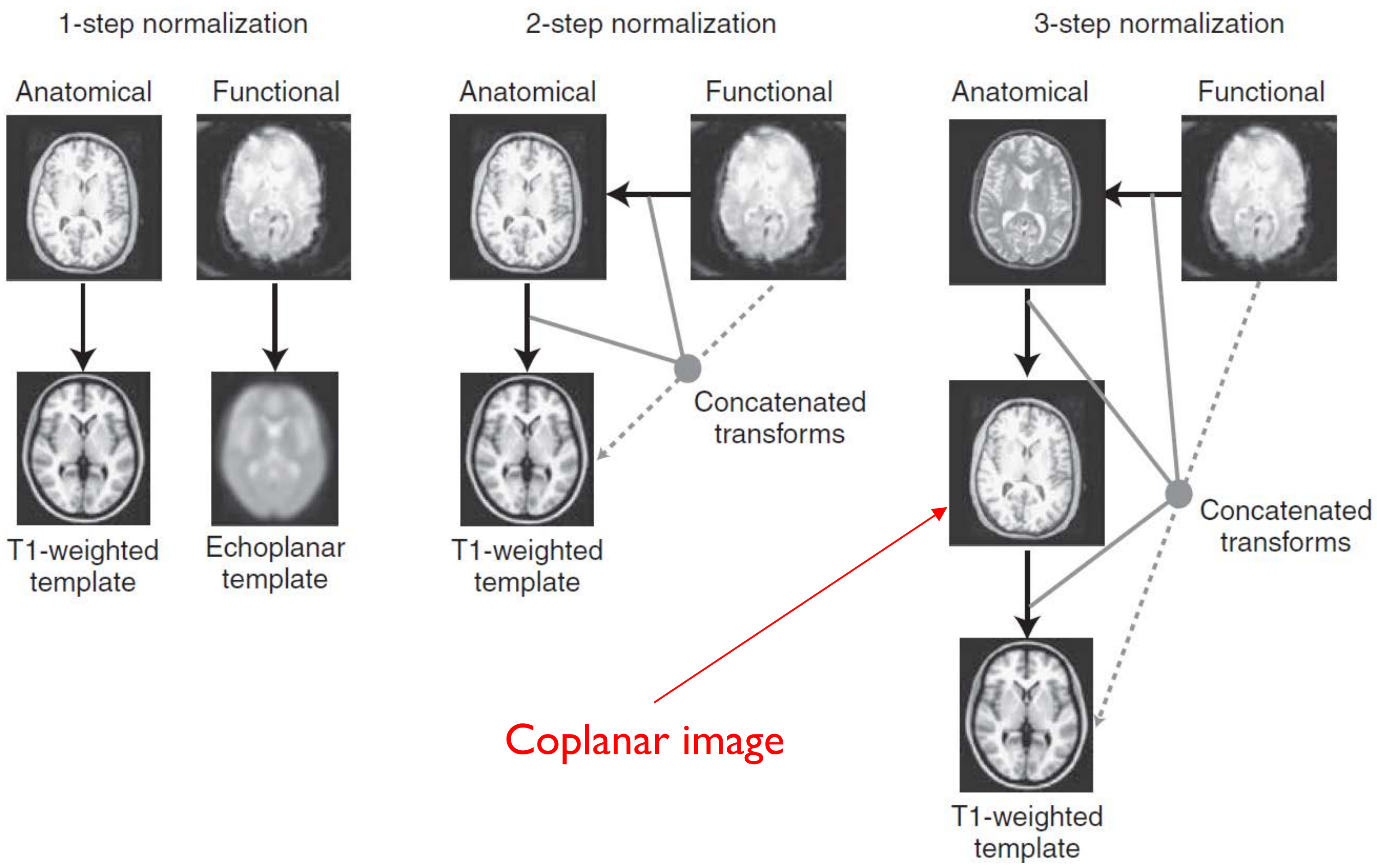
Spatial Normalization

- Process of spatially transforming data into a common space for analysis
  - Aka. Intersubject registration
- Necessary for integrating results from multiple individuals
  - generalization

Prestatistics Approach

- Compute the 1st level GLM for all subjects
  - Preprocessing → GLM
- Spatial normalization of 1st level statistical outputs

Approaches in Operation of Spatial Normalization

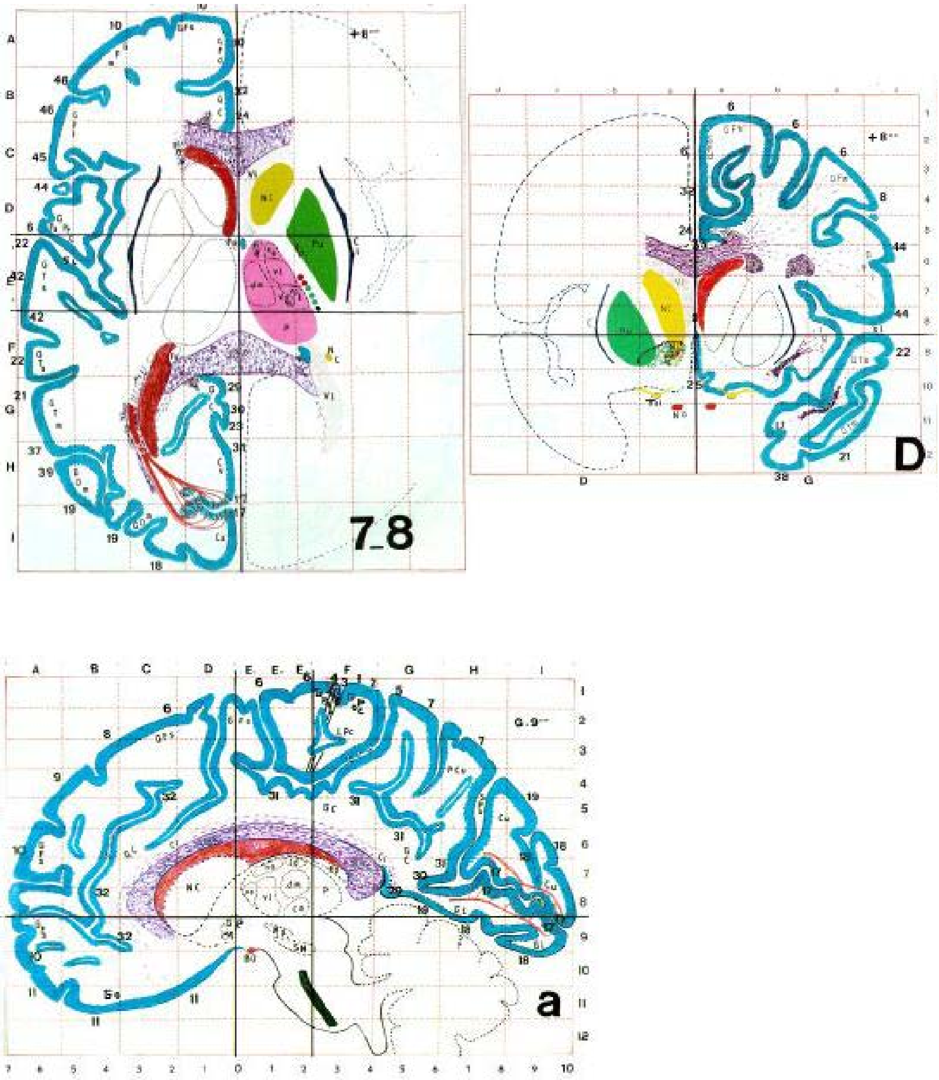


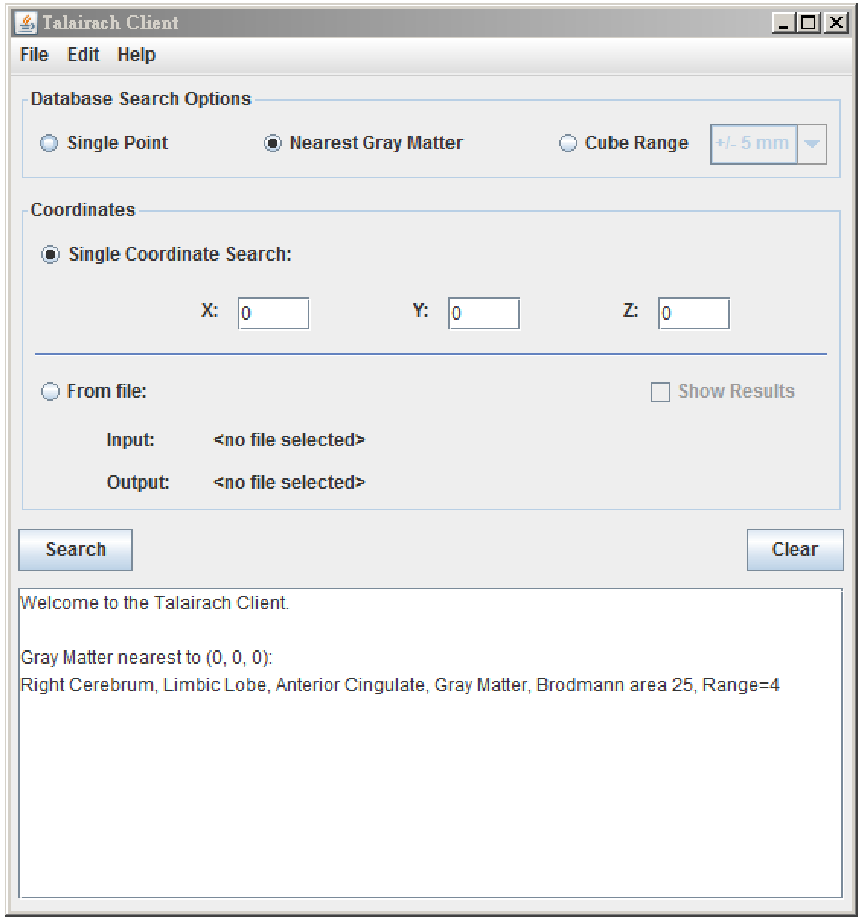
Atlas and Templates

- Atlas
  - Providing a guide to the location of anatomical features in a coordinate space
  - E.g., Talairach atlas

- Template
  - An image representative of the atlas
  - Providing a target where individual image can be aligned
  - E.g., MNI305, ICBM-152, … etc

Talairach Atlas

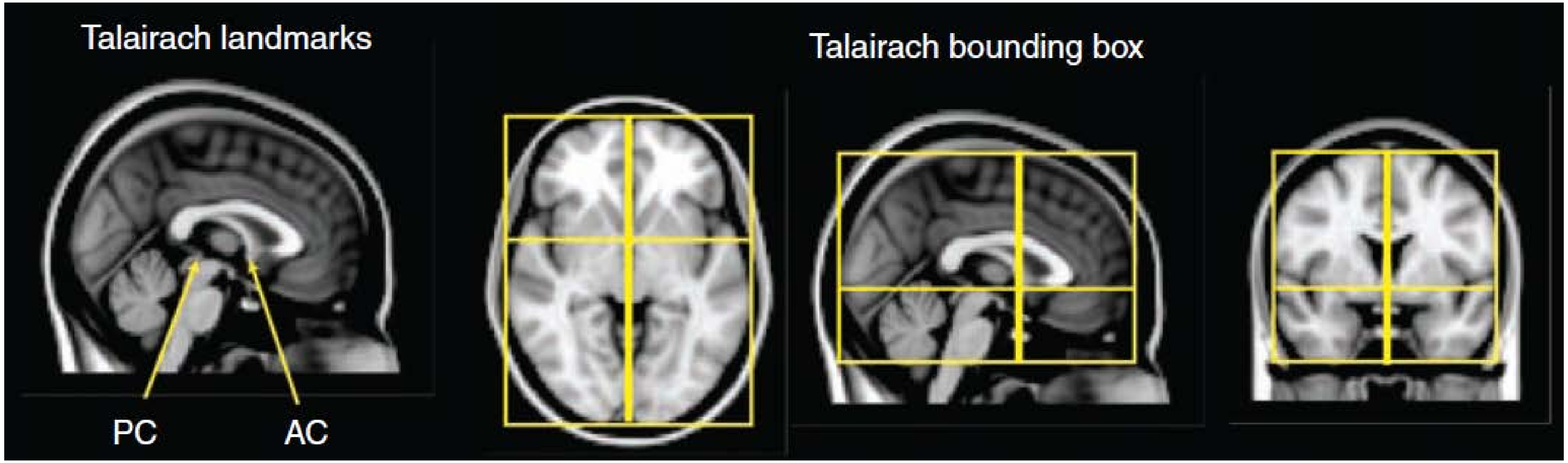


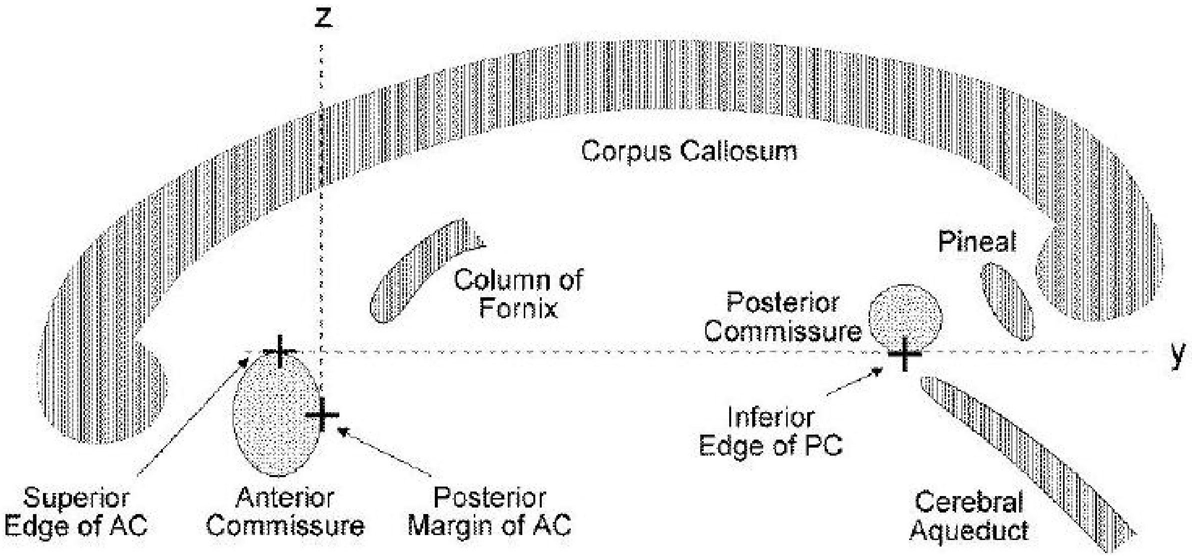


Spatial Normalization Methods

- Landmark-based
  - E.g., Talaraich Landmarks
    - Anterior and posterior commisures, midline sagittal plane, and the exterior boundaries of the brain in each direction
- Volume-based
- Surface-based

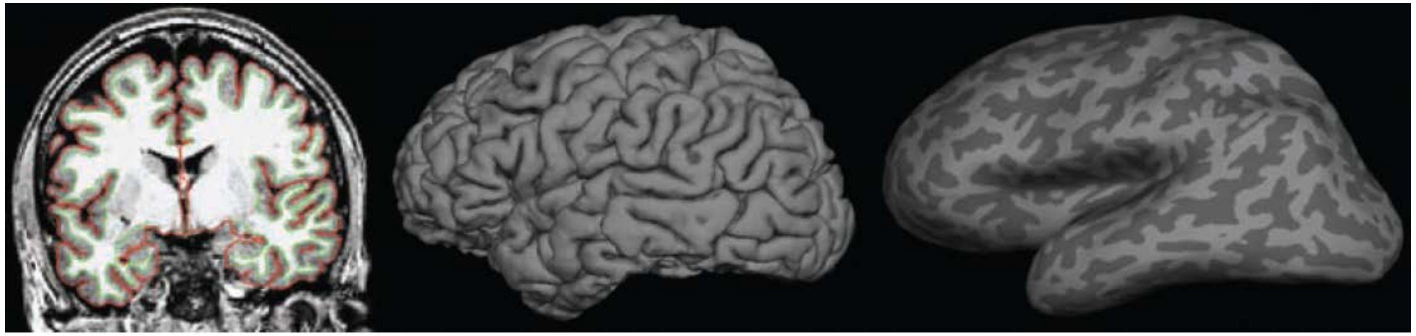
Landmark Based





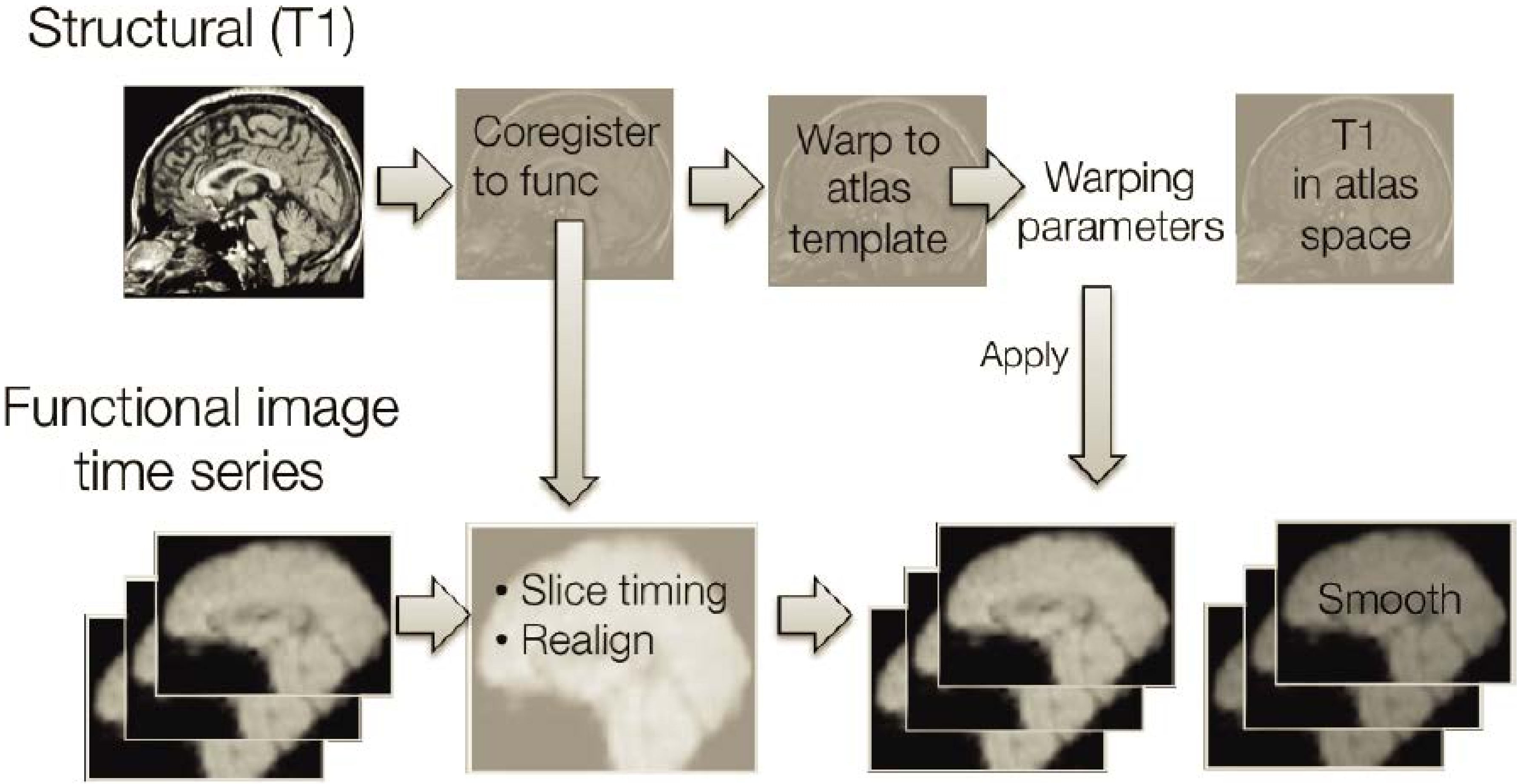
Surface based

- Extraction of cortical surface



- Registration to surface atlas
  - More accurate registration of cortical features
  - Not ready for subcortical structures yet

Summary


Structural (T1)
Coregister
to func
Warp to
atlas
template
Warping
parameters
T1
in atlas
space
Apply
Functional image
time series
• Slice timing
• Realign
Smooth

Questions?

Matrix Expression of GLM

$\boldsymbol{Y = \quad X \cdot \beta \ + \varepsilon}$

- Write out equation for each observation of variable Y from 1 to J:

$$Y_1 = X_{11}\beta_1 + \ldots + X_{1l}\beta_l + \ldots + X_{1L}\beta_L + \varepsilon_1$$

$$Y_j = X_{j1}\beta_1 + \ldots + X_{jl}\beta_l + \ldots + X_{jL}\beta_L + \varepsilon_j$$

$$Y_J = X_{J1}\beta_1 + \ldots + X_{Jl}\beta_l + \ldots + X_{JL}\beta_L + \varepsilon_J$$

Can turn these simultaneous equations into matrix form to get a single equation:

$$\begin{pmatrix} Y_1 \\ Y_j \\ Y_J \end{pmatrix} = \begin{pmatrix} X_{11} \ldots X_{1l} \ldots X_{1L} \\ X_{j1} \ldots X_{1l} \ldots X_{1L} \\ X_{11} \ldots X_{1l} \ldots X_{1L} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_j \\ \beta_J \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_j \\ \varepsilon_J \end{pmatrix}$$

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

Observed data | Design Matrix | Parameters | Residuals/Error

Solution to the Equation

$$\mathrm{X}'\mathrm{Y} = \mathrm{X}'\mathrm{X}\boldsymbol{\beta}$$

$$\hat{\sigma}^2 = \frac{\mathrm{e}'\mathrm{e}}{T-(p+1)}$$

Any $\boldsymbol{\beta}$ satisfies the normal equation minimizes the sum of the squares of residuals (e'e)

$$\hat{\boldsymbol{\beta}} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{Y}$$

Assuming this is invertible

Hypothesis Testing: Contrast t-test

$$\mathbf{c}\hat{\boldsymbol{\beta}} \sim N(0, \mathbf{c}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}'\sigma^2)$$

$$t = \frac{\mathbf{c}\hat{\boldsymbol{\beta}}}{\sqrt{\mathbf{c}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}'\hat{\sigma}^2}}$$

df: $T-(p+1)$

$$H_A : \mathbf{c}\boldsymbol{\beta} > 0$$

$$P(T_{T-(p+1)} \geq t)$$

$$H_A : \mathbf{c}\boldsymbol{\beta} \neq 0$$

$$P(T_{T-(p+1)} \geq |t|)$$

$$t = \frac{c\hat{\beta}}{\sqrt{\boxed{c(X'X)^{-1}c'}\boxed{\hat{\sigma}^2}}}$$

Design matrix & Contrast Vector; depending on your experimental design

Residual error unaccounted for by your design; depending on the quality of data