

fMRI Course, Day 8:
Design Optimization and
2nd-Level Analysis
August 8th, 2021

Questions from previous lecture?

How did the lab go?

Did the contrasts make sense?

What was the most confusing part of the lab?

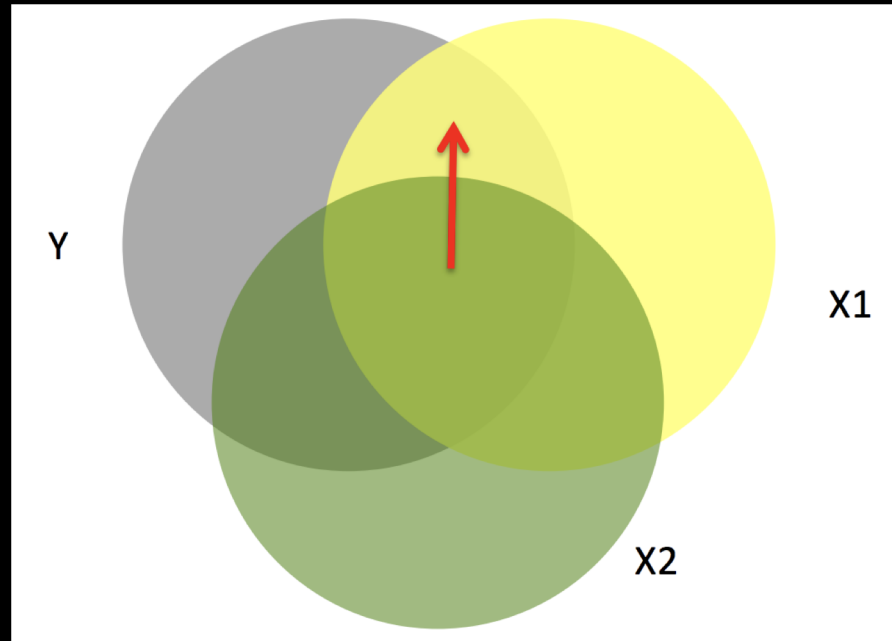
Questions:

1. How to increase the size of the t-statistic?
2. How to deal with low-frequency noise?
3. Should we lowpass filter fMRI data?

How did the lab go?

What is collinearity? How can we reduce it?

If I orthogonalize X_2 with respect to X_1 , what does that mean?



Today's Lecture

Design optimization and power analysis

OptimizeX, Gpower, optseq

Group-level analysis options

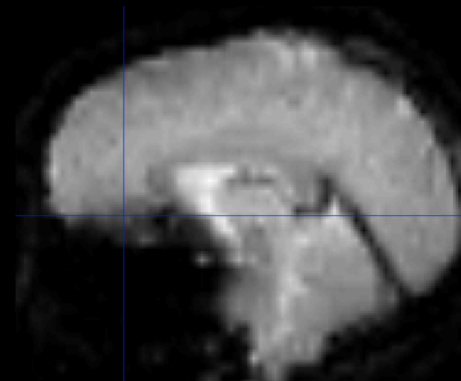
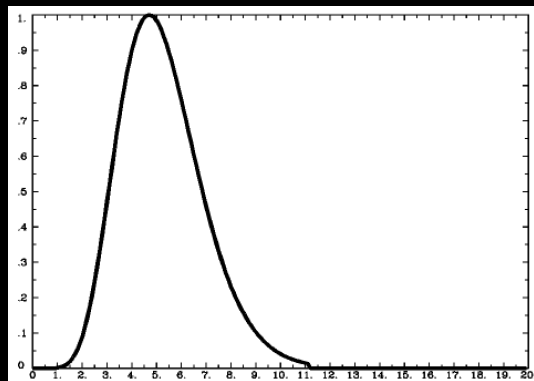
Other Statistical Scenarios

Review So Far

From stimulus to the BOLD response

How tissue properties, blood flow, and magnetic properties interact

Creating contrast images from T1- and T2-weightings



Preprocessing Steps

Brain extraction (or “skull stripping”)

Motion Correction

Slice Timing Correction

Smoothing

Registration

Normalization

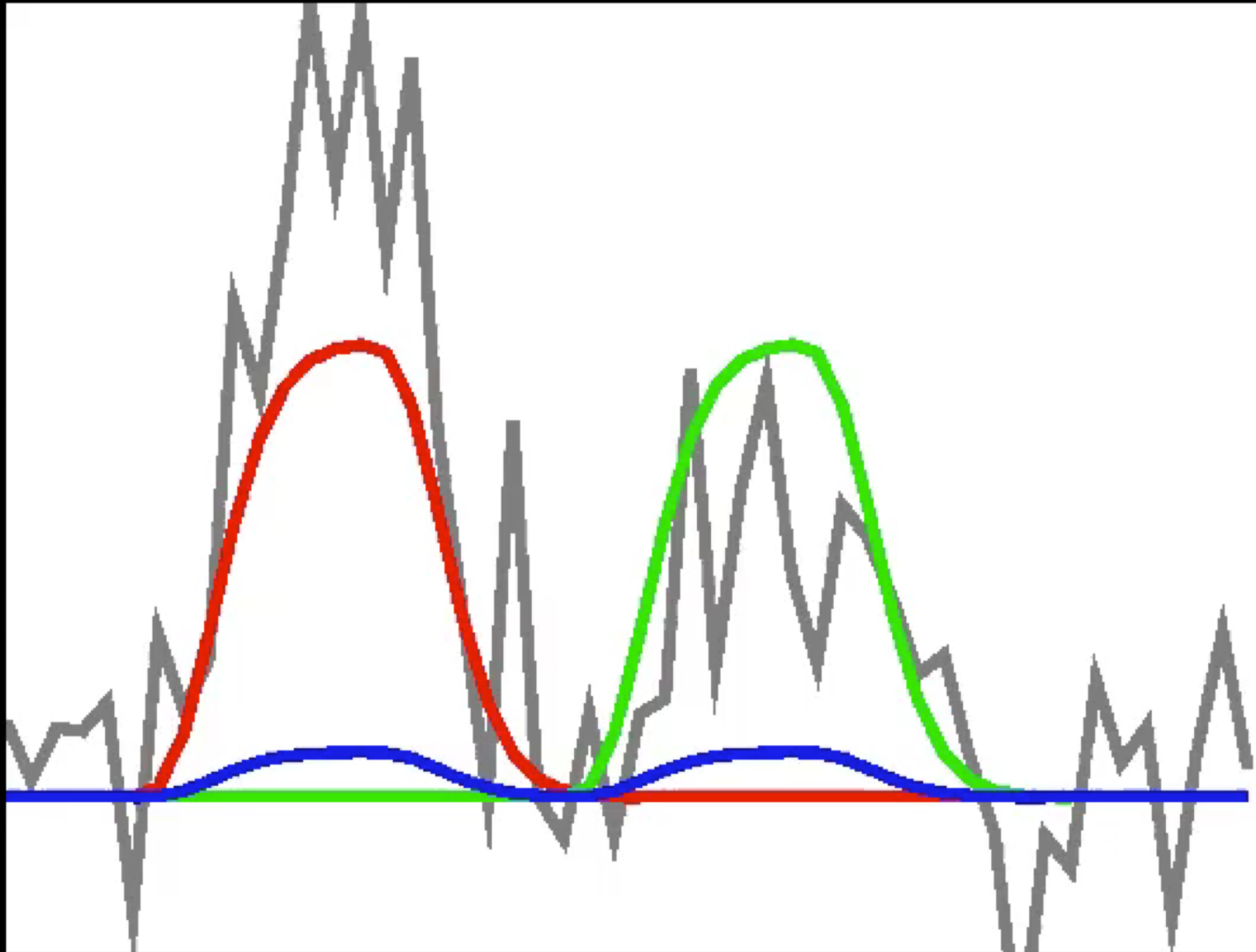
Temporal Filtering

Review So Far

Overview of the General Linear Model

Parameter Estimates

Creating beta maps and contrast maps

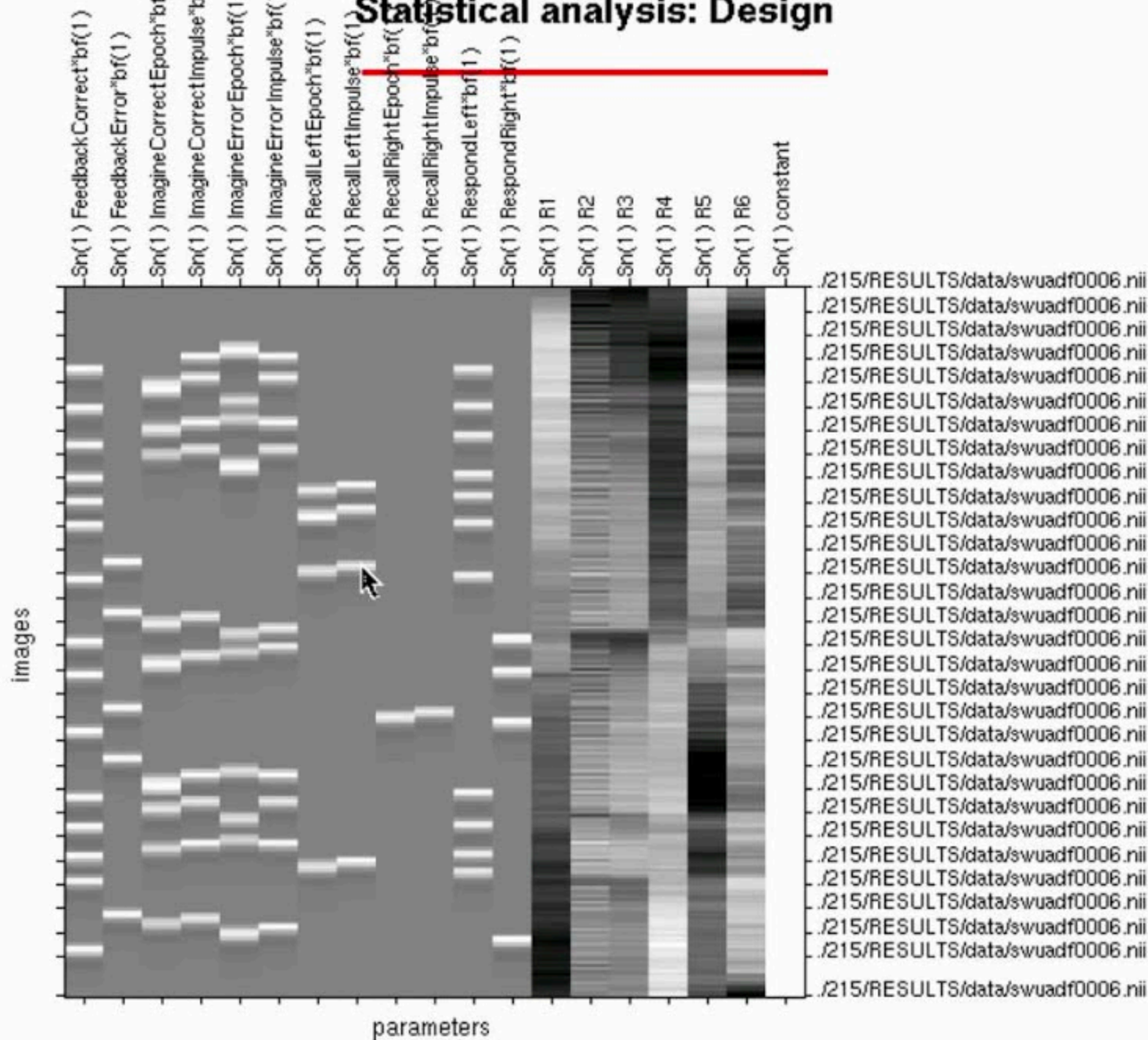


Fit at each voxel (“mass univariate” approach)

Questions

Let's look at some GLMs, and see if you can identify each part!

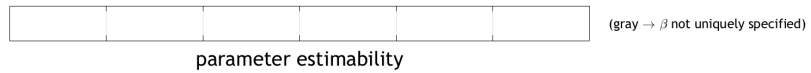
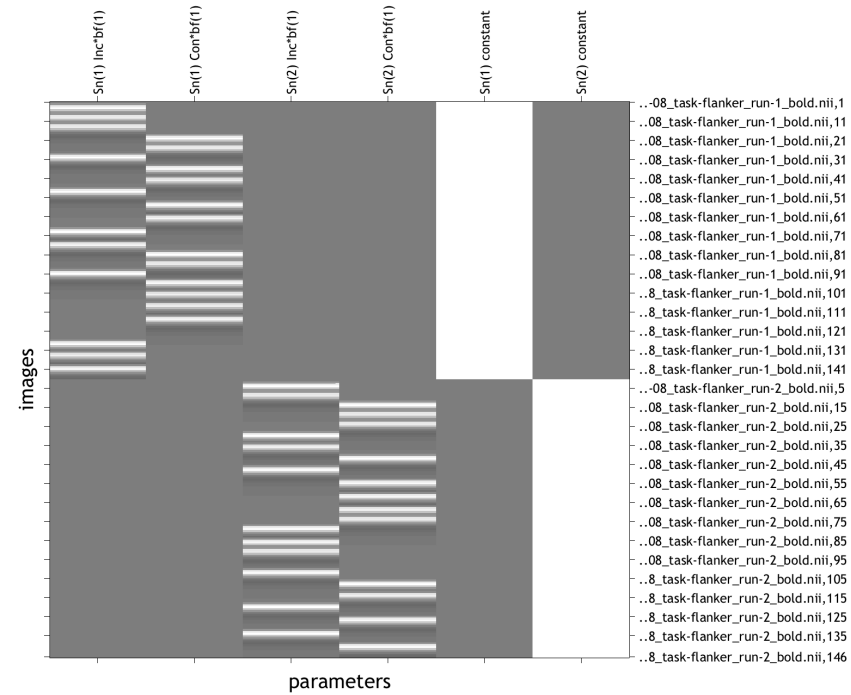
Statistical analysis: Design



ultis_2

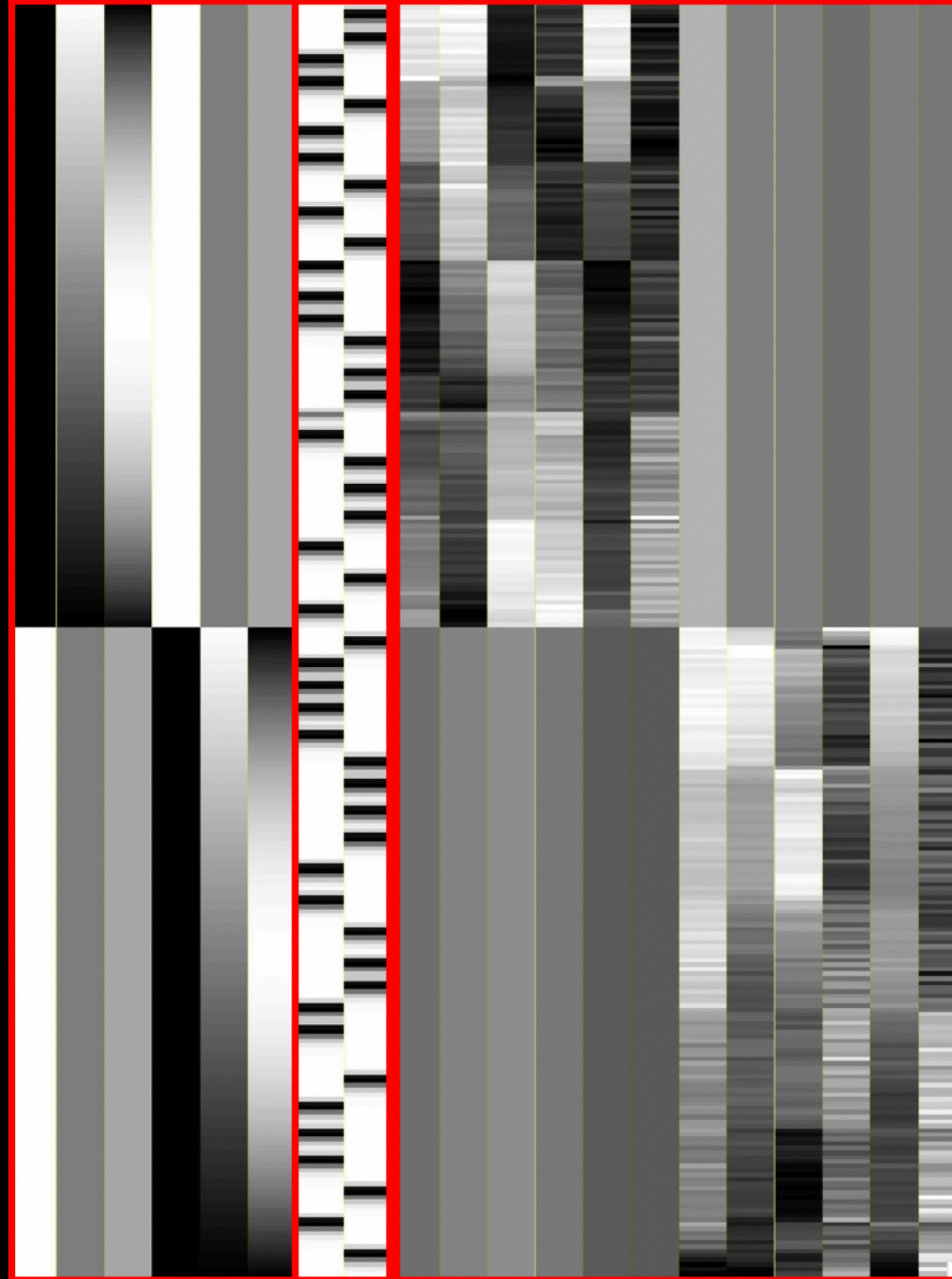
ultis_2/

Statistical analysis: Design



Design description...

Basis functions : hrf
 Number of sessions : 2
 Trials per session : 2 2
 Interscan interval : 2.00 [s]
 High pass Filter : [min] Cutoff: 128 [s]
 Global calculation : mean voxel value
 Grand mean scaling : session specific
 Global normalisation : None



Review: SPM Terms for Analysis

1st-Level Analysis: Individual subject (all trials across runs)

**2nd-Level Analysis: Group-Level Analysis
(all subjects within the experiment)**

Review of Collinearity

Last week, we looked at a correlation matrix

SPM12 (7771): Menu

Realign (E... ▾) Slice timing Smooth

Coregiste... ▾ Normalise... ▾ Segment

Select your SPM.mat file

Specify 1st-level Review

Specify 2nd-level Estimate

Results

Dynamic Causal Modelling

SPM for functional MRI

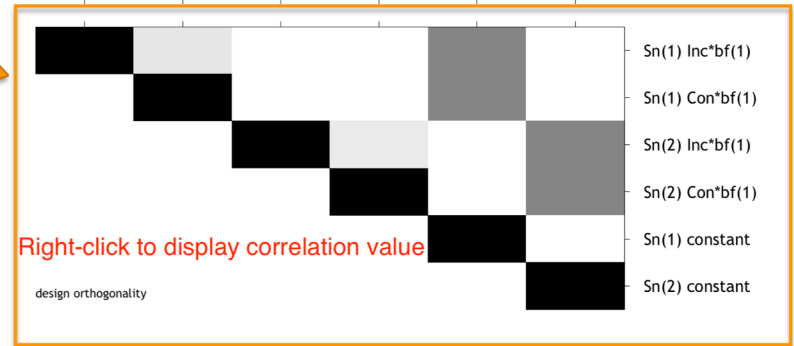
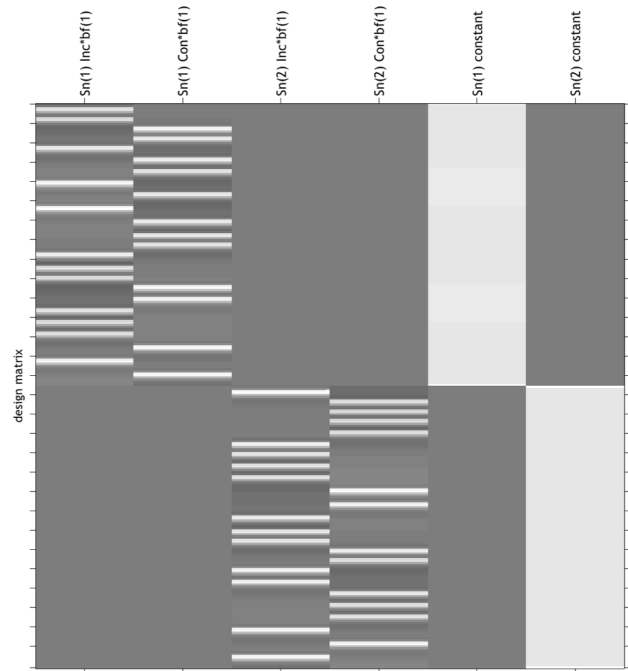
Display Check Reg Render... ▾ FMRI ▾

Toolbox: PPIs ImCalc DICOM Import

Help Utils... ▾ Batch Quit

Design

Statistical analysis: Design orthogonality



Measure : abs. value of cosine of angle between columns of design matrix
 Scale : black - colinear (cos=+1/-1)
 white - orthogonal (cos=0)
 gray - not orthogonal or colinear


```

DR PUBLISH VIEW
insert fx fx
ment % %
dent % %
EDIT BREAKPOINTS RUN
Run Section
Run Run and Advance Run and Time
Advance Run and Time
Flanker ▶ sub-01 ▶ 1stLevel
Editor - /Users/ajahn/Downloads/fmridata/LIPREAD/Subjects/Decoding_audOnly_RSA_ROI.m
convertOnsetTimes.m Decoding_audOnly_RSA_ROI.m Haxby_RSA.m Haxby_MVPA_ROI.m decoding.m decoding_generate
% If you like to combine multiple designs in one cfg.
%% Decide whether you want to see the searchlight/ROI/... during
cfg.plot_selected_voxels = 500; % 0: no plotting, 1: every step
%% Add additional output measures if you like
% See help decoding_transform_results for possible measures
% cfg.results.output = {'accuracy_minus_chance', 'AUC'}; % 'acc
% You can also use all methods that start with "transres_", e.g
% cfg.results.output = {'SVM_pattern'};
% will use the function transres_SVM_pattern.m to get the patte
% linear svm weights (see Haufe et al, 2015, Neuroimage)
%% Nothing needs to be changed below for a standard leave-one-r
%% validation analysis.
% The following function extracts all beta names and correspond
% numbers from the SPM.mat
regressor_names = design_from_spm(['/Users/ajahn/Downloads/fmri
% Extract all information for the cfg.files structure (labels w
cfg = decoding_describe_data(cfg,labelnames,labels,regressor_na
% This creates the leave-one-run-out cross validation design:
cfa_design = make_design_cv(cfg):

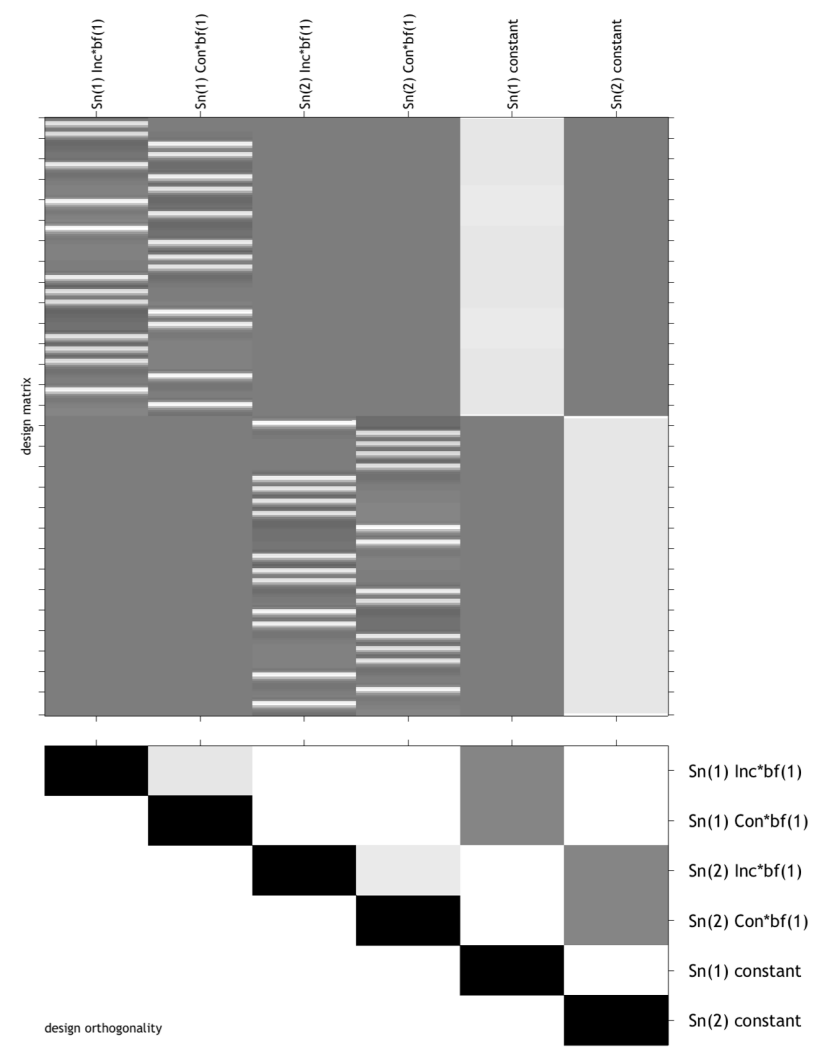
```

Command Window

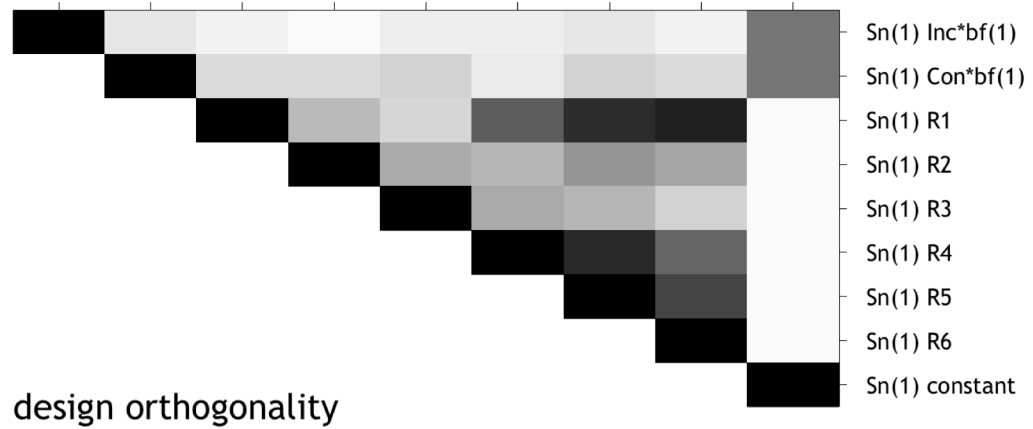
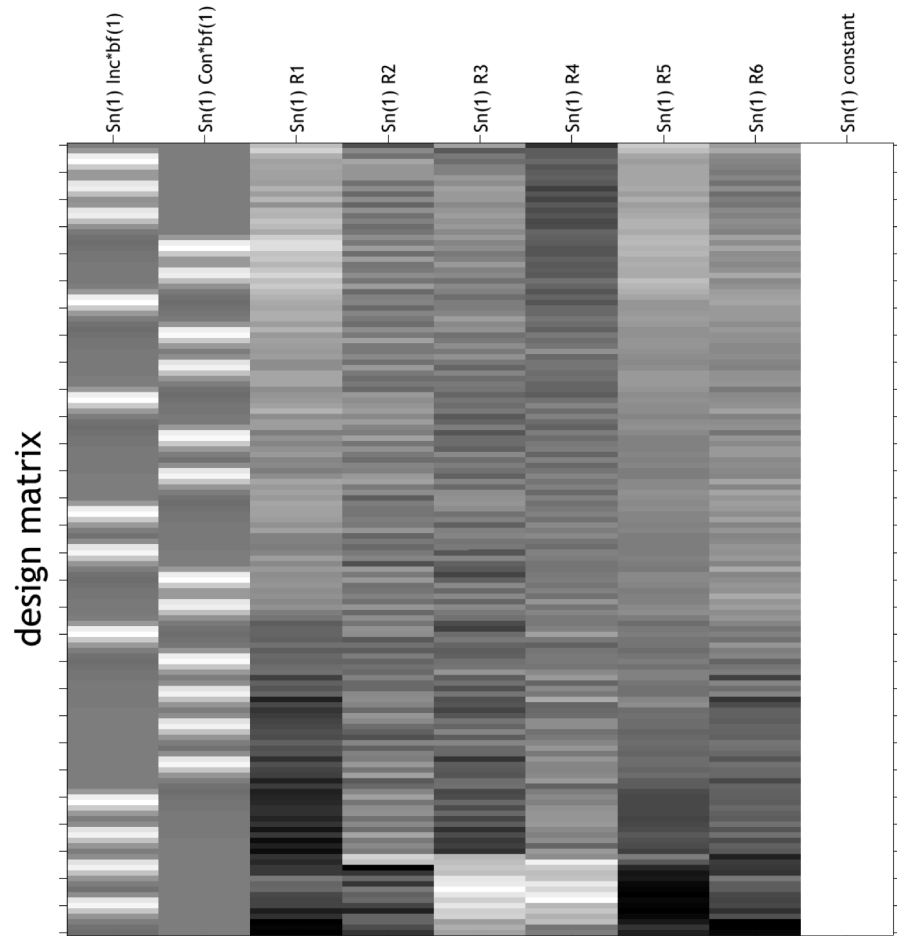
ans =						
1.0000	-0.0925	0	0	0.4635	0	
-0.0925	1.0000	0	0	0.4625	0	
0	0	1.0000	-0.0655	0	0.4588	
0	0	-0.0655	1.0000	0	0.4655	
0.4635	0.4625	0	0	1.0000	0	
0	0	0.4588	0.4655	0	1.0000	

fx >>

Statistical analysis: Design orthogonality



Measure : abs. value of cosine of angle between columns of design matrix
Scale : black - colinear (cos=+1/-1)
white - orthogonal (cos=0)
gray - not orthogonal or colinear



Review of Collinearity

Rule of thumb: Correlations of 0.4 or greater are considered “moderate” (source: AFNI command xmat_tool.py)

However, a high correlation between one set of regressors may not matter, given the regressors you are focused on

The challenge is to include as many regressors as is reasonable, without overfitting or introducing collinearity

Variance Inflation Factor

What what if one regressor is a linear sum of two or more other regressors? Pairwise correlations don't show this

e.g., does $X_1 = X_2 + X_3$?

Variance Inflation Factor

$$X_1 = \beta_0 + \beta_1 X_2 + \beta_2 X_3 + \epsilon$$

$$\mathbf{R} = \text{Cor}(X_1, \hat{\beta}_0 + \hat{\beta}_1 X_2 + \hat{\beta}_2 X_3) = \text{Cor}(X_1, \widehat{X}_1)$$

Example

X_1	X_2	X_3
1	1	0
2	2	0
3	3	0
4	4	0
5	5	0
1	0	1
2	0	2
3	0	3
4	0	4
5	0	5

$$\text{cor}(X_1, X_2) = \text{cor}(X_1, X_3) = .39$$

$$X_1 = \beta_1 X_2 + \beta_2 X_3$$

What's the solution for the betas?

$$X_1 = \widehat{X}_1 \implies \text{cor}(X_1, \widehat{X}_1) = 1$$

Example

$$VIF = \frac{1}{1 - R^2} = \frac{1}{1 - [\text{cor}(X_1, \widehat{X}_1)]^2}$$

Goal: VIF < 5

What is a cutoff? VIF no greater than 10

Repeat the same process for each regressor in the model

VIF: Summary

Detects any collinearity from any combination of regressors

Matlab code: `vif = diag(inv(corrcoef(X)))'`;

Where “X” is the design matrix

VIF: Summary

Useful for checking whether a design has high collinearity

Solutions: Remove the regressor, or change the design

These edits can be done before scanning

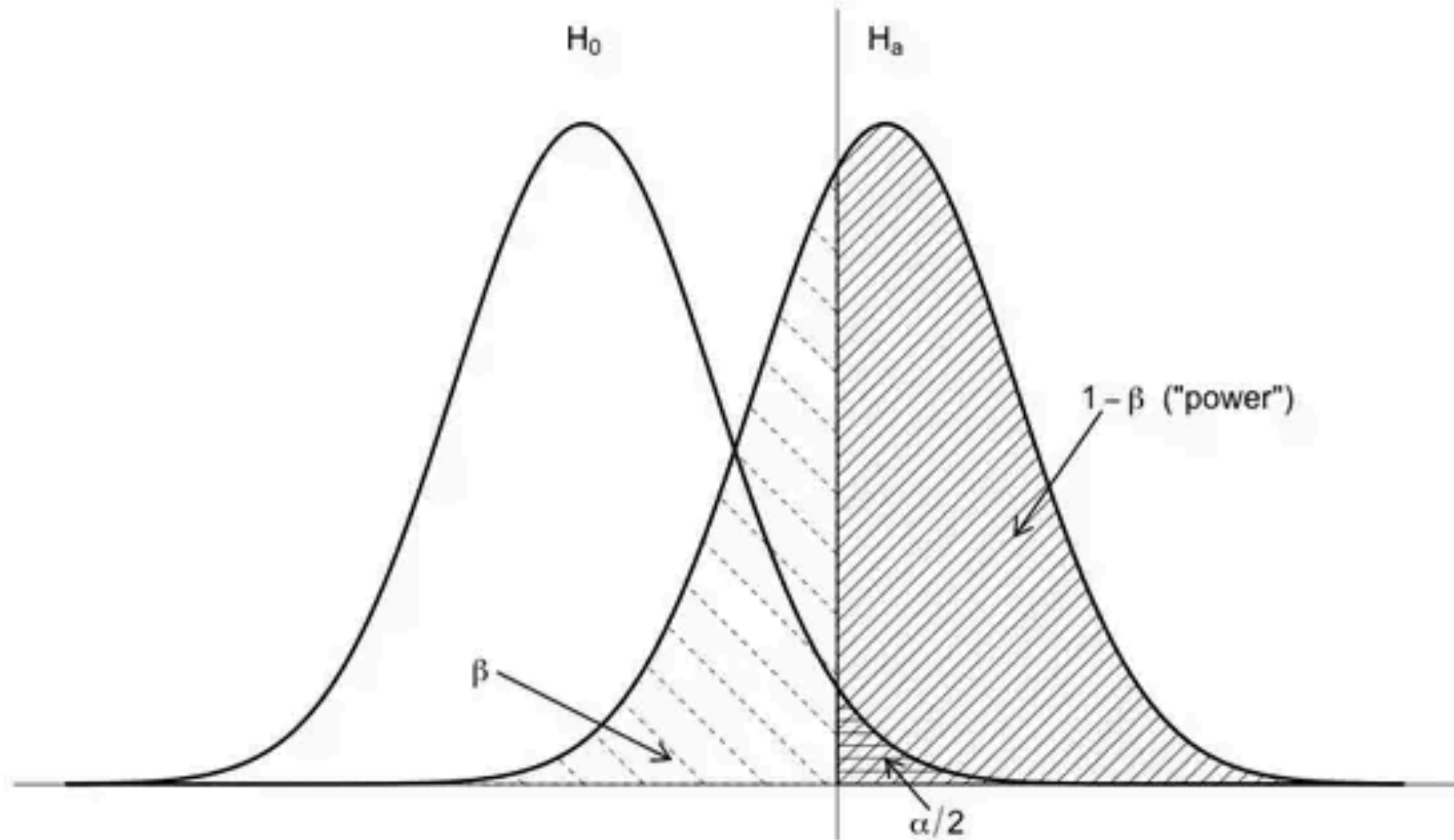
Efficiency vs. Power

Two terms you will come across are Efficiency and Power

Let's begin with Power: Can you detect an effect if it is actually there?

As variance increases, power decreases

Statistical Power



Efficiency vs. Power

Efficiency is inversely proportional to variance

$$\frac{1}{c(X'X)^{-1}c'}$$

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

$$\hat{\sigma}^2 = \frac{e'e}{N-p}$$

where $e = Y - X\hat{\beta} = Y - \hat{Y}$

All measures of efficiency are relative

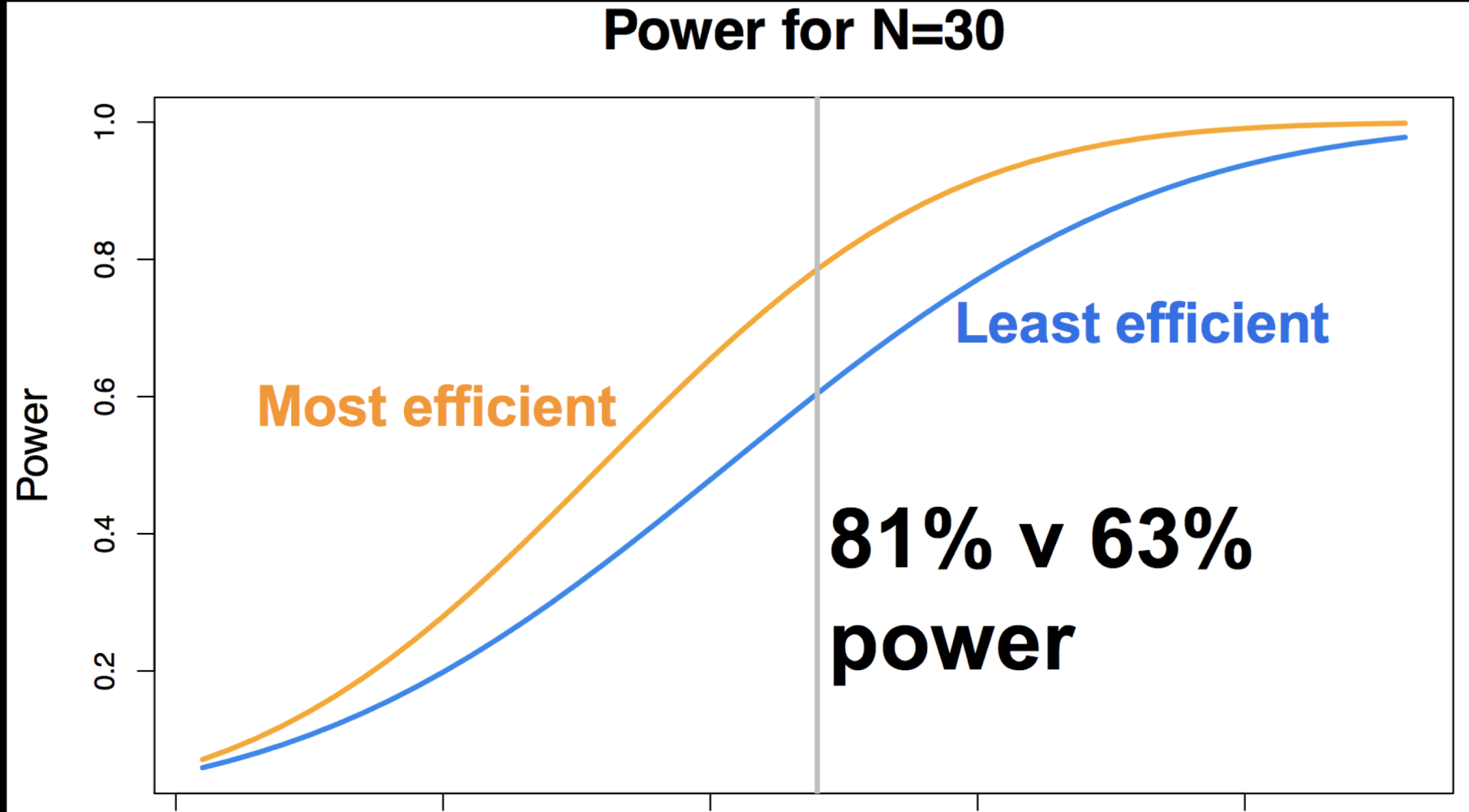
Efficiency vs. Power

Example: You have a fixed number of subjects that you can scan (due to budget, population, etc.)

Options:

- 1. Scan longer**
- 2. Include more trials**
- 3. Increase ITI**
- 4. Create a more efficient experiment**

Efficiency vs. Power



Efficiency vs. Power

Figure 2:
Slow Event-Related Design -
Fixed Inter-Stimulus Interval

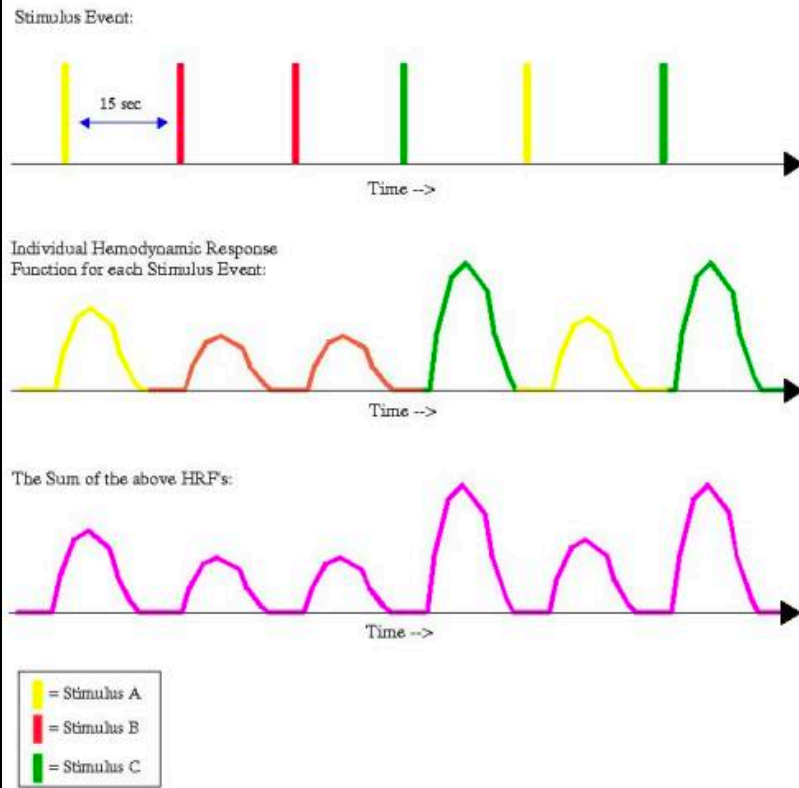


Figure 3:
Rapid Event-Related Design -
Fixed ISI and Nonrandom Stimulus Presentation

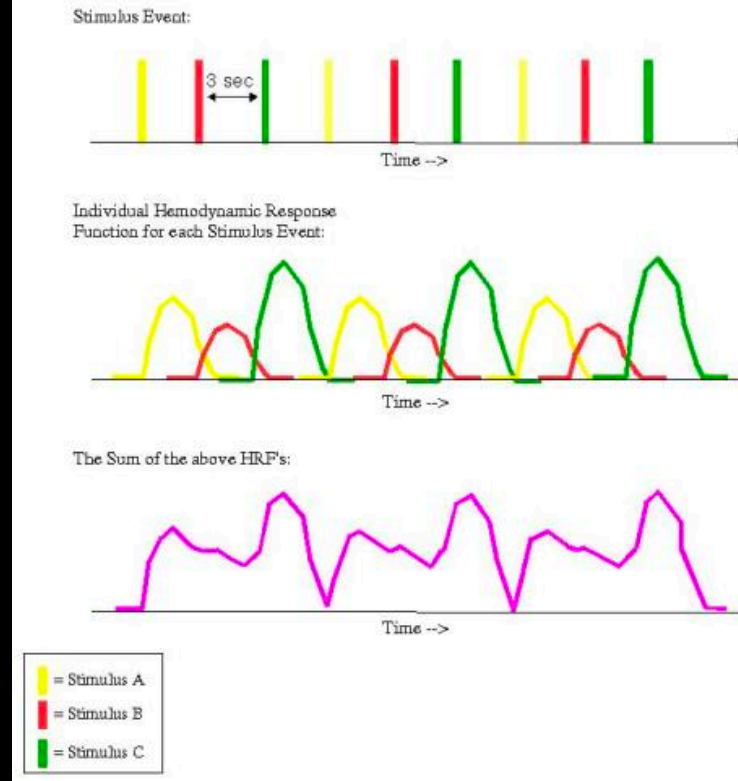
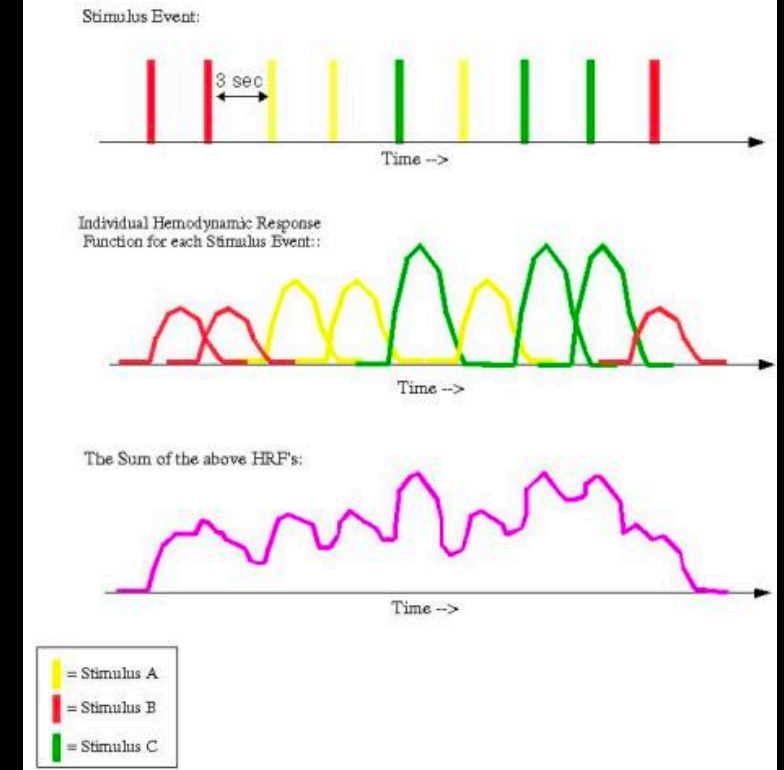


Figure 4:
Rapid Event-Related Design -
Fixed ISI and Randomized Stimulus Presentation



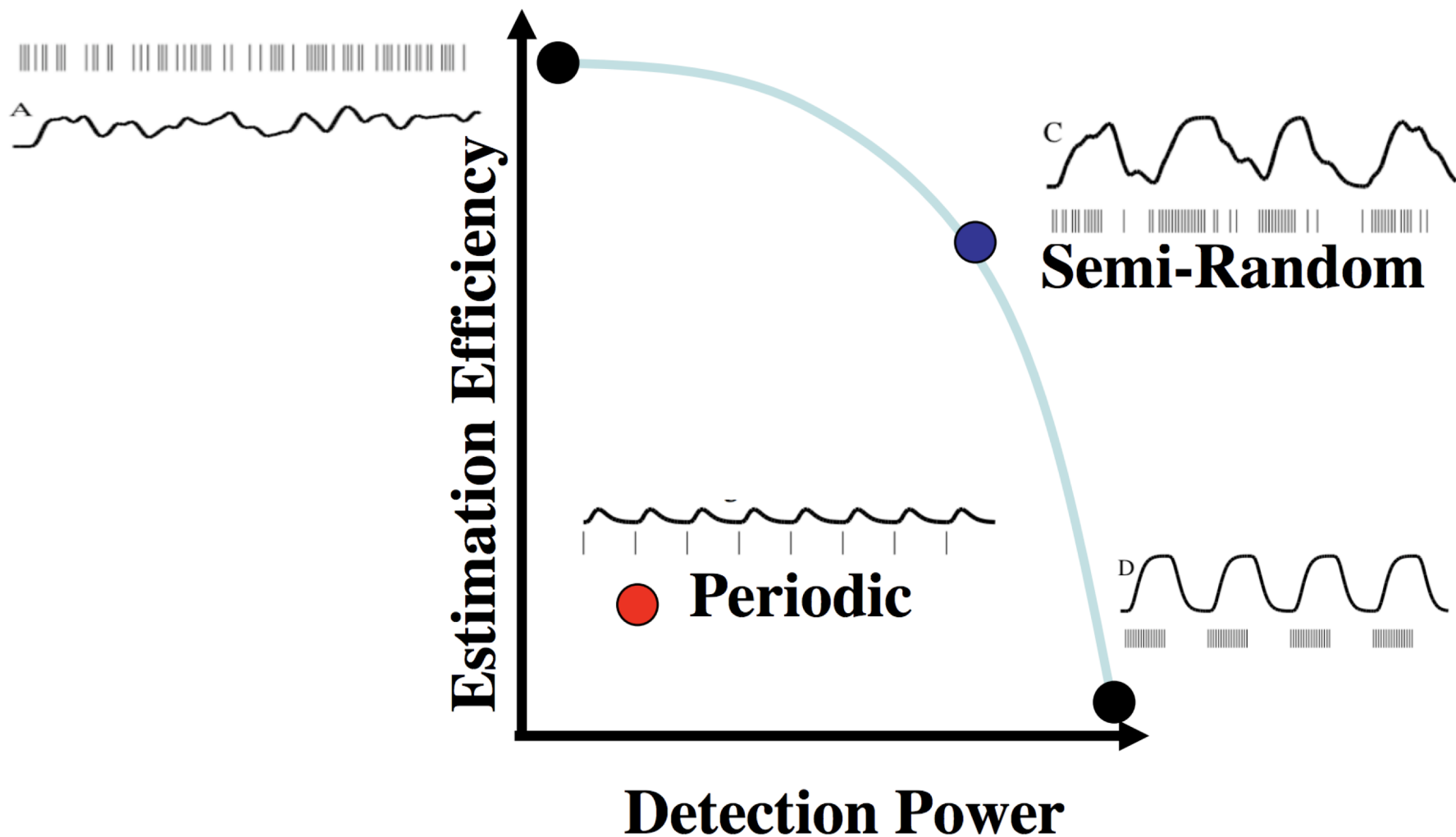
Estimation vs. Detection

Detection is observing a signal if it is really there

Estimation is the analysis of the finer details of the signal, such as the shape of the BOLD response

There tends to be a tradeoff between the two

Fundamental Trade-off



Optimization Strategies

**Let's say we just want to increase our efficiency;
how to choose?**

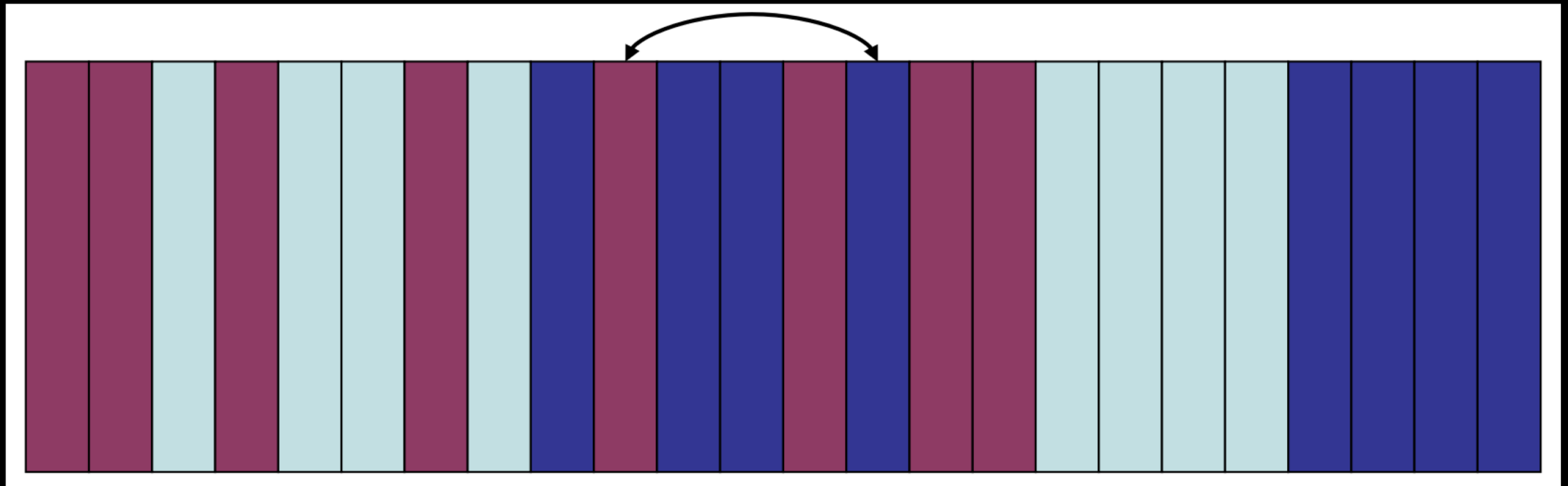
You could just create designs, calculate efficiency, and repeat

Drawbacks of this approach?

Optimization Strategies

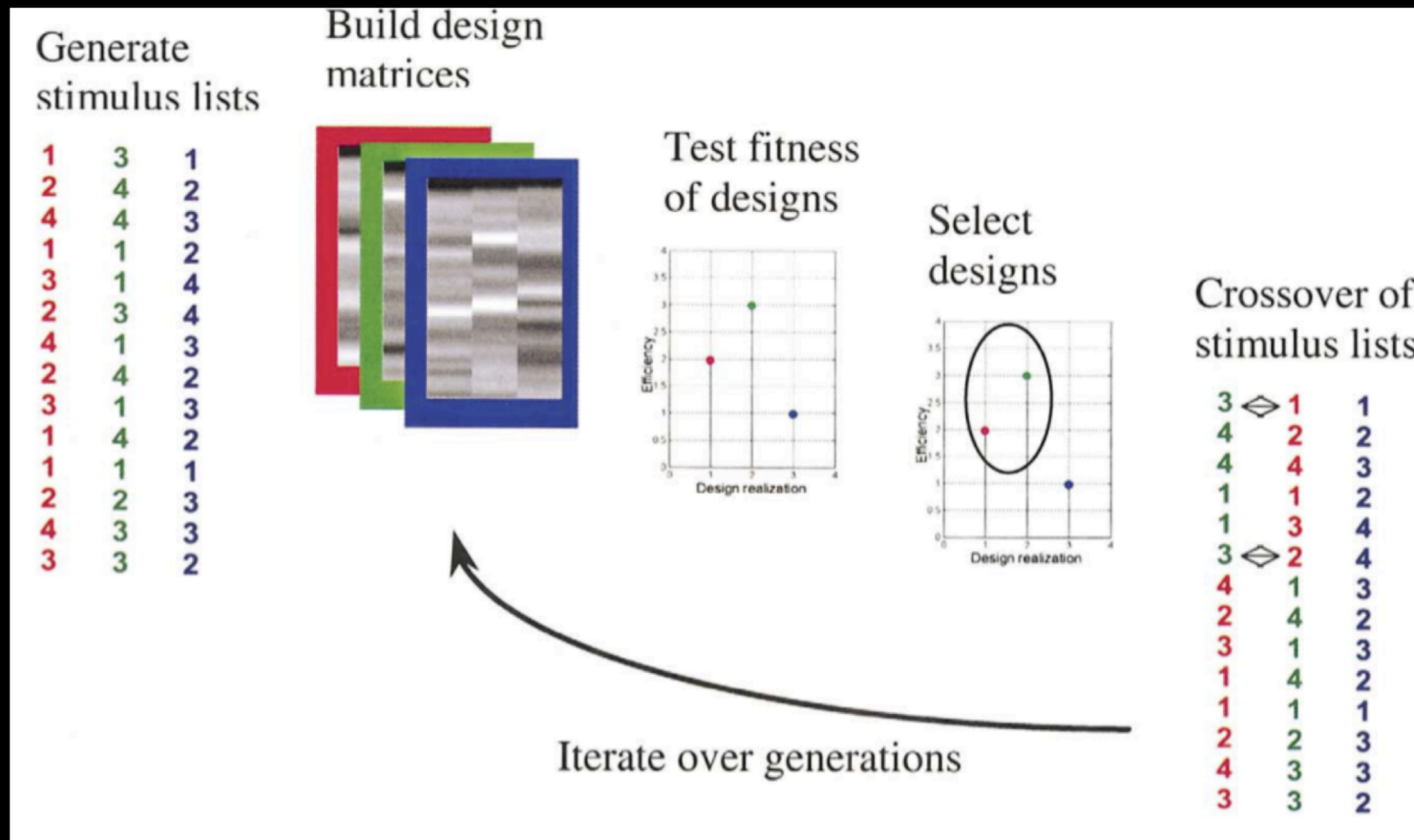
Permuted block design

Start with stimuli blocked and then randomly permute



Optimization Strategies

Genetic algorithm



Optimization Strategies

Resources

Welcome to the Optseq Home Page

optseq2 is a tool for automatically scheduling events for rapid-presentation event-related (RPER) fMRI experiments (the schedule is the order and timing of events). Events in RPER are presented closely enough in time that their hemodynamic responses will overlap. This requires that the onset times of the events be jittered in order to remove the overlap from the estimate of the hemodynamic response. RPER is highly resistant to habituation, expectation, and set because the subject does not know when the next stimulus will appear or which stimulus type it will be. RPER is also more efficient than fixed-interval event related (FIER) because more stimuli can be presented within a given scanning interval at the cost of assuming that the overlap in the hemodynamic responses will be linear. In SPM parlance, RPER is referred to as 'stochastic design'.

The flexibility of RPER means that there are a huge number of possible schedules, and they are not equal. optseq2 randomly samples the space of possible schedules and returns the 'best' one, where the user can control the definition of 'best'. Cost functions include: average efficiency, average variance reduction factor (VRF), and a weighted combination of average and stddev of the VRF. The user can also specify that the first order counter-balancing of the sequence of stimuli be pre-optimized.

Download the [Linux version](#) of optseq2.

Download the [Linux x86_64 version](#) of optseq2.

Download the [MacOSX-PowerPC version](#) of optseq2.

Download the [MacOSX-Intel version](#) of optseq2.

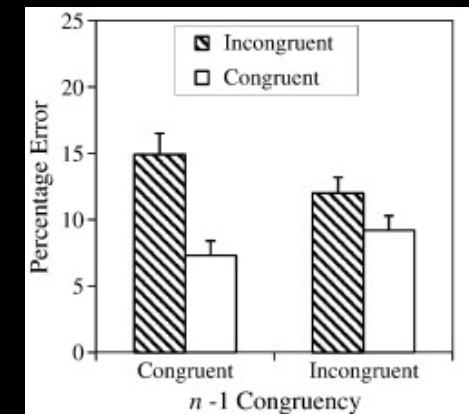
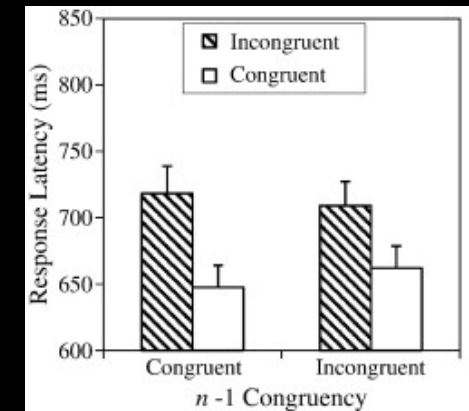
Download the [Cygwin version](#) of optseq2.

Optimization Strategies

```
(base) ajahn:~/Desktop/Flanker/2ndLevel_Inc-Con$ optseq2 --ntp 160 --tr 2 --psdwin 0 20 2
--ev disgustingPic 2 20 --ev attractivePic 2 15 --ev neutralPic 2 30 --evc 1 -1 0 --nkeep
3 --o IAPS --tnullmin 2 --tnullmax 8 --nsearch 1000
```

Why not just use
the best one?

0.0000	2	2.000	1.0000	attractivePic
2.0000	0	2.000	1.0000	NULL
4.0000	2	2.000	1.0000	attractivePic
6.0000	0	2.000	1.0000	NULL
8.0000	1	2.000	1.0000	disgustingPic
10.0000	0	2.000	1.0000	NULL
12.0000	3	2.000	1.0000	neutralPic
14.0000	0	4.000	1.0000	NULL
18.0000	1	2.000	1.0000	disgustingPic
20.0000	0	2.000	1.0000	NULL
22.0000	3	2.000	1.0000	neutralPic
24.0000	0	2.000	1.0000	NULL
26.0000	1	2.000	1.0000	disgustingPic
28.0000	0	2.000	1.0000	NULL
30.0000	3	2.000	1.0000	neutralPic
32.0000	0	2.000	1.0000	NULL



Optimization Strategies

Resources

NeuroPowerTools

NeuroPower ▾

NeuroDesign ▾

OVERVIEW

MAIN INPUT

CONTRASTS AND PROBABILITIES

REVIEW

CONSOLE

RESET

SETTINGS

NeuroDesign

This toolbox helps researchers with the planning of experimental designs for fMRI experiments. In short: depending on the exact time and order of stimulus presentations, a study can achieve higher statistical power or efficiency for estimating the brain signal. As such, depending on the design criteria, you'll have more power with fewer subjects or with shorter experiments. For more details about the methods, please see [the methods section](#), or take a look at the [step-by-step tutorial](#) on this page. Do you want to know how to run an optimisation on your computer without the GUI, go to the [package information page](#).

Help Wanted

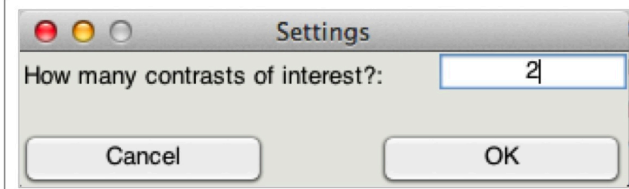
Optimization Strategies

Resources

TR you will use to acquire images →	TR (s):	<input type="text" value="1"/>
High-pass filter you will use to analyze images →	High-Pass Filter Cutoff (s):	<input type="text" value="128"/>
Task Settings		
# of conditions in your design →	N Conditions:	<input type="text" value="4"/>
# of trials per condition (unbalanced OK, e.g. 25 20 15 25) →	N Trials Per Condition:	<input type="text" value="20 20 20 20"/>
blocks = trials from same condition occurring in a row →	Maximum Block Size:	<input type="text" value="3"/>
Timing (s)		
duration of your trials (0 for purely event-related) →	Trial Duration:	<input type="text" value="3"/>
mean interstimulus interval →	Mean ISI:	<input type="text" value="3"/>
minimum value for interstimulus interval →	Min ISI:	<input type="text" value="2"/>
maximum value for interstimulus interval →	Max ISI:	<input type="text" value="6"/>
“rest” interval to add to beginning of scan →	Time before first trial:	<input type="text" value="10"/>
“rest” interval to add at end of scan →	Time after last trial:	<input type="text" value="10"/>
Optimization Settings		
number of “optimal” designs to save →	N Designs to Save:	<input type="text" value="5"/>
number of generations to test →	N Generations to Run:	<input type="text" value="50"/>
# of designs to include in each generation →	N Designs Per Generation:	<input type="text" value="1000"/>
maximum amount of time to run the program →	Max Time to Run (minutes):	<input type="text" value="2"/>
		<input type="button" value="Cancel"/> <input type="button" value="OK"/>

Optimization Strategies

Resources

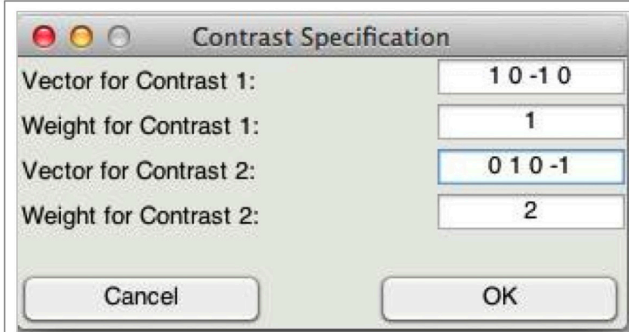


Settings

How many contrasts of interest?:

Cancel OK

This is me telling the software that I am looking for a design that maximizes the efficiency of two contrasts among my conditions.



Contrast Specification

Vector for Contrast 1:

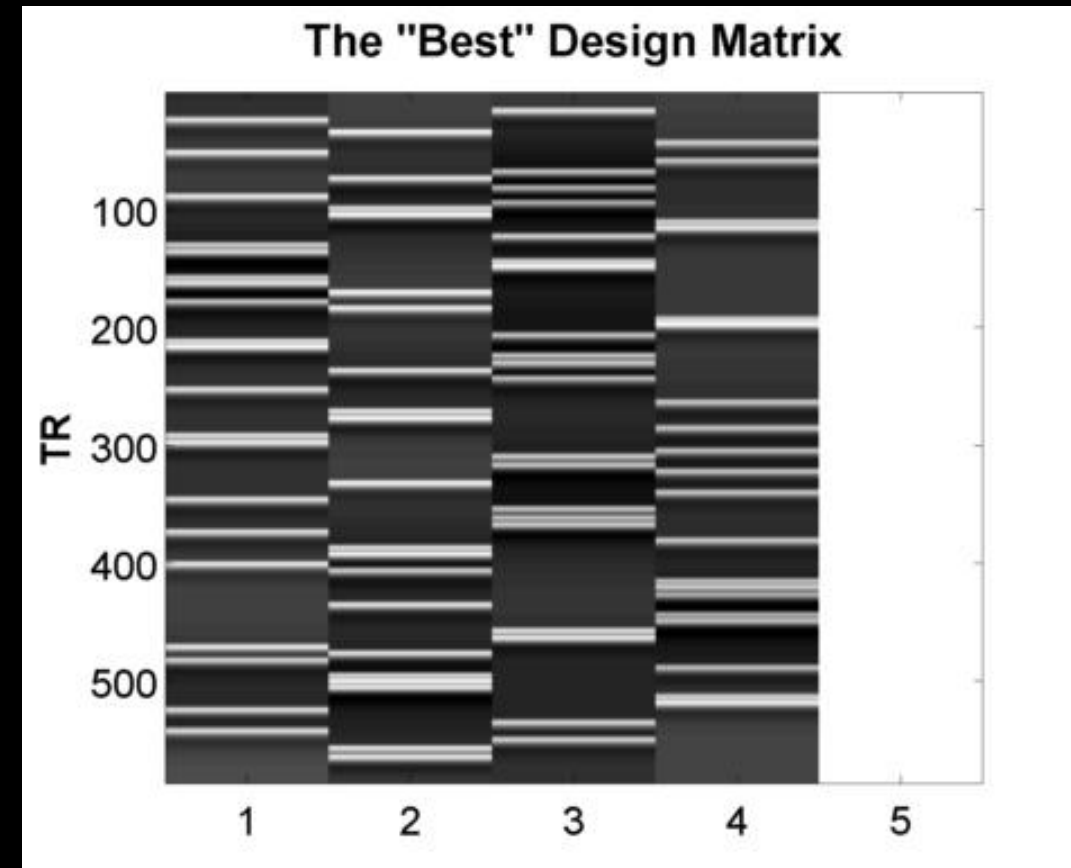
Weight for Contrast 1:

Vector for Contrast 2:

Weight for Contrast 2:

Cancel OK

This is me telling the software that although I do care about the comparison among predictors 1 and 3 (**Contrast 1**), I actually care *more* about the comparison of predictors 2 and 4 (**Contrast 2**).



Optimization Strategies

Resources

	A	B	C	D	E
1	Trial	Condition	Onset	Duration	ISI
2	1	3	10.00	3.00	4.77
3	2	1	17.77	3.00	7.32
4	3	2	28.10	3.00	6.03
5	4	4	37.13	3.00	5.20
6	5	1	45.33	3.00	4.04
7	6	4	52.37	3.00	6.31
8	7	3	61.68	3.00	2.69
9	8	2	67.37	3.00	4.92
10	9	3	75.29	3.00	4.45
11	10	1	82.74	3.00	2.13
12	11	3	87.88	3.00	2.04
13	12	2	92.91	3.00	2.87
14	13	2	98.78	3.00	2.56
15	14	4	104.35	3.00	2.59
16	15	4	109.93	3.00	3.53
17	16	3	116.46	3.00	4.43
18	17	1	123.88	3.00	3.10
19	18	1	129.99	3.00	4.77

Optimization Strategies: Summary

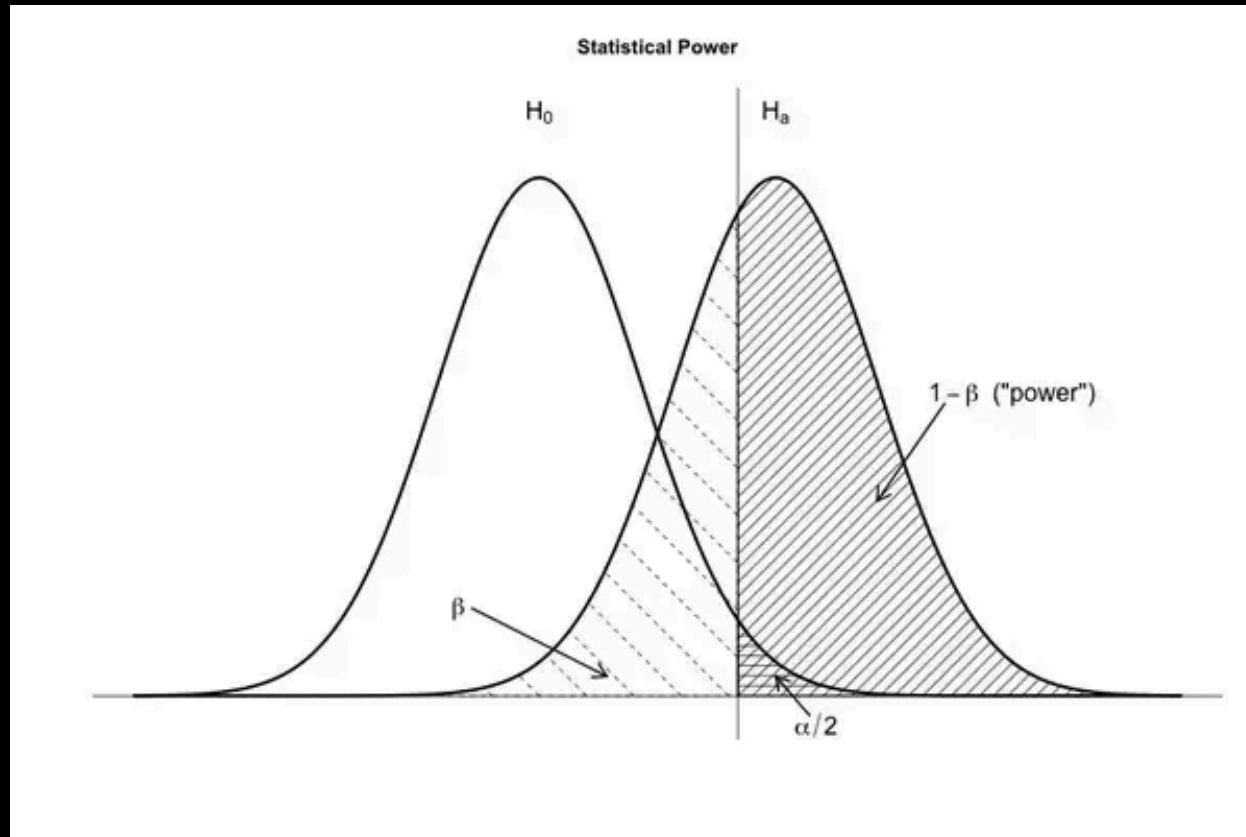
Each resource generates multiple designs

There is no best design; all efficiencies are relative!

Also need to consider whether the design “feels” right

Power Analysis

Remember this?



Power Analysis

Power analyses for behavioral studies are simpler:

Easier to recruit large N

Relatively inexpensive to run lots of subjects

Behavioral effects can be very strong

Power Analysis

Now for imaging studies:

More difficult to recruit large N (e.g., >50 per study)

Expensive to run lots of subjects (\$500-\$900 per hour)

Imaging effects can be very weak

Several sources of noise

Power Analysis

What happens during grant writing?

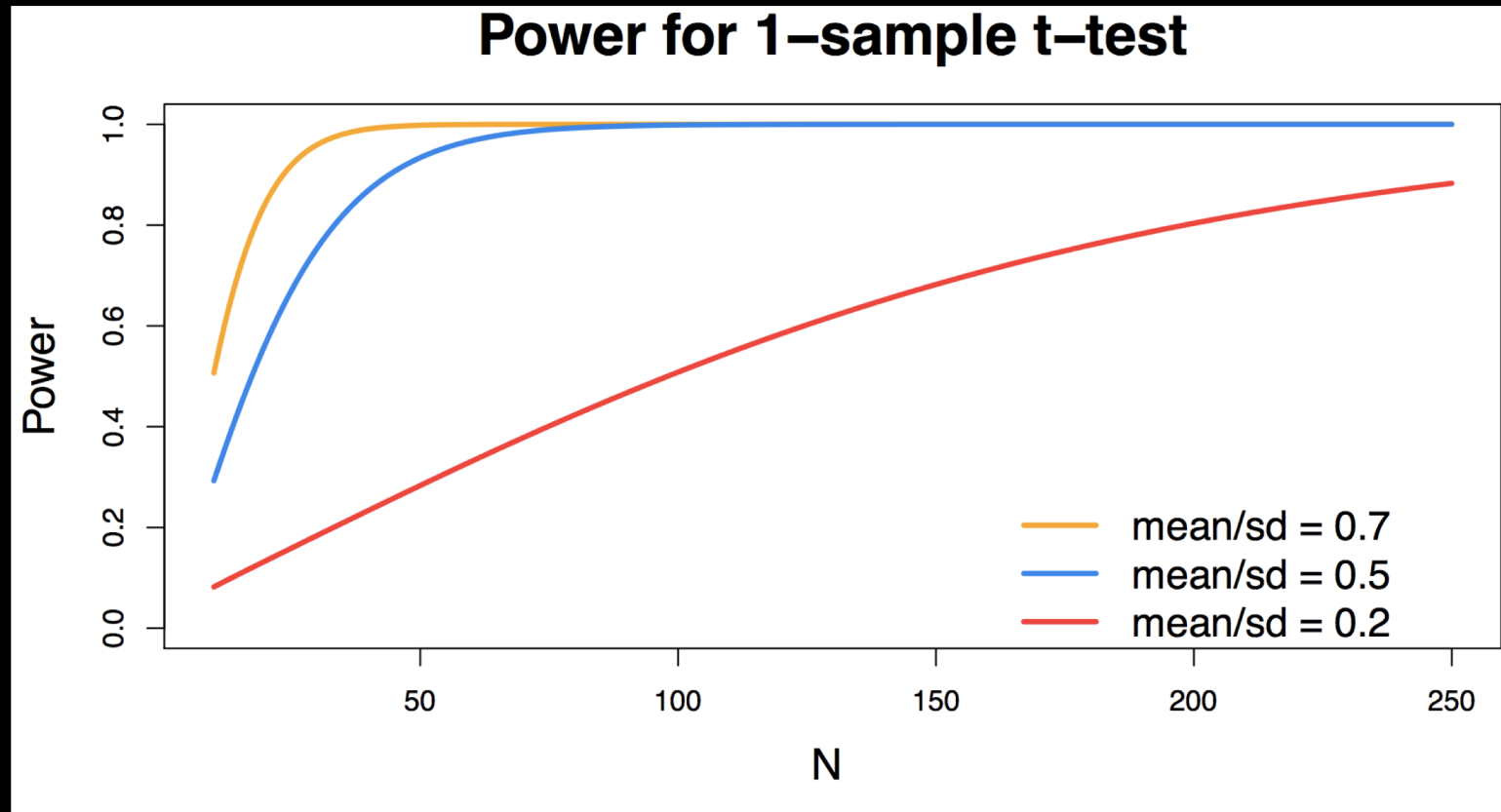
Becomes a hunt for 80% power

**Most fMRI studies won't have this kind of power,
for the reasons listed above**

Does that mean we shouldn't even do it in the first place?

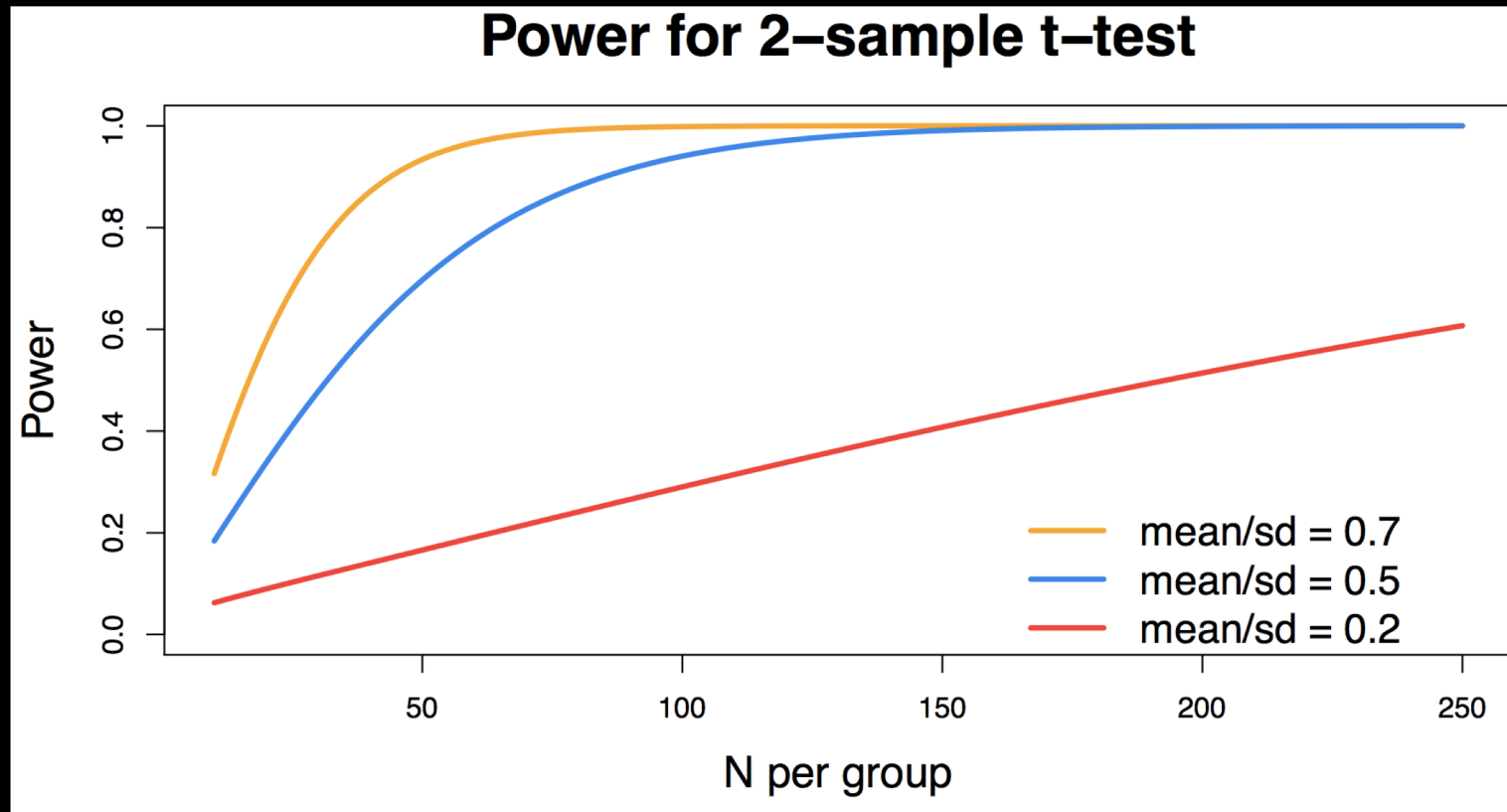
Power Analysis

What are reasonable ranges for power, given effect size?



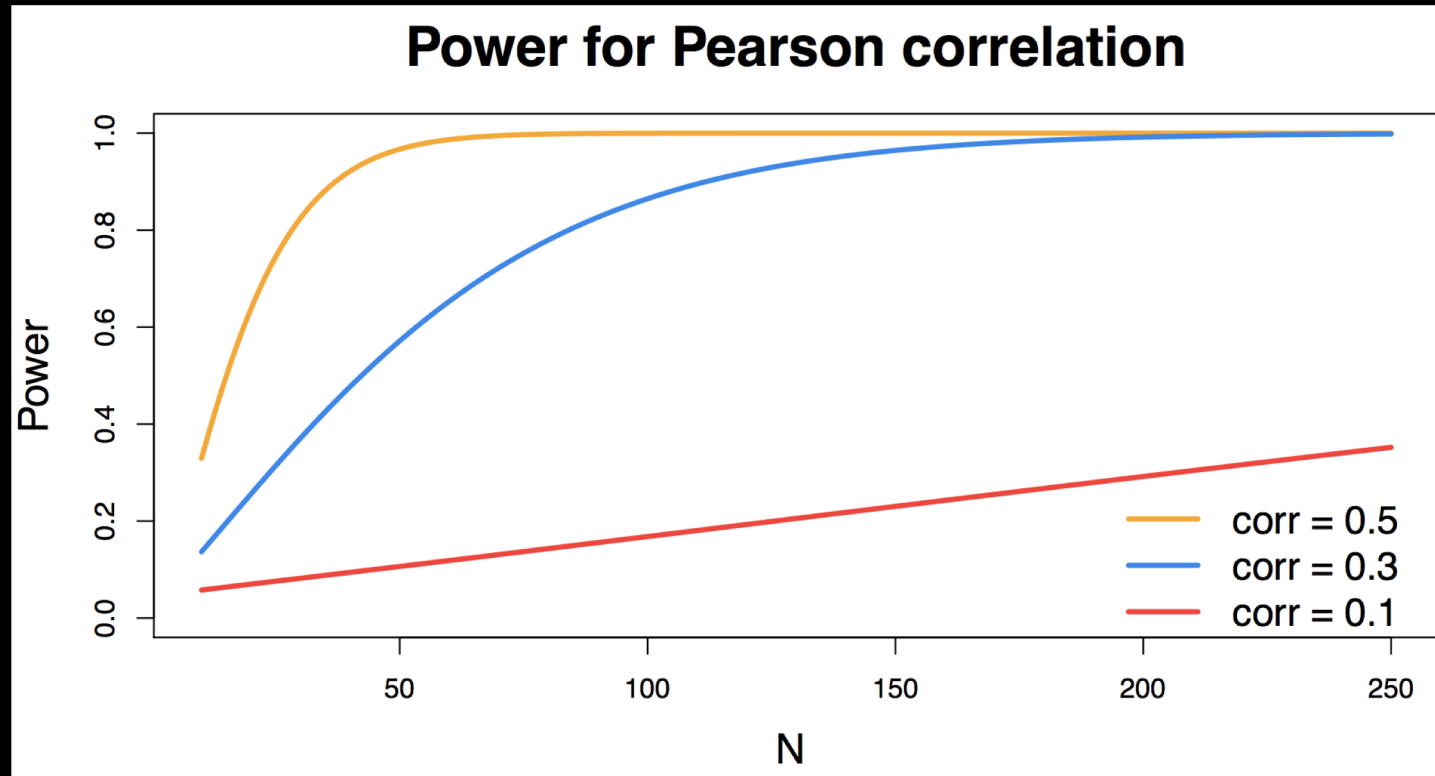
Power Analysis

What are reasonable ranges for power, given effect size?



Power Analysis


What are reasonable ranges for power, given effect size?



Power Analysis

What is the best way to estimate power?

Home > Education > Teaching Tools > G*Power




G*Power for Mac

3.1.9.6

31 March 2020

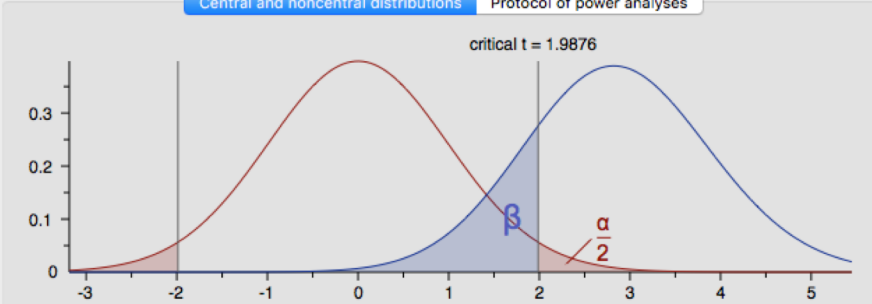
Statistical power analysis program.

[Follow this app](#) [Developer website](#)



Power Analysis

Central and noncentral distributions Protocol of power analyses



critical t = 1.9876

Test family: t tests

Statistical test: Linear bivariate regression: One group, size of slope

Type of power analysis: A priori: Compute required sample size - given alpha, power, and effect size

Input parameters

Tail(s): Two

Determine

Slope H1: 0.1

alpha err prob: 0.05

Power (1-beta err prob): 0.8

Slope H0: 0

Std dev sigma_x: 1

Std dev sigma_y: 0.3464102

Output parameters

Noncentrality parameter delta: 2.8444520

Critical t: 1.9876083

Df: 87

Total sample size: 89

Actual power: 0.8031595

Input mode: p, residual sigma, sigma_x => slope, sigma_y

Correlation rho: 0.5

Std dev residual sigma: 0.3

Std dev sigma_x: 1

Std dev sigma_y: 0.3464102

Calculate

Slope H1: 0.1732051

Calculate and transfer to main window

Close effect size drawer

X-Y plot for a range of values Calculate

Power Analysis

What about estimating power from another published study?

Keep in mind that only significant results are usually published; this may just contribute to the file drawer problem (to be discussed more on Tuesday)

Power Analysis

What about calculating power after a study is run?
(e.g., post-hoc power analysis?)

This is a statistical fallacy, since the null hypothesis has already been either rejected or not rejected; there is no “power” to calculate!

Power Analysis

Statistical Practice

The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis

John M. HOENIG and Dennis M. HEISEY

It is well known that statistical power calculations can be valuable in planning an experiment. There is also a large literature advocating that power calculations be made whenever one performs a statistical test of a hypothesis and one obtains a statistically nonsignificant result. Advocates of such post-experiment power calculations claim the calculations should be used to aid in the interpretation of the experimental results. This approach, which appears in various forms, is fundamentally flawed. We document that the problem is extensive and present arguments to demonstrate the flaw in the logic.

calculations as a matter of policy (Anon. 1995; Anon. 1998). We emphasize that these calculations are sought primarily with the thought that they are useful for explaining the observed data, rather than for the purpose of planning some future experiment. We even found statistical textbooks that illustrate the flawed approach (e.g., Rosner 1990; Winer, Brown, and Michels 1991; Zar 1996). Researchers need to be made aware of the shortcomings of power calculations as data analytic tools and taught more appropriate methodology.

It is important to understand the motivation of applied scientists for using power analysis to interpret hypothesis tests with nonsignificant results. The traditional, widely ac-

Power Analysis

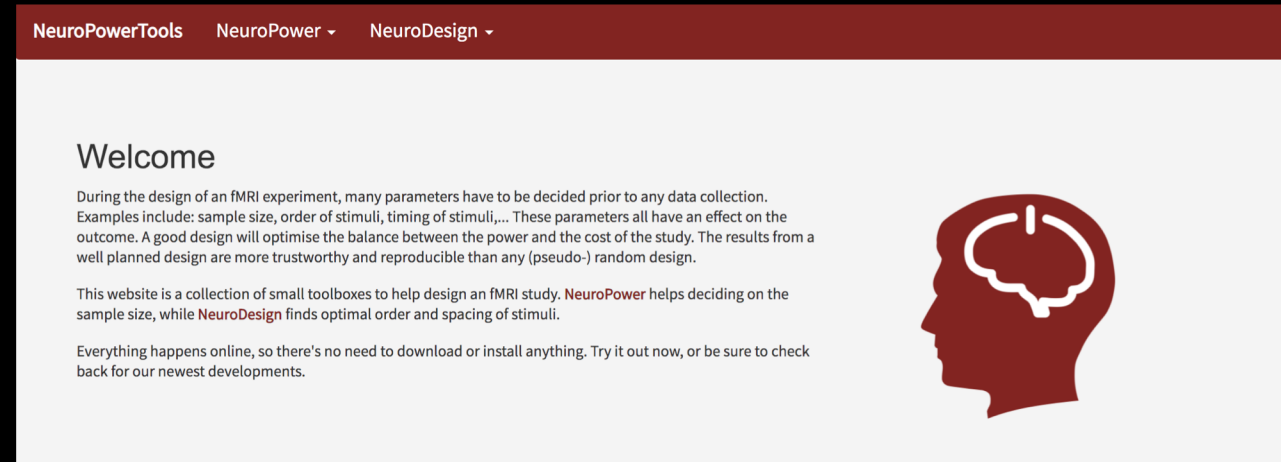
Tools for power analysis

Fmripower

What happened to the fmripower tool?

The fMRIPower tool hit old age and was starting to have quite a few issues. I have chosen to replace the tool with a set of instructions so you can carry out power analyses on your own. I will post that information soon! In a nutshell, all you really need to do is perform an ROI analysis, get your effect size and carry out a power analysis using that effect size. Of course, don't forget, you must use an a priori ROI. I realize it is super frustrating, but you cannot use the ROI that was active in the same data set from which you are running the power analysis.

Hopefully that's enough for now, but I will make a video and blog post with instructions soon.



The screenshot shows the homepage of the NeuroPowerTools website. At the top, there is a dark red navigation bar with the text "NeuroPowerTools", "NeuroPower" with a dropdown arrow, and "NeuroDesign" with a dropdown arrow. Below the navigation bar, the main content area is white. On the left side, there is a "Welcome" section with the following text: "During the design of an fMRI experiment, many parameters have to be decided prior to any data collection. Examples include: sample size, order of stimuli, timing of stimuli,... These parameters all have an effect on the outcome. A good design will optimise the balance between the power and the cost of the study. The results from a well planned design are more trustworthy and reproducible than any (pseudo-) random design." Below this, it says: "This website is a collection of small toolboxes to help design an fMRI study. NeuroPower helps deciding on the sample size, while NeuroDesign finds optimal order and spacing of stimuli." At the bottom of the welcome section, it states: "Everything happens online, so there's no need to download or install anything. Try it out now, or be sure to check back for our newest developments." On the right side of the page, there is a dark red silhouette of a human head in profile, facing left, with a white outline of a brain inside.

Power Analysis

Data location

Either paste a link to the online nifti-file OR upload your statistical map.

URL

Upload

Currently: [/maps/spmT_0001.nii](#)

Change:

No file chosen

Mask location (optional)

Upload a full brain mask or a Region-of-Interest mask. If no mask is selected, all non-null voxels are used.

Maskfile

No file chosen

Design specifications

Are the data Z- or T-values?*

What is the screening threshold, also known as the clusterforming threshold or the excursion threshold (either p-value or z-value units)?*

How many subjects does the group map represent?*

Is this a one-sample or a two-sample test?*

At which alpha-level are the statistical tests carried out?

Do you want to manually specify the smoothness or estimate from the data?

Note though that estimating smoothness on statistical maps leads to [biases](#). It is preferable to manually specify the data.*

Manual

Estimate

What is the smoothness of the data in mm?

What is the voxel size in mm?

Power Analysis

OVERVIEW START PEAK TABLE MODEL FIT **POWER CALCULATION** POWER TABLE RESET

There is not enough power to estimate a threshold for BH.

Power Hover over the lines to see detailed power predictions

To see the power for a certain sample size or vice versa, please fill out either the minimal power or the sample size.

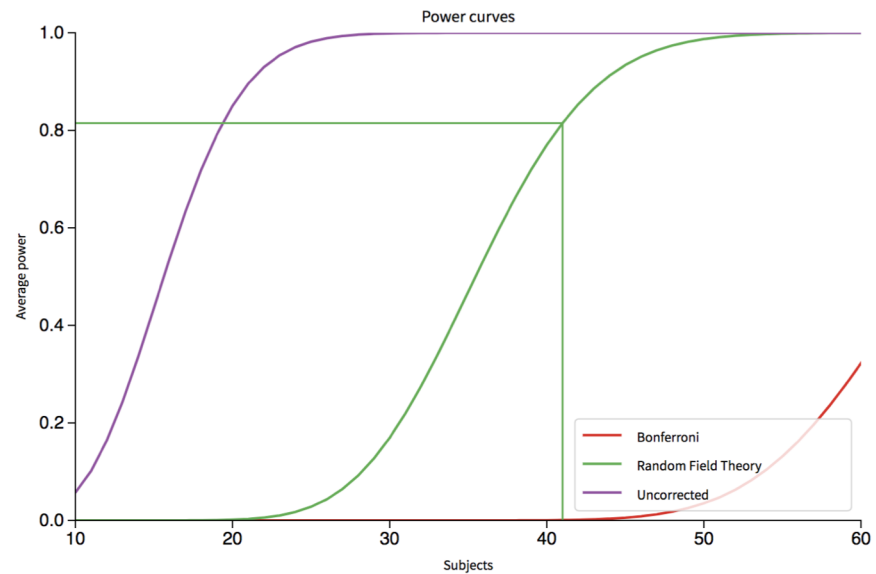
MCP*

Random Field Theory ▾

Sample size

Power

Submit parameters



save svg

To obtain a statistical power of 0.8 this study would require a sample size of 41 subjects.

Power Analysis

When evaluating a power analysis, make sure that the estimates seem to fall within reasonable bounds, and that the parameters were clearly defined

Questions?

Group-Level Analysis

Once we have estimated a model for each individual subject (1st-level analysis), we combine them into a 2nd-level analysis

In SPM: Usually focus on just the mean of the parameter estimate; variance is discarded at the group level

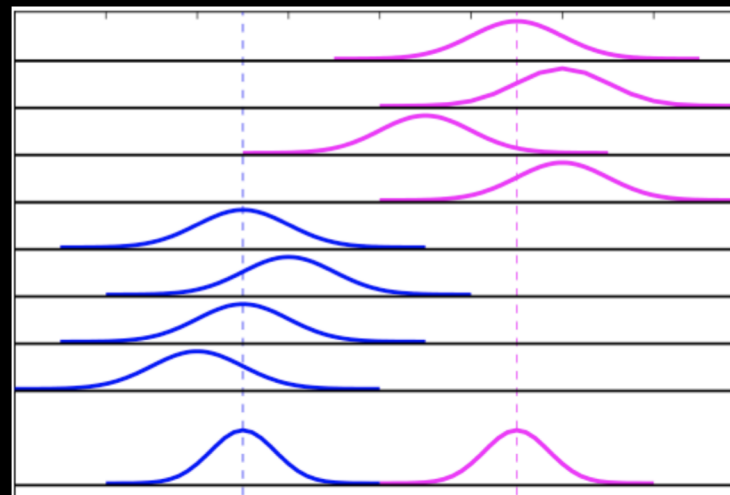
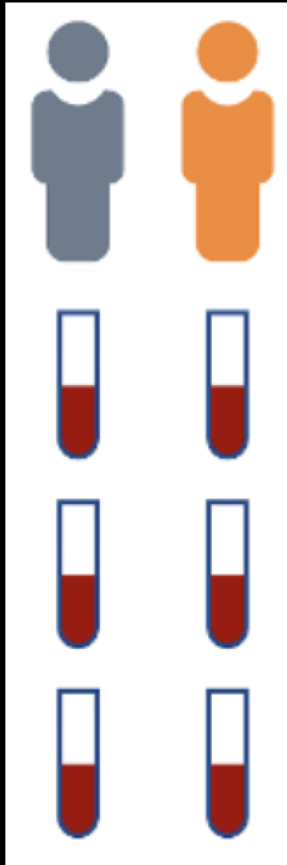
Group-Level Analysis

N.B.: The way that betas are calculated in the 1st-level is different than how they are estimated at the 2nd-level

**The difference can be expressed as “Fixed-Effects” vs.
“Random-Effects”**

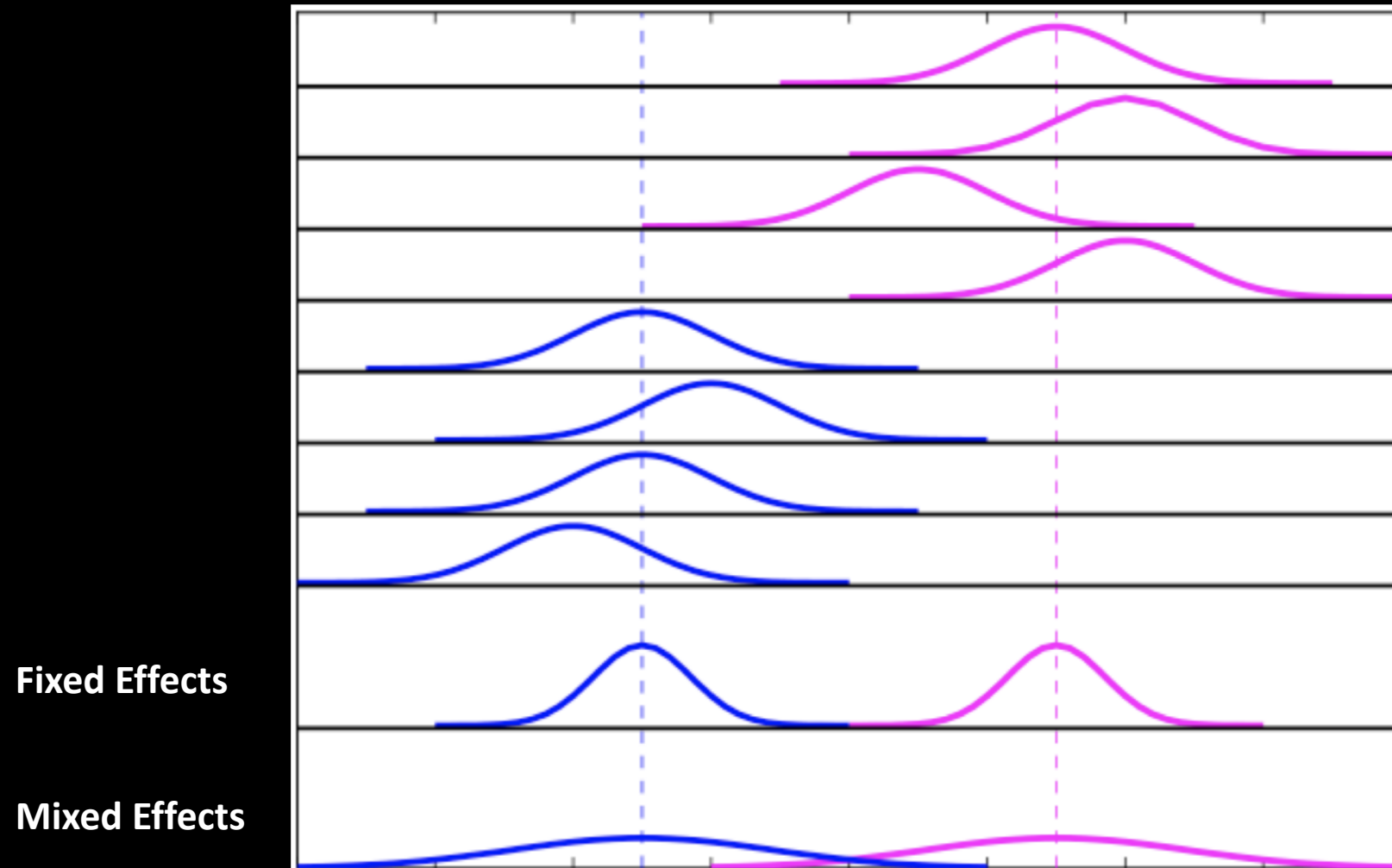
Combining both generates “Mixed-Effects”

Group-Level Analysis



Fixed Effects

Group-Level Analysis



Group-Level Analysis

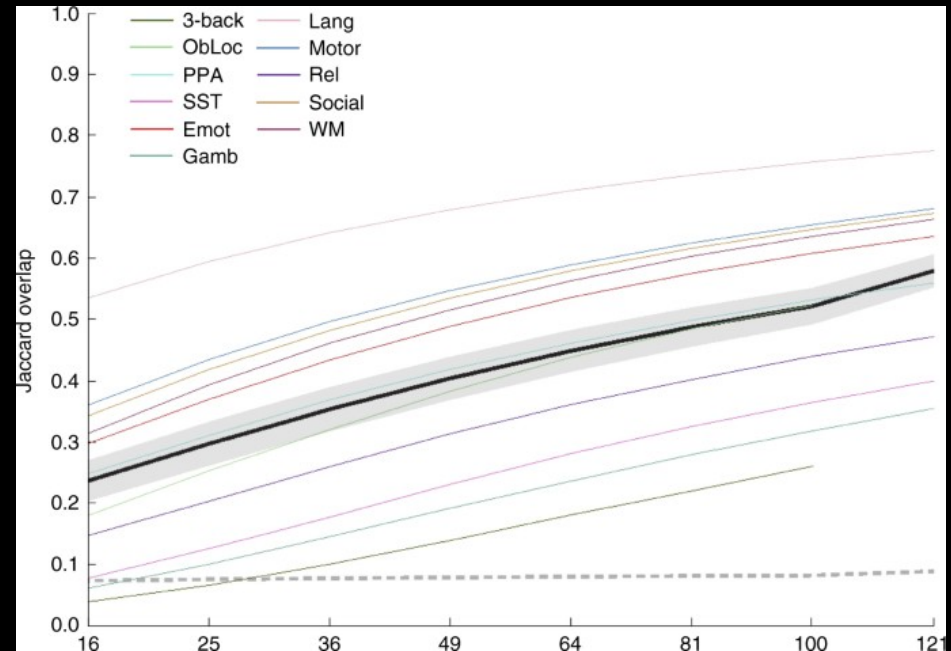
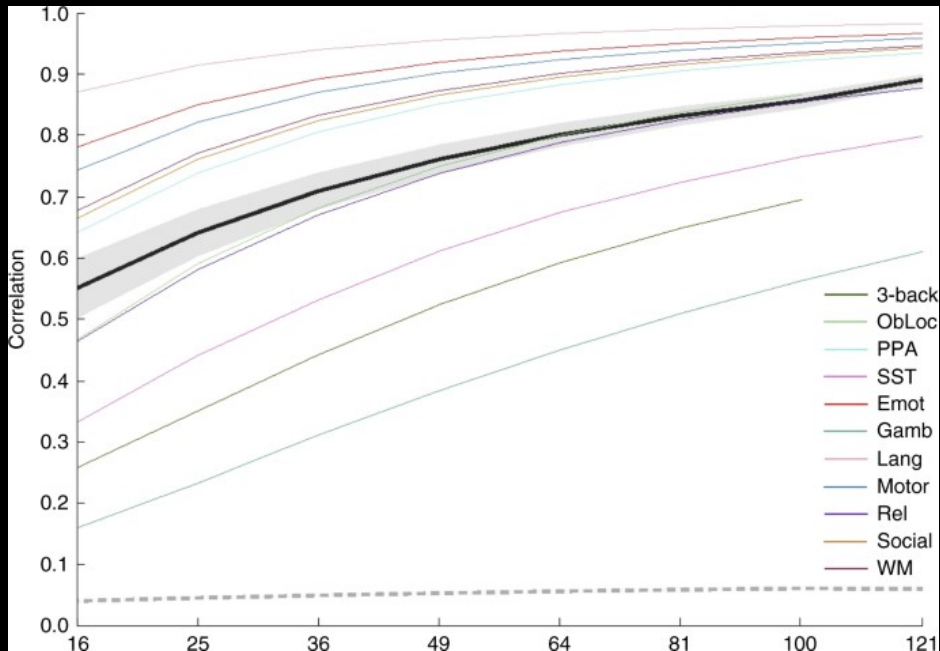
Summary: Fixed Effects applies only to the subjects you sampled

Random effects assumes that the subjects were randomly sampled from the population, and that you're trying to make an inference about the population (i.e., parameters)

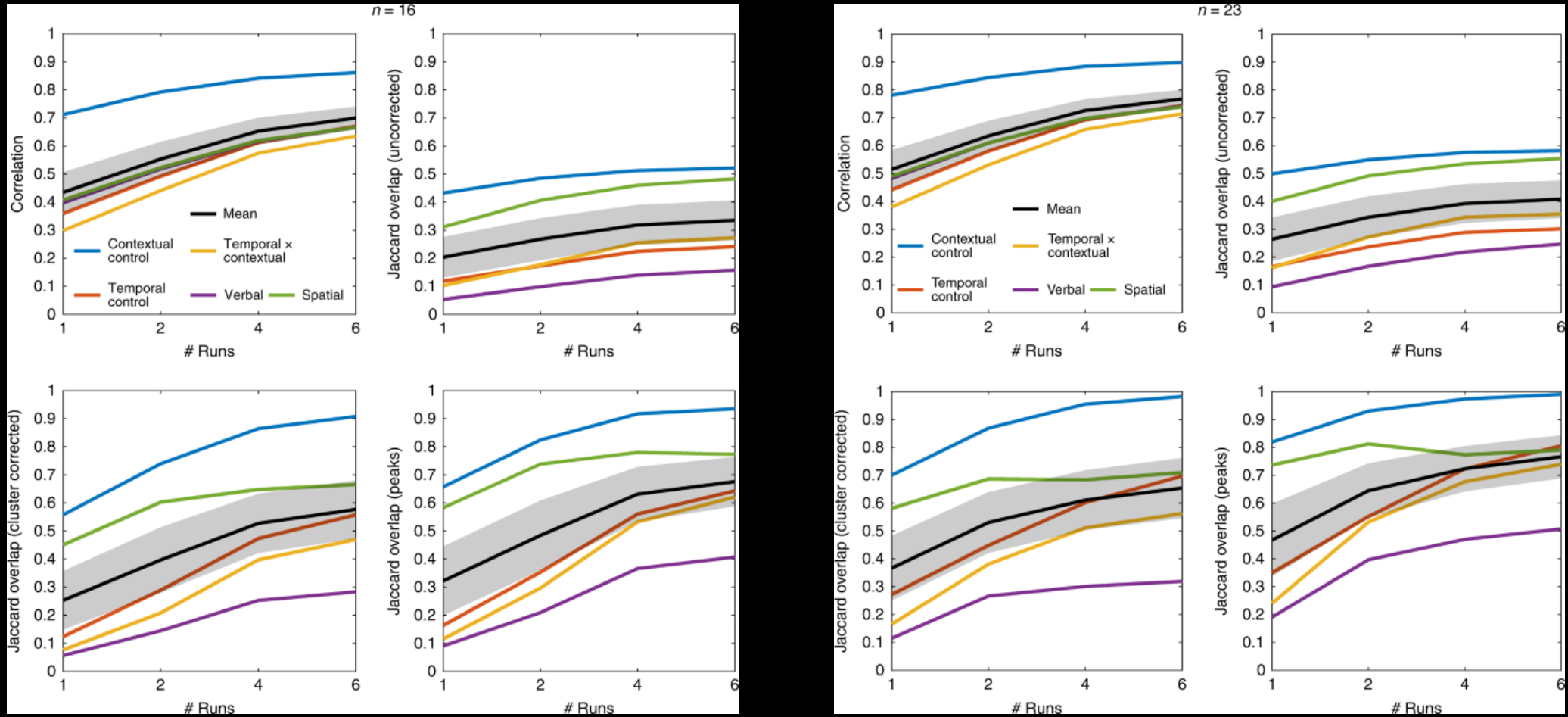
Mixed effects combines the two

Question: To reduce overall variance, should we collect more samples, or more subjects?

Group-Level Analysis



Group-Level Analysis



Group-Level Analysis

Simplest model: 1-sample t-test

A parameter estimate (or contrast of parameter estimates, also called a contrast estimate) is submitted to a t-test

Is the average of the parameter (or contrast) estimates significantly different from zero?

Group-Level Analysis

However, this typically isn't very interesting

**Since fMRI signal is arbitrary, it is more useful
to contrast one condition to another**

**You can then run a one-sample t-test on this contrast
(to be discussed in a little bit)**

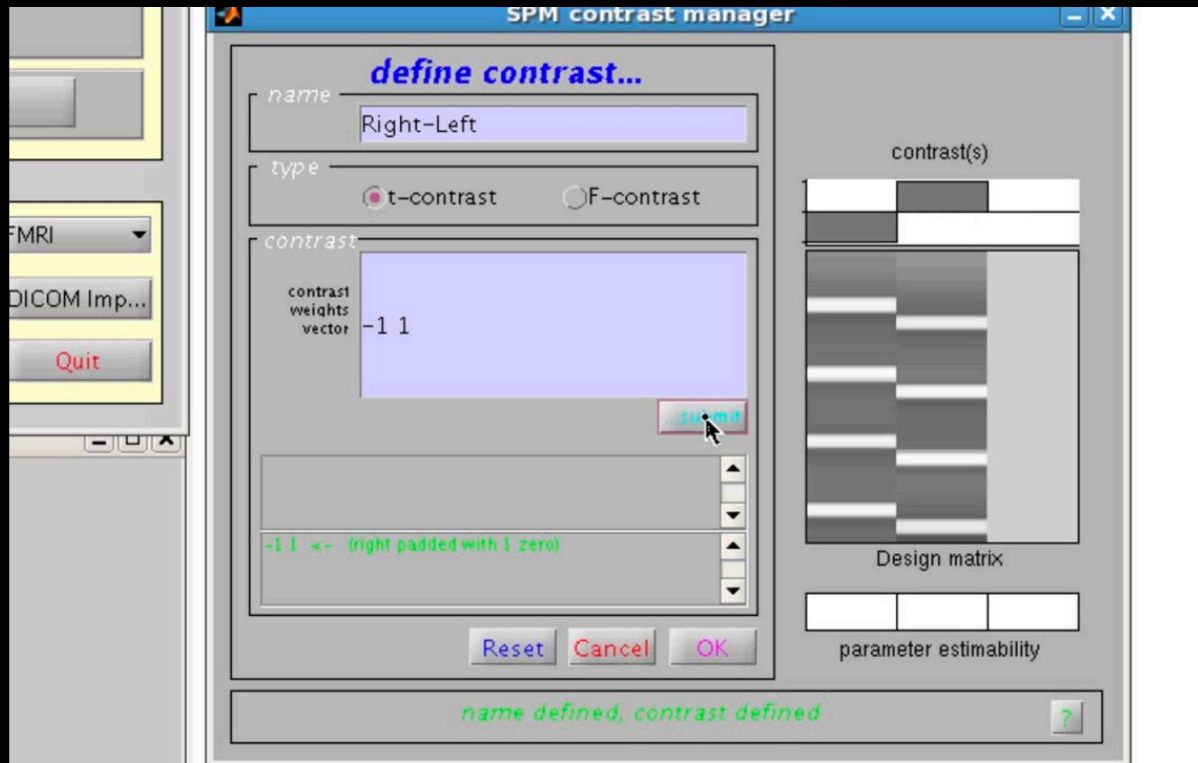
Group-Level Analysis

e.g., button presses



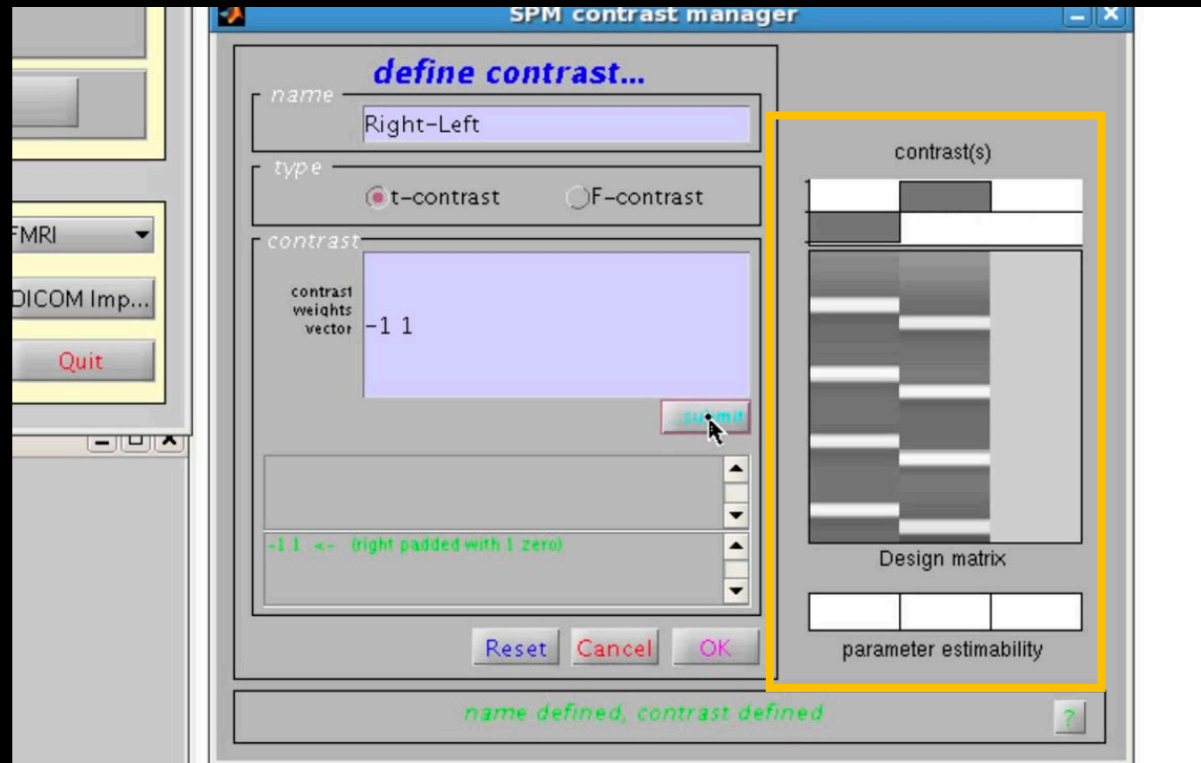
Group-Level Analysis

Very simple contrast (and useful to check whether your timings are correct!)



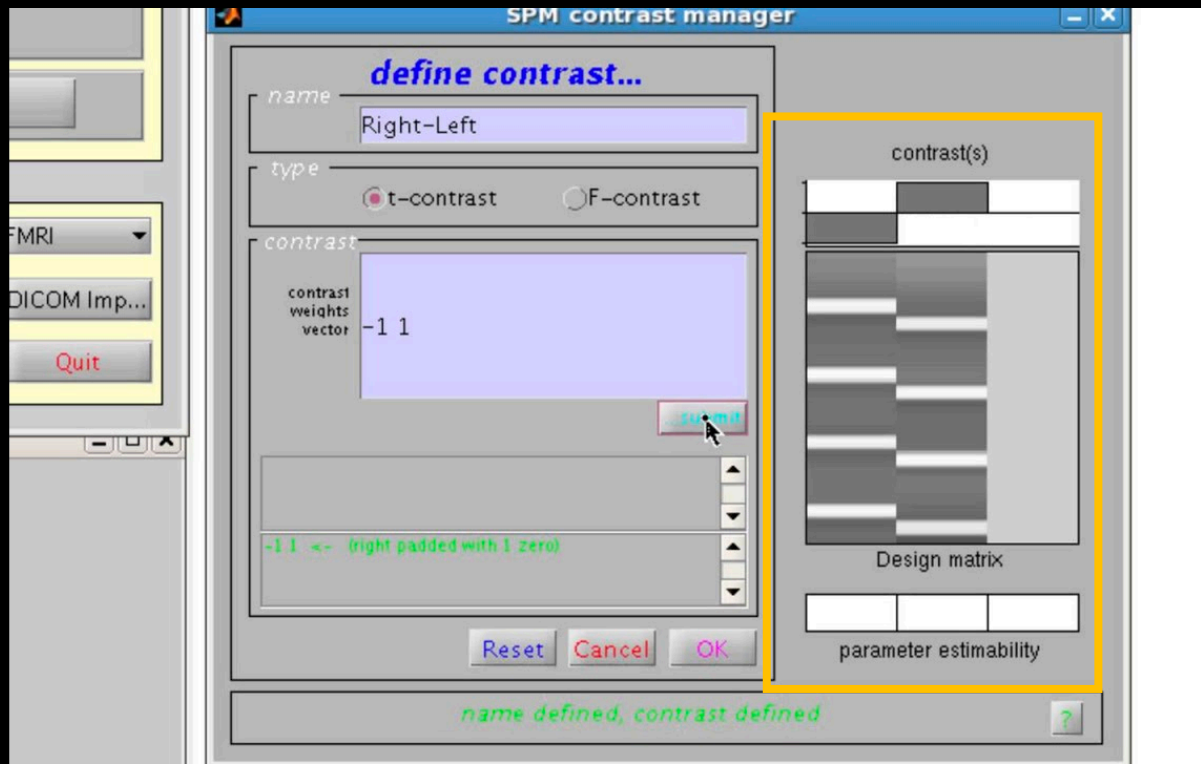
Group-Level Analysis

What is this?



Group-Level Analysis

How you specify the contrast weights depends on the order of the regressors

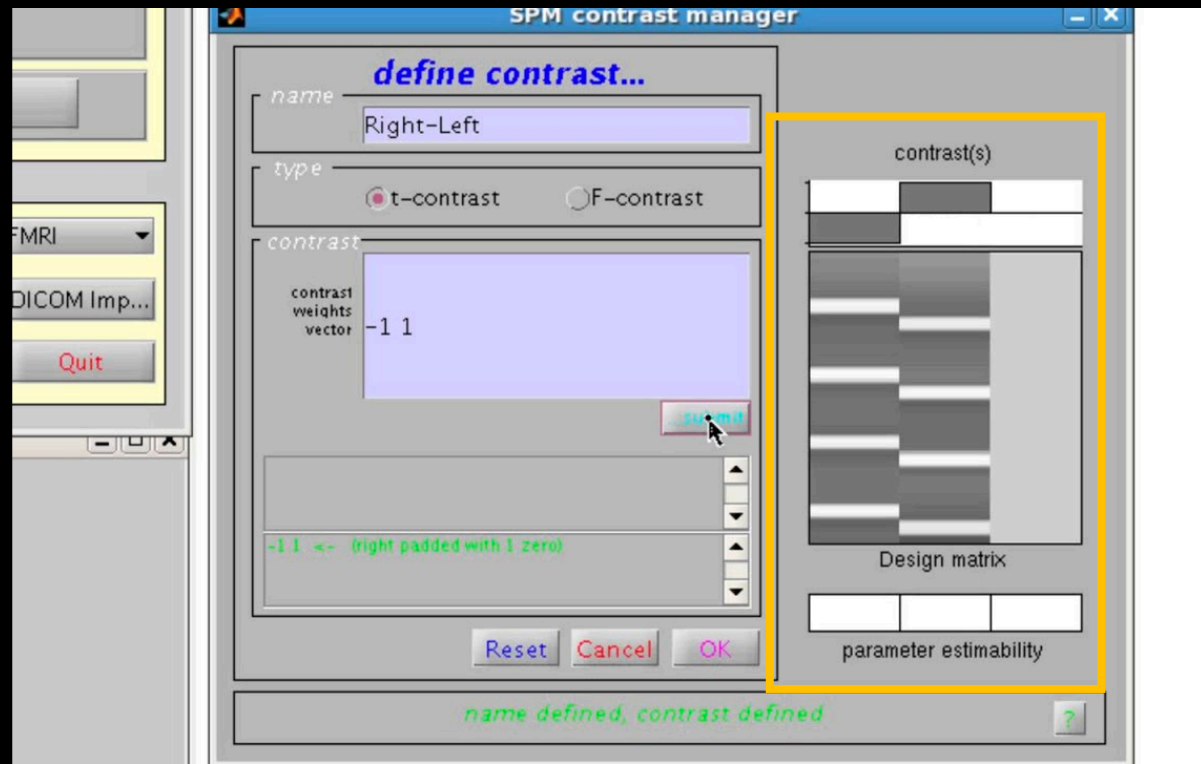


$$c = [-1 \ 1]$$

Group-Level Analysis

In this example: Left was first, Right was second

How would I specify the contrast of Left-Right?



$$c = [-1 \ 1]$$

Group-Level Analysis

In a group-level context

One-sample:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_0 \\ \vdots \\ \beta_0 \end{bmatrix} + \epsilon$$

Group-Level Analysis

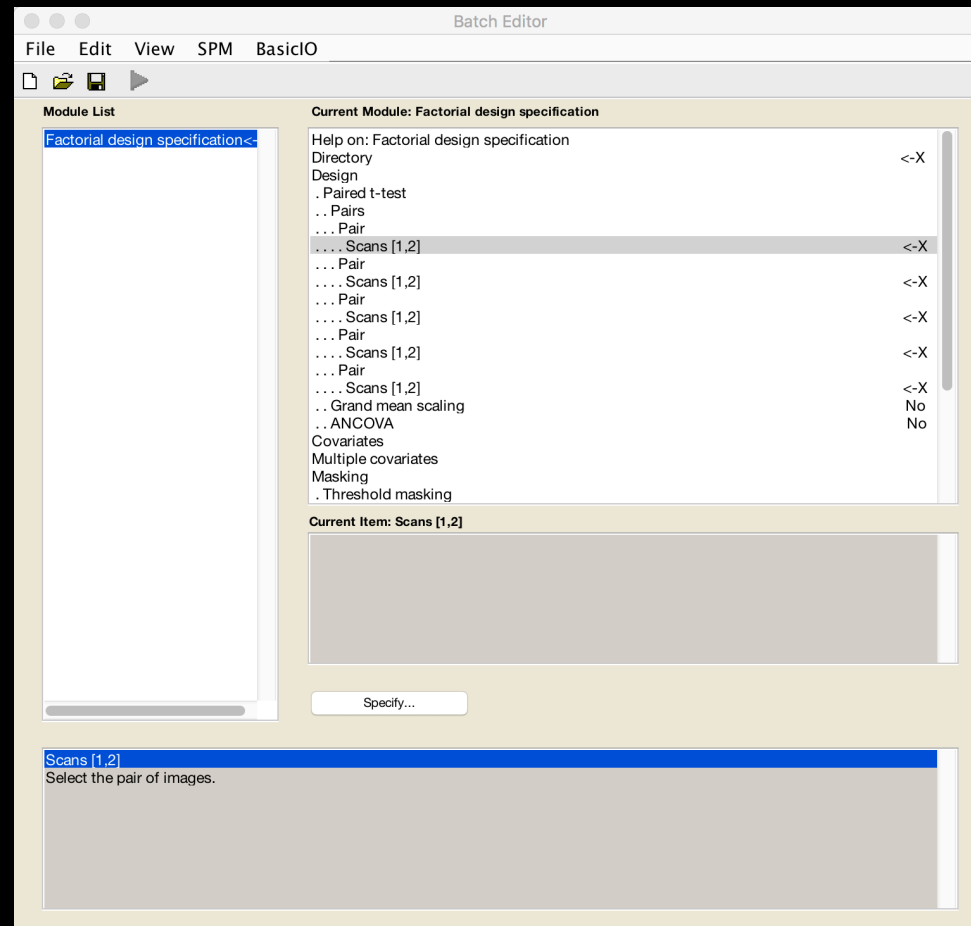
In a group-level context

Two-sample:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \epsilon$$

Group-Level Analysis

What about a paired t-test?



Group-Level Analysis

What about an interaction?

If each subjects has two conditions (A,B) with two levels (1,2) then you can do the following contrast:

The image shows a software interface for defining a contrast. The dialog box is titled "define contrast..." and contains the following fields:

- name:** Interaction Contrast
- type:** t-contrast F-contrast
- contrast weights vector:** .5 -.5 -.5 .5

To the right of the dialog box is a visualization of the contrast matrix, labeled "contrast(s)". The x-axis represents the contrast weights (0, 0.2, 0.4, 0.6, 0.8, 1) and the y-axis represents the contrast values (0, 0.5, 1). The matrix shows a pattern of weights for the contrast vector [0.5, -0.5, -0.5, 0.5].

contrast(s)	0	0.2	0.4	0.6	0.8	1
1	0.5	-0.5	-0.5	0.5	0	0
0.5	0	0.5	0.5	0	0	0
0	0	0	0	0	0	0

Group-Level Analysis

What if we perform a one-sample t-test on contrasts?

e.g., calculate $A-B$ for each of 10 subjects, to create 10 contrasts

Is this valid?

Summary Statistics

The above is called a “summary statistics” approach

Valid if the intra-subject variabilities are relatively similar across subjects

For most studies, this assumption is true
(Penny & Holmes, 2004)

Summary Statistics

Pros: Easy to implement, simplifies interactions

Cons: Assumptions may not be valid; check whether the variance and number of runs is similar for each subject

Example: Using Summary Statistics to run a one-sample t-test on an interaction term



Module List

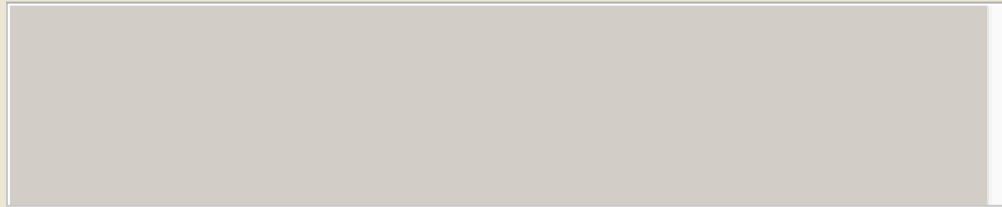
Factorial design specification<

Current Module: Factorial design specification

Help on: Factorial design specification

Directory	<-X
Design	
. Paired t-test	
.. Pairs	
... Pair	
... Scans [1,2]	<-X
... Pair	
... Scans [1,2]	<-X
... Pair	
... Scans [1,2]	<-X
... Pair	
... Scans [1,2]	<-X
... Pair	
... Scans [1,2]	<-X
.. Grand mean scaling	No
.. ANCOVA	No
Covariates	
Multiple covariates	
Masking	
. Threshold masking	

Current Item: Scans [1,2]



Scans [1,2]

Select the pair of images.





Module List

Factorial design specification<-

Current Module: Factorial design specification

Help on: Factorial design specification

Directory <-X

Design

. One-sample t-test

. Scans <-X

Covariates

Multiple covariates

Masking

. Threshold masking

. . None

. Implicit Mask Yes

. Explicit Mask

Global calculation

. Omit

Global normalisation

. Overall grand mean scaling

. . No

. Normalisation None

Enter Contrast estimates here

Current Item: Scans

Specify...

Scans

Select the images. They must all have the same image dimensions, orientation, voxel size etc.

F-tests

Also called "omnibus" tests

Tests whether one or more contrasts is significant

Question: Does this maximize Detection or Estimation?

F-tests

Instead of a contrast vector, F-tests require contrast matrices

define contrast...

name

type t-contrast F-contrast

contrast

contrast	1 0 0 0
weights	0 1 0 0
matrix	0 0 1 0
	0 0 0 1

or eye(4)

or

columns for reduced design

contrast(s)

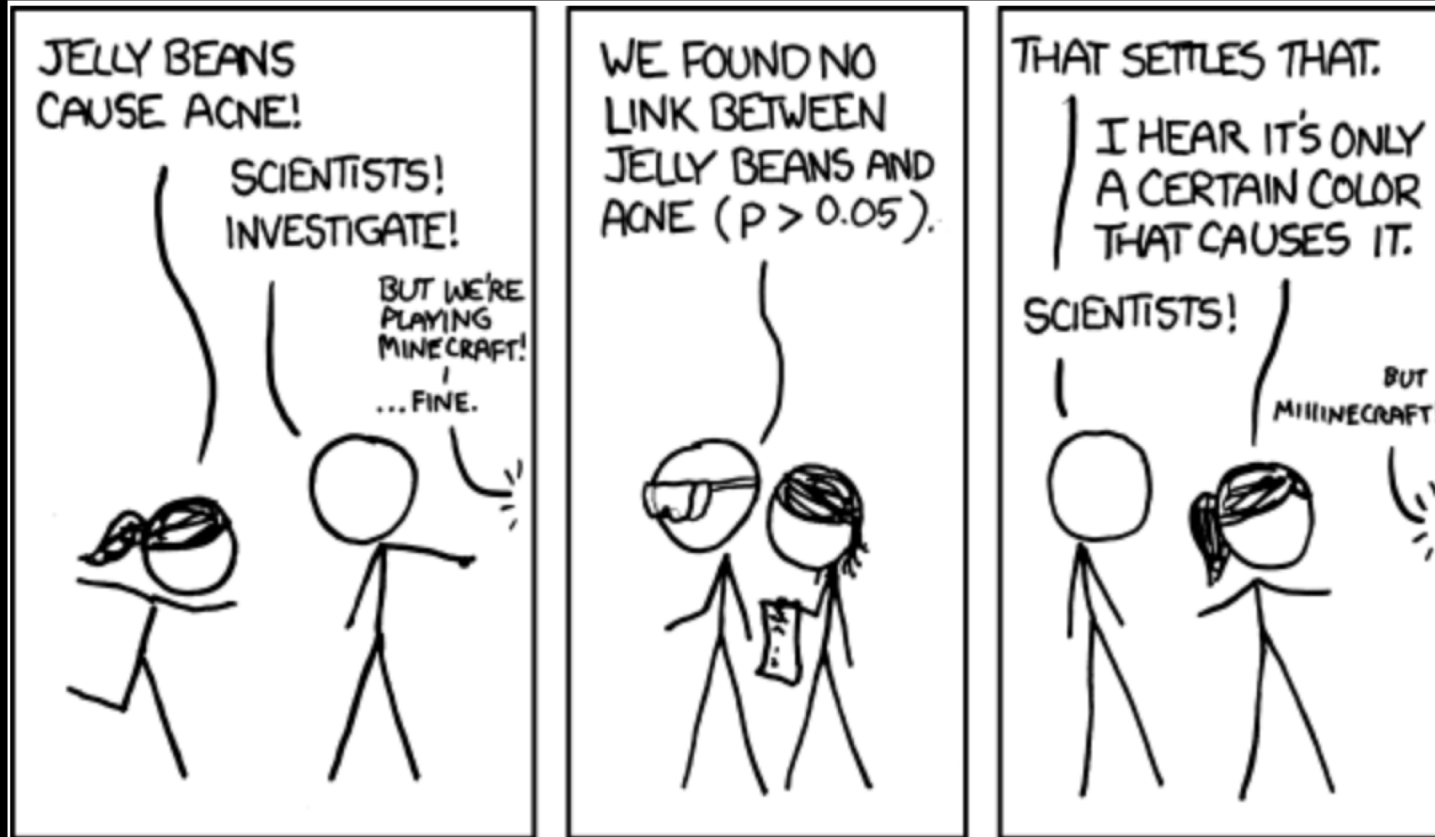
F-tests

You can also specify multiple contrasts, e.g.:

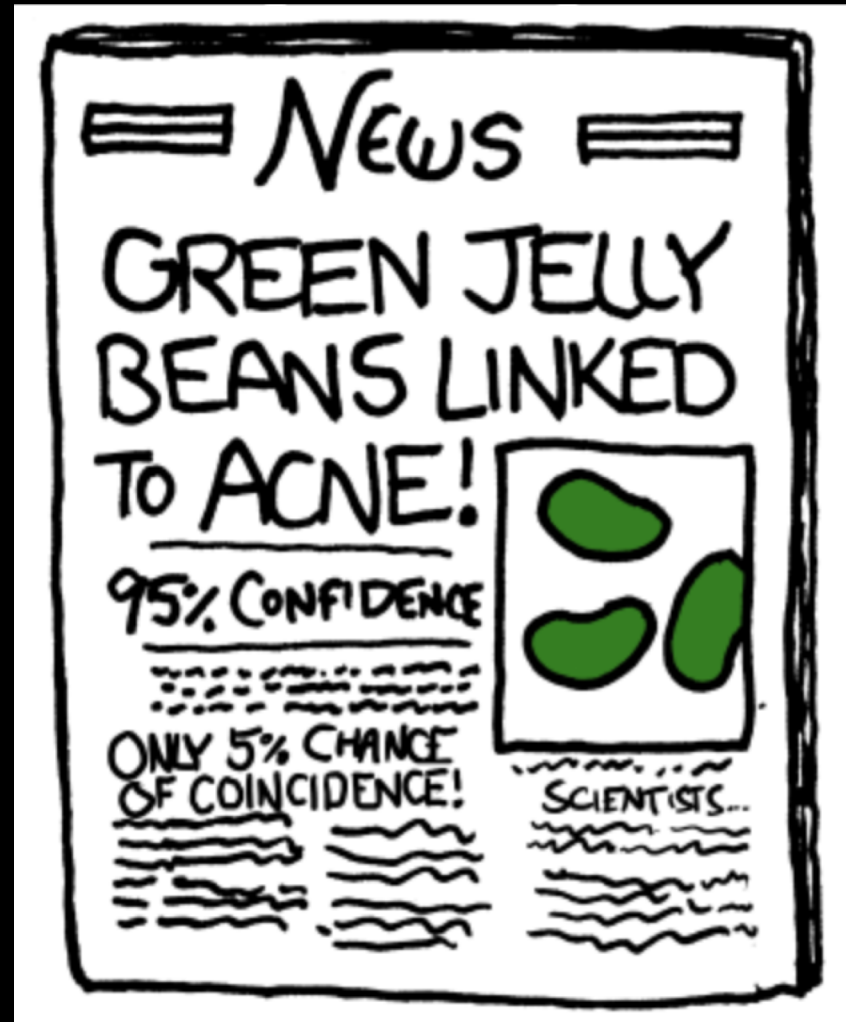
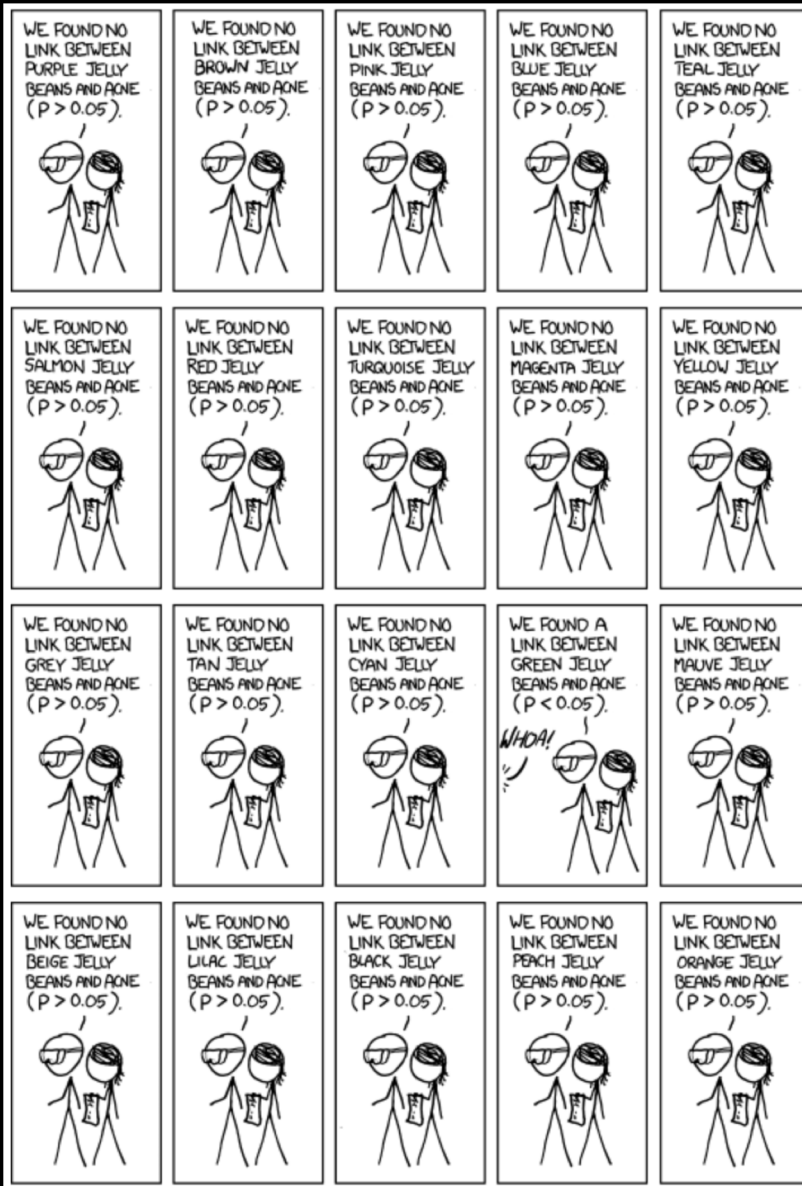
$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$

Demonstration of Group-Level Analysis

Correcting for Multiple Comparisons



Correcting for Multiple Comparisons



Correcting for Multiple Comparisons

COMING SOON!

BLUE-EMU[®]

Pain Relief *Micro-Foam*

SOFT & SOOTHING

SOFT AND SMOOTH

- 🌿 Skin Nourishing Pain Relief
- 🌿 Enriched with Aloe-vera
- 🌿 Advanced Micro-Foam™

AMERICA'S
ODOR FREE
EMU OIL FORMULA

BLUE-EMU[®]

America's Number One
Emu Oil Formula

The Blue-Emu[®] product line is the #1 selling over-the-counter emu oil brand in the United States. Our products, loved by both customers and Hall of Famers like Rusty Wallace and Johnny Bench, are made with soothing Aloe Vera and penetrating Emu Oil right here in the USA. Our products allow you to get back in the game without smelling like a locker room.

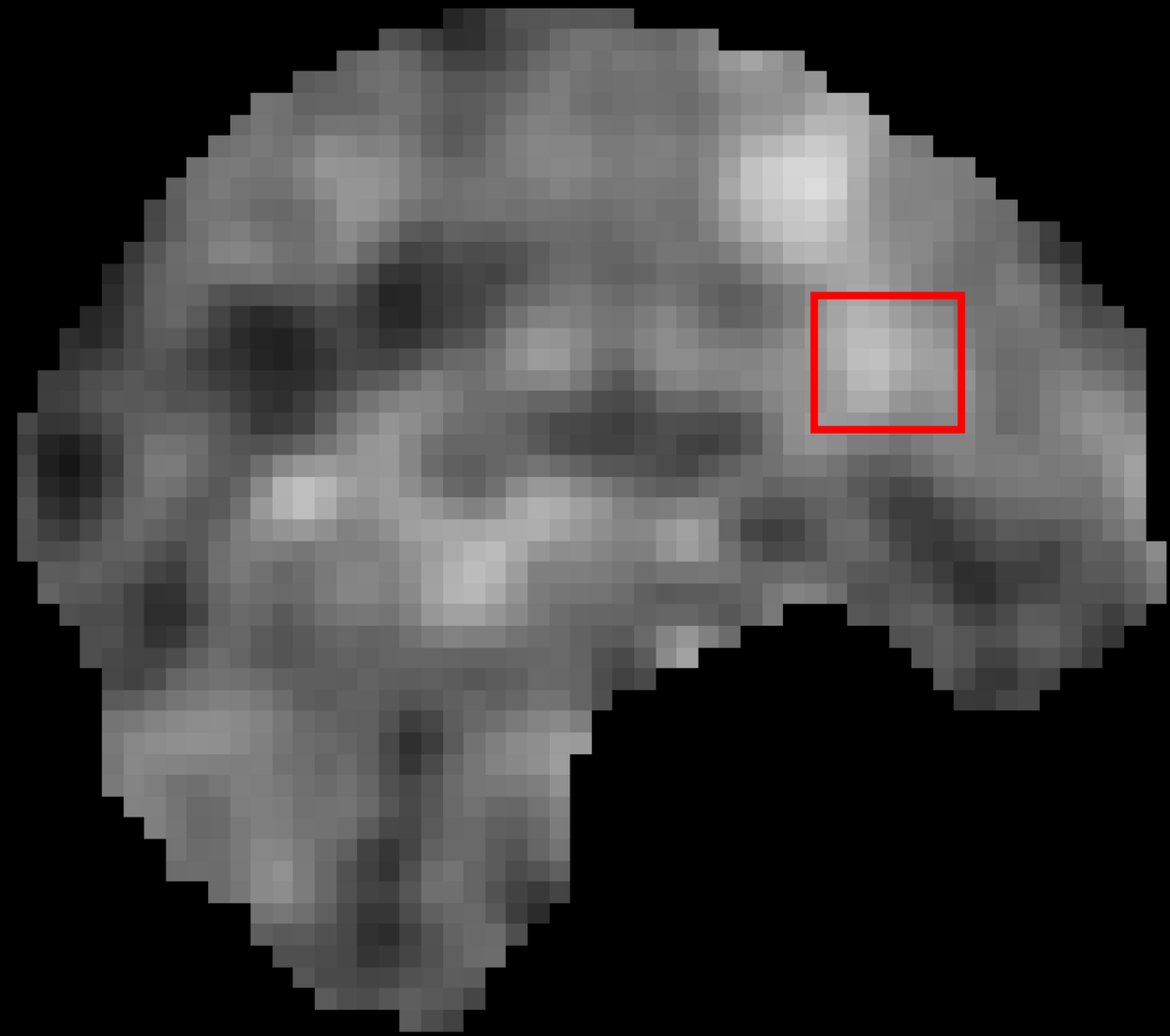
Our product line includes Original Blue Emu[®] cream, Blue-Emu[®] Continuous Pain Relief Spray, Blue-Emu[®] Maximum Arthritis Pain Relief Cream, and Blue-Emu[®] Maximum Strength Lidocaine Numbing Pain Relief Cream. You can feel 100% confident in using our products to help soothe tired muscles and joints.

Correcting for Multiple Comparisons

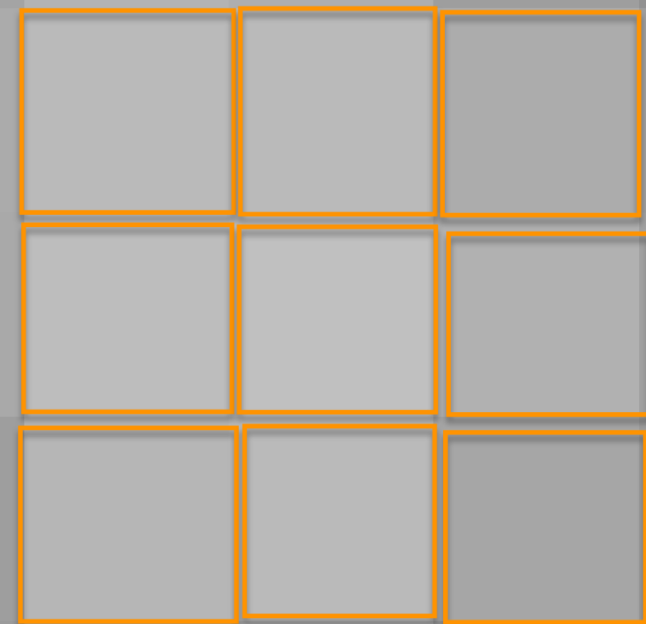
Now consider an fMRI dataset

Quick poll: How many voxels in a typical volume?

Depends on voxel size, but usually between 100k-300k



Test for significance in each voxel



Correcting for Multiple Comparisons

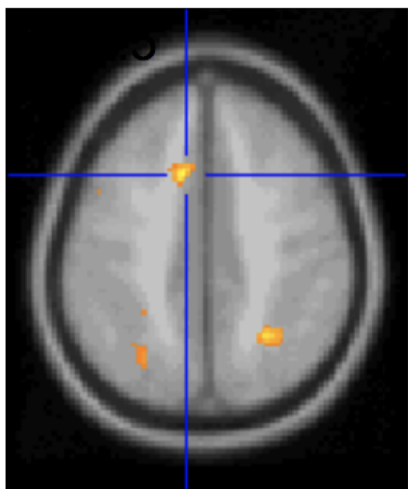
Balancing act between Type I and Type II errors

Legal Errors

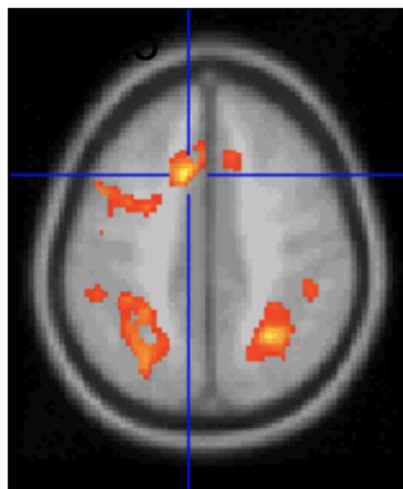
		Actual Criminal	
		Yes Alternate Hypo	No Null Hypothesis
Decision on the basis of Case Trial	Punish (Criminal)	Good Decision	Convicting the Innocent
	Acquit (Innocent)	Acquit Guilty	Good Decision

Correcting for Multiple Comparisons

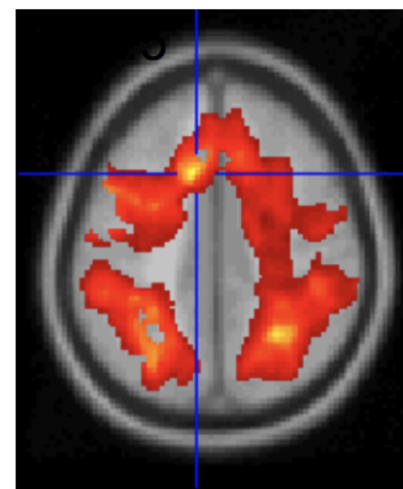
High Threshold



Med. Threshold



Low Threshold



Good Specificity

Poor Power

(risk of false negatives)

Poor Specificity
(risk of false positives)

Good Power

Correcting for Multiple Comparisons

What can be done?

Bonferroni correction

FDR correction

Cluster correction

Bonferroni Correction

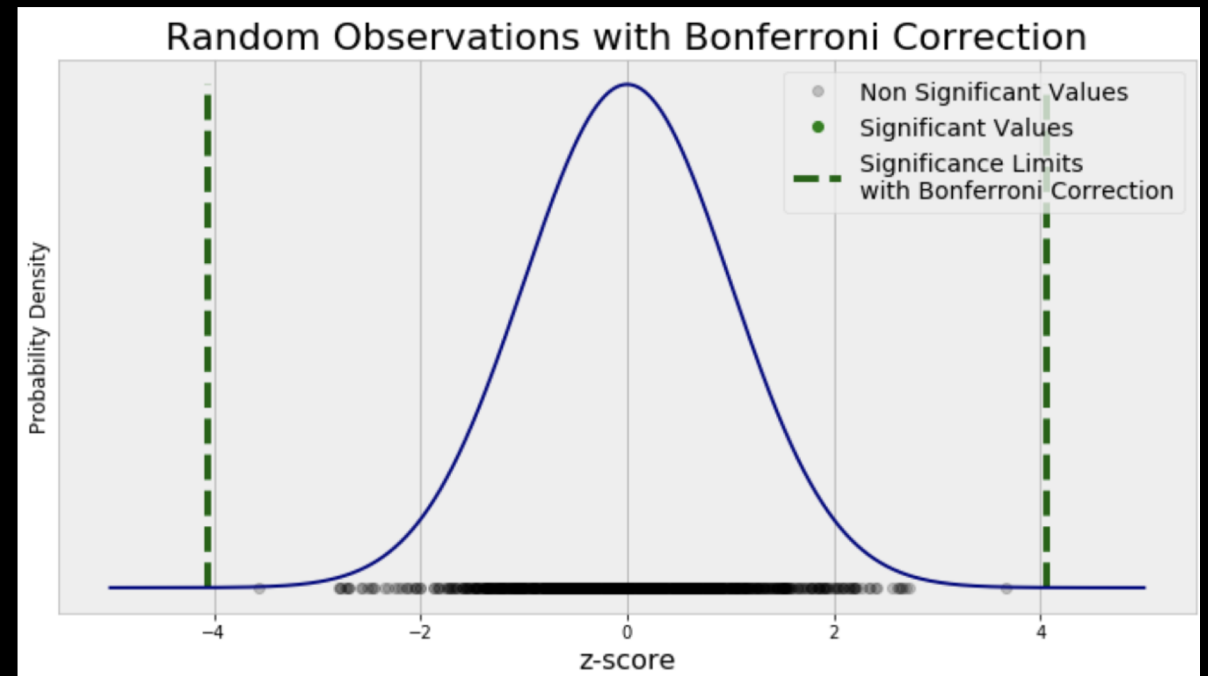
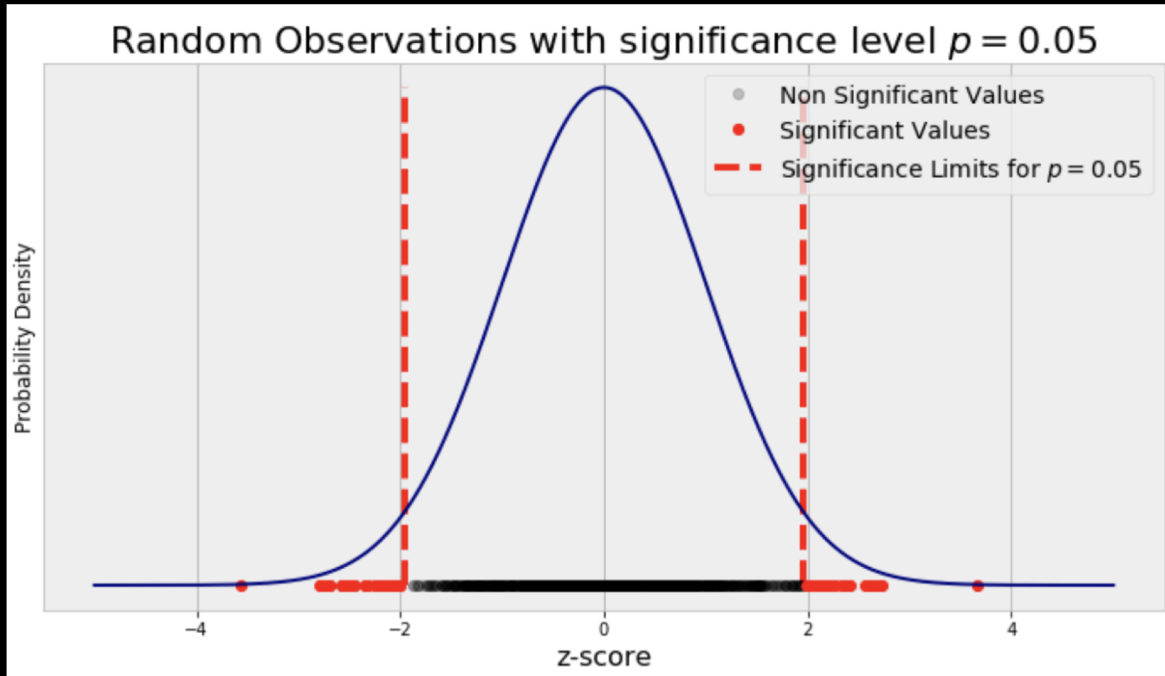
Simplest correction method to understand and calculate

Given an alpha level α and number of tests n ,
the corrected alpha level can be found by α/n

Example: $\alpha=0.05$, $n=10$

$$\alpha = 0.005$$

Bonferroni Correction



Bonferroni Correction

Example: $\alpha=0.05$, $n=100,000$

$$\alpha = 0.000005 (!)$$

This revised alpha is then used at each voxel in the analysis

Bonferroni Correction

Pros: Easy to understand, easy to use

Excellent for guarding against Type I error

Cons: Conservative, too severe for fMRI

Inflates the probability of Type II errors

Alternative: False Discovery Rate (FDR)

Bonferroni and other correction methods control for the Probability of observing a single false positive

FDR: Control the fraction of false positives

i.e.: You know there will be a certain percentage of false positives, but you can live with it

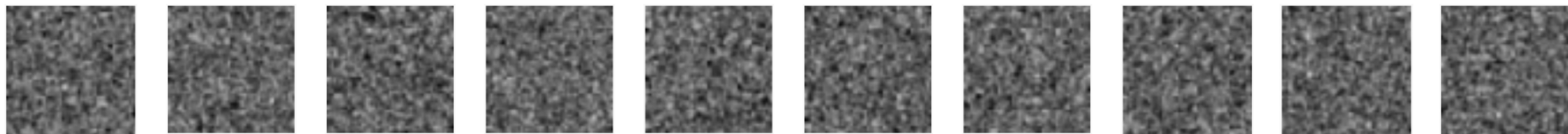
Alternative: False Discovery Rate (FDR)

- FDR

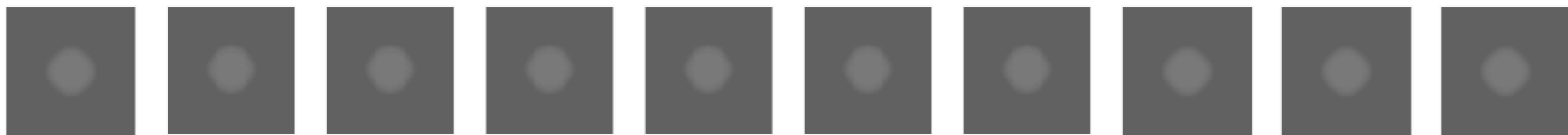
- E (# of true null declared active / # voxels declared active)

	Declared active	Fail to Declare active	Total
Non-active	50	950	1000
Active	80	20	100
Total	130	970	1100

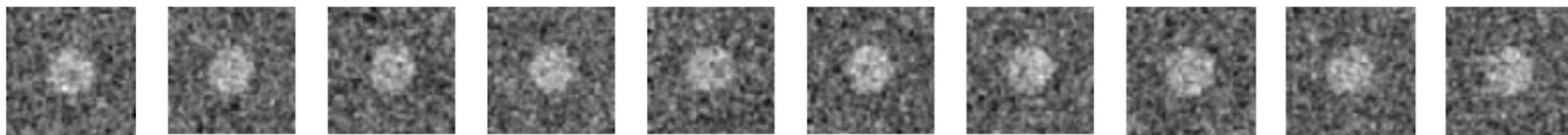
Noise



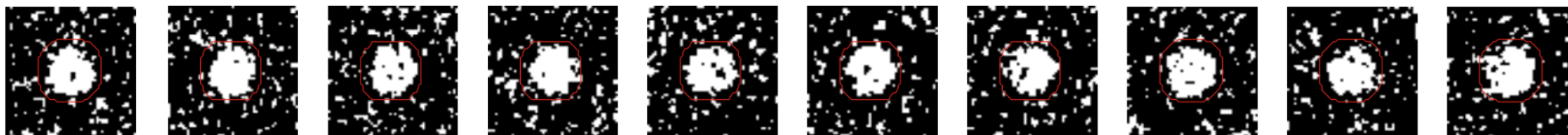
Signal



Signal+Noise



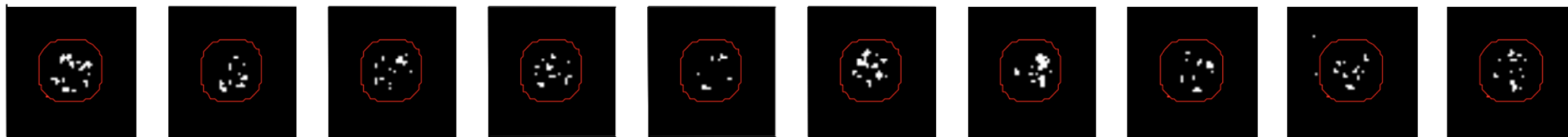
Control of Per Comparison Rate at 10%



11.3% 11.3% 12.5% 10.8% 11.5% 10.0% 10.7% 11.2% 10.2% 9.5%

Percentage of Null Pixels that are False Positives

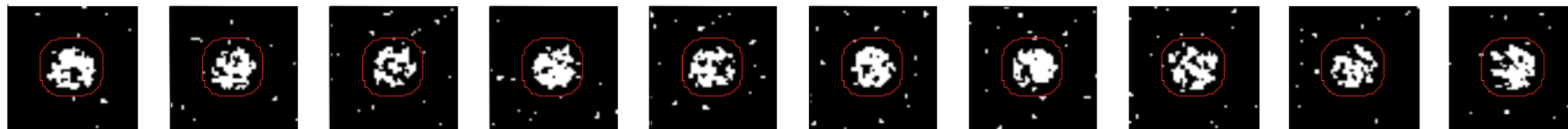
Control of Familywise Error Rate at 10%



FWE

Occurrence of Familywise Error

Control of False Discovery Rate at 10%



6.7% 10.4% 14.9% 9.3% 16.2% 13.8% 14.0% 10.5% 12.2% 8.7%

Percentage of Activated Pixels that are False Positives

Cluster Correction

Bonferroni might be appropriate if each voxel were independent

But are they? Consider how the brain is designed

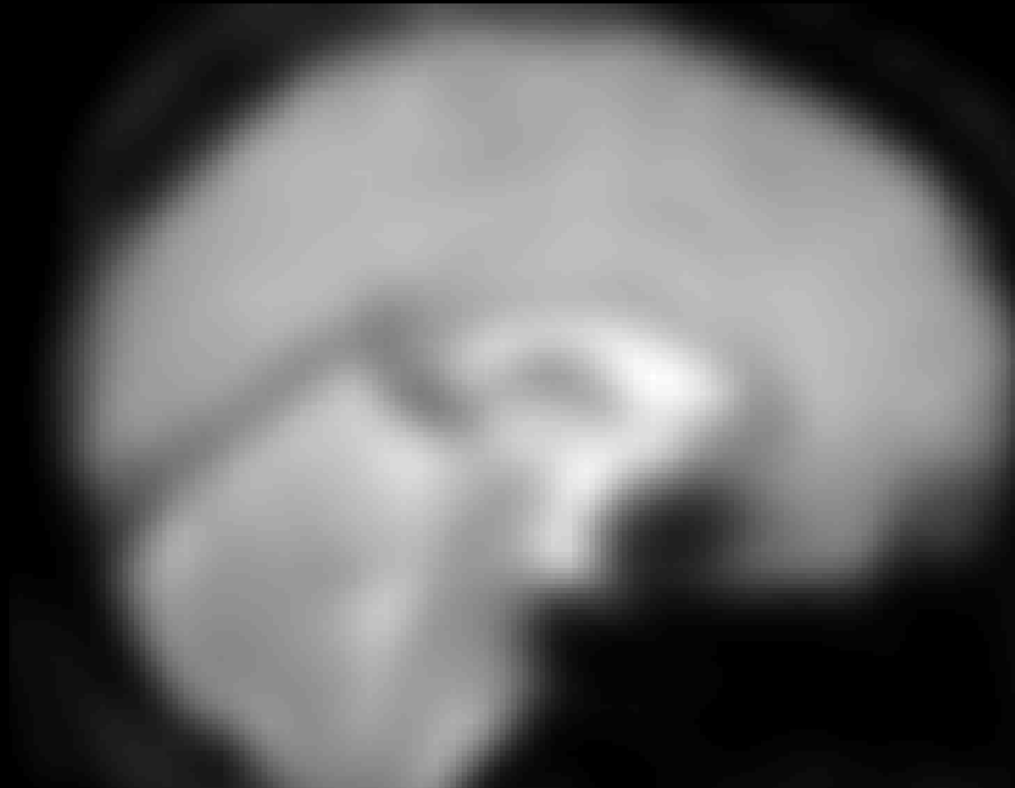
Cluster Correction

Functional image, unsmoothed



Cluster Correction

Functional image, smoothed



Cluster Correction

In SPM, cluster correction thresholds are calculated with
Random Field Theory (RFT)

Accounts for the spatial smoothness of the data

Based on the estimated FWHM_x (not the same as
applied FWHM_x!)

Cluster Correction

Example after 8mm smoothing kernel

table shows 3 local maxima more than 8.0mm apart

Height threshold: $T = 2.34$, $p = 0.010$ (1.000)
Extent threshold: $k = 30$ voxels, $p = 0.107$ (0.999)
Expected voxels per cluster, $\langle k \rangle = 11.965$
Expected number of clusters, $\langle c \rangle = 7.21$
FWEp: 4.869, FDRp: Inf, FWEc: 198, FDRc: 198

Degrees of freedom = [1.0, 278.0]
FWHM = 10.5 10.5 10.2 mm mm mm; 3.5 3.5 3.4 {voxels}
Volume: 1811403 = 67089 voxels = 1431.6 resels
Voxel size: 3.0 3.0 3.0 mm mm mm; (resel = 41.85 voxels)

Cluster Correction

Statistics: *p-values adjusted for search volume*

set-level		cluster-level			
p	c	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	k_E	p_{uncorr}
0.846	5	0.026	0.024	198	0.000
		0.537	0.175	85	0.011
		0.999	0.794	31	0.102
		0.489	0.175	89	0.010
		0.105	0.050	146	0.002

Multiple Comparisons Correction: Summary

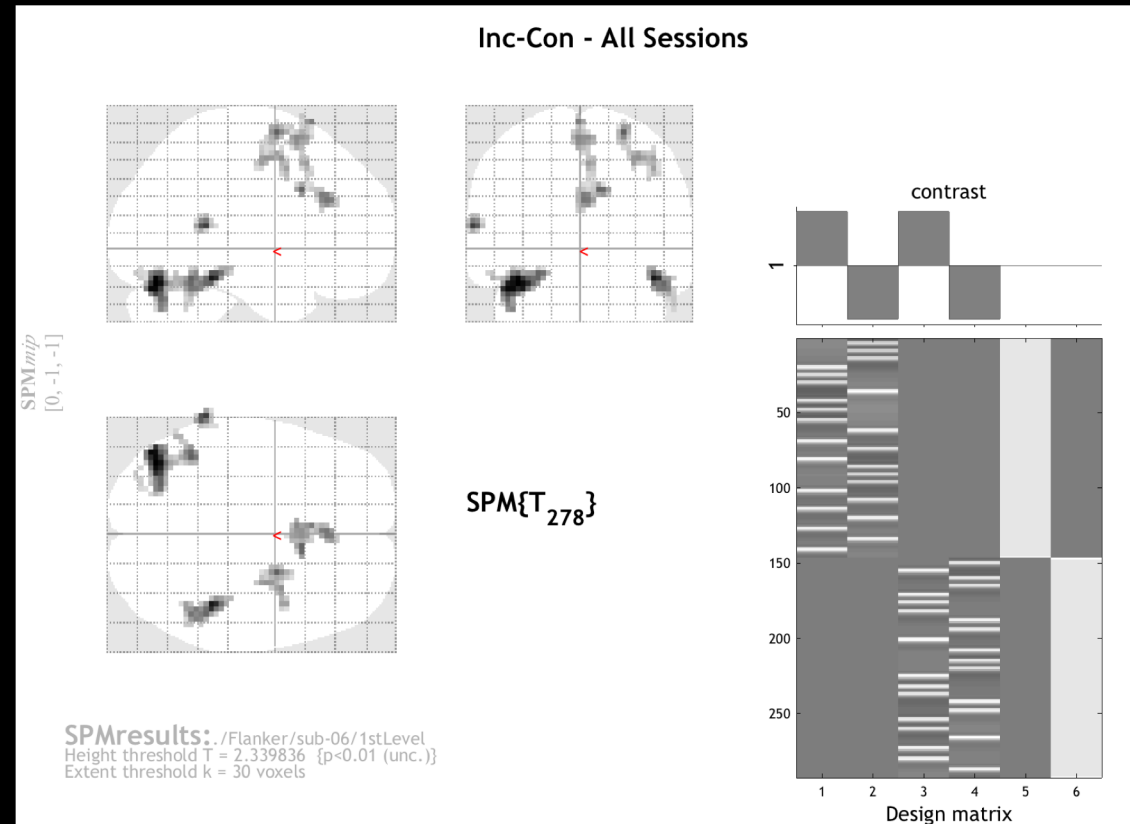
Most common method is cluster-wise thresholding

Cons: Loss of spatial specificity

As we will see tomorrow, you should use a
Cluster-forming threshold of $p=0.001$ for most experiments

Non-parametric options seem to be getting more popular

Applying this to a dataset



Statistics: *p-values adjusted for search volume*

set-level		cluster-level				peak-level					mm mm mm		
p	c	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	k_E	p_{uncorr}	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	T	(Z_E)	p_{uncorr}			
0.846	5	0.026	0.024	198	0.000	0.716	0.996	4.00	3.94	0.000	-39	-73	-22
						1.000	0.996	3.26	3.23	0.001	-42	-52	-19
						1.000	0.996	3.24	3.20	0.001	-27	-70	-16
		0.537	0.175	85	0.011	0.912	0.996	3.79	3.74	0.000	42	-40	-19
						1.000	0.996	3.12	3.09	0.001	48	-52	-25
						0.999	0.794	31	0.102	0.998	0.996	3.48	3.44
		0.489	0.175	89	0.010	1.000	0.996	3.26	3.22	0.001	24	-1	65
						1.000	0.996	2.91	2.89	0.002	33	2	50
						1.000	0.996	2.89	2.86	0.002	30	-7	50
		0.105	0.050	146	0.002	1.000	0.996	3.24	3.20	0.001	9	14	32
						1.000	0.996	3.11	3.08	0.001	0	29	26
						1.000	0.996	3.04	3.01	0.001	-3	20	71

Applying this to a dataset

Set-level: Probability of finding that many clusters

Cluster-level: Probability of finding a cluster of a given size

Peak-level: Probability of a statistic that size in that voxel

Applying this to a dataset

Statistics: *p-values adjusted for search volume*

set-level		cluster-level				peak-level					mm	mm	mm
p	c	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	k_E	p_{uncorr}	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	T	(Z_E)	p_{uncorr}			
0.846	5	0.026	0.024	198	0.000	0.716	0.996	4.00	3.94	0.000	-39	-73	-22
						1.000	0.996	3.26	3.23	0.001	-42	-52	-19
						1.000	0.996	3.24	3.20	0.001	-27	-70	-16
		0.537	0.175	85	0.011	0.912	0.996	3.79	3.74	0.000	42	-40	-19
						1.000	0.996	3.12	3.09	0.001	48	-52	-25
						0.999	0.794	31	0.102	0.998	0.996	3.48	3.44
		0.489	0.175	89	0.010	1.000	0.996	3.26	3.22	0.001	24	-1	65
						1.000	0.996	2.91	2.89	0.002	33	2	50
						1.000	0.996	2.89	2.86	0.002	30	-7	50
						0.105	0.050	146	0.002	1.000	0.996	3.24	3.20
1.000	0.996	3.11	3.08	0.001	0	29	26						
					1.000	0.996	3.04	3.01	0.001	-3	20	71	

Demonstration

Other Statistical Scenarios

Once you calculate a contrast, are you done?

Consider this: My brother and I both play basketball. If I tell you that I am slightly better than he is, does that mean:

We are both really good, but I'm just a little better?

I'm a little above average, and he's a little below average?

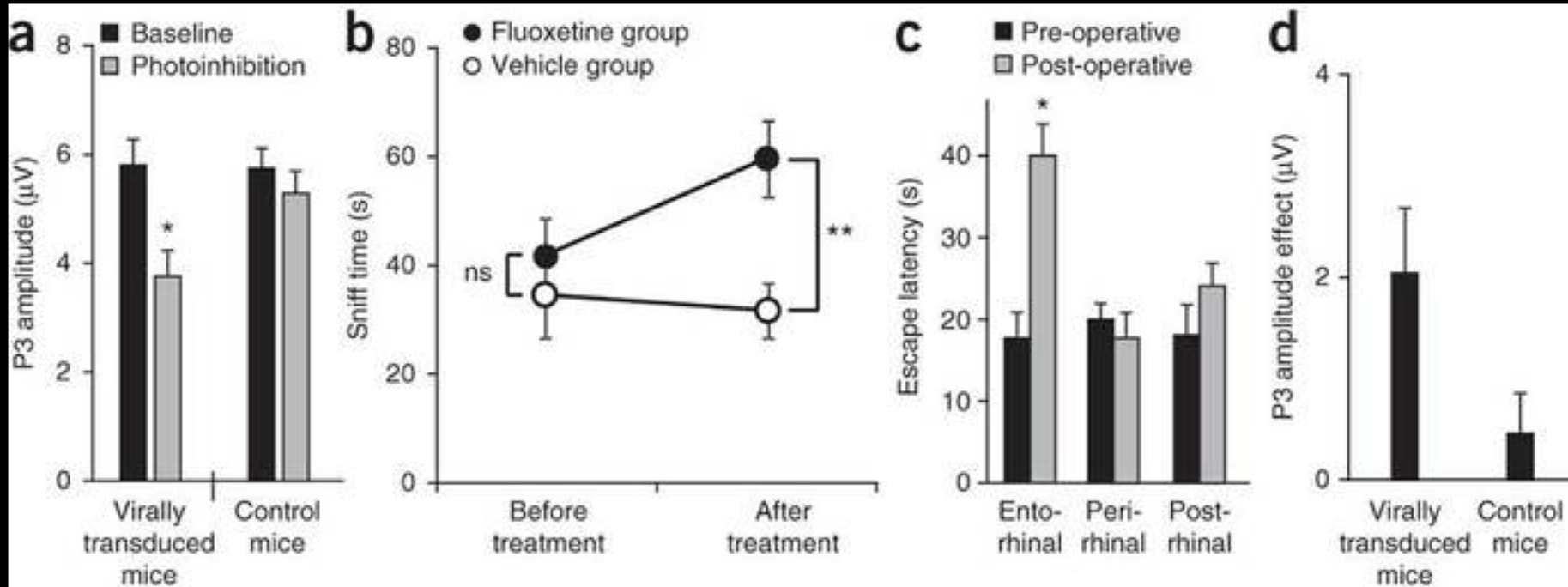
Maybe we're both terrible, and I'm just a little better than he is

Double Dissociations

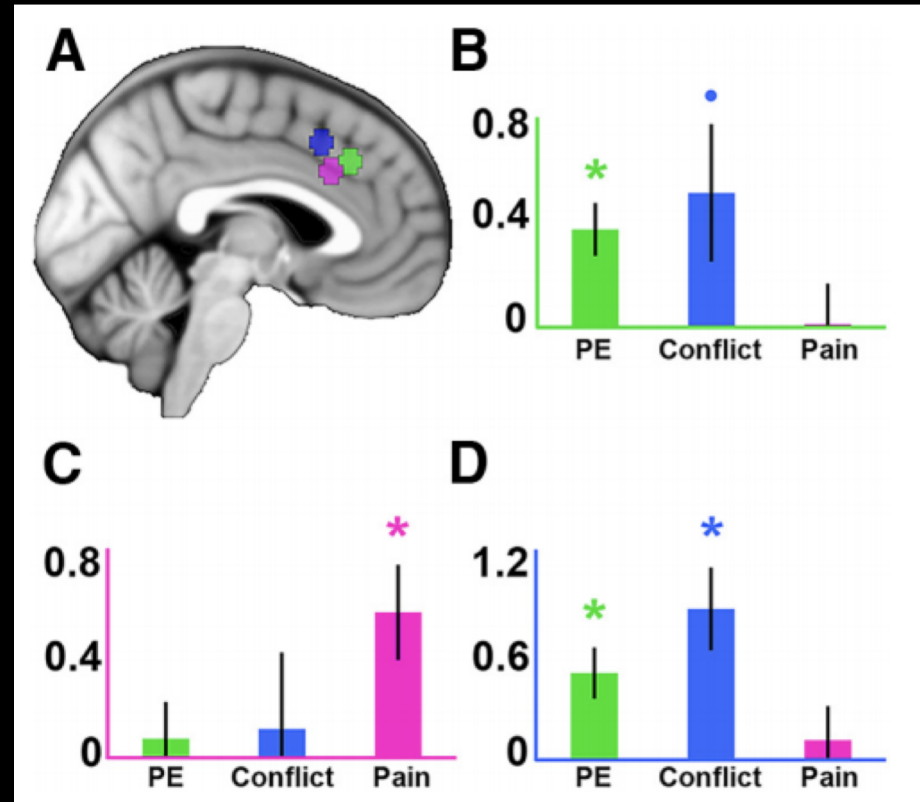
Condition A is significant in region A but not region B, and condition B is significant in region B but not region A

**Remember to run a paired t-test within each region,
and also a Region x Condition interaction**

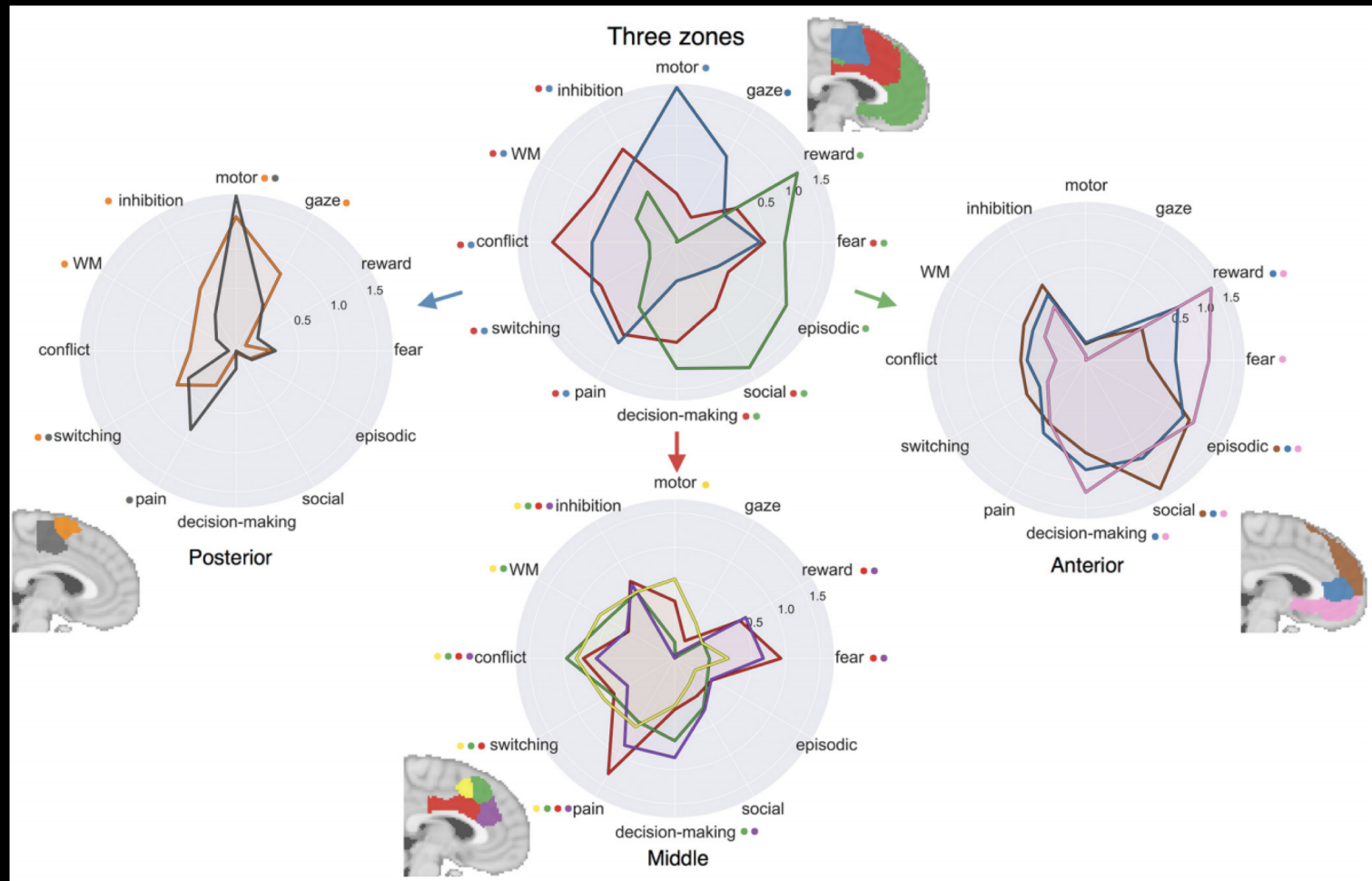
Double Dissociations



Double Dissociations



Triple Dissociations (!)



Preview: ROI Analysis

**In the examples just shown, the data was extracted from
Regions of Interest (ROIs)**

That is, subsets of voxels that we are interested in

- HUMAN ATLAS->TD brodmann areas+
- ..
 - brodmann area 1
 - brodmann area 2
 - brodmann area 3
 - brodmann area 4
 - brodmann area 5
 - brodmann area 6
 - brodmann area 7
 - brodmann area 8
 - brodmann area 9
 - brodmann area 10
 - brodmann area 11
 - brodmann area 12
 - brodmann area 13
 - brodmann area 14
 - brodmann area 15
 - brodmann area 16
 - brodmann area 17
 - brodmann area 18
 - brodmann area 19
 - brodmann area 20
 - brodmann area 21
 - brodmann area 22
 - brodmann area 23
 - brodmann area 24
 - brodmann area 25
 - brodmann area 26
 - brodmann area 27
 - brodmann area 28
 - brodmann area 29
 - brodmann area 30
 - brodmann area 31
 - brodmann area 32
 - brodmann area 33
 - brodmann area 34
 - brodmann area 35
 - brodmann area 36
 - brodmann area 37
 - brodmann area 38
 - brodmann area 39
 - brodmann area 40
 - brodmann area 41
 - brodmann area 42
 - brodmann area 43
 - brodmann area 44
 - brodmann area 45
 - brodmann area 46
 - brodmann area 47
 - Amygdala
 - Anterior Commissure
 - Caudate Body
 - Caudate Head
- Atlas Information

BASIC ADVANCED

ADD ->

MOVE ALL ->>

<- REMOVE SELECTED

<<- REMOVE ALL

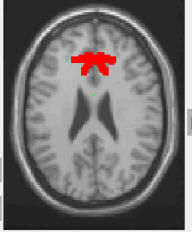
2D 3D

DILATE: 1

Flip Lock L/R U/D

Left Left + Right Right

Display: Neurologic 49



L R

ITKSnap

Mag

ANALYSIS RESULTS

Write Independent Regions

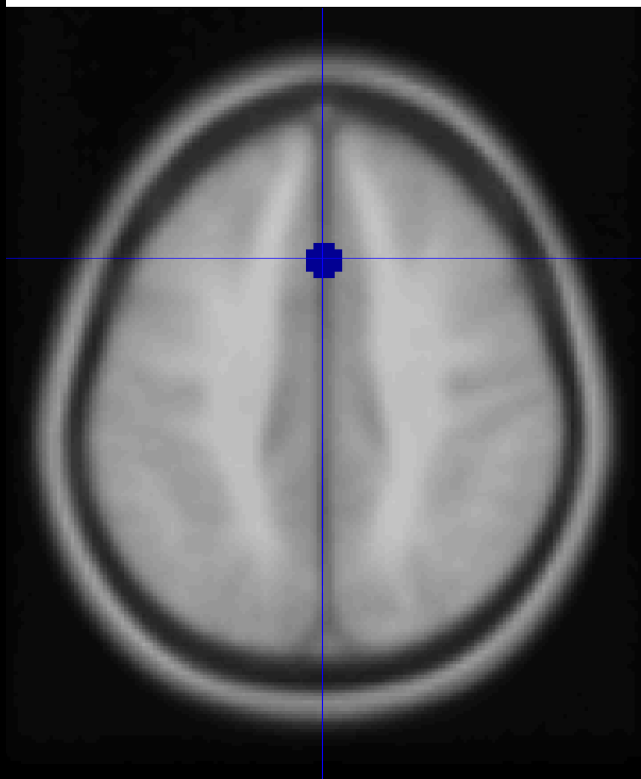
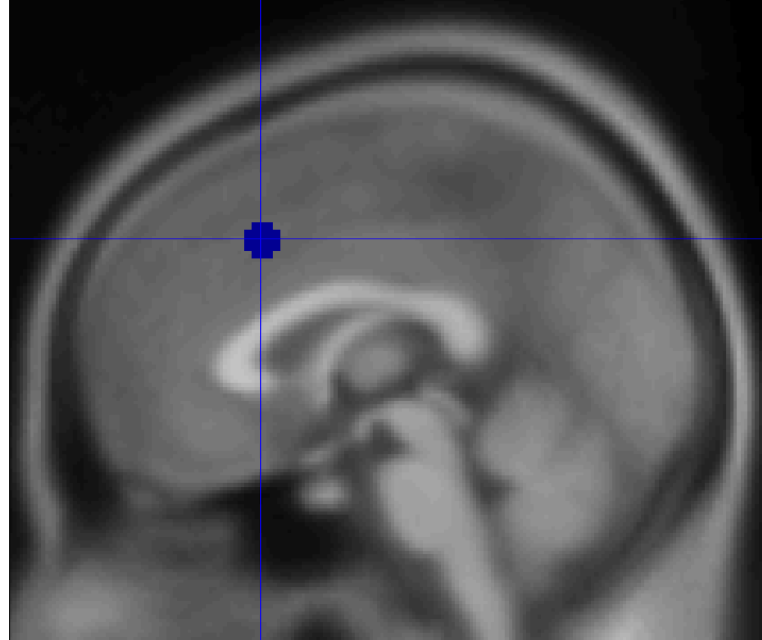
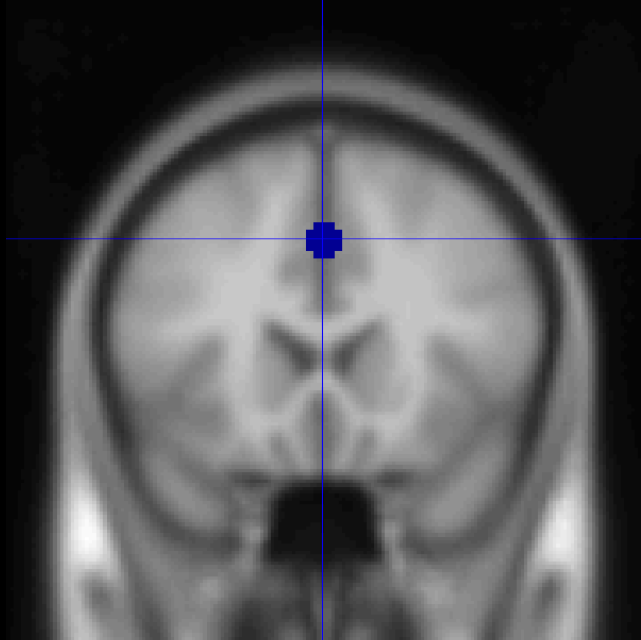
SAVE MASK

DONE CANCEL

WORKING REGION1

brodmann area 32

	CUBE	46	64	37	GO	ATLAS	SUBREGION	VALUE
MNI	0	0	0	GO	TD brodmann areas+	NA	1000	
Tal	0	0	0	GO	TD Lobes	NA	1500	



1: dACC_Sphere

Label: dACC_Sphere

Centre of mass: 0 20 40

Volume (mm): 648.00

Max/min X(mm): -4 4

Max/min Y(mm): 16 24

Max/min Z(mm): 36 44

Questions?

Lab Preview

Create an ideal experimental design, *before* collecting data

**For AFNI users: Similar to using the `-nodata` option
in `3dDeconvolve`, calculating correlations**

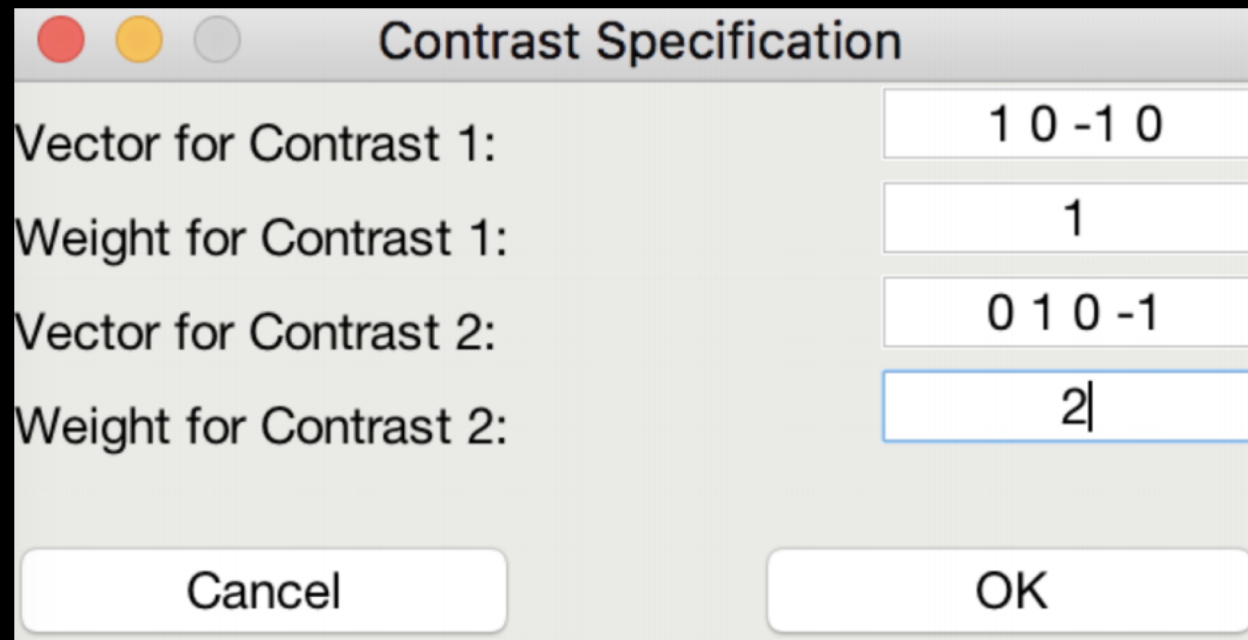
Parameters to change: ISI, number of trials, and contrasts

Lab Preview

```
ts = output.X0;           % get the raw timeseries from 'output'
tr = output.TR;          % do the same for the TR
irf = spm_hrf(tr);       % create the impulse response function
cts = conv(ts, irf);     % convolve the timeseries with the IRF
cts = cts(1:length(ts)); % ensure that the long tail of the hemodynamic response hasn't
lengthened our image timeseries
```

Lab Preview

Using OptimizeX to specify which contrasts you are interested in



A dialog box titled "Contrast Specification" with a standard macOS window title bar (red, yellow, and grey buttons). The dialog contains four input fields for specifying contrasts. The first two fields are for Contrast 1, and the next two are for Contrast 2. The "Weight for Contrast 2" field is currently selected with a blue border. At the bottom, there are "Cancel" and "OK" buttons.

Vector for Contrast 1:	1 0 -1 0
Weight for Contrast 1:	1
Vector for Contrast 2:	0 1 0 -1
Weight for Contrast 2:	2

Buttons: Cancel, OK

Lab Preview

Create several design matrices, calculate VIF

We will provide the code for this

Lab Preview

Exploring the SPM.mat file

details on experiment:

SPM.xY.RT - TR length (RT ="repeat time")

SPM.xY.P - matrix of file names

SPM.xY.VY - # of runs x 1 struct array of mapped image volumes (.img file info)

SPM.modality - the data you're using (PET, FMRI, EEG)

SPM.stats.[modality].UFp - critical F-threshold for selecting voxels over which the non-sphericity is estimated (if required) [default: 0.001]

SPM. stats.maxres - maximum number of residual images for smoothness estimation

SPM. stats.maxmem - maximum amount of data processed at a time (in bytes)

SPM.SPMid - version of SPM used

SPM.swd - directory for SPM.mat and img files. default is pwd

basis function:

SPM.xBF.name - name of basis function

SPM.xBF.length - length in seconds of basis

SPM.xBF.order - order of basis set

SPM.xBF.T - number of subdivisions of TR

SPM.xBF.T0 - first time bin (see slice timing)

SPM.xBF.UNITS - options: 'scans'!'secs' for onsets

SPM.xBF.Volterra - order of convolution

SPM.xBF.dt - length of time bin in seconds

SPM.xBF.bf - basis set matrix