

Experimental Design for Imaging

Susan Bookheimer
UCLA School of Medicine

fMRI experimental design: A basic plan

Define mental process
to examine



Define tasks to manipulate
that process



Measure fMRI data
during tasks



Compare fMRI data
between tasks

Replace “fMRI data” with “RT”
and you have cognitive
psychology!

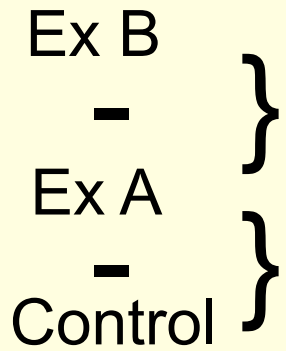
Conceptual and methodological aspects of experimental design

- There are two aspects of fMRI design that are important to distinguish
- Conceptual design
 - How do we design tasks to properly measure the processes of interest?
 - The issues here are very similar to those in cognitive psychology
- Methodological design
 - How can we construct a task paradigm to optimize our ability to measure the effects of interest, within the specific constraints of the fMRI scanning environment?

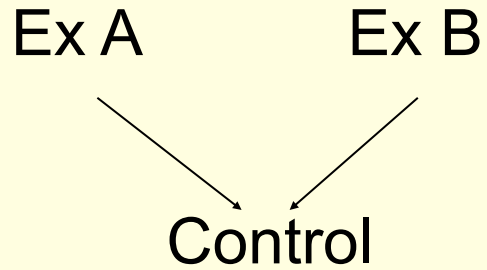
IV's and contrasts: basics

- There are (almost always) two or more conditions in activation imaging
- We make a series of assumptions about the **cognitive** and the **neural** processes involved, and their relation to each other, in every experiment
- The logic involved and choosing tasks and contrasting them, and the problems of assumptions in these choices, spans all experimental designs
- In this context, makes no difference whether we use event related or blocked designs, eg. “Null” events in ER designs often = “rest” in block designs

Hierarchical



Common baseline



Parallel Comparisons

Ex B > Ex A

Ex A > Ex B

Tailored Baseline

Ex A > Ctl A
Ex B > Ctl B } >

Parametric

A < A < A < A

Selective attention

A B C
A B C
A B C

Factorial Designs

Ex A Ex B
 AxB

Conjunction Designs

	A	B		D	
			C	D	
		B			E

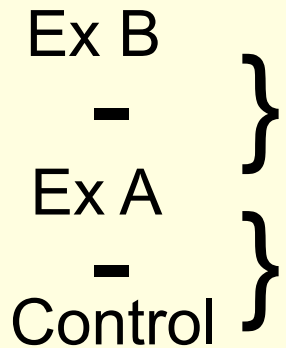
Priming/Adaptation Designs

A → A
A → B

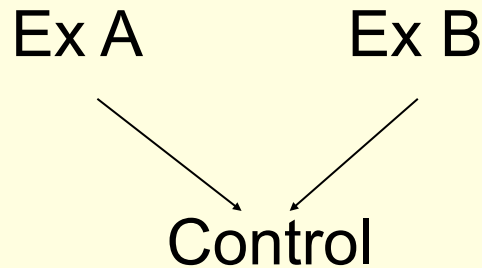
ROI based approaches

Mixed, Nested Designs

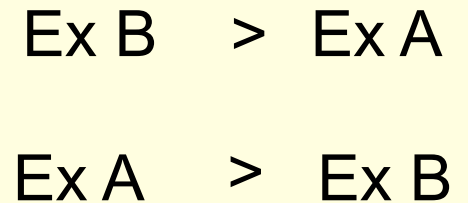
Hierarchical



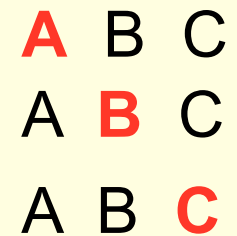
Common baseline



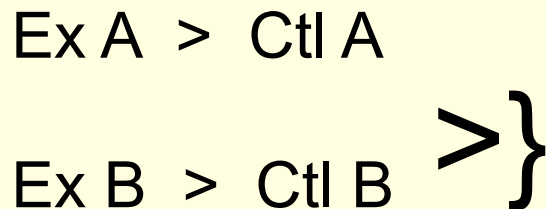
Parallel



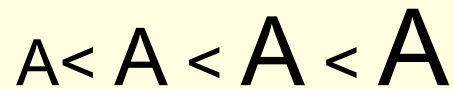
Selective attention



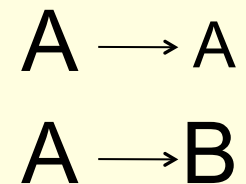
Tailored Baseline



Parametric



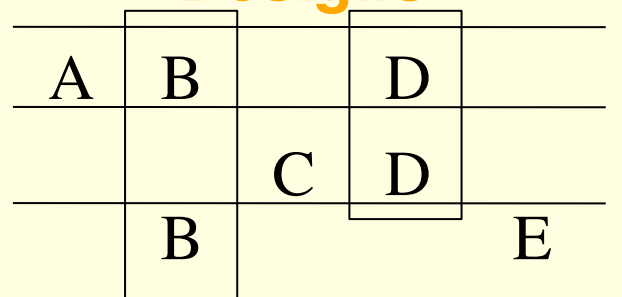
Priming/Adaptation Designs



Factorial Designs



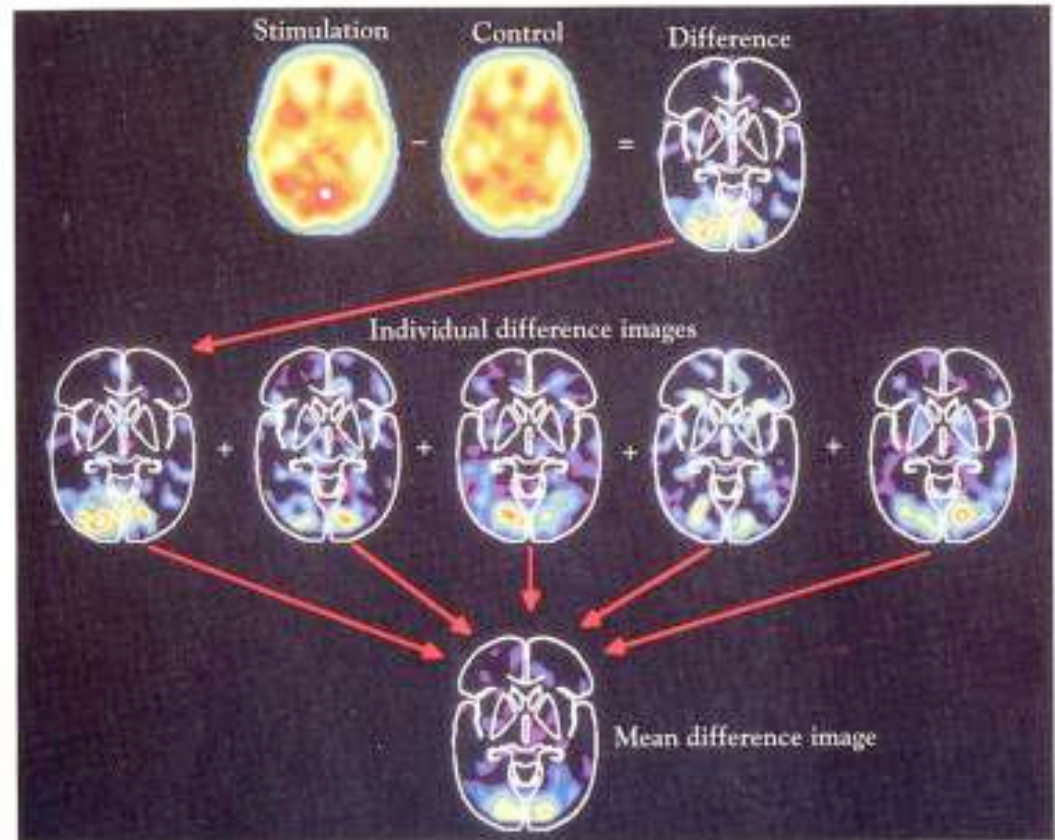
Conjunction Designs



Mixed, Nested Designs

The subtraction method

- Acquire data under two conditions
 - These conditions putatively differ only in the cognitive process of interest
- Compare brain images acquired during those conditions
- Regions of difference reflect activation due to the “subtracted” process of interest



Petersen et al., 1988

Hierarchical subtraction

example from Petersen, 1991

- Rest Control - } Sensory
- Auditory words vs. rest: A1, word recognition centers
- Visual words vs rest: visual areas, word form areas - } Motor
- Reading or repeating words vs passive words: motor areas - } Semantic
- Generating words vs. repeating: semantic (language) areas

PI assumptions in hierarchical designs

- In additional hierarchical levels, “controlled” variables do not change
 - Eg, reading words aloud vs silently does not change visual processes
- We know this to be false
- Multiple hierarchical contrasts accentuate the fallacies

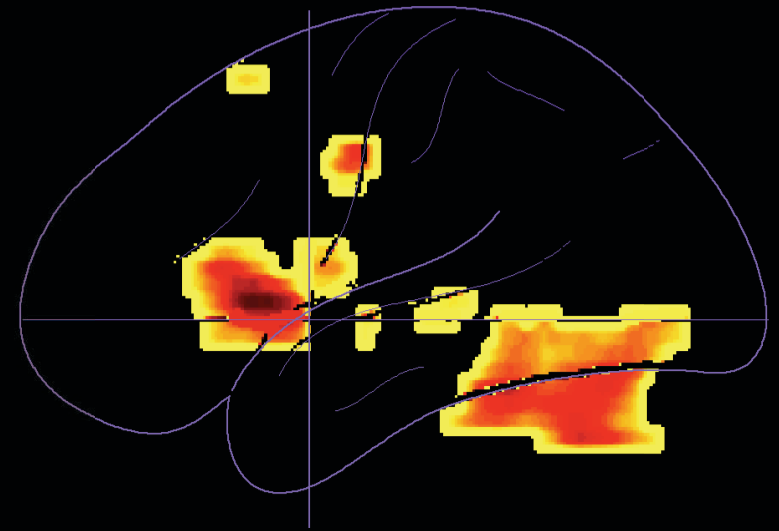
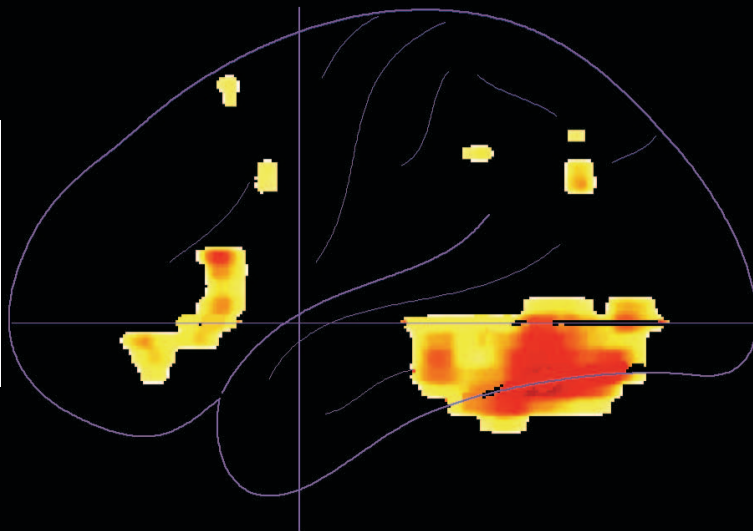
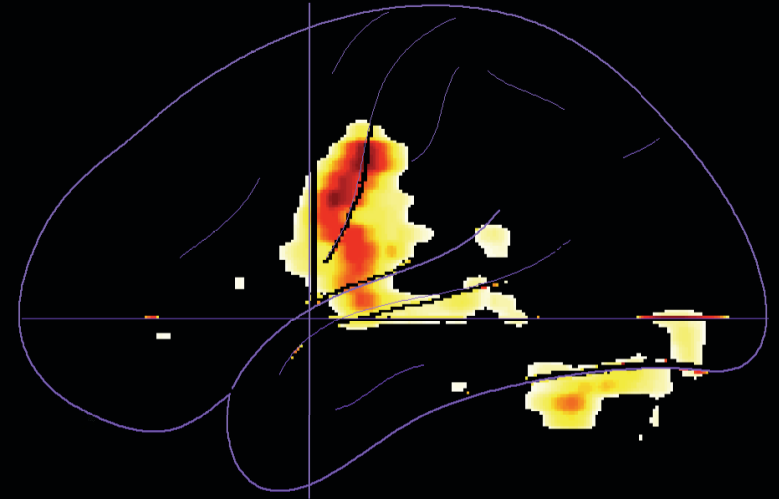
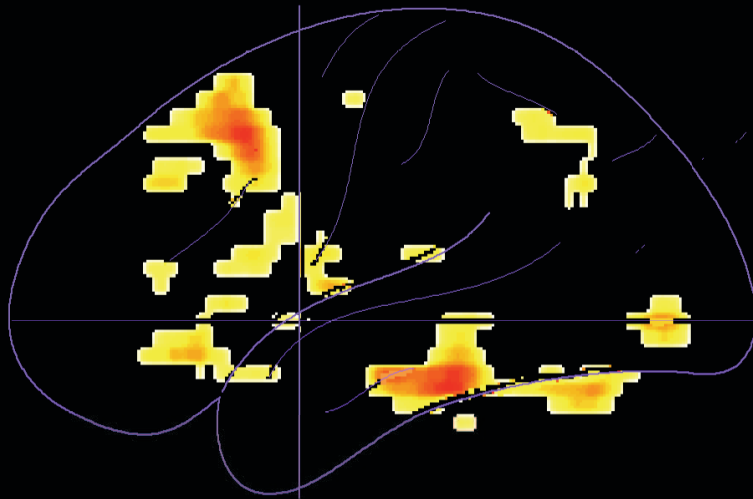
The pure insertion assumption

- Subtraction requires a strong assumption of “pure insertion”
 - Insertion of a single cognitive process does not affect any of the other processes (no *interactions*)
- Failure of PI means that the results cannot be interpreted with regard to the specific cognitive process of interest
- PI must hold at both neural and cognitive levels
- Also make assumptions about equivalence of task effort and difficulty level

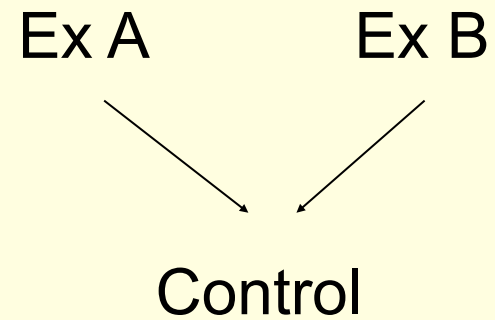
'House'

"House"

Read
"HOUSE"



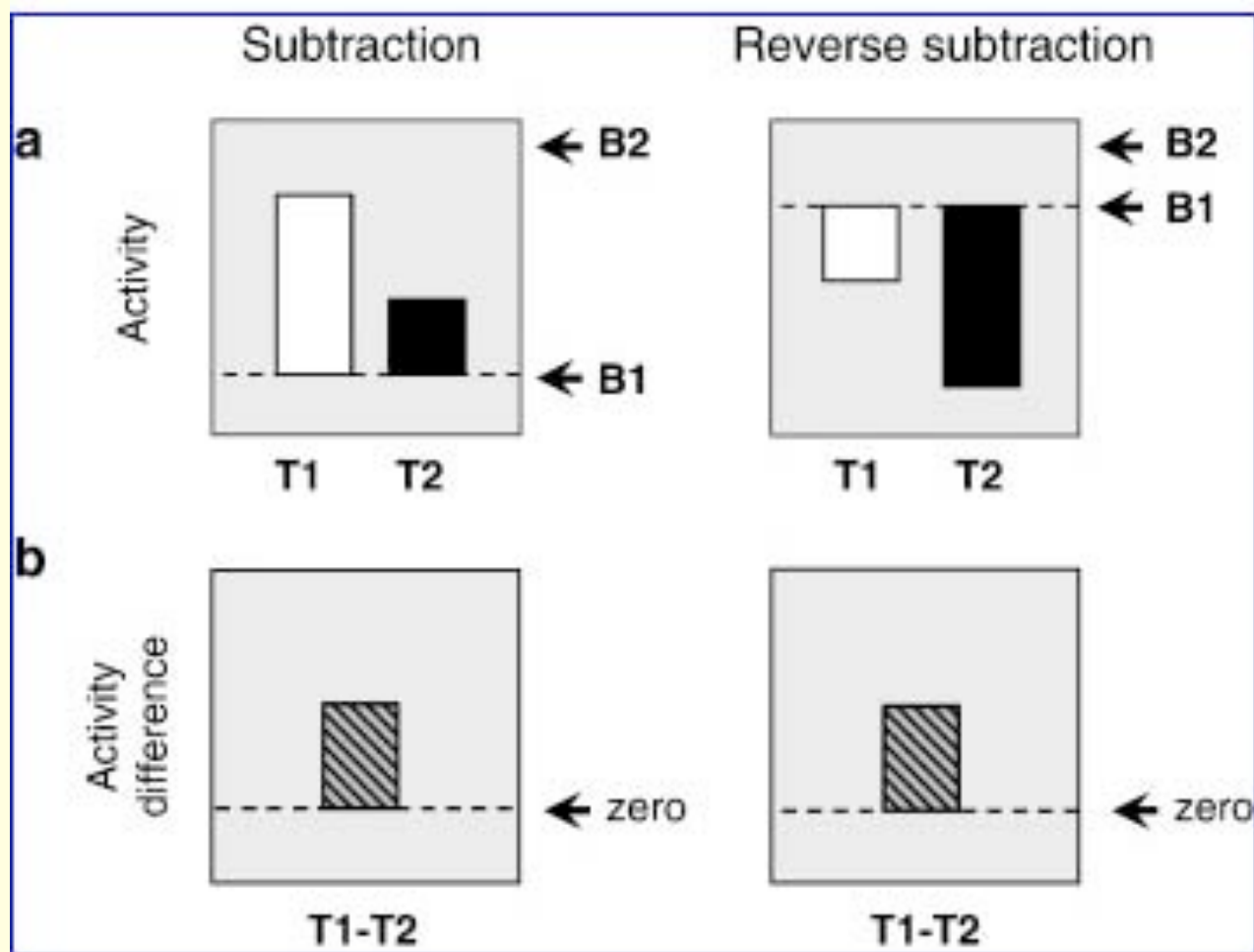
Common Baseline



- One level of hierarchy
- Still have assumptions of pure insertion, but only at 1 level in hierarchy
- Assumes A and B have similar psychometric properties (ie, level of difficulty, variability)
- Test for violation of additivity assumption
- Allows you to see common areas active for A and B
- Need additional approach to see unique areas

Simple subtraction

- Two conditions
- No assumptions of linearity (any relationship will be linear by definition)
- Make assumptions about
 - What your tasks are doing- do they tap into the processes of interest
 - How they differ (what variables are shared, what are unique)
- Often assume (mistakenly) that differences are due to increases in one condition- that which is the “higher order” task or the experiment (vs. control) task



From Morcom and Fletcher, NeuroImage, 2006

Factorial design

- A factorial design involves multiple concurrent subtractions
- Allows for testing of interactions between components
- Still requires pure insertion assumption and task decomposition
 - But additivity can be tested for the specific factors that are manipulated

Ex B > Ex A

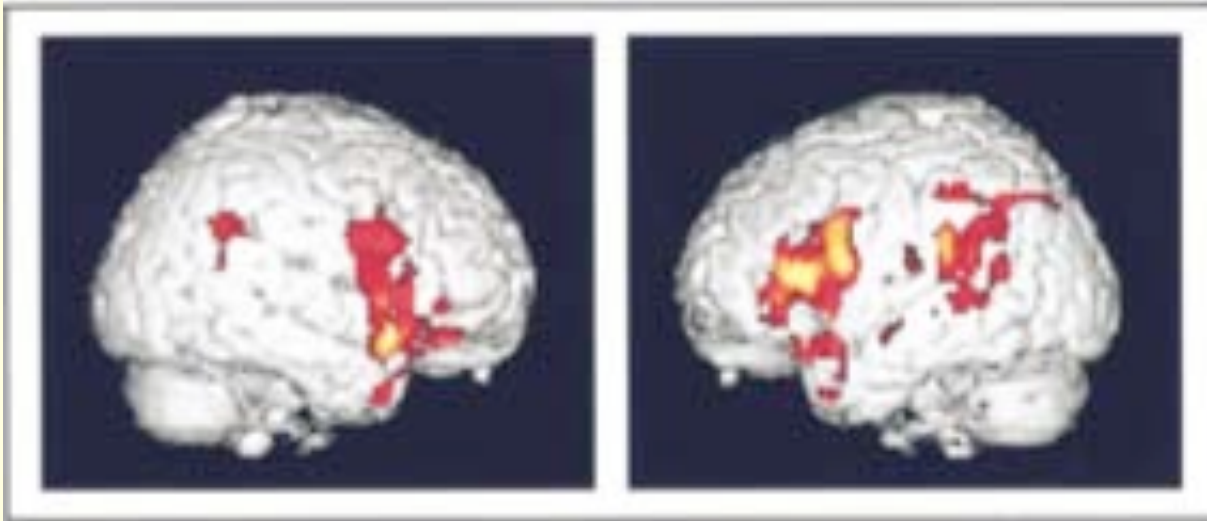
Ex A > Ex B

- **Parallel Comparisons**

- Want to compare 2 tasks with no notion of a hierarchical relationship
- Task A vs B; B vs A
- EG: silent vs. oral reading and reverse
- EG: Seeing words vs. hearing words
- Alone, see no common areas; add common baseline
- Use common baseline as mask to reduce errors and increase power in likely areas
- Strong assumption of similar psychometric properties

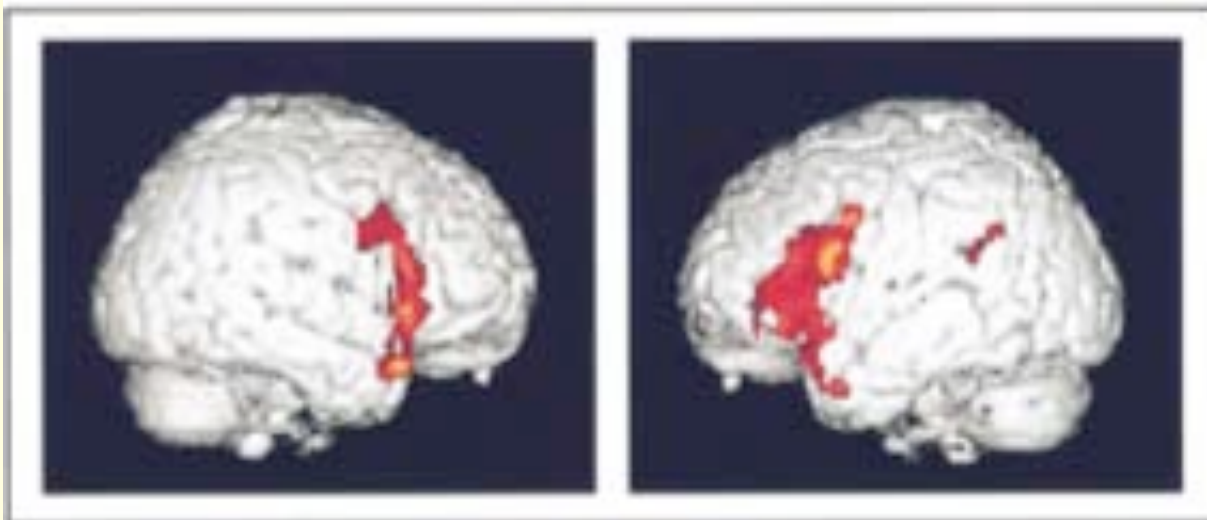
Example: Sentence judgment task: semantic vs. syntactic (Dapretto, 1999)

A



Word order change

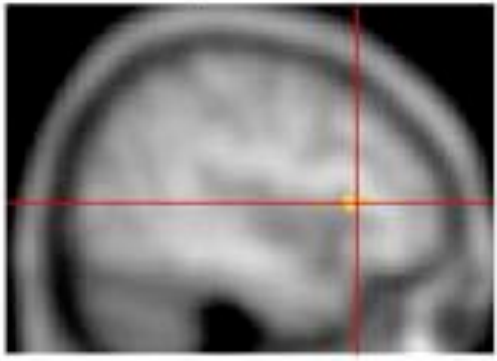
B



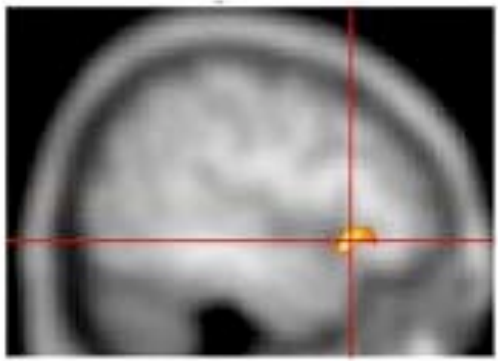
Word change

45

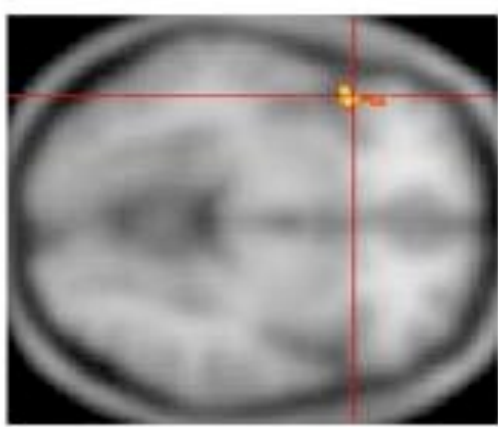
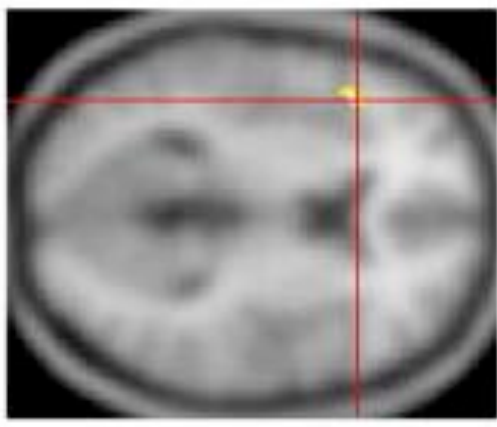
Syntax



Semantic



47



Dapretto and Bookheimer, Neuron, 1999

Tailored baseline

- Two experimental tasks that differ in several dimensions, interested in only 1
- Want a tailored control task for each experimental task
- EG: processing printed words vs. heard words: Controls are false fonts and nonsense speech

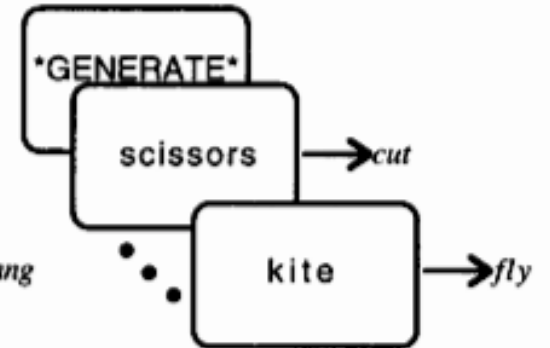
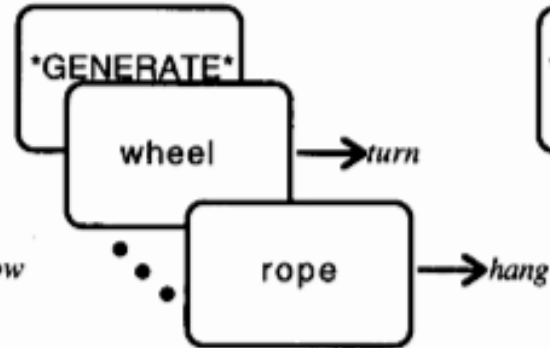
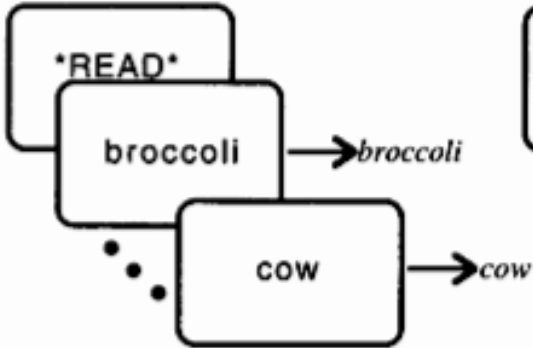
Conditions

Baseline

High Selection

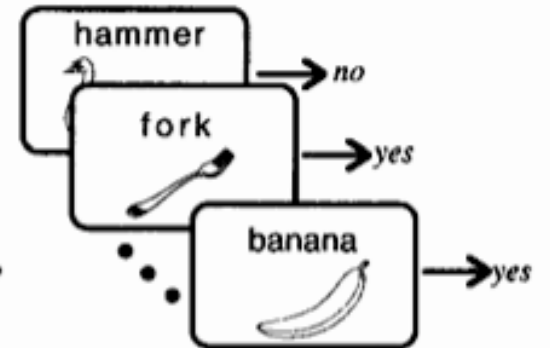
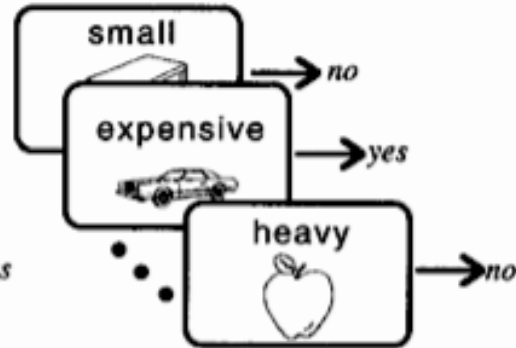
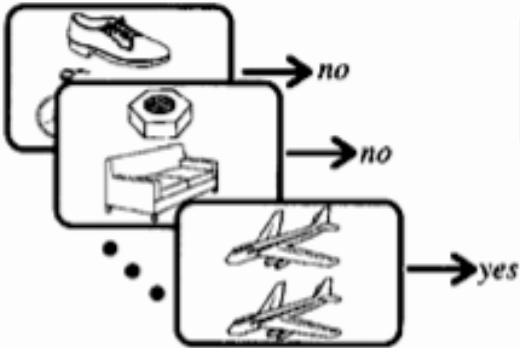
Low Selection

Generation



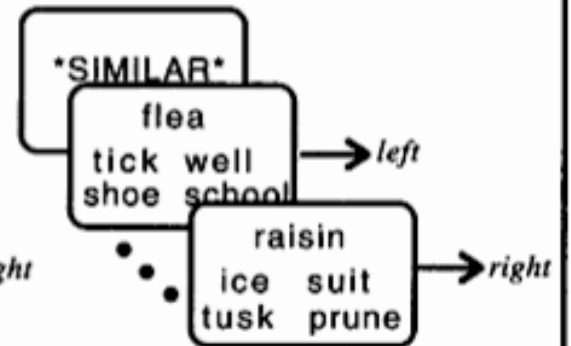
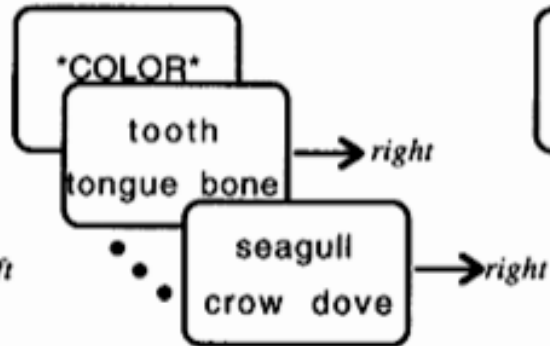
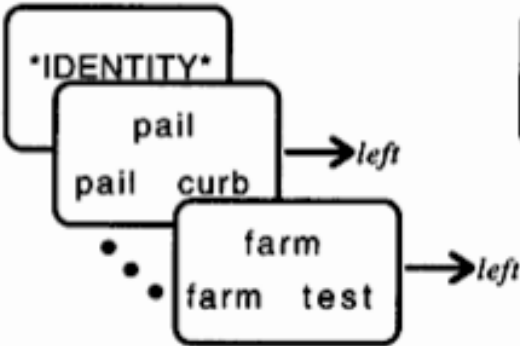
A

Classification



B

Comparison



C

Semantic Tasks

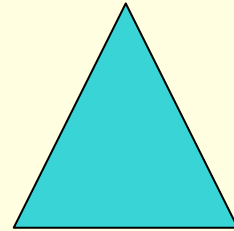
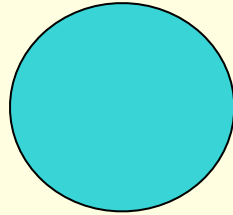
Tailored baseline troubles

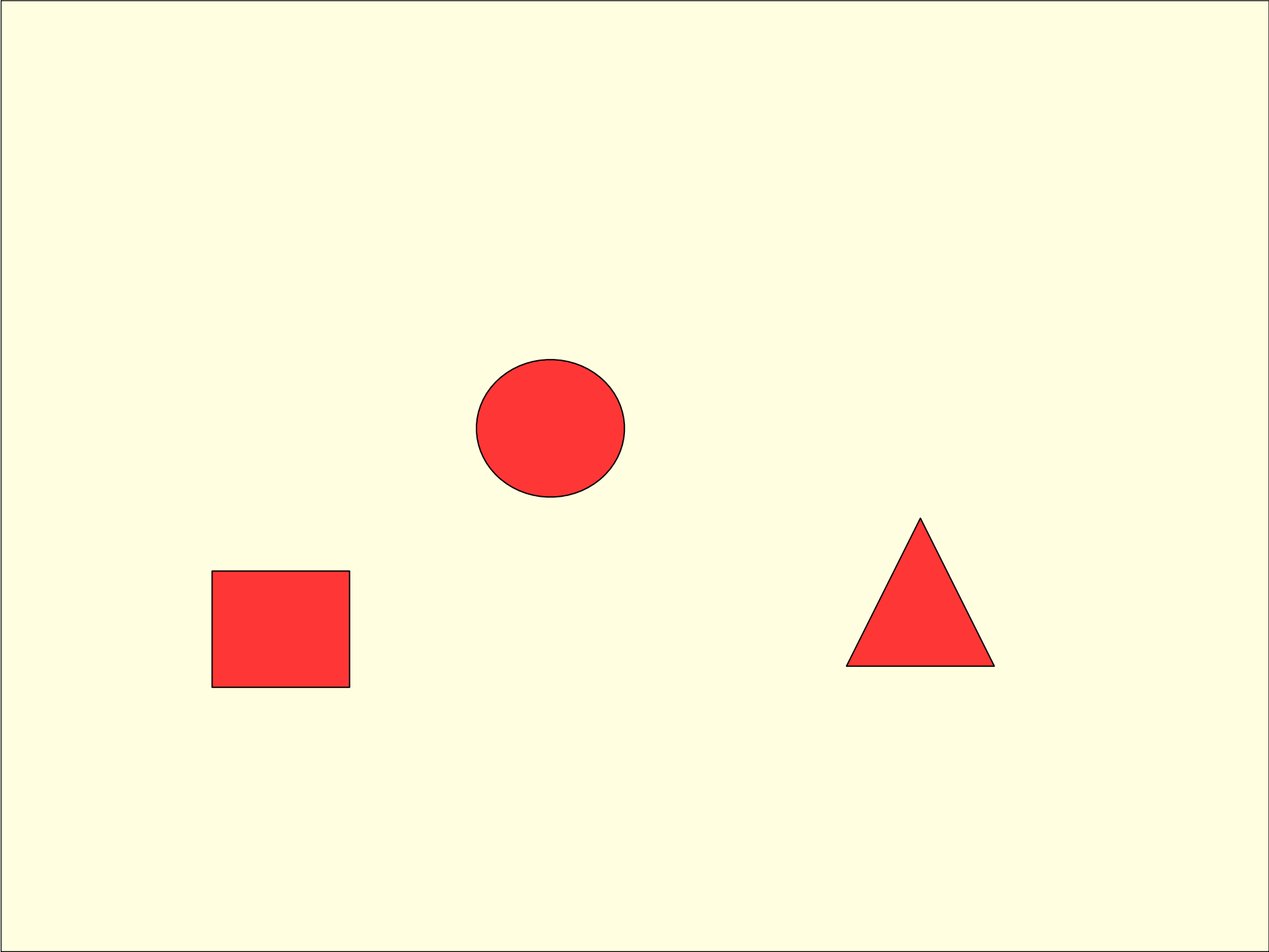
- Multiple layers of assumptions
- Assumes baseline tasks control for (sensory input) equally
- Assumes similar psychometric properties of both experimental and both control tasks: rather unlikely
- Add an additional common baseline

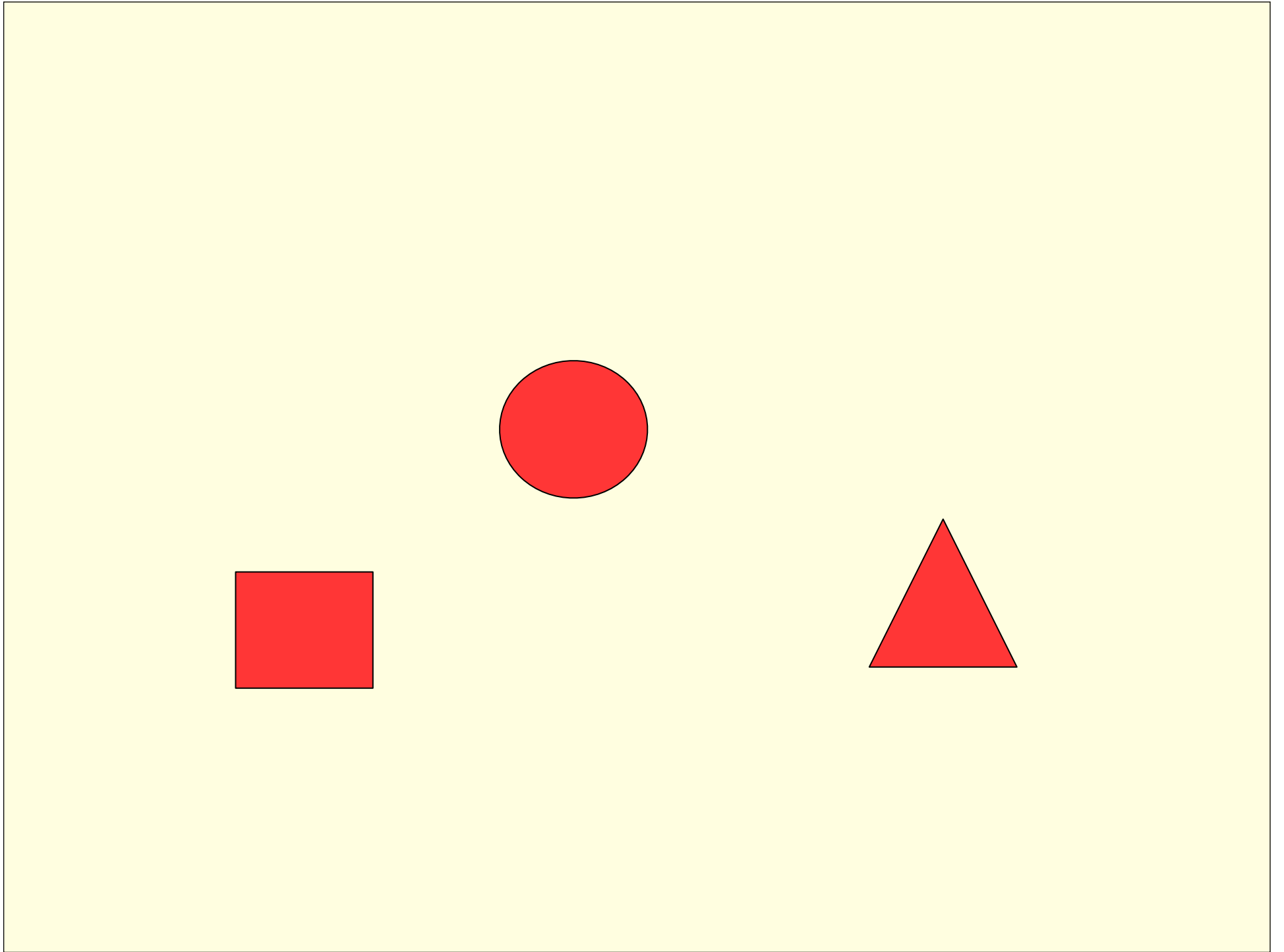
Directed Attention Models

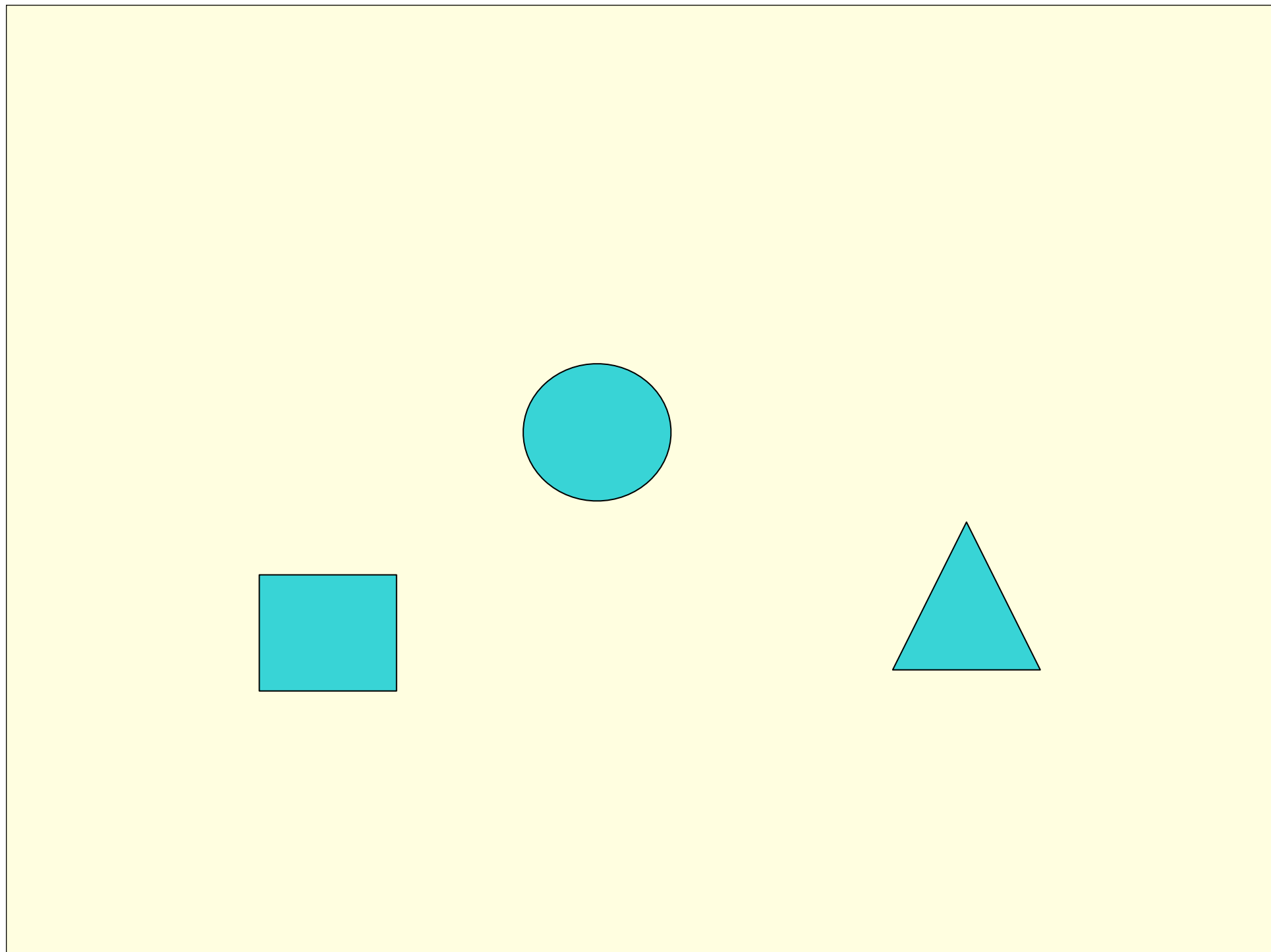
- All stimuli identical in all conditions
- Direct attention towards different features
- Implicit or explicit
- Assumes process is modified by directed attention
- Assumes passive processing does not capture your variable of interest

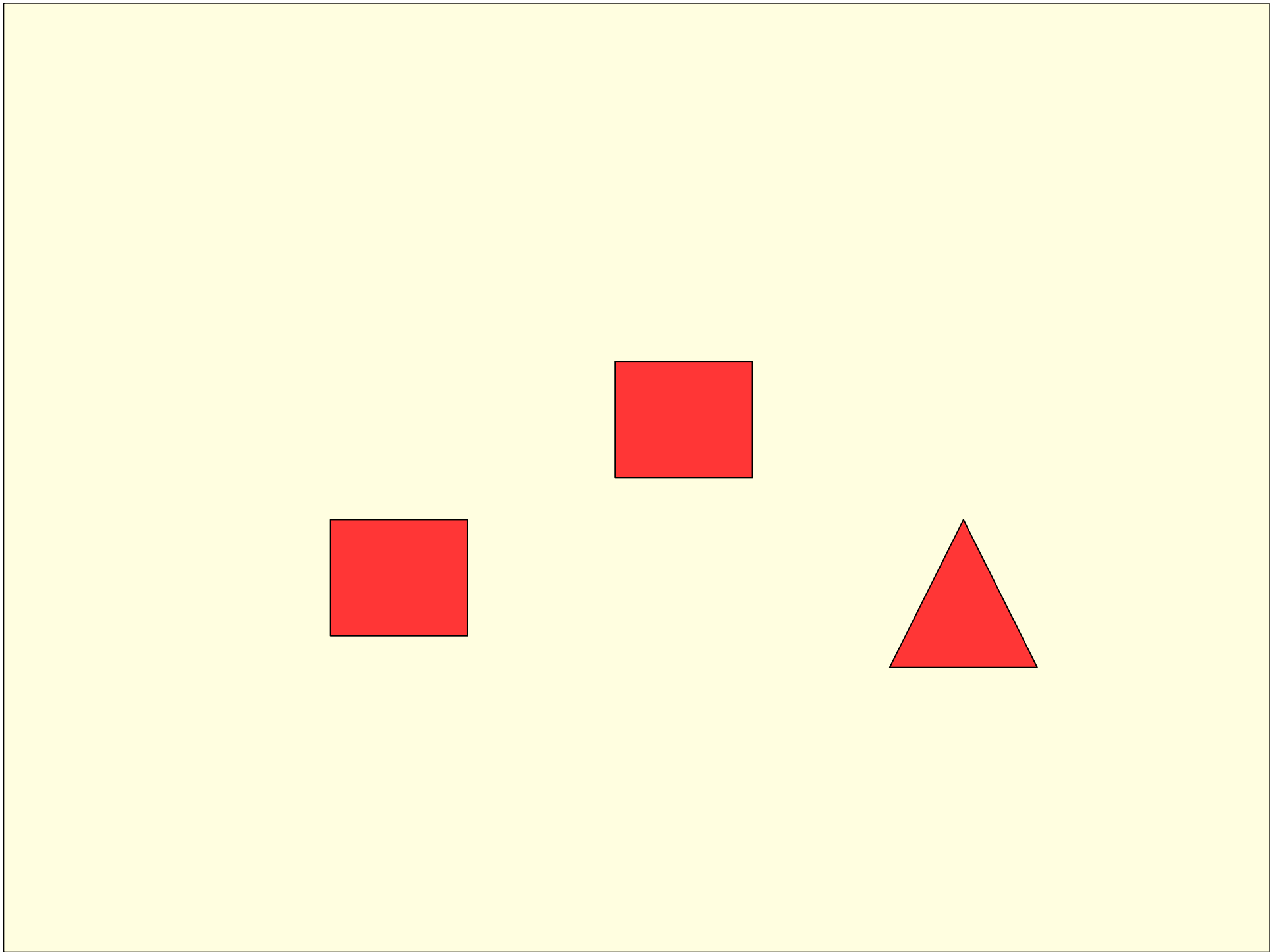
EG Corbetta et al







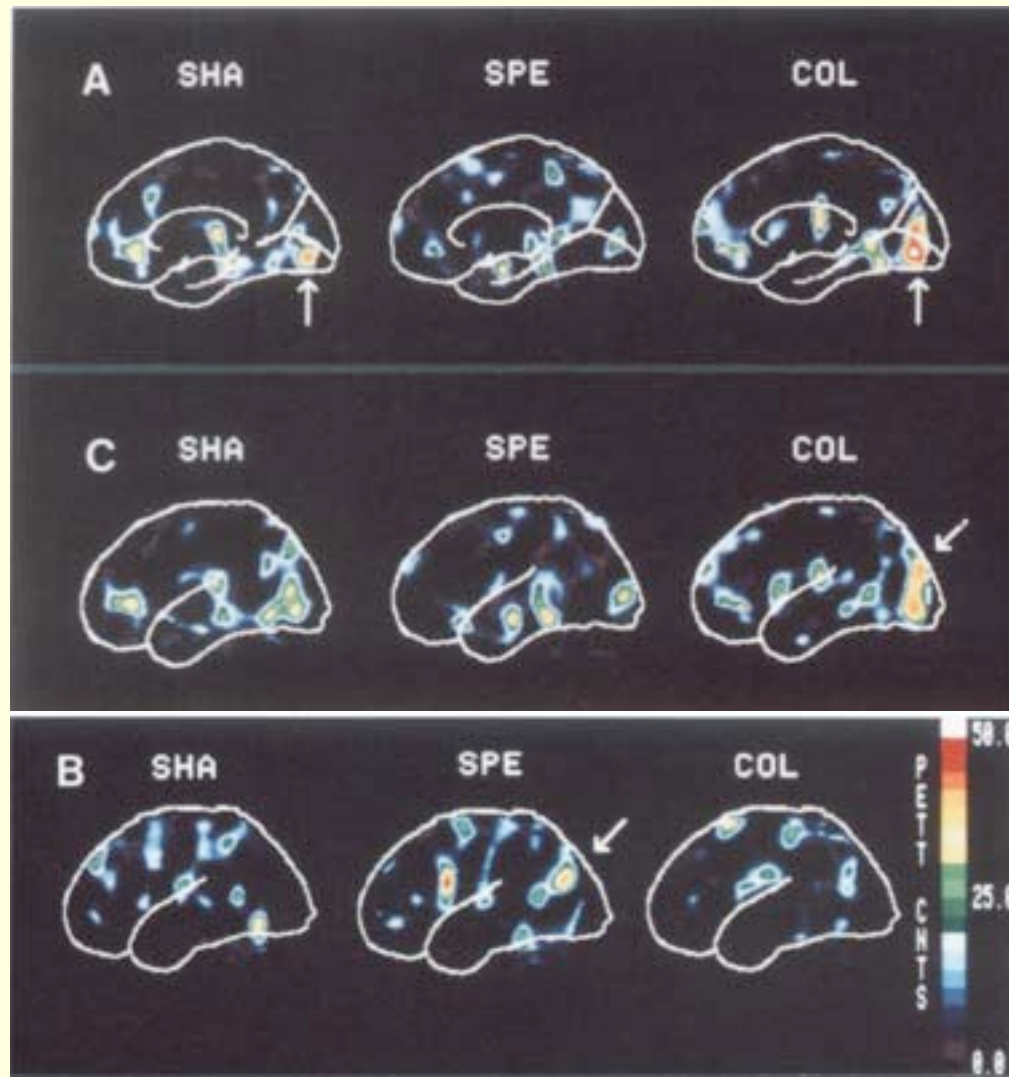




Direction attention models

- In every condition, all three variables change; all conditions are identical
- Told to respond to a shape, color or movement change in different blocks
- Selectively activates form, color, motion centers
- Attention to features substantially enhances activity in those areas

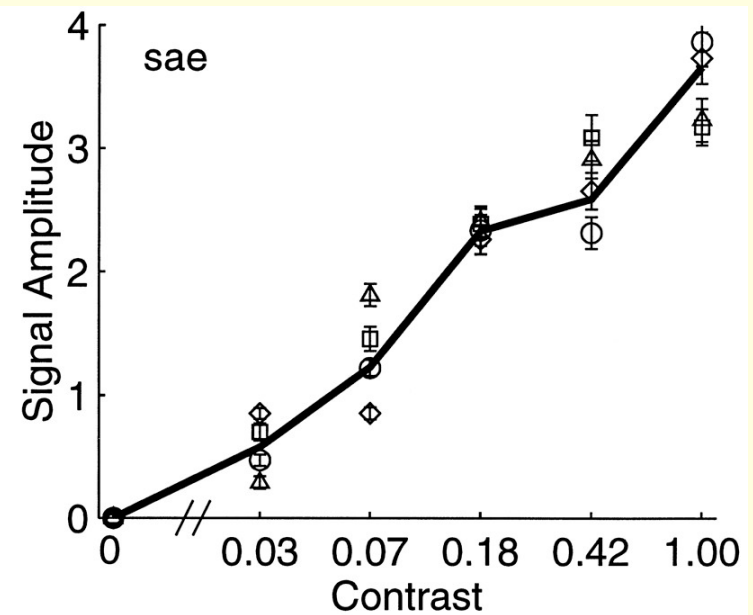
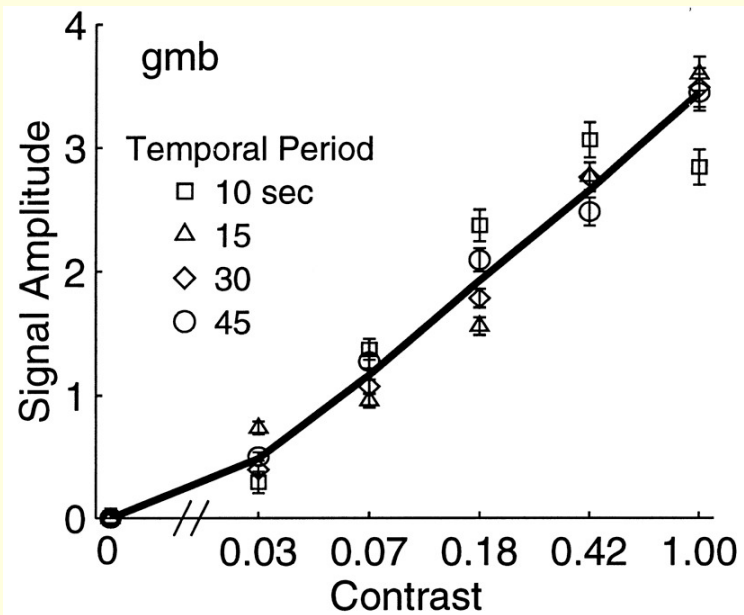
Selective attention to shape, color, motion



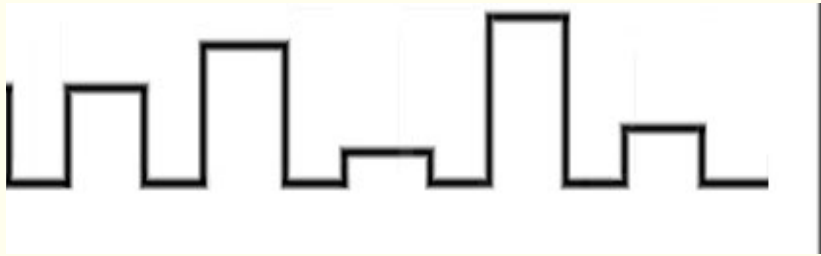
Parametric designs

- Can't or don't want to make an assumption of pure insertion; don't want to tailor a control task
- Employs continuous variation in a stimulus/task parameter
 - E.g., working memory load, stimulus contrast
- Inference:
 - Modulation of activity reflects sensitivity to the modulated parameter

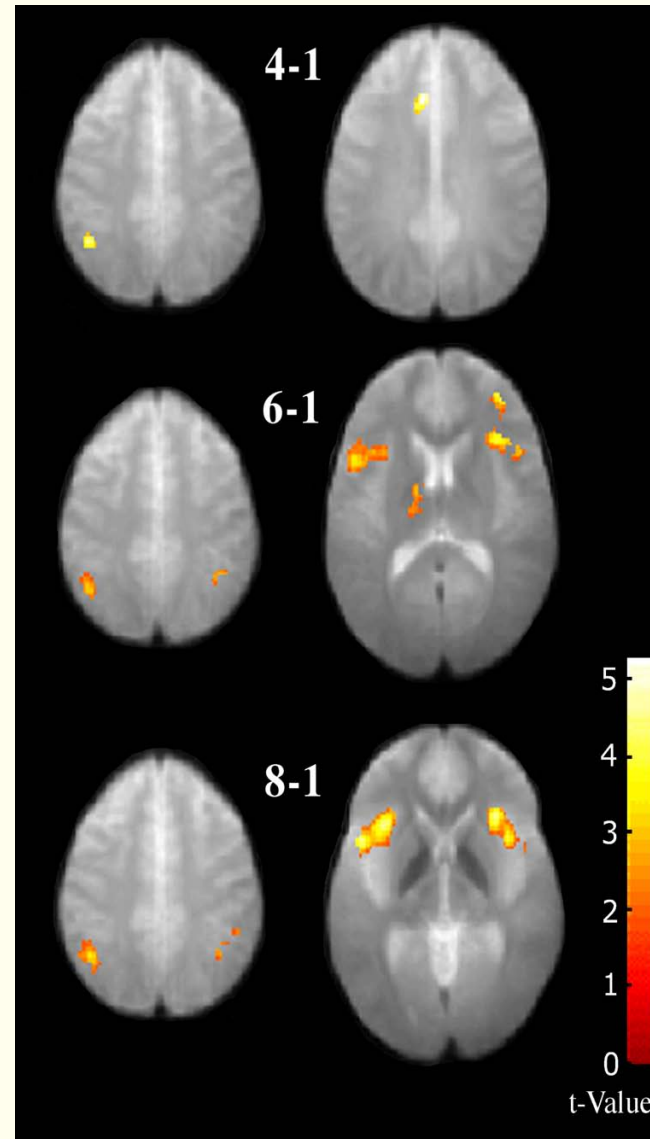
Boynton et al., 1996



Difficulty Effect: Digit Span Task



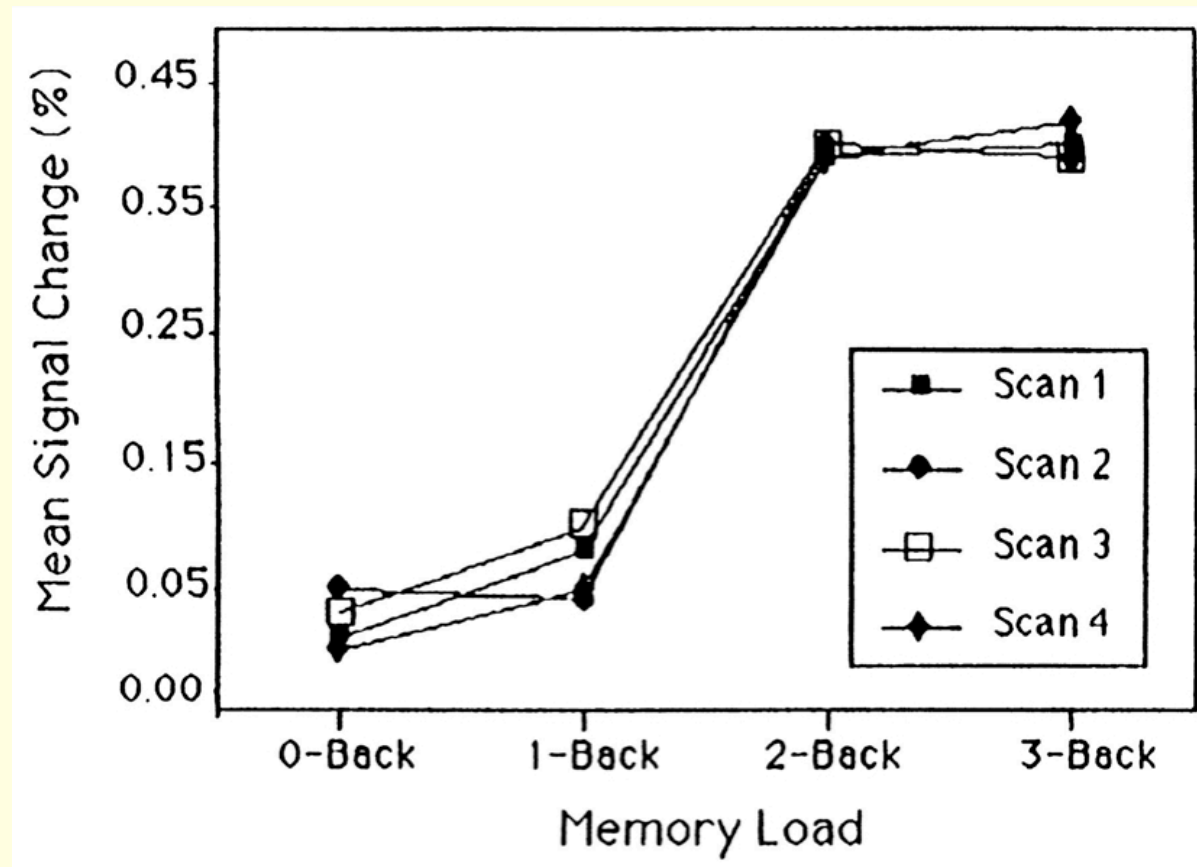
Burggren et al, 2001



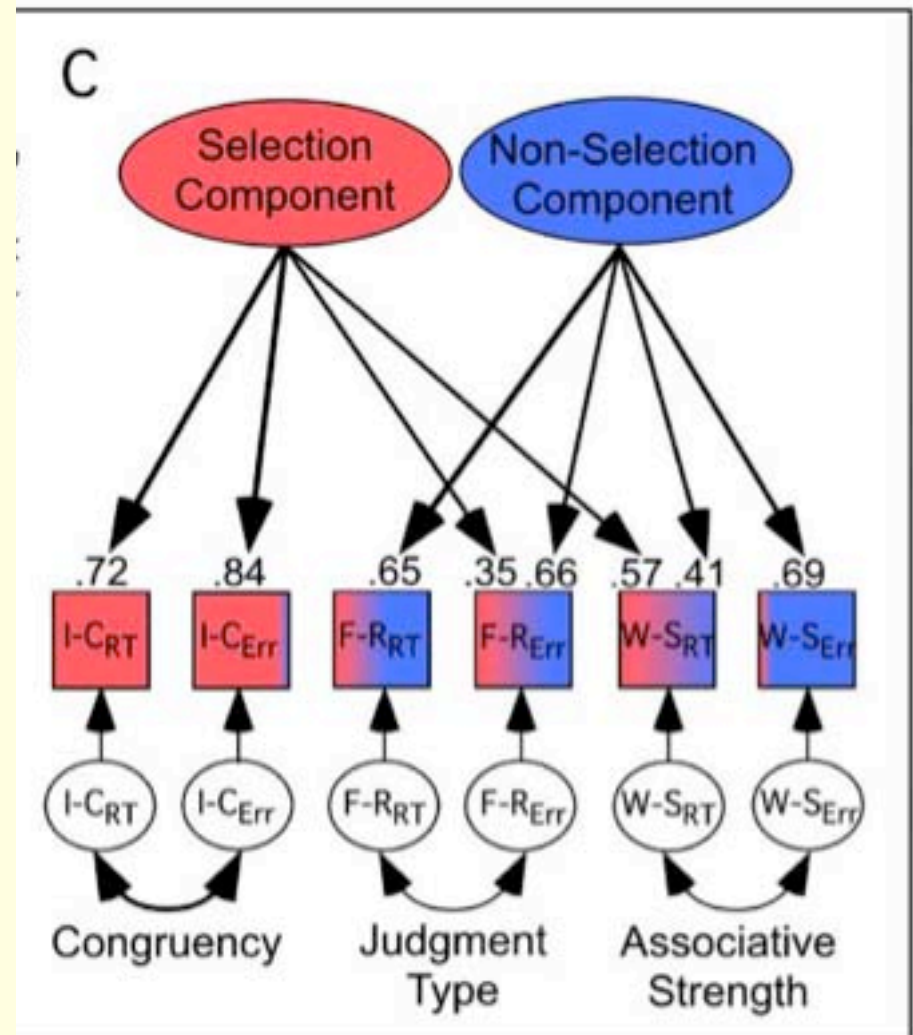
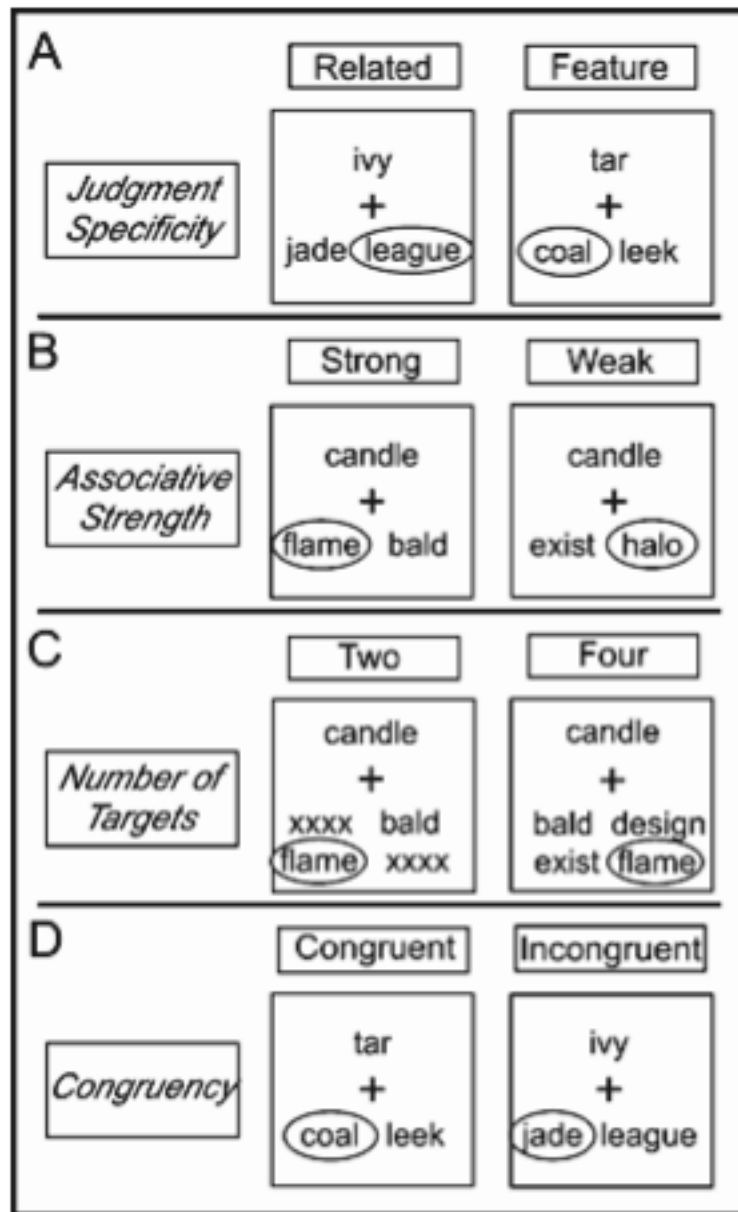
Assumptions of parametric designs

- Assumption of pure modulation
 - Each level of the task differs quantitatively in the level of engagement of the process of interest, rather than qualitatively
 - Assumes you can define the magnitude differences across levels (usually assumes equality)
 - How might this fail?

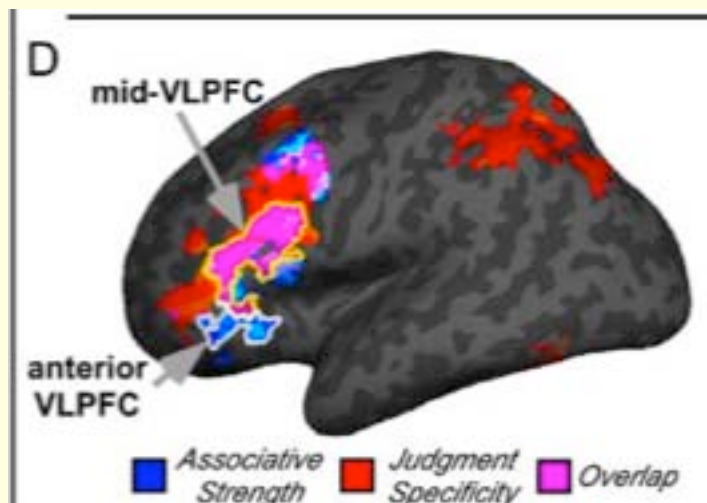
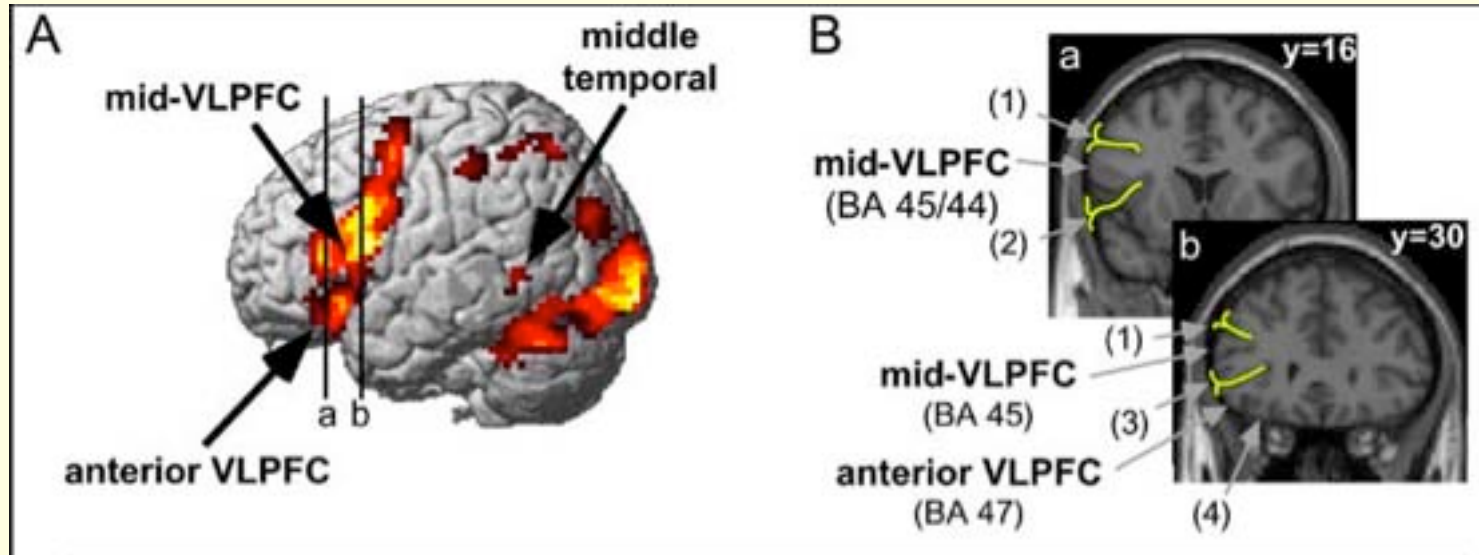
When task manipulations are not parametric: eg, Cohen et al., 1996



Factor-determined component classification: Badre, Poldrack et al 2005



IFG dissociations



Badre, Poldrack etc 2005

Priming/adaptation designs

- Presentation of an item multiple times leads to changes in activity
 - Usually decreased activity upon repetition (magnitude and/or spatial extent)
- Based on behavioral priming literature (repetition priming, semantic priming) and single unit studies
 - Eg, NURSE—DOCTOR vs XXXX-DOCTOR
- Inference:
 - Regions showing decreased activity are sensitive to (i.e. represent) whatever stimulus features were repeated
- Requires version of pure modulation assumption
 - Assumes that processing of specific features is reduced but that the task is otherwise qualitatively the same

Adaptation Designs

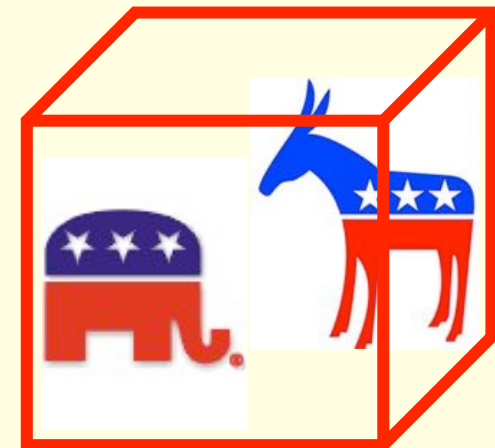


Two stimuli: can neurons tell the difference?



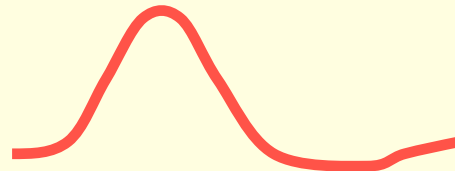
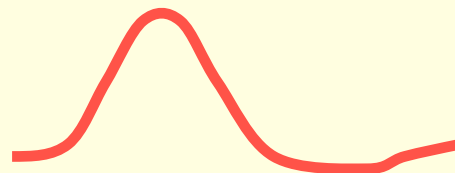
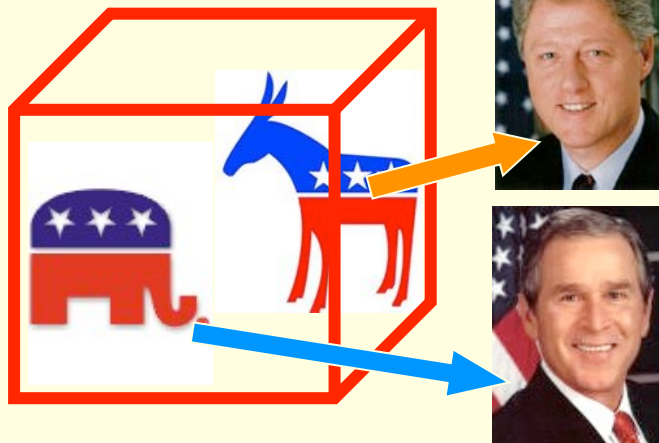
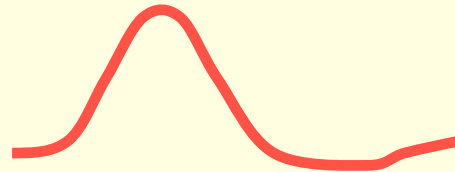
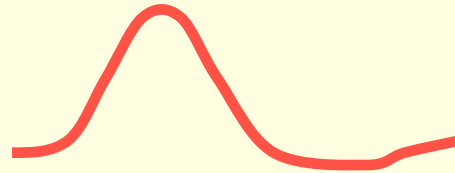
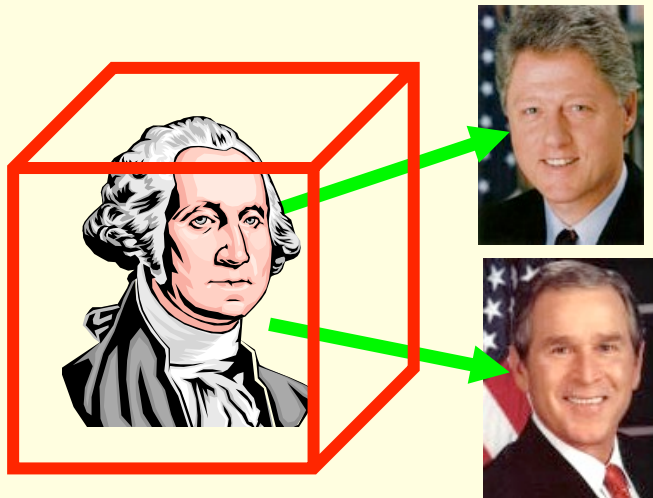
- A voxel containing neurons that respond to all politicians, irrespective of party

- A voxel containing some specifically Democratic neurons, and other specifically Republican neurons.



From R. Raizada

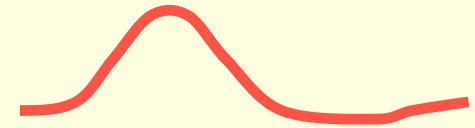
Responses to individual stimuli do not show whether neurons can tell the



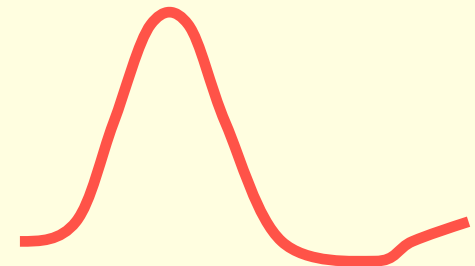
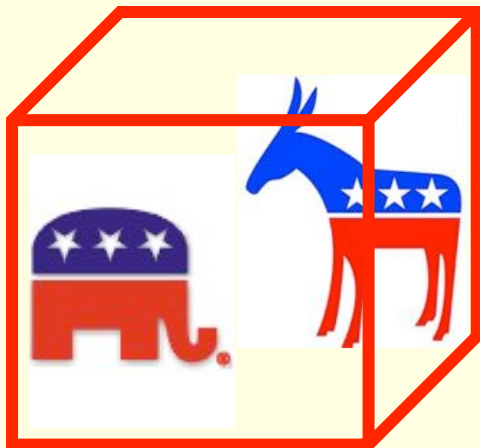
- Different sets of neurons are active within the voxel, but overall fMRI responses are indistinguishable

From R. Raizada

Neural adaptation to repeated stimuli does show the difference:
What counts as repetition for neurons in a voxel?



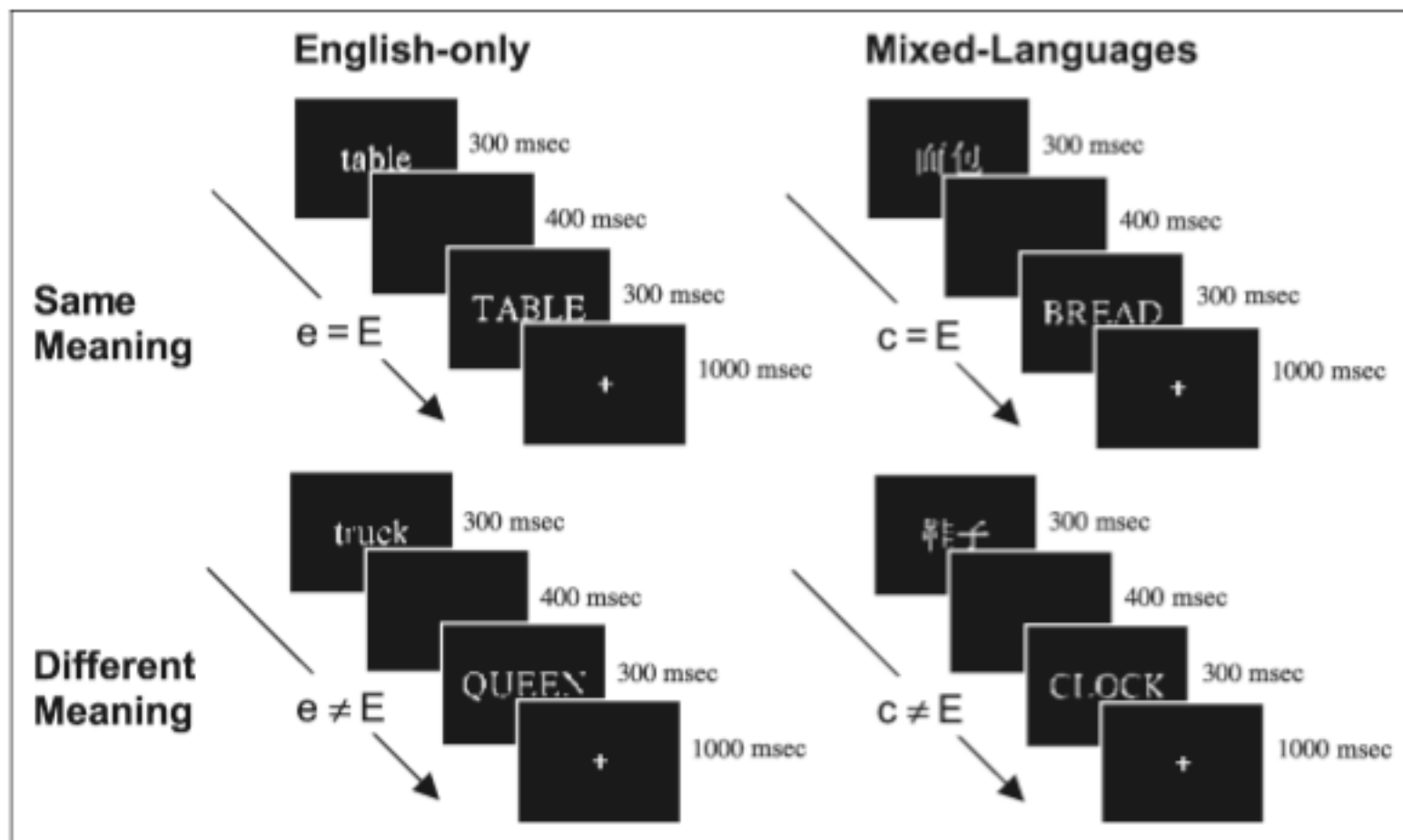
Same neurons, adapting:
It's a politician again



Different, fresh neurons:
It's a Democrat

From R. Raizada

Adaptation in bilingual subjects

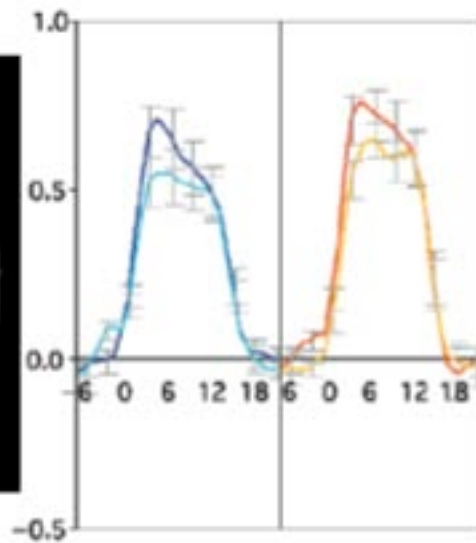
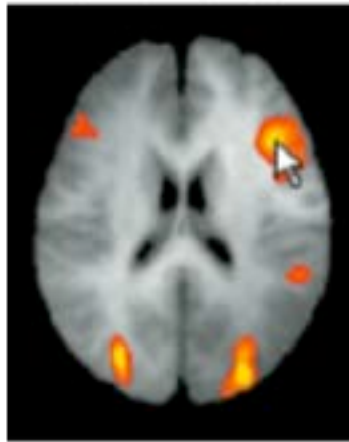


Chee et al 2003

Left Ventrolateral Prefrontal

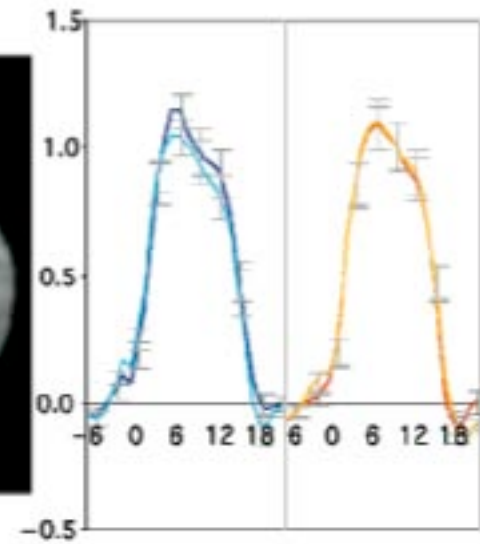
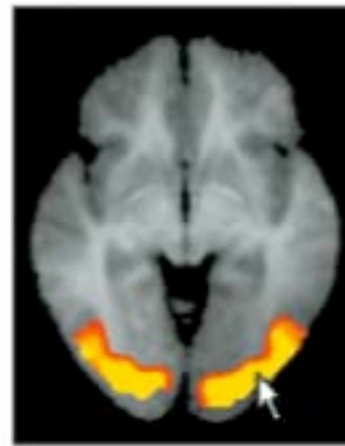
†

(-40, 25, 20)



Left Lateral Occipital

(-32, -79, -3)



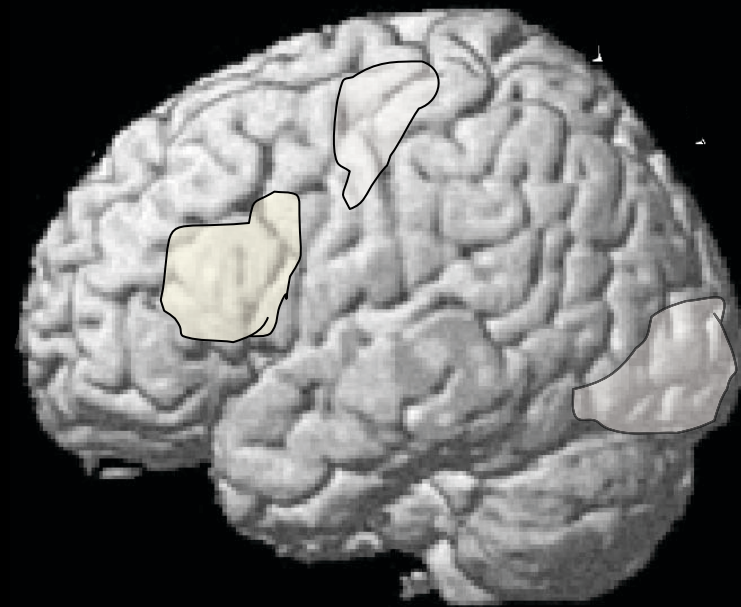
$e \neq E$ ———
 $e = E$ - - - -
 $c \neq E$ ———
 $c = E$ - - - -

Main effect for meaning (adaptation) in LIFG, not LOcc

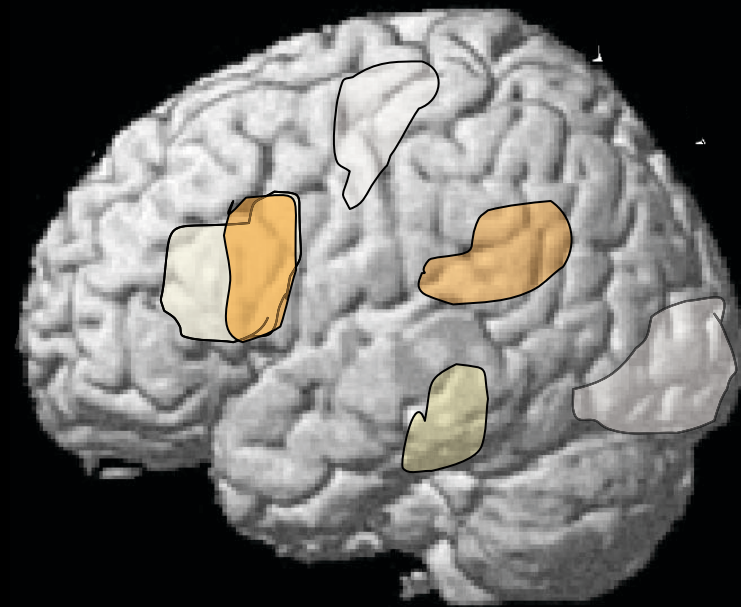
Conjunction analysis (Price & Friston, 1997)

- Perform several parallel subtractions
 - Each of which isolates only the process of interest
- Find regions that show common activation across all of these

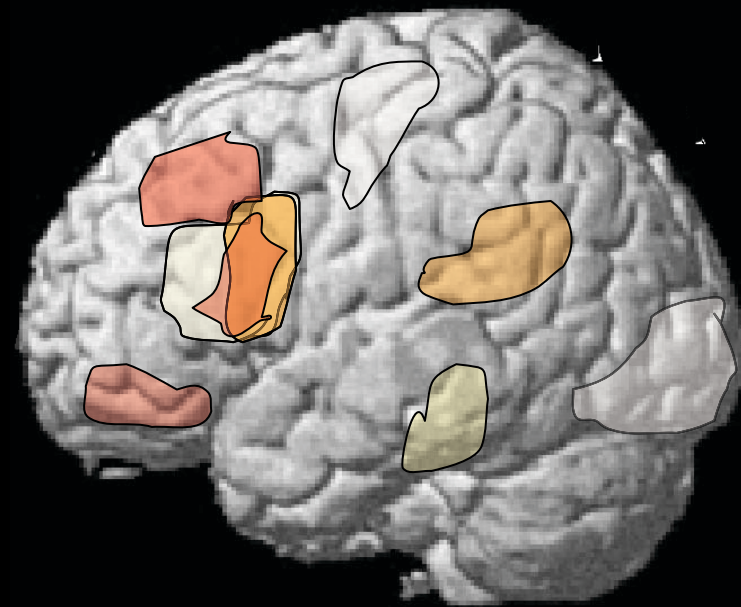
Ex A - Ctl

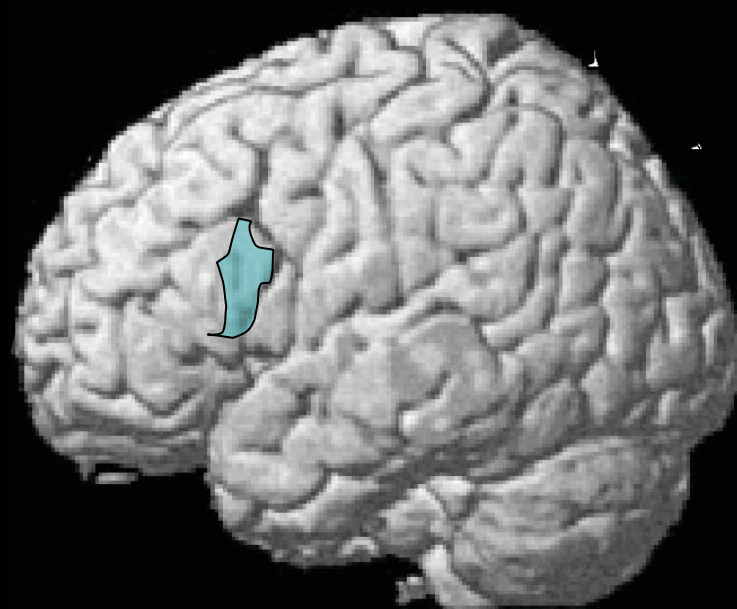


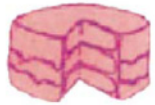



Ex B - Ctl

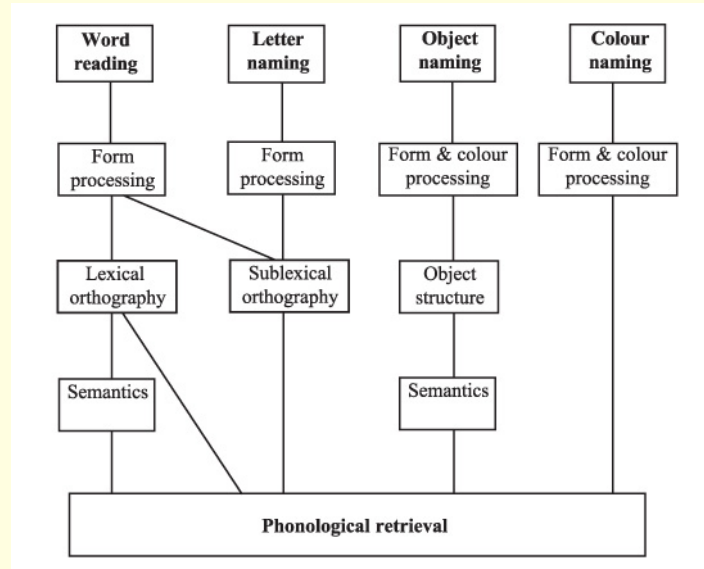


Ex C - Ctl





	Name (A)	Say "YES" (B)
Words:	1 badge	2 Znbkqo
Letters:	3 r	4 n
Objects:	5 	6 
Colours:	7 	8 



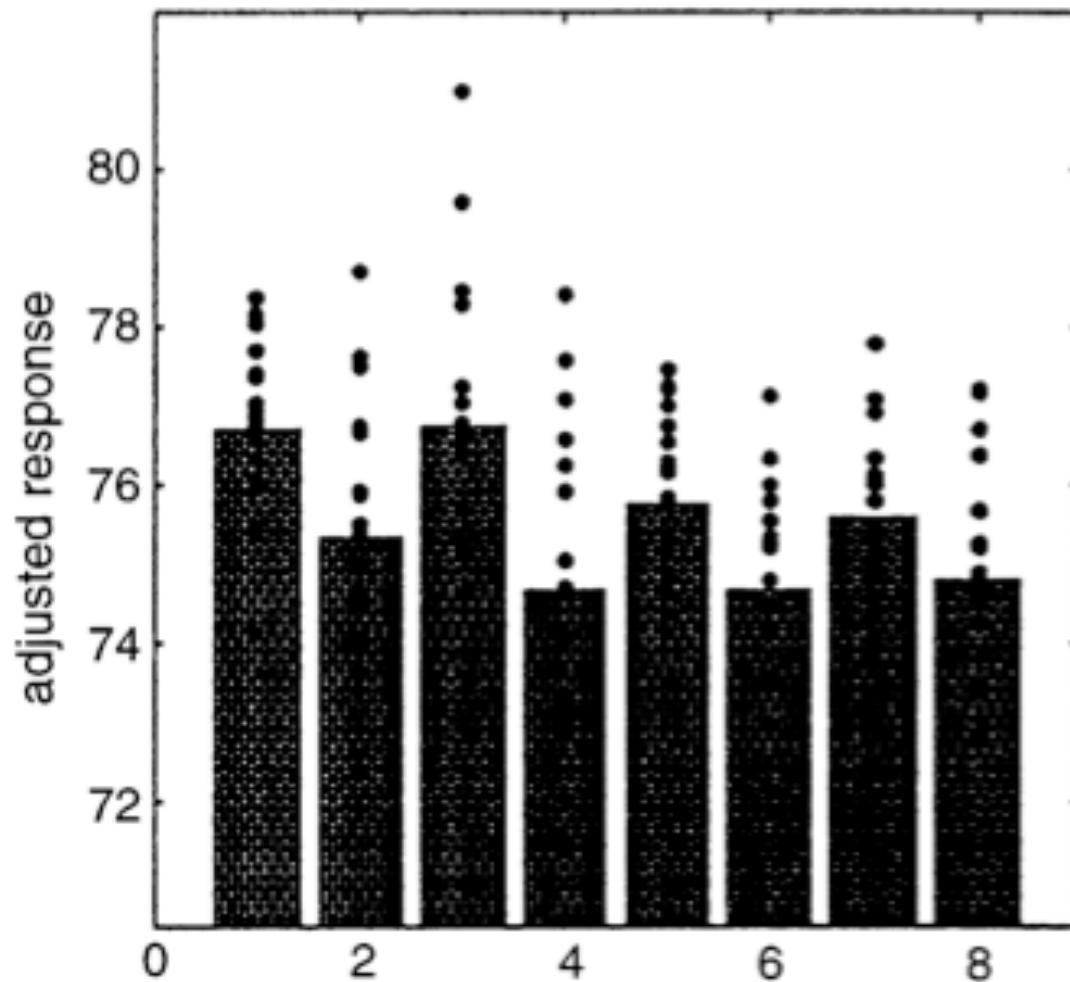
Tasks:	Task Pair I		Task Pair II		Task Pair III		Task Pair IV	
	Words		Letters		Objects		Colours	
	A	B	A	B	A	B	A	B
	1	2	3	4	5	6	7	8
<u>Cognitive Processes</u>								
Form processing	█	█	█	█	█	█	█	█
Colour processing					█	█	█	█
Lexical orthography	█							
Sublexical orthography	█		█					
Object structure					█	█		
Semantics	█				█	█		
Phonology	█		█		█		█	
Articulation	█	█	█	█	█	█	█	█

from Price & Friston, 1997

Left Posterior basal temporal lobe (BA37)

$x = -44, y = -62, z = -16$

$Z = 4.66, p = 0.000 (0.016 \text{ corrected})$



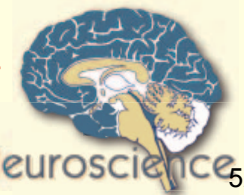
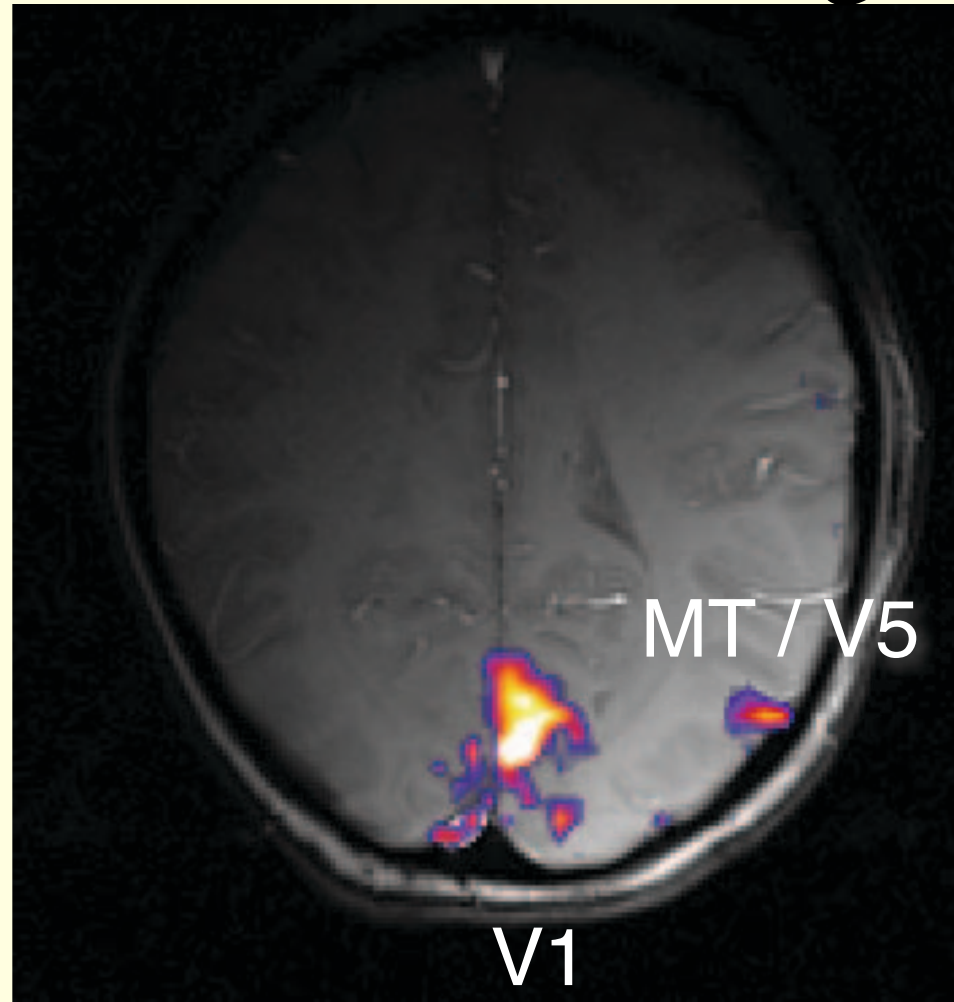
Problems with conjunction analysis (Caplan & Moo, 2003)

- Implicit processing
 - Subjects may engage processes that are not necessary for the task
- Interactions between processing stages
 - Conjunction only gets rid of interactions if they do not activate the same regions to the same degree across tasks
 - E.g., dual-route model

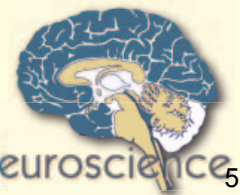
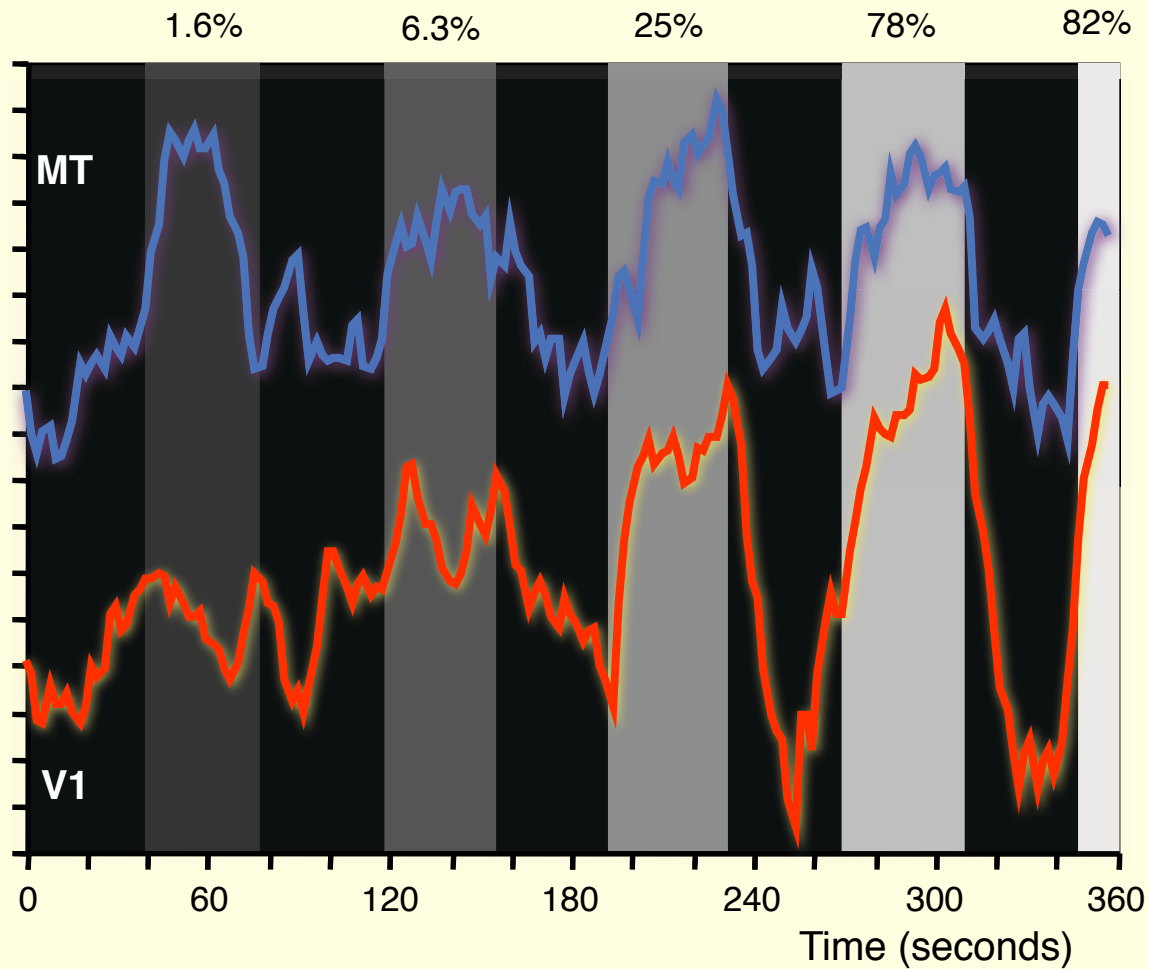
The functional characterization approach

- Many studies take a “functional mapping” approach
 - The final goal is to say that area X is active in relation to process A
- Another approach is to first identify an area of interest
 - Often based on a subtraction design, along with other knowledge from neuroscience
- Then, use multiple task manipulations to characterize the function of the area

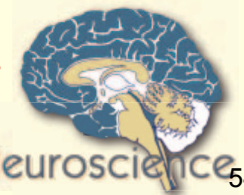
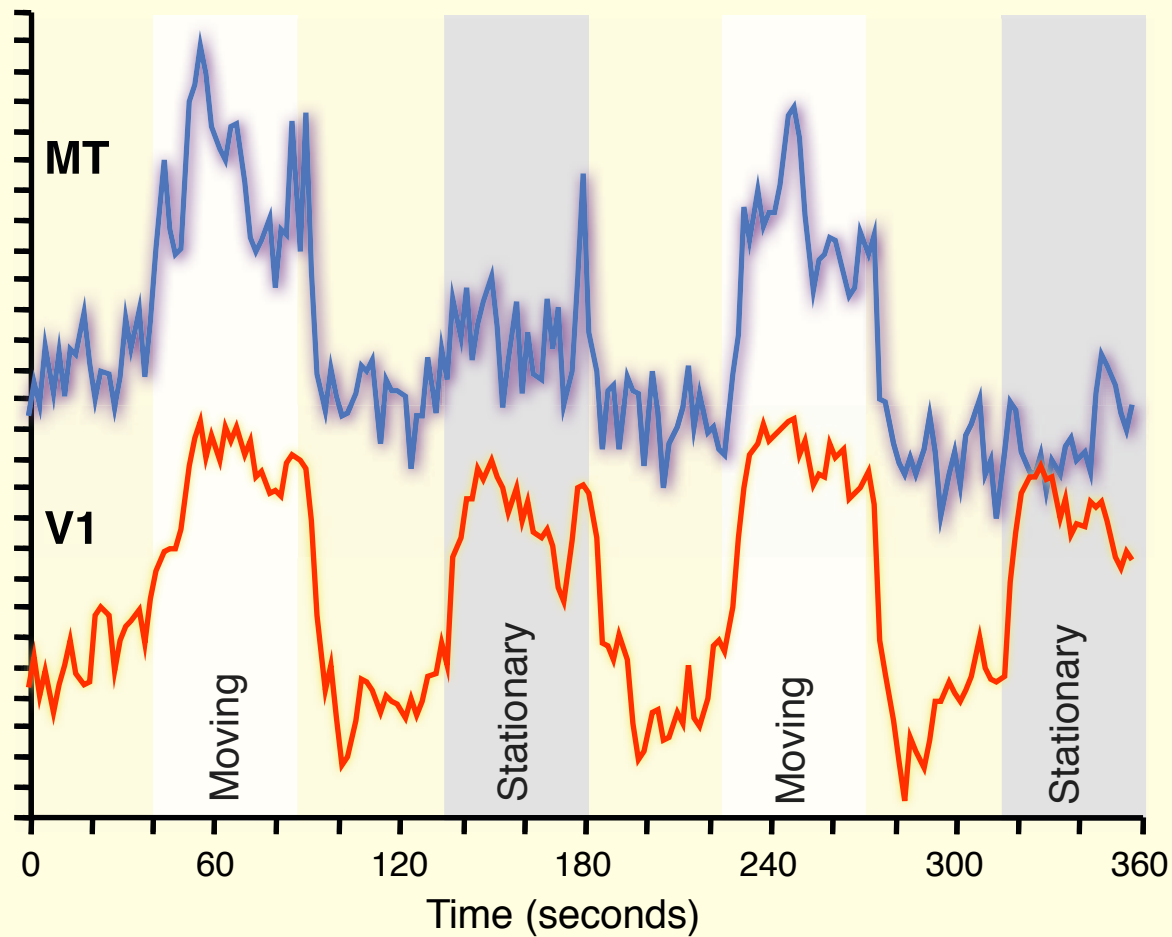
Activation with Moving Visual



Contrast Response Test



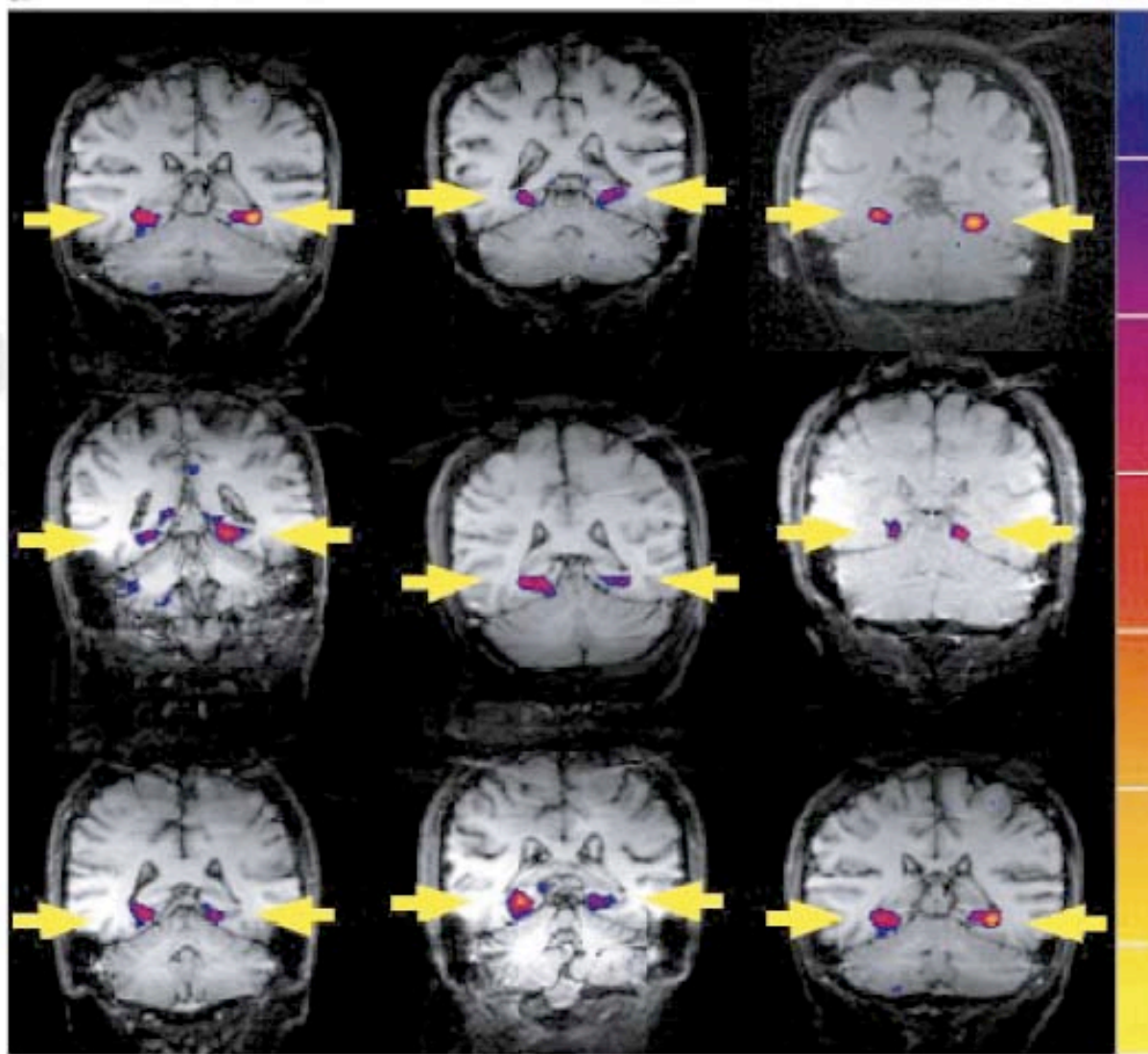
Motion Sensitivity Test



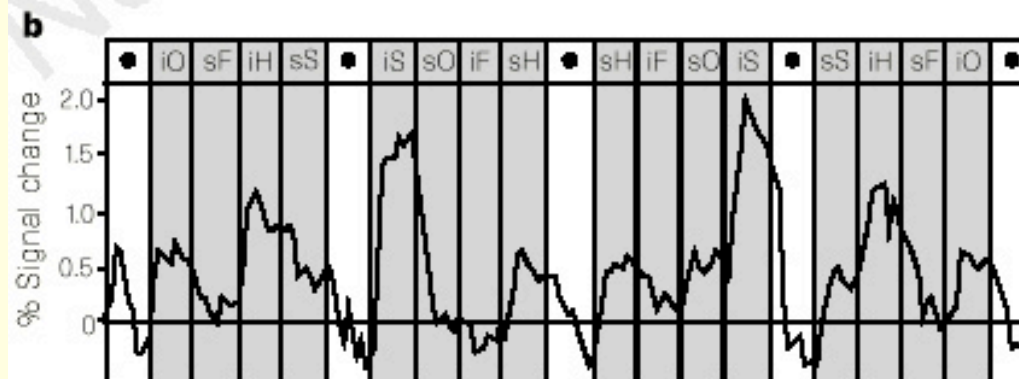
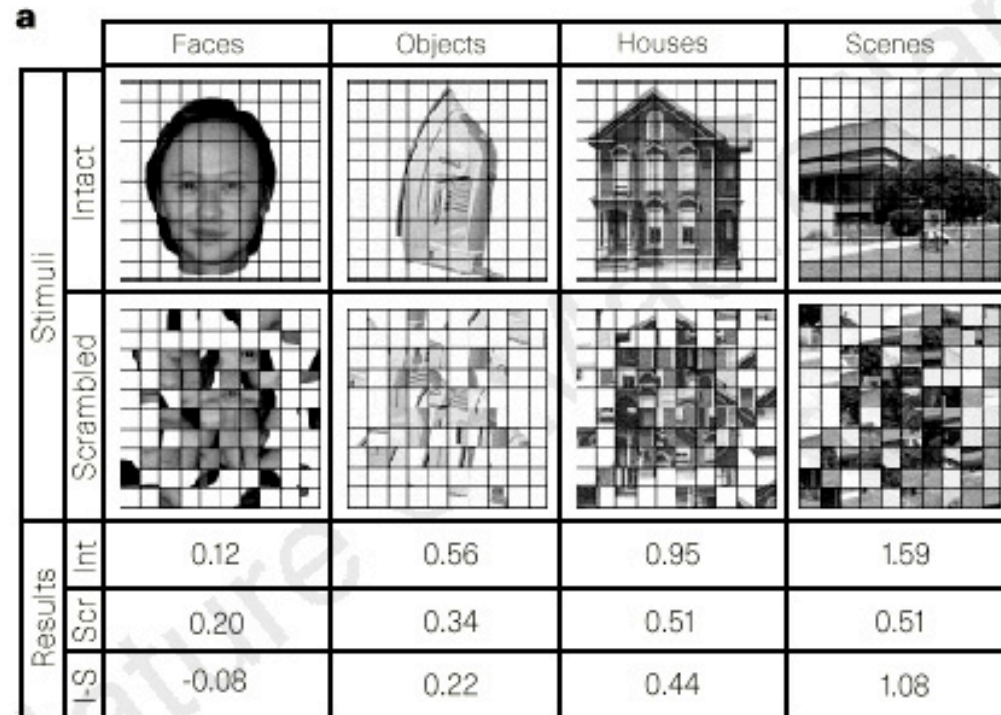
Parahippocampal place area (Epstein and Kanwisher 1998)

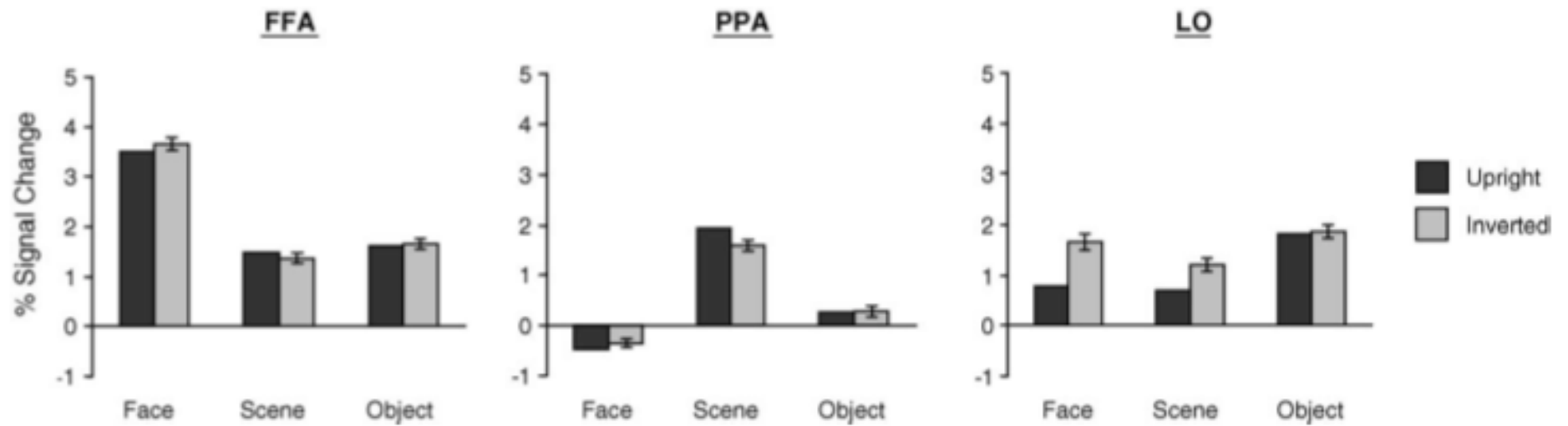
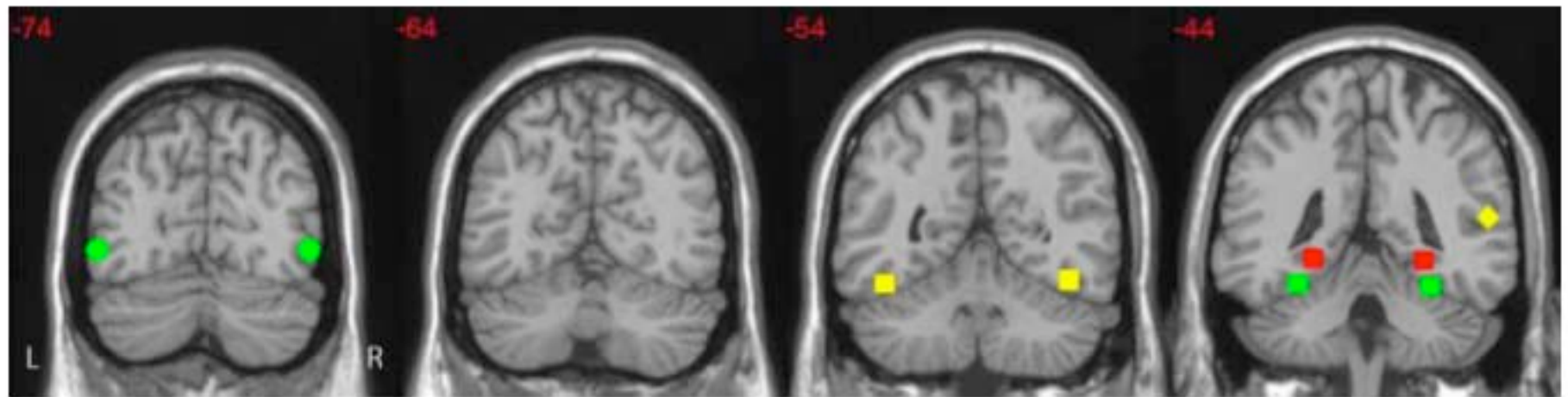
- Blocks: Faces, objects, houses, scenes
- Intact and scrambled
- Analysis:
 - ROIs = pixels in target area **from an independent scan**
 - Compute signal diff for scram. vs. intact
 - Diff sig > for scenes than for others
- Alternatives and control exp.s
 - Can you think of other alternatives or controls?

Epstein et al: Localizer



Epstein et al: Functional characterization





2-group designs

- Build on any of the prior designs
- Additional between group comparisons
- Hypothesis sounds something like: the differences between experimental and control task in my patient group differs from that difference in controls
- Assumes baseline task performance is equal

Counterbalancing

- With more than 2 conditions- essential
- EG: Lo, medium and high stress conditions
 - Habituation
 - Order effects eg High carry-over
- Complete counterbalancing
 - 1 2 3 132 231 213 312 312
- Latin Square
 - 123 231 312
 - Each condition in each serial order; assumes no task-task order interactions

Summary

- No design is perfect; all make assumptions that are not fully verifiable; know them!
- Use that which is most consistent with your specific research question; freely admit weaknesses
- Multiple “baseline” conditions help interpretation
- Beware of your assumptions

ROI Analysis

Why ROI?

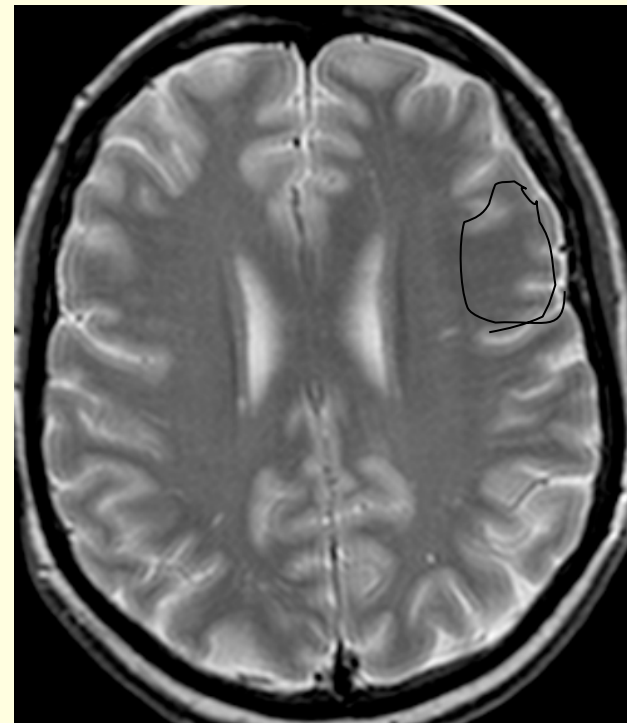
- Your research is hypothesis driven
- In voxel based approaches you have to correct for multiple comparisons all over the brain
- With ROI approach, can test an ROI of any size as a single comparison

Determining ROIs

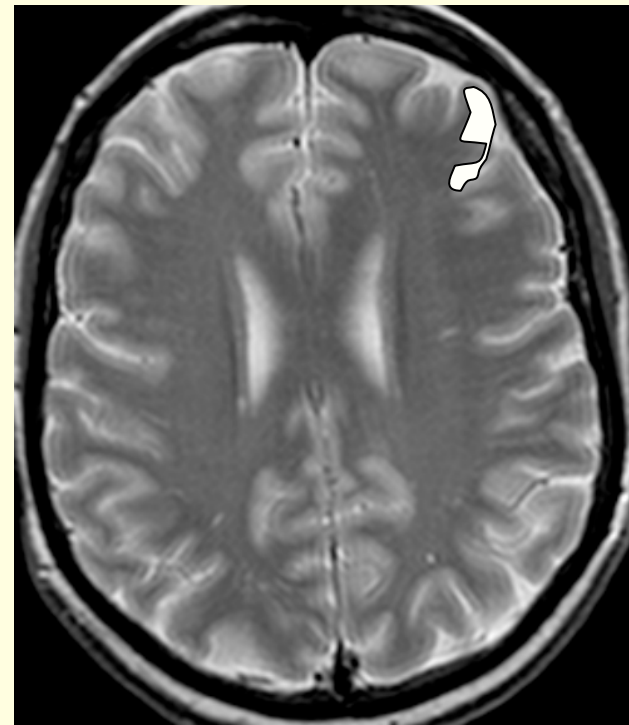
- Anatomically based Rois
- Functionally based Rois
 - Danger! Use with extreme caution

Anatomical ROIs

- Where? How big? How to select?
 - Individuals anatomy; by lobe, region

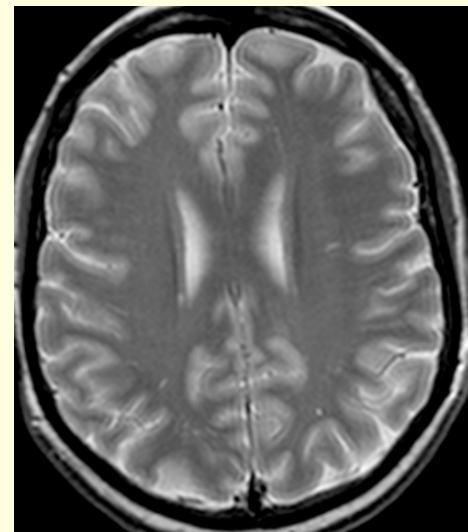


- Exclude WM; segment first



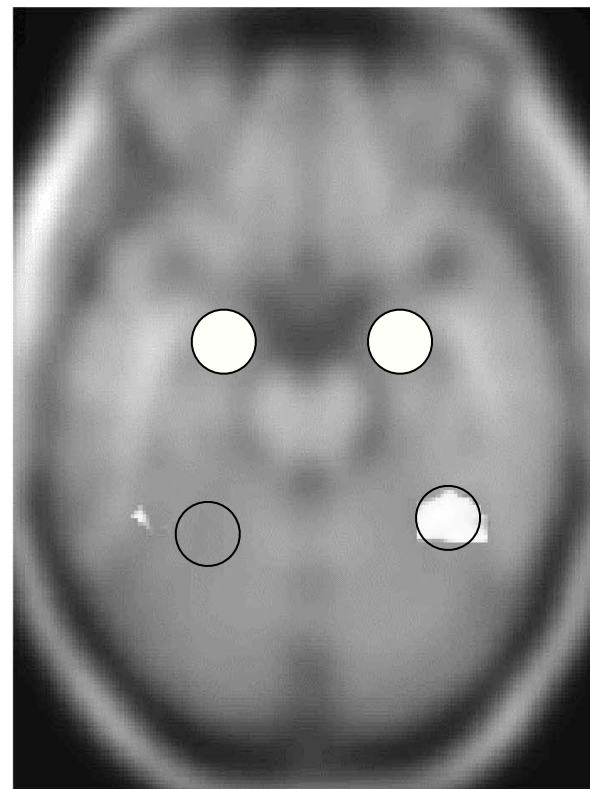
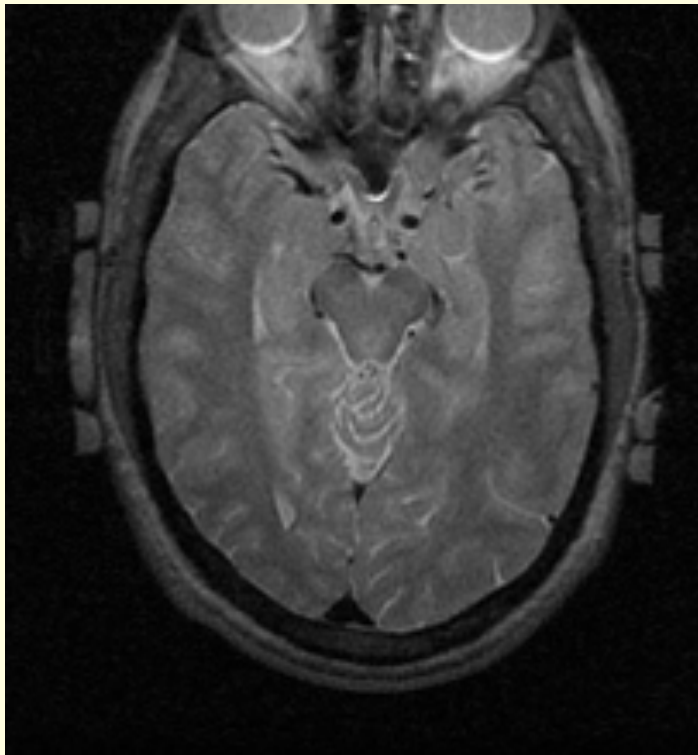
Things you must do:

- Get anatomical training
- Establish inter-rater reliability
- Watch for “drift”
- Remain blind to group, if relevant
- *** Remain blind to functional activation



RoI on normalized brain

- Draw 1 ROI instead of N Roi's
- Register data very well
- Smooth data; can lose signal if region is small



I

D

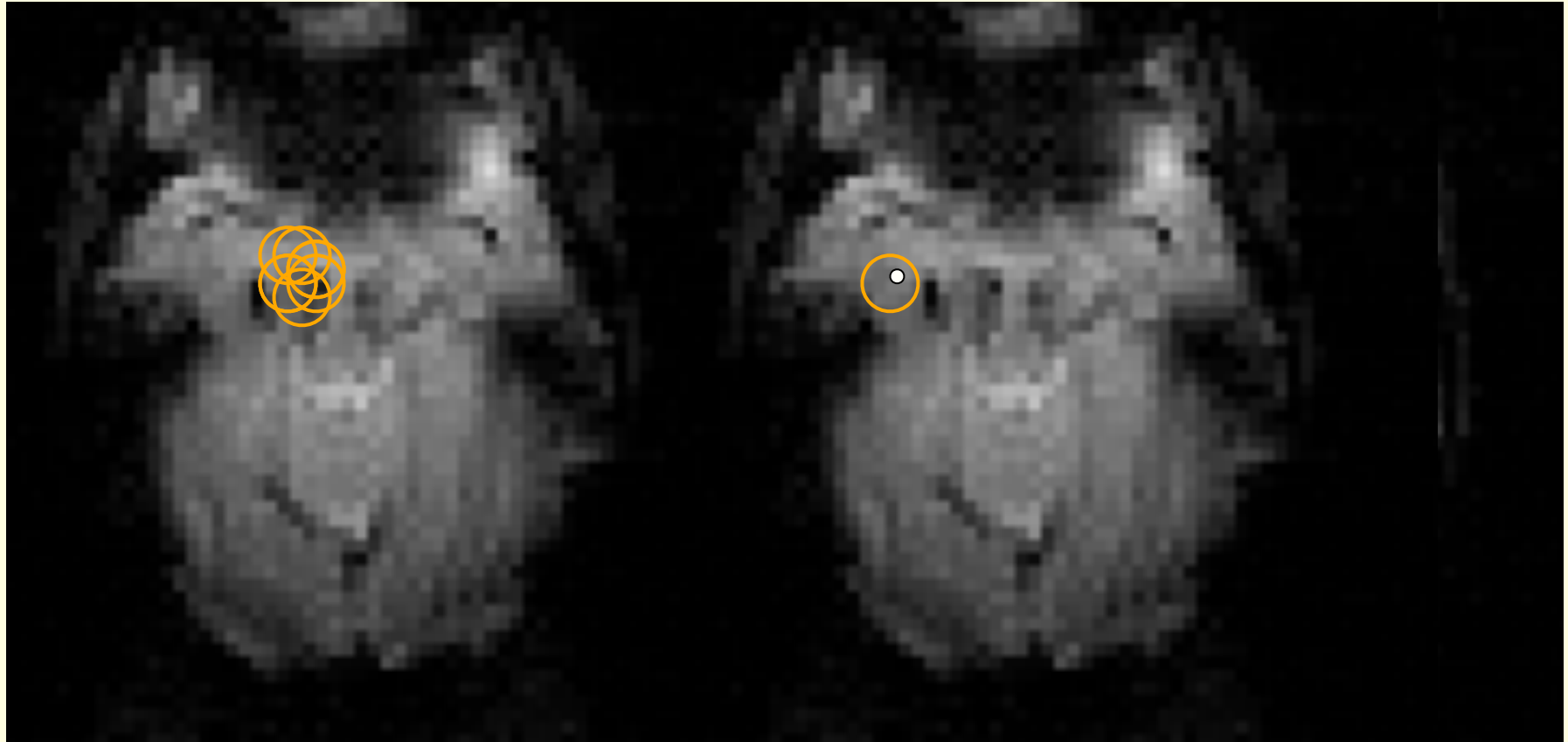
Bases of ROI choices

- Anatomical structure (amygdala); gyrus
- Functional structure: DLPFC (where are the boundaries)
- Brodmann's area
- Sphere or point (center on a Talairach coordinate)
 - Requires smooth data

Functional ROI

- Base ROI on some functional MRI data
- Advantage: don't average in areas that are not functionally relevant
- Questions to ask: how is this region modulated by certain task parameters

The location problem



Safe functional ROI practices

- Use roi/maximum from another sample (ok if groups are effectively the same)
- From literature
- Use individualized localizer on one scan; apply to a second scan/
second condition

Eg FFA localizers

- Kanwisher
 - Run a face scan, find maximum consistent with face processing; choose a subject-specific functional ROI
 - Apply ROI to a NEW set of stimuli from a new scan

“Voodoo” Correlations

- Perform a between condition or between-group GLM analysis
- Extract the functional ROI from this analysis
- Do a t-test on the ROI, or correlate the data with a variable used to separate the groups
- This approach is non-independent; the second analysis will ALWAYS be significant, by definition

Example of the wrong thing

nature
neuroscience

ARTICLES

High-resolution imaging reveals highly selective nonface clusters in the fusiform face area

- Obtained an ROI by identifying all voxels with a higher magnitude positive activity for one stimulus and averaged across them
- Applied that ROI to compare responses from the selected category to others

Corrigendum: High-resolution imaging reveals highly selective nonface clusters in the fusiform face area

Baker *et al.*¹ have drawn to our attention flaws in our analysis of voxel selectivity arising because noise can generate similar results and because this analysis excluded voxels whose amplitudes were negatively correlated across runs, which the Methods failed to state. In addition, Simmons *et al.*² note that our selectivity index incorrectly assigns high values to voxels with low or negative activations. A re-analysis of our data to address these concerns shows that our assertion that many voxels in the fusiform face area (FFA) show high selectivity to nonface objects was incorrect (Fig. 1). In the re-analysis, we used the standard general linear model³ to identify active voxels and cross-validated the results (**Supplementary Methods** online). Owing to these errors, Figures 4 and 8 in the original paper are not valid. The other conclusion remains correct: the fusiform face area is heterogeneous, in that regions of high selectivity for faces are intermingled with regions of lower selectivity (Fig. 1b). These flaws do not affect reproducibility analyses that included all FFA voxels (Fig. 3 and Supplementary Fig. 2), principal component analysis (PCA; Fig. 6), pattern analyses (Supplementary Fig. 5) or comparison of high-resolution and standard-resolution fMRI data (Fig. 7), which remain valid. We sincerely regret these errors.

Erratum in:

Nat Neurosci. 2007 Jan;10(1):133.

Comment in:

Nat Neurosci. 2006 Sep;9(9):1084-6.

Nat Neurosci. 2007 Jan;10(1):3-4.

Nat Neurosci. 2007 Jan;10(1):4-5.

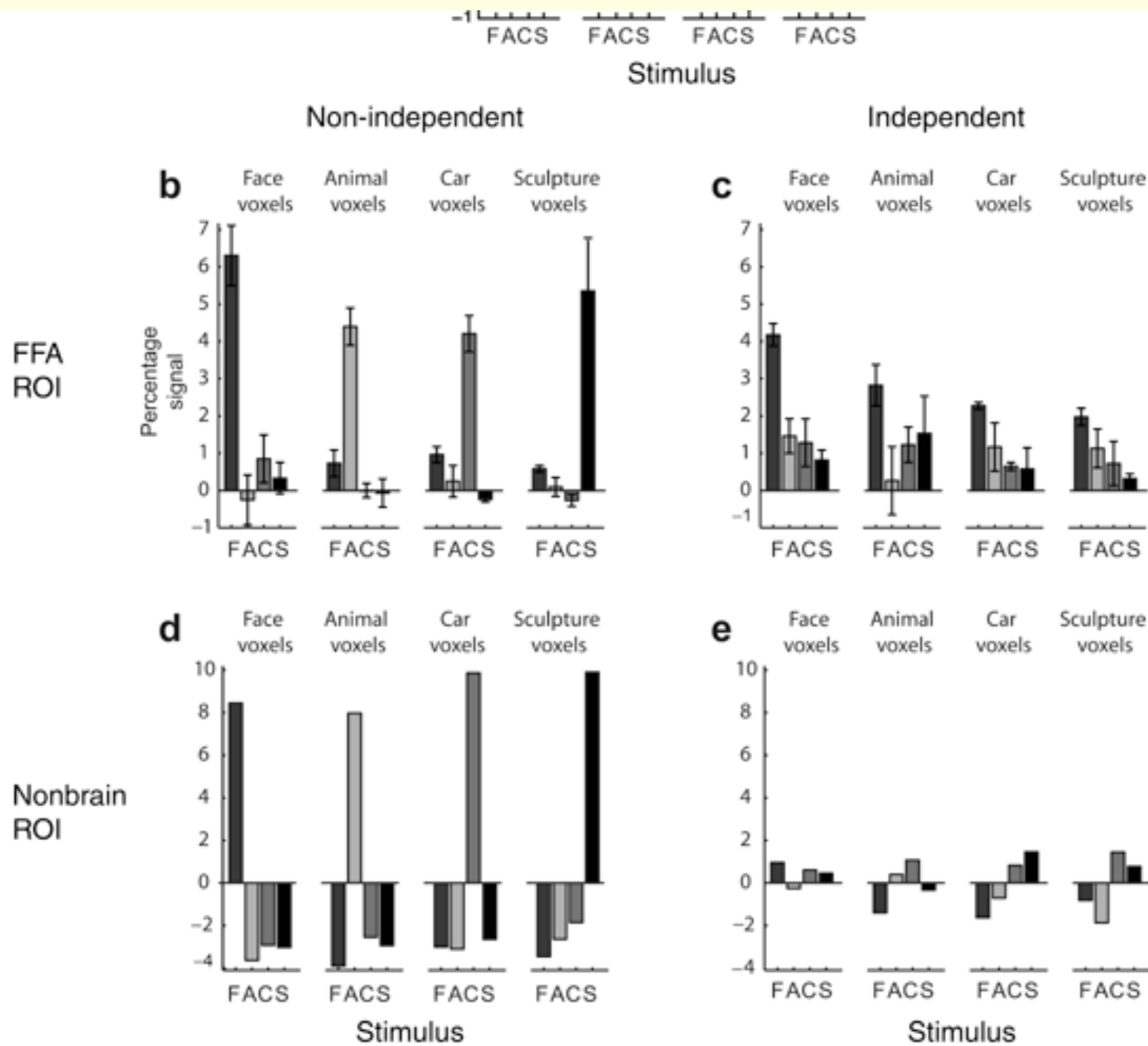
Comment (Kanwisher et al)

To the editor:

(The authors) used high-resolution fMRI (1 mm isotropic

To quantify the selectivity of voxels within the FFA at high resolution, (the authors first determined the preferred category for each voxel (based only on magnitude of response, not on

The problem with this analysis is that it is not independent—the same data were used to select voxels and to produce the selectivity profiles. Even if the data consisted entirely of



While (the authors) reported strong reliability of their selectivity index, this analysis is flawed (see Supplementary Note 1). One standard way to demonstrate reliability of

- **Independence** of the functional ROI for hypothesis tested is essential

2-group problems

- Contrast 2 groups in an fMRI study
- Find they differ in area X (controls have more activation)
- Apply that region to new data
- Conclude that pt group has less activity (including on a separate task).
- What is wrong?

- ROI is optimized for one group-you might find other voxels- same region- optimized for pt group
- Using an average functional ROI across groups- not sufficient (given magnitude diffs)
- Best used to compare within group on different tasks