



# Experimental Design

---

Sonia Bishop

fMRIB, NDCN, University of Oxford &  
HWNI & Dept. Psychology, UC Berkeley

With thanks to:  
Kate Watkins  
Heidi Johansen-Berg  
Joe Devlin



# Outline

---

- Choices for experimental paradigm
  - Subtraction
  - Factorial } *covered by ED*
  - Parametric }
  - Conjunction }
- Choices for fMRI protocol
  - Blocked vs. Event-related
    - Efficiency (choice of block length, fixed vs random ISIs for event related designs, trial order, use of null trials)
    - Sampling of HRF
  - Mixed designs
  - Sparse Sampling



# Outline

---

- Other things to think about
  - between subject covariates (e.g. anxiety, age)
  - controlling for (unwanted) effects of arousal, task difficulty etc.
  - counterbalancing
- Recent advances
  - Examining representational similarity
    - Adaptation, MVPA, calculating voxel 'tuning-curves'



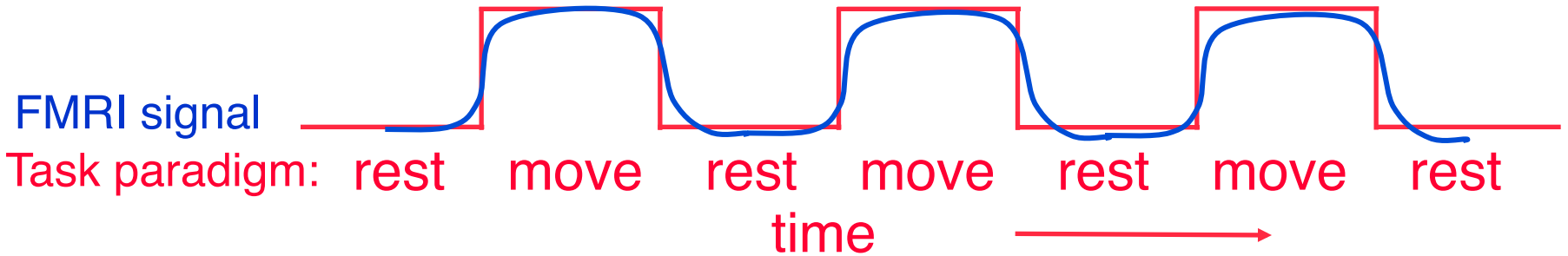
# Outline

---

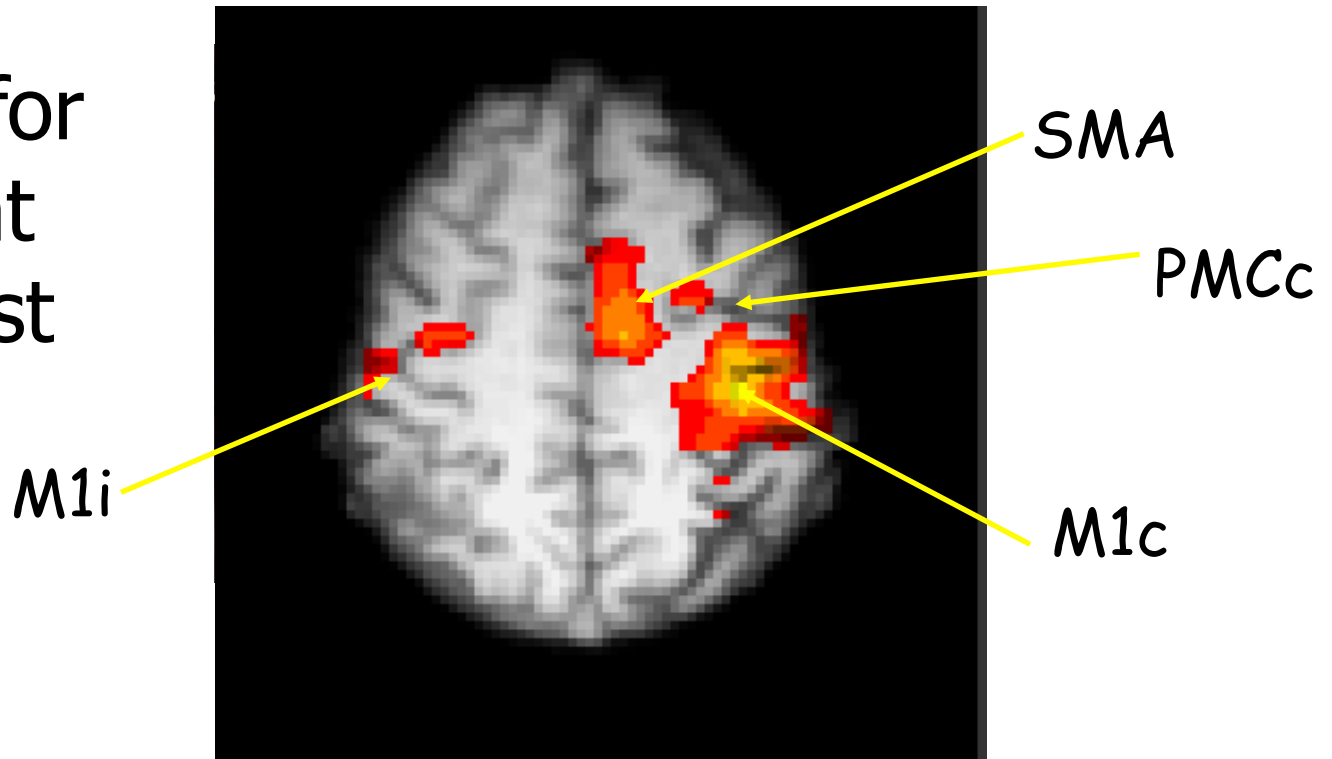
- Choices for experimental paradigm
  - **Subtraction**
  - Factorial } *covered by ED*
  - Parametric }
  - Conjunction }
- Choices for fMRI protocol
  - Blocked vs. Event-related
    - Efficiency (choice of block length, fixed vs random ISIs for event related designs, trial order, use of null trials)
    - Sampling of HRF
  - Mixed designs
  - Sparse Sampling



# A very simple experiment



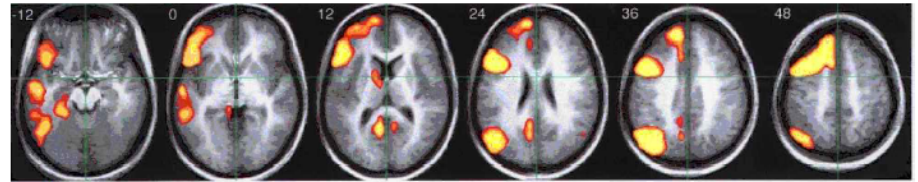
Activation for  
movement  
versus rest



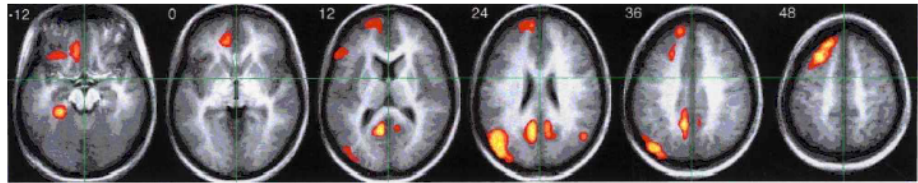
# Choosing Baselines

---

- Rest may not be truly rest
- Need to control as much as possible to isolate component of interest
- Even if a task does not explicitly involve a particular component, subjects may engage in it anyway



Rest vs tones

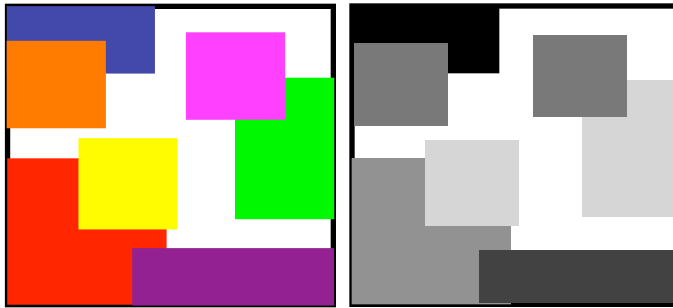
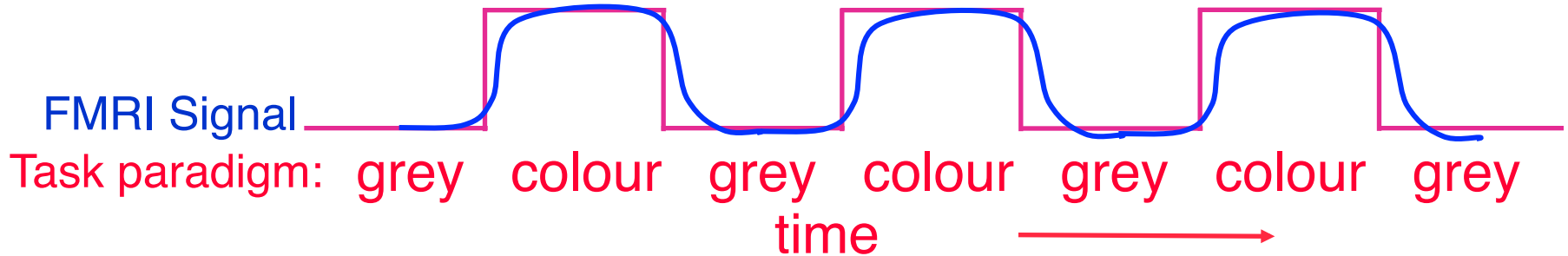


Semantic processing vs tones

*Binder et al, 1999*



# Simple experiment: subtraction



Colour > No colour  
Luminance = Luminance  
Contrast = Contrast  
Shape = Shape

Brain mapping



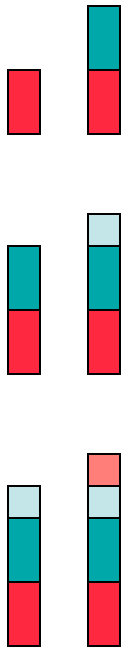
Isolate functional area



# Hierarchical processing

## Single-Word processing

<b>Subtraction</b>	<b>Control</b>	<b>Task</b>	<b>Hypothetical Cognitive Operations</b>
Sensory	Fixation	Passive words	Passive sensory processing
Production	Passive words	Repeat words	Motor programming and output
Association	Repeat words	Generate words	Semantic association, selection

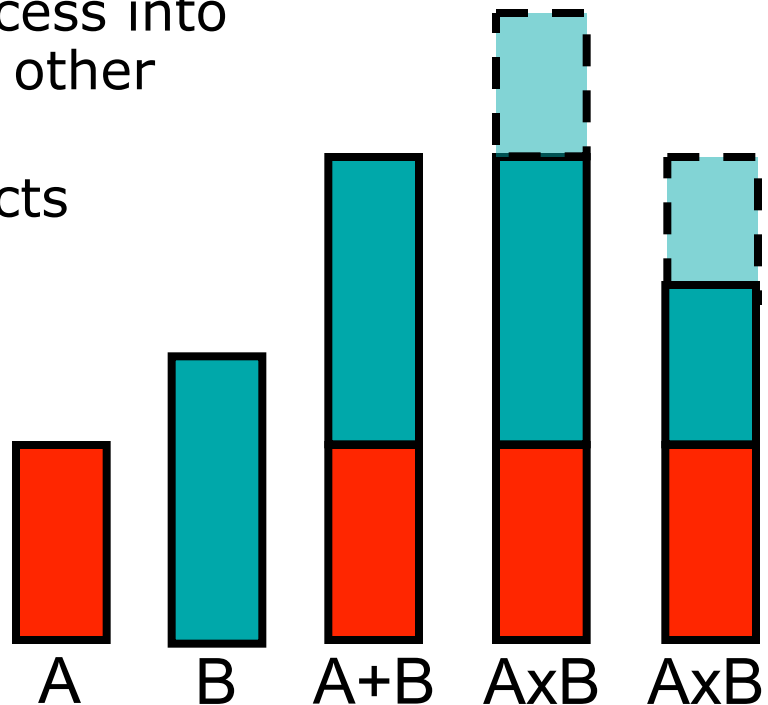


Petersen et al., Nature 1998

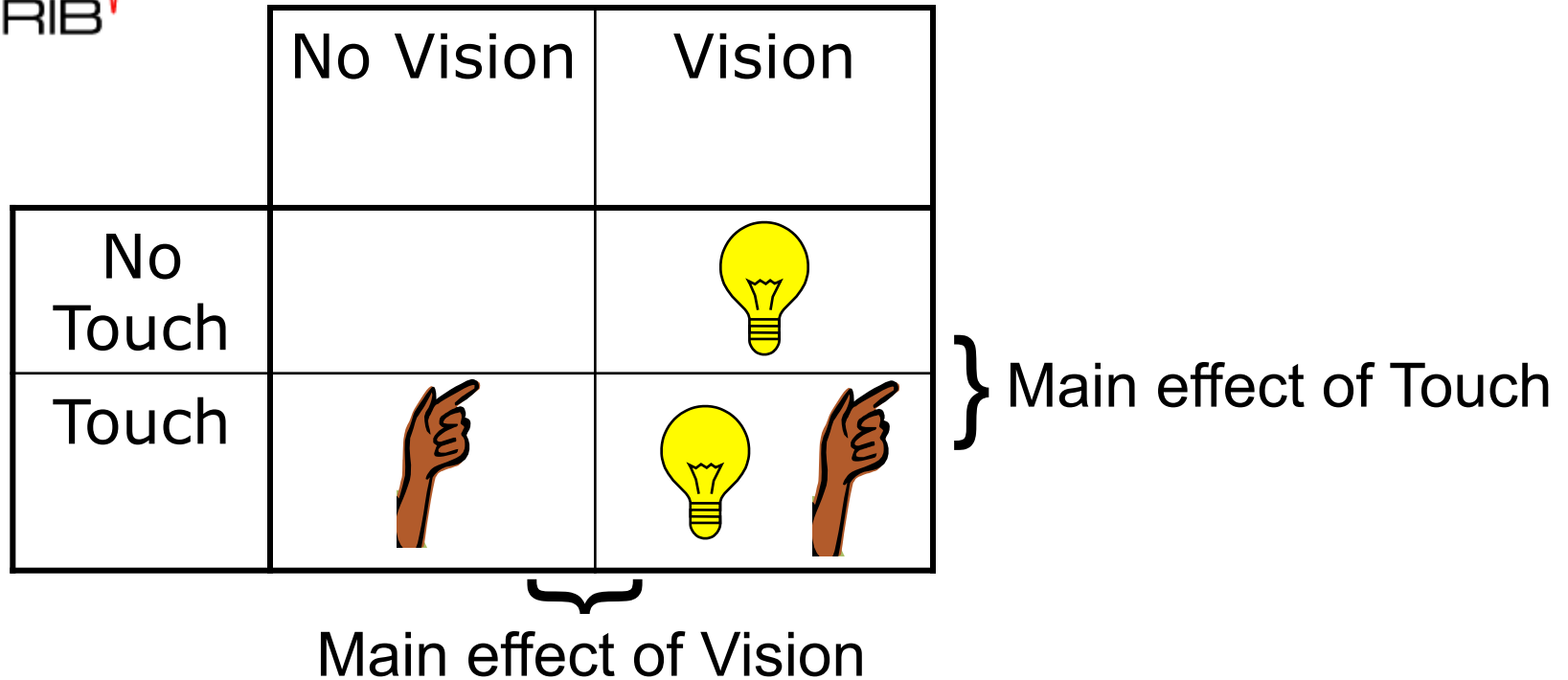


# Problems with subtractive designs

- Depends on the assumption of **'Pure Insertion'**
  - i.e. the idea that you can insert a single component process into a task without affecting other processes
  - Can get interactive effects



# Factorial designs



Allows you to characterise **interactions** between component processes – i.e., effect that one component has on another (does not make assumption of pure insertion)

# Conjunction analyses

## Cognitive subtraction

	Task A	Task B
Process 1	Grey	Grey
2	Grey	Grey
3	Grey	Grey
4 (PI)	Black	White
5	Grey	Grey

Activation vs baseline  
= component of interest

## Cognitive conjunction

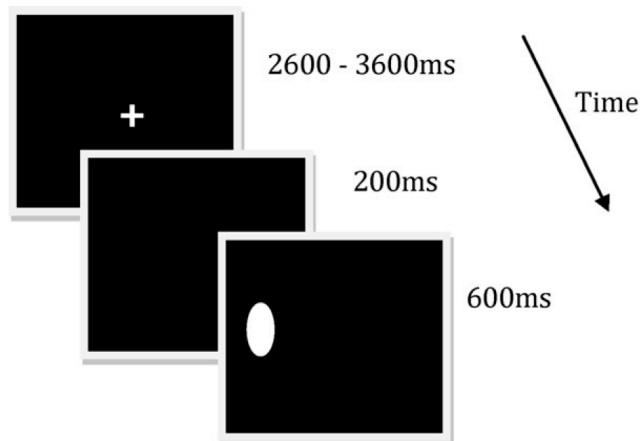
	Task Pair I		Task Pair II	
	A	B	A	B
Process 1	Grey	Grey	White	White
2	Grey	White	Grey	Grey
3	White	White	Grey	Grey
4 (PI)	Black	White	Black	White
5	Grey	Grey	White	White

Commonalities in  
activation vs baseline  
across task pairs  
= component of interest

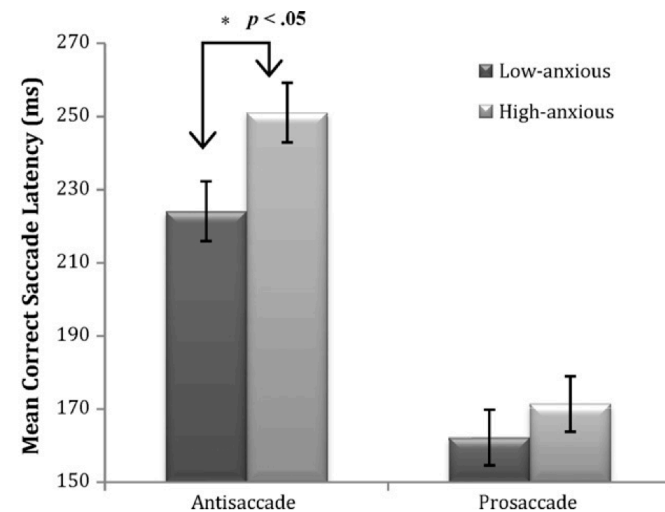
- Does not assume pure insertion
- Does not depend on perfect baseline

# An initial example

- Often we may want to adapt experiment from the psychology or EEG literature for use in fMRI – this may involve certain changes being needed ...
- The antisaccade task

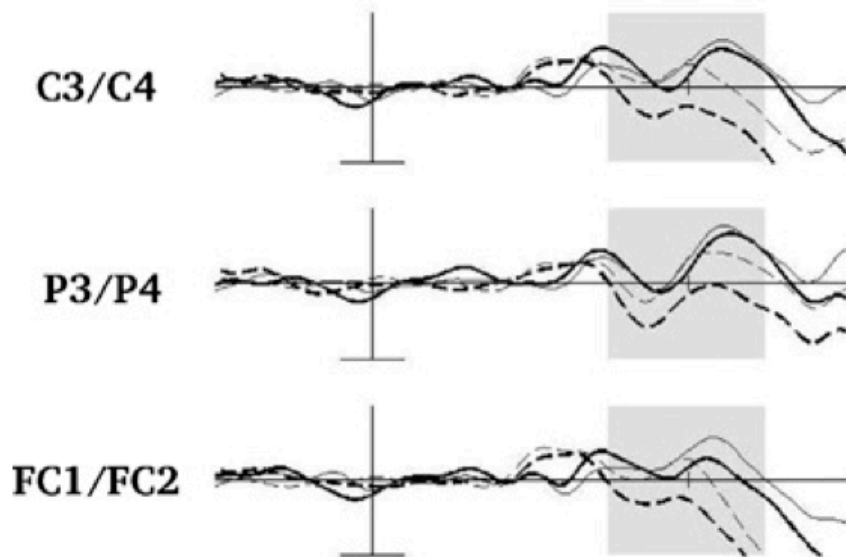


**Fig. 1.** Schematic representation of a typical trial: on antisaccade trials, participants were instructed to look AWAY from the oval and on prosaccade trials they were instructed to look TOWARD the oval as soon as it appeared.



# An initial example

- Often we may want to adapt experiment from the psychology or EEG literature for use in fMRI – this may involve certain changes being needed ...
- The antisaccade task



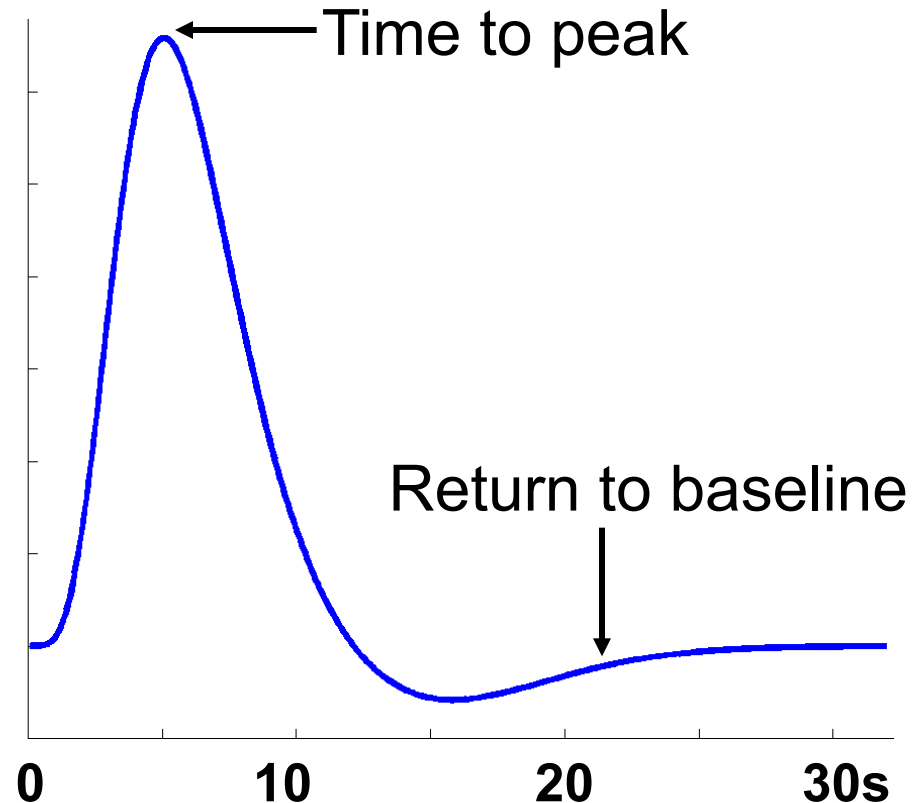
antisaccade (black)  
 prosaccade (grey)  
 Low anxious (dotted line)  
 show preparatory negative  
 deflection on antisaccade  
 trials unlike high anxious  
 (solid line)

■ = 50ms pre target to 50ms post target



# An initial example

- The antisaccade task
- *So what are the issues with adapting for fMRI*
- It will be hard to separate BOLD response to preparatory activity just before target from BOLD response to seeing the oval or making the saccade





# An initial example

---

- Solution?
- 'Half' trials ... on 1/3<sup>rd</sup> of trials subjects prepare but instead of the oval the screen remains blank until fixation returns signalling next trial (i.e. only preparation, no saccade)
- Using *subtraction* logic activity to half antisaccade trial vs half prosaccade trial should reflect preparation for antisaccade (if you know it is an antisaccade trial)
- Also using *subtraction* can examine anti-saccade activity linked to the making of the saccade by subtracting half antisaccade trial activity from full antisaccade trial activity
- Using *conjunction* analysis can examine common preparatory response to half antisaccade and prosaccade trials



# Outline

---

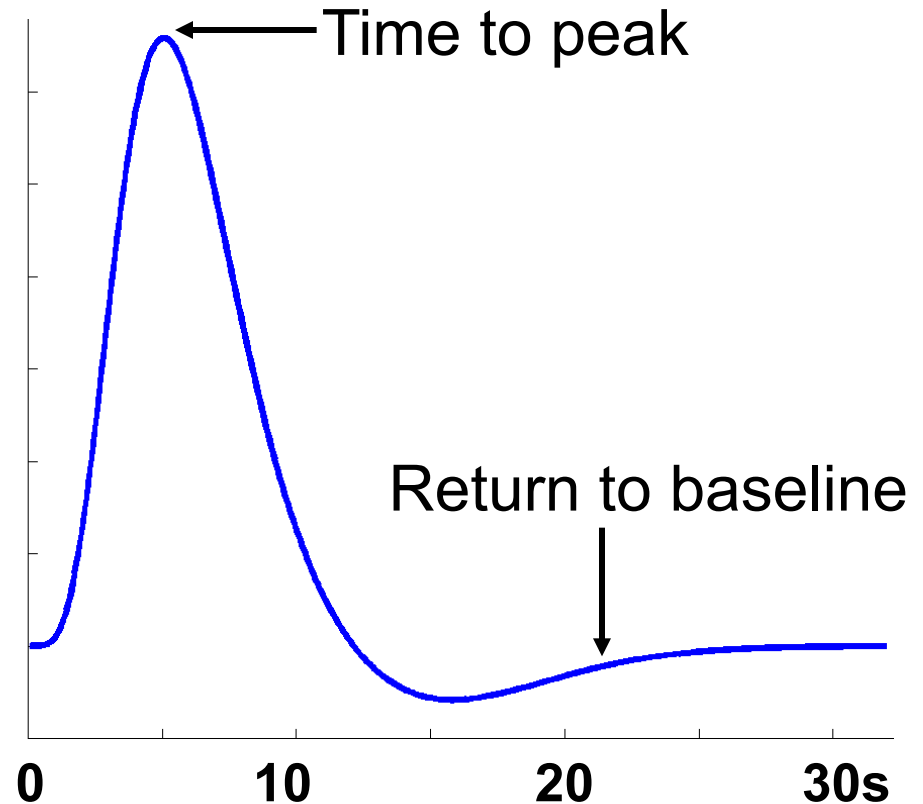
- Choices for experimental paradigm
  - Subtraction
  - Factorial } *you will hear more*
  - Parametric } *about these tomorrow!*
  - Conjunction }
- Choices for fMRI protocol
  - Blocked vs. Event-related
    - Efficiency (choice of block length, fixed vs random ISIs for event related designs, trial order, use of null trials)
    - Sampling of HRF
  - Mixed designs
  - Sparse Sampling





# Haemodynamic response

- HRF is slow to peak
- Peak response comes 4-6s after stimulus onset
- Can vary in time-to-peak across brain areas and across subjects
- Returns to baseline about 21s after stimulus ends

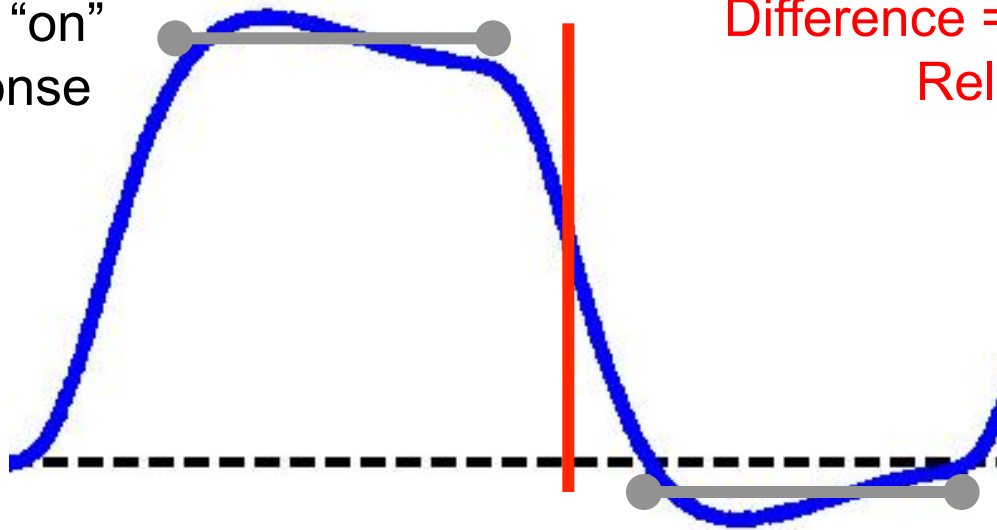




# BOLD is a relative measure

---

Mean "on"  
response



Difference =  
Relative change

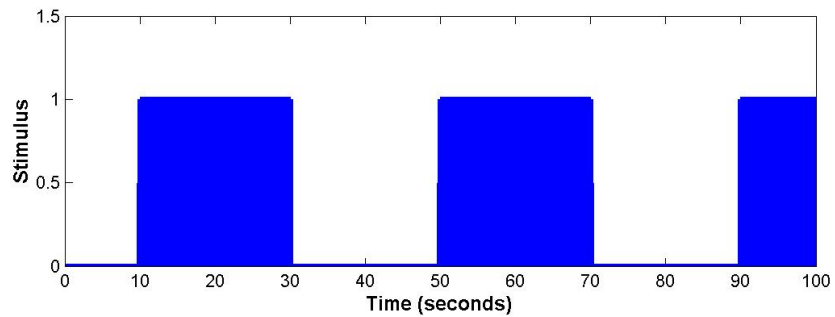
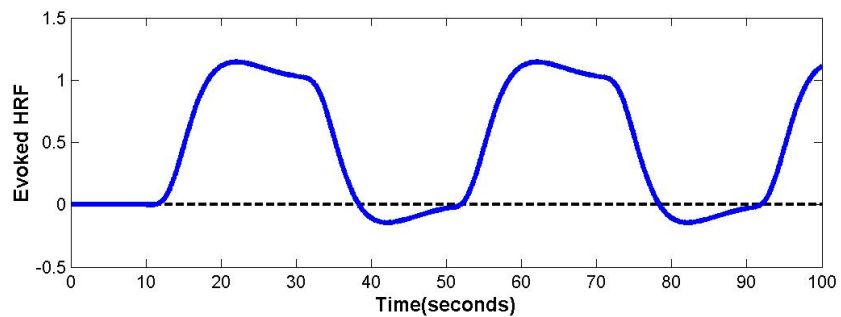
Mean "off"  
response

Sensitivity depends on **maximizing** relative change

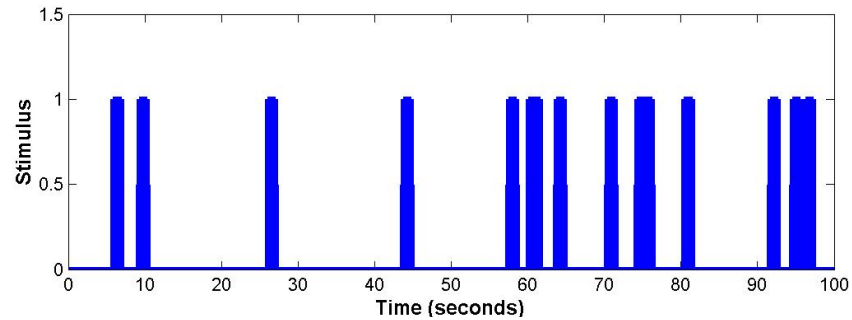
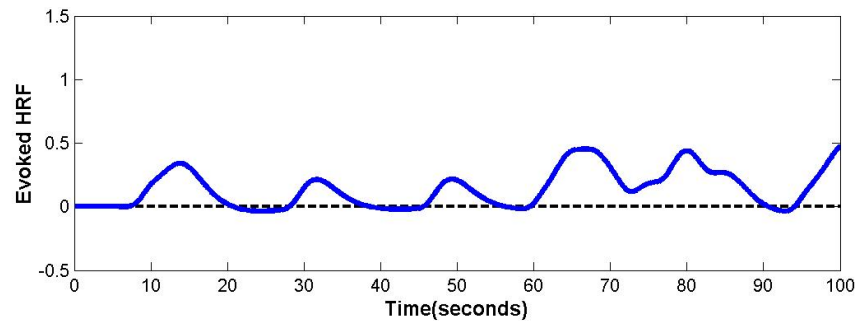


# Basic designs

## Blocked

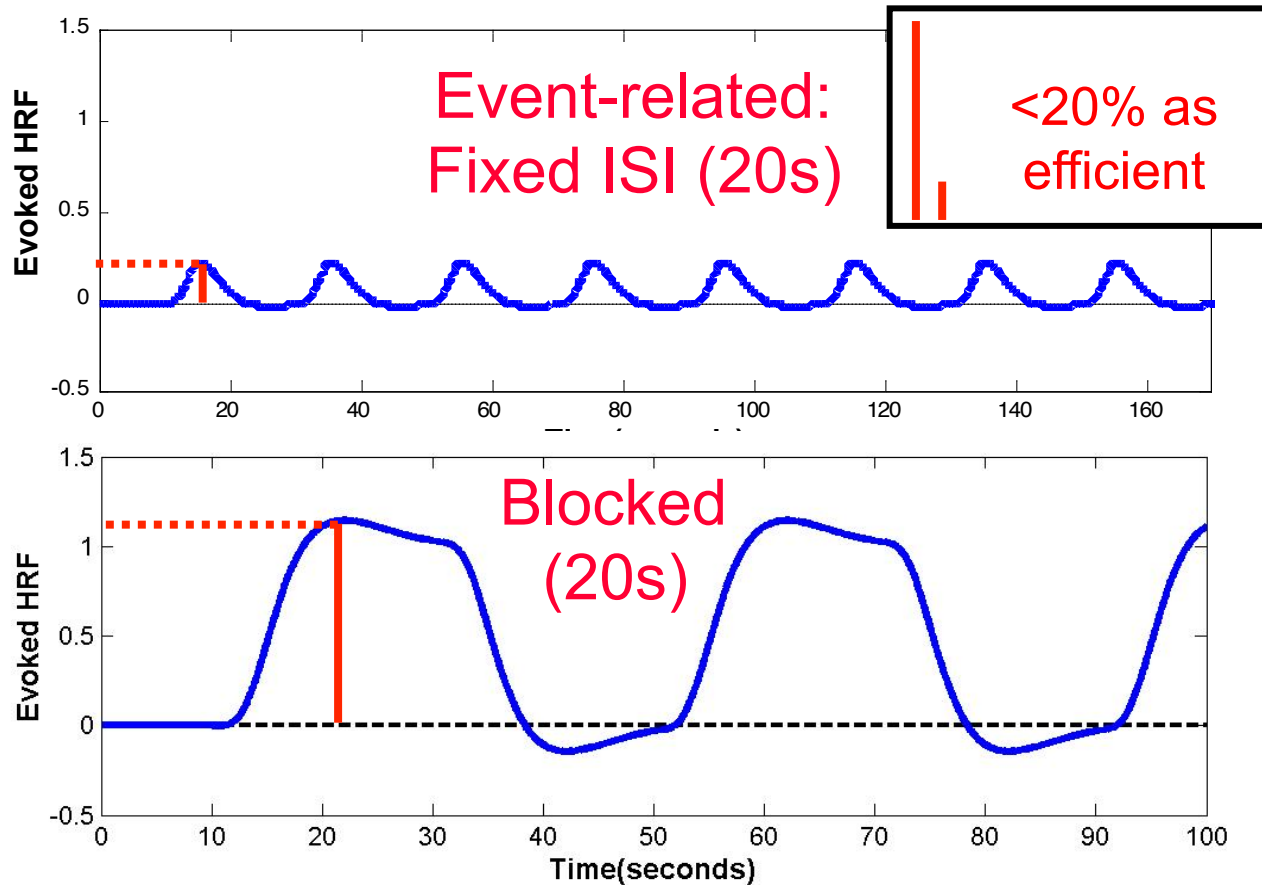


## Event related





# Efficiency: event-related vs blocked

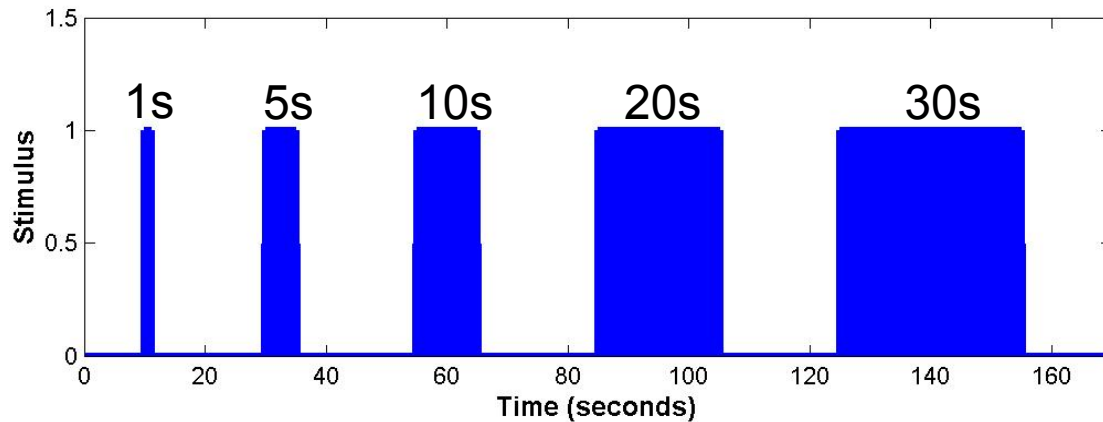
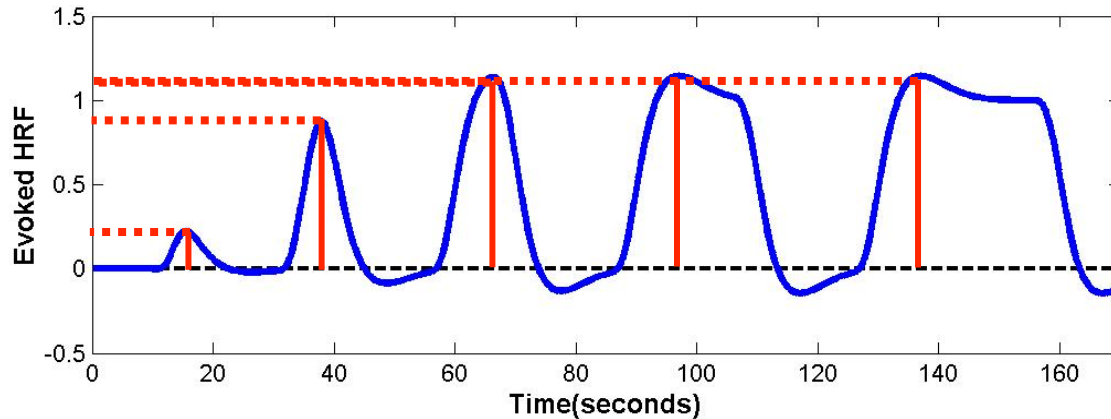


Blocked designs are more efficient than (slow) event-related designs

Higher efficiency -> less time doing task for same power to detect effect



# Block length: efficiency, switching and task set



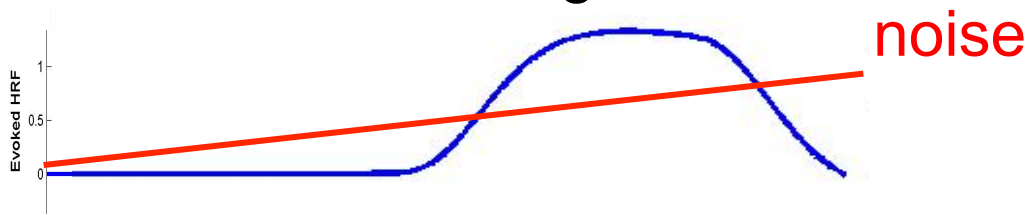
No further efficiency benefit to increasing block length once reach 10-20s

*Other considerations:*  
shorter the block,  
more task switching,  
harder to establish  
'attentional' or 'task'  
set

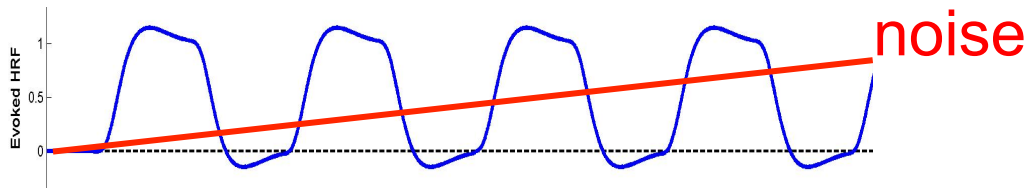


## Block length: noise ( $1/f$ ), scanner drift

One long block



Multiple short blocks



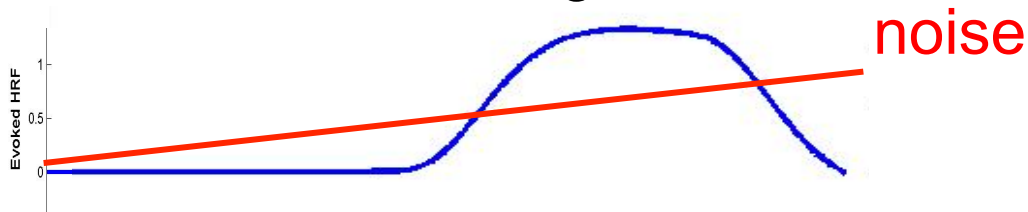
No further efficiency benefit to increasing block length once reach 10-20s

*Other considerations:*  
longer the block,  
more overlap  
between design  
frequency and noise  
frequency / impact of  
scanner drift

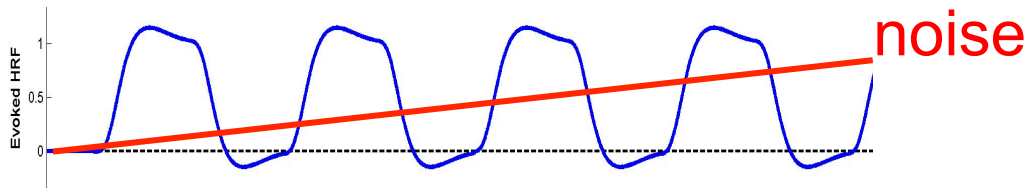


# Block length: take home

## One long block



## Multiple short blocks



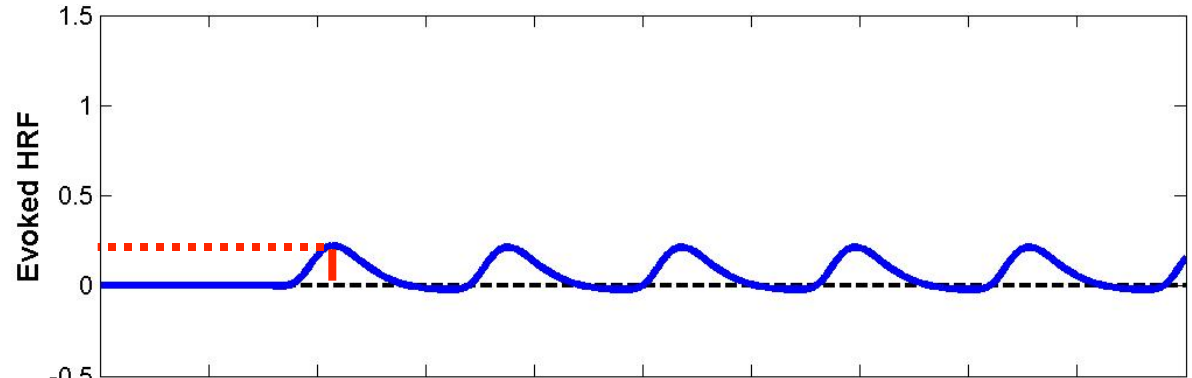
No further efficiency benefit to increasing block length once reach 10-20s

*Take home:*  
Blocks of ~20s often good, max efficiency, not so short can't get into 'set', not so long problem with scanner noise



# ER design efficiency: fixed ISI

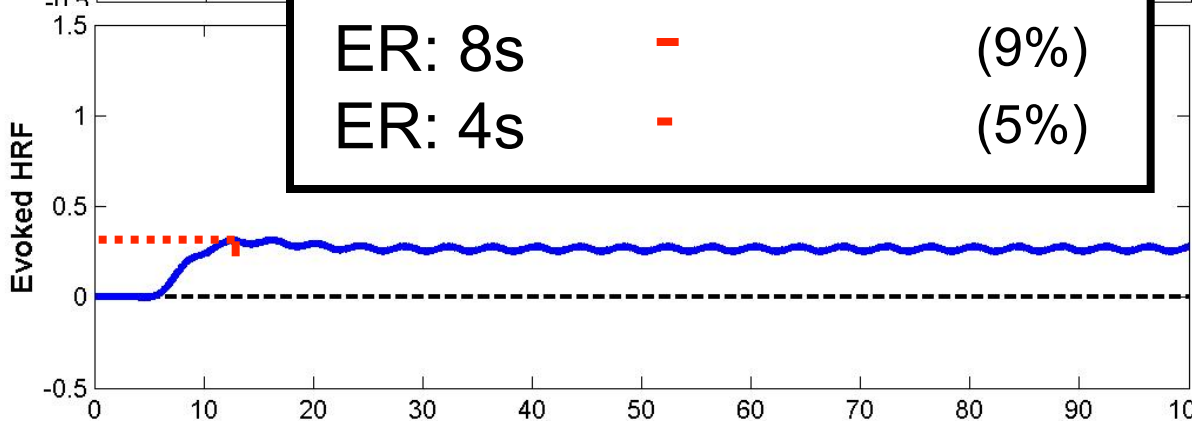
**Fixed ISI (16s)**



**Fixed ISI (8s)**



**Fixed ISI (4s)**

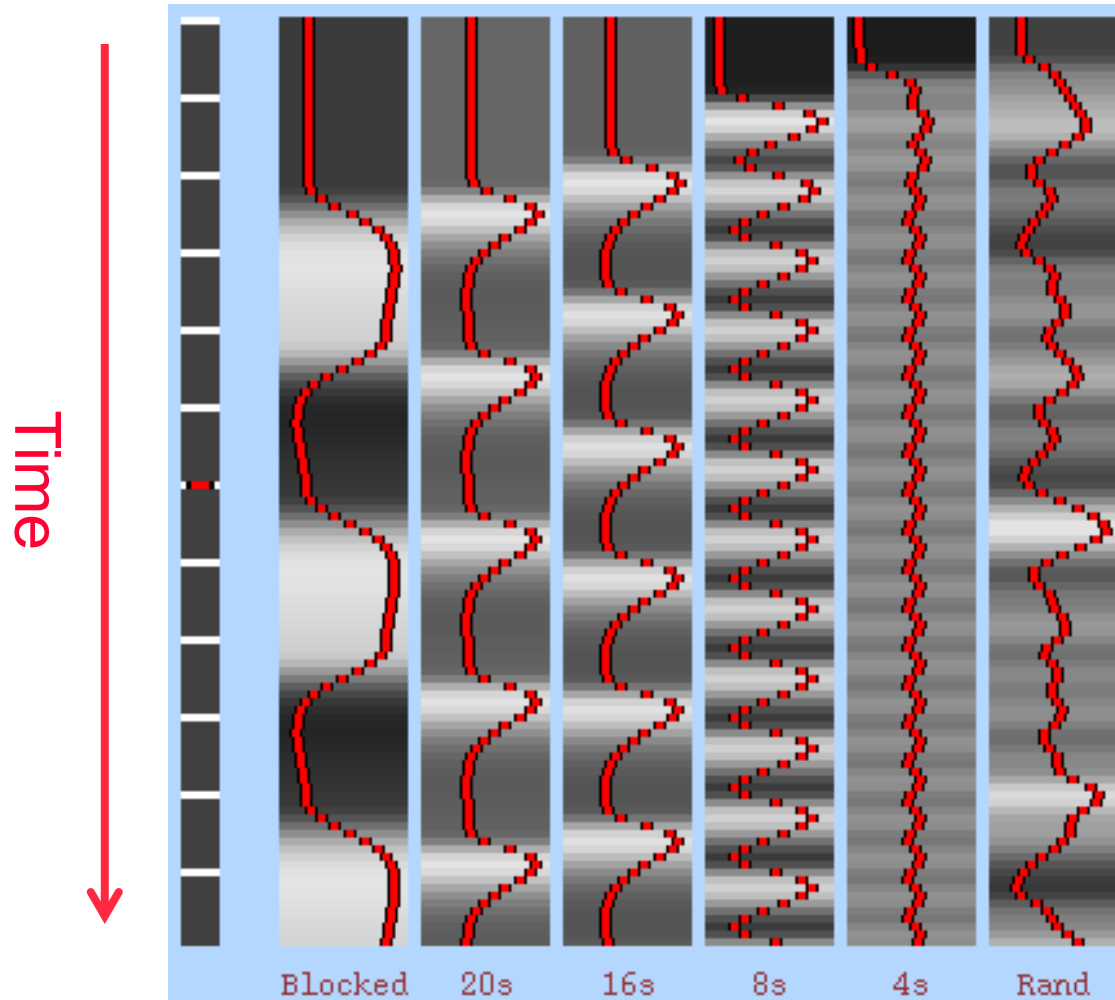


<u>Design</u>	<u>Efficiency</u>
Blocked	— (100%)
ER: 20s	- - (18%)
ER: 16s	- - - (13%)
ER: 8s	- - - - (9%)
ER: 4s	- - - - - (5%)





# ER designs: random vs fixed ISI



By 'jittering' ISI  
can increase  
efficiency



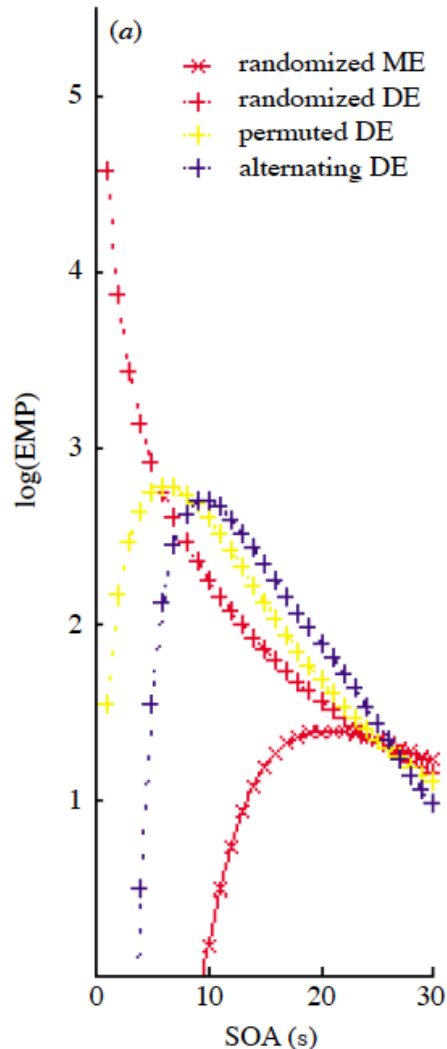
# Event-related design, random isi

---

- Same random isi design can be more or less efficient due to chance differences in random order
- Programs for optimising random presentation
  - OptSeq: Greve
  - Genetic algorithm: Wager & Nichols
- Optimise designs *before* scanning



# ER designs: trial sequence



*Classic paper: Josephs & Henson, 1999*

*ME (condition A or B vs baseline)*

*DE (condition A vs Condition B)*

*ABBABBBAAABA ...* Random

*ABABABABABAB ...* Alternating

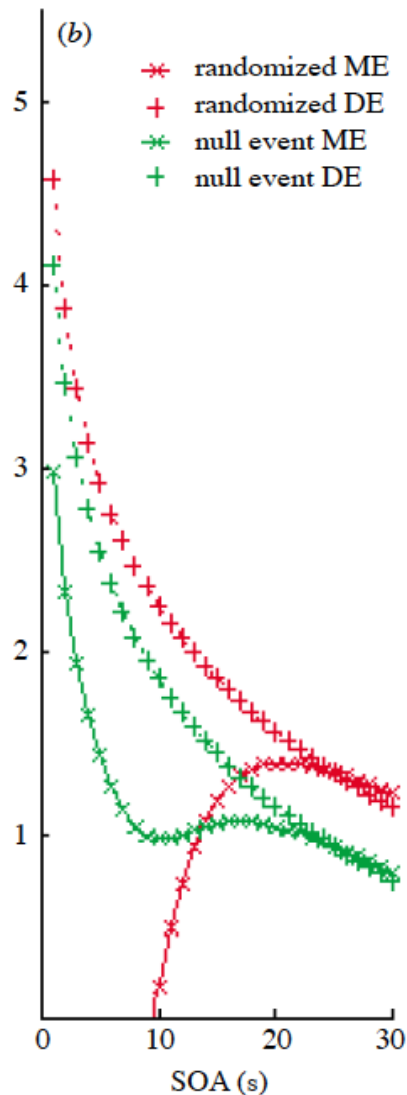
*ABBABAABBABA ...* Permuted

*Can see as soa (isi) gets shorter, fixed (alternating) order power\* drops off rapidly; random A B design also cannot detect A vs baseline or B vs baseline at short soas (isis) (ME)*

*\* Estimated measurable power (sum of squared signal / nu scans)*



# ER designs: trial sequence



*Classic paper: Josephs & Henson, 1999*

*ME (condition A or B vs baseline)*

*DE (condition A vs Condition B)*

*random A B design also cannot detect  
A vs baseline or B vs baseline at short  
isis (ME)*

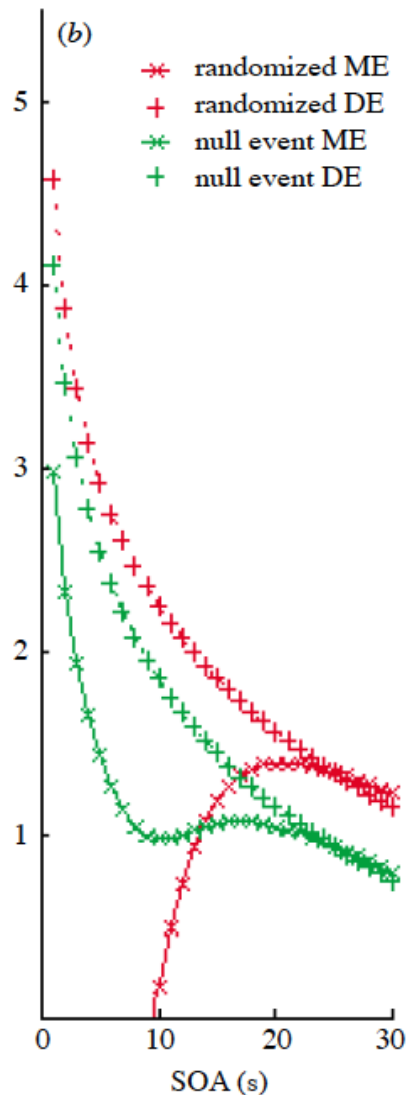
*If you add in null trials (in random order)  
this can be overcome*

*(this approximates to jittering isi)*

***You need to know if you want to  
detect just A-B or also A vs baseline,  
B vs baseline***



# ER designs: trial sequence



*Classic paper: Josephs & Henson, 1999*

*ME (condition A or B vs baseline)*

*DE (condition A vs Condition B)*

*random A B design also cannot detect A vs baseline or B vs baseline at short SOA (ME)*

***Reviewers do not always understand you cannot look at A vs baseline or B vs baseline in fast randomized designs!***

***You may need to be able to explain this ....***

*If you have 3 + trial types, and try look at A vs baseline in such a design baseline will really just capture all trials other than A*



# Blocked design

---

- **Advantages**

- Simple (for you and for subject)
- Minimise task switching
- Maximum efficiency
- Does not depend on accurate HRF model
- Robust to uncertainty in timing
- Straightforward analysis

- **Disadvantages**

- Not all tasks can be blocked
- Subjects can anticipate conditions - order and duration
- Does not allow separation of response to individual trials
- No timing information



# Event-related designs

---

- **Advantages**

- Flexible – removes anticipation, allows for surprises
- Good estimate of time course of HRF
- Post hoc sorting of trial types, e.g. correct vs. incorrect; remembered vs. forgotten stimuli
- Can separate our response to task components – e.g., cue, target, response
- High temporal resolution

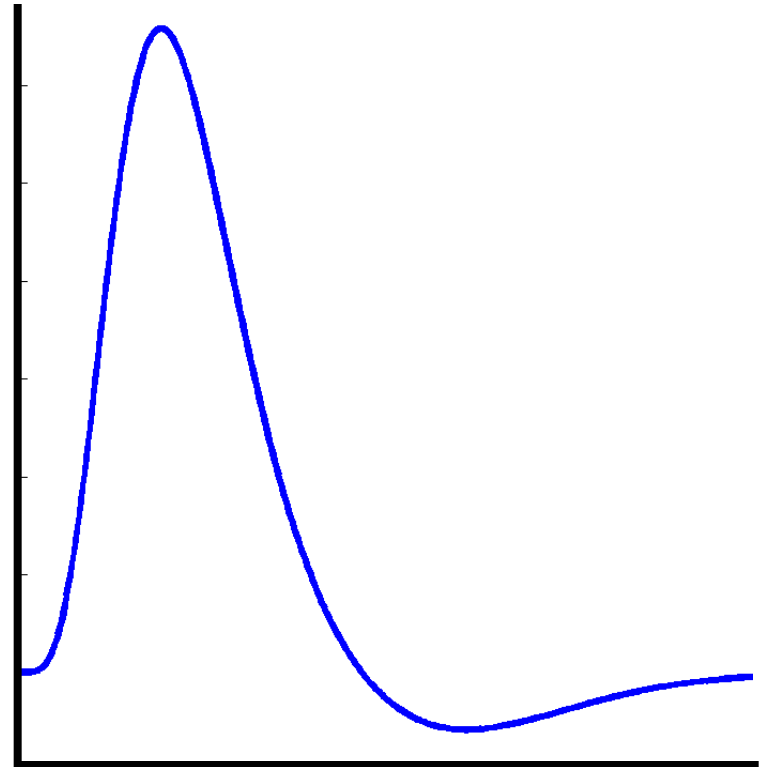
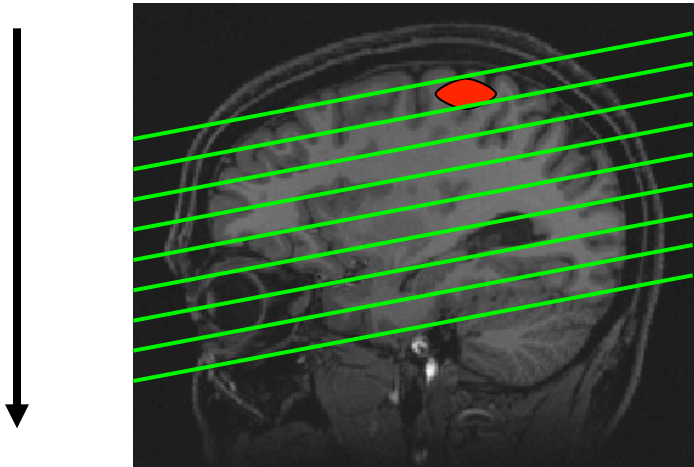
- **Disadvantages**

- More things can go wrong
- Reduced efficiency
- Typically results in longer experiments
- More dependent on accurate HRF modelling
- Increased task switching



# Sampling bias

---



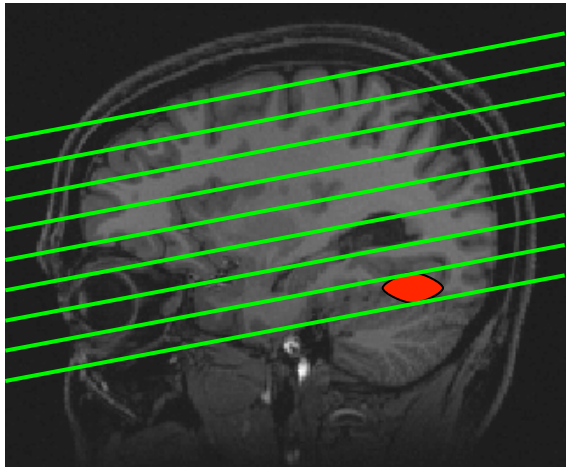
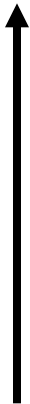
Only sample the HRF once per TR (3s)



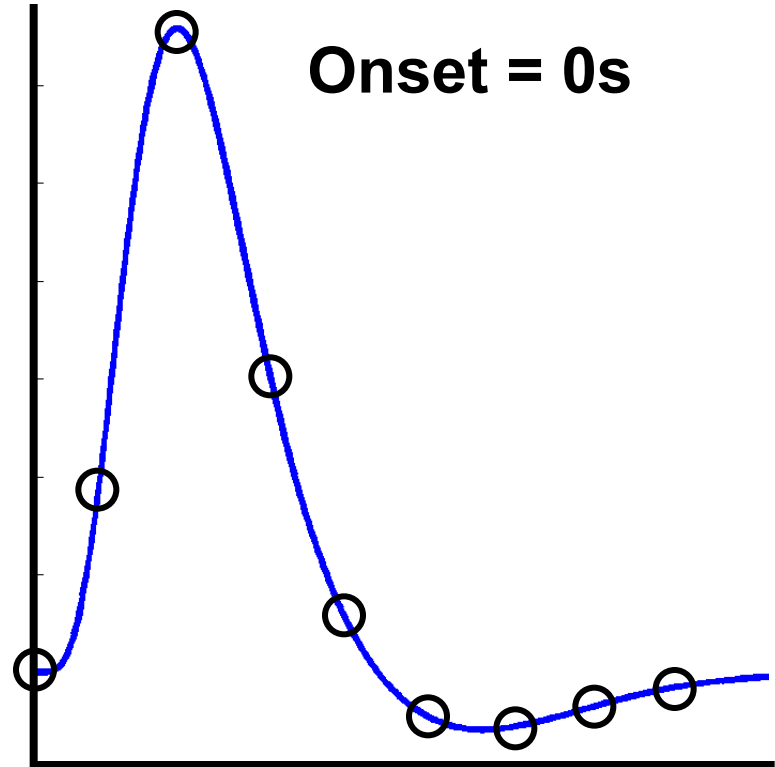


# Sampling bias

Regionally specific



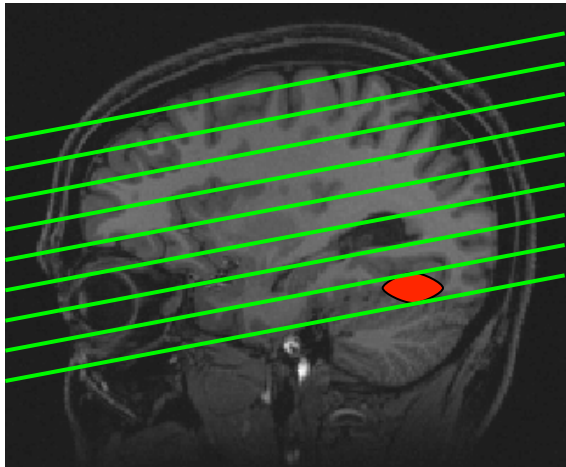
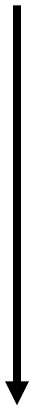
Good estimate



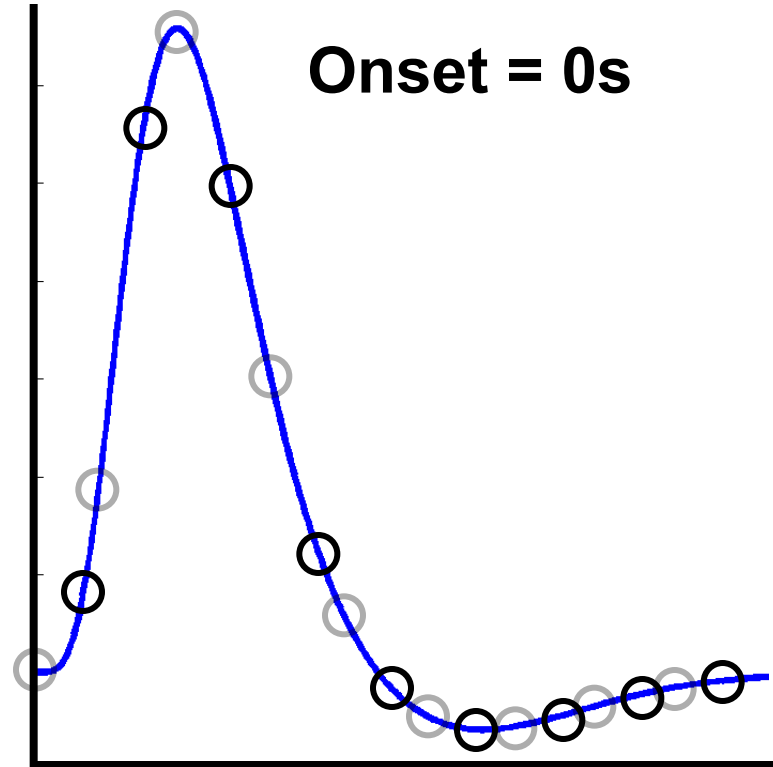


# Sampling bias

Regionally specific



Poor estimate



Can underestimate effect sizes

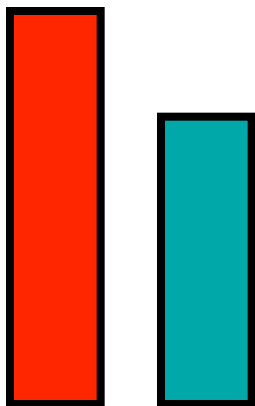


# Effects of sampling bias

Contrasts can over- or under-estimate effects

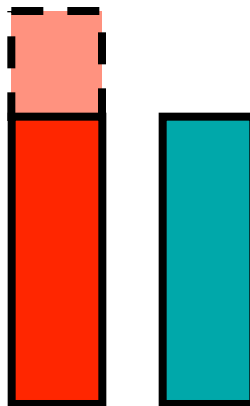
Reality:

$A > B$



Underestimate A:

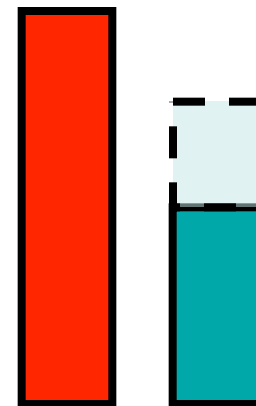
$A = B$



Underestimates  
difference

Underestimate B:

$A \gg B$

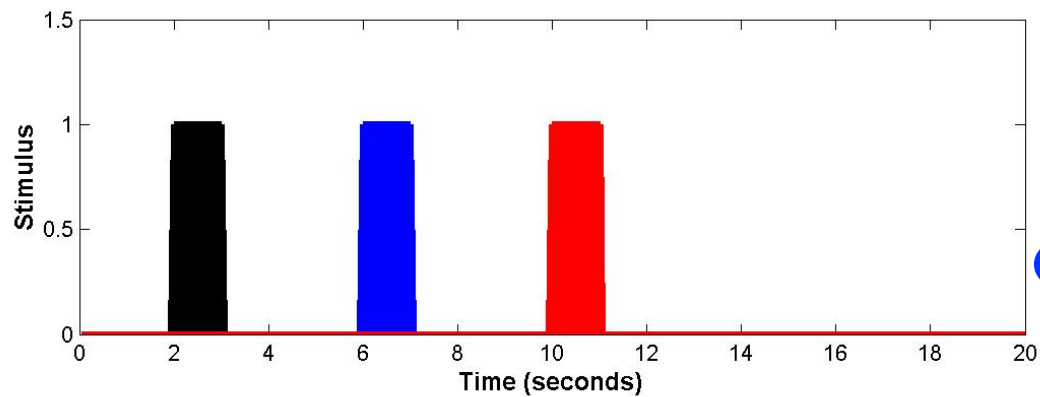
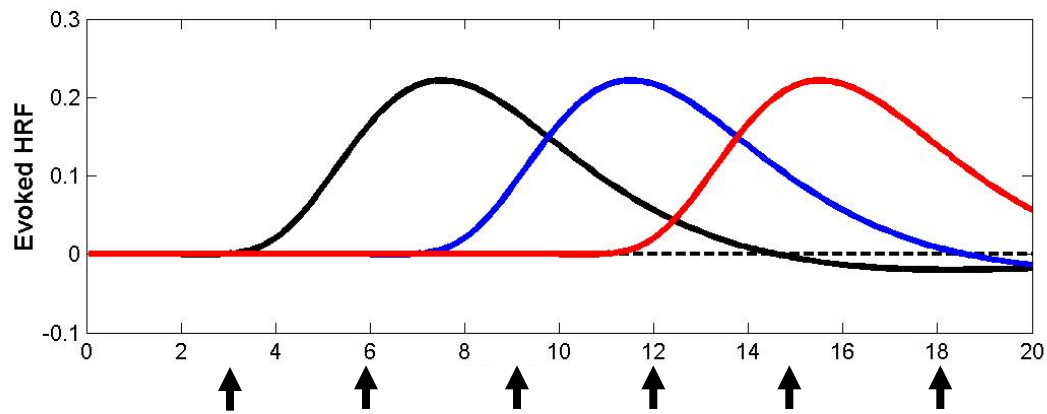


Over estimates  
difference

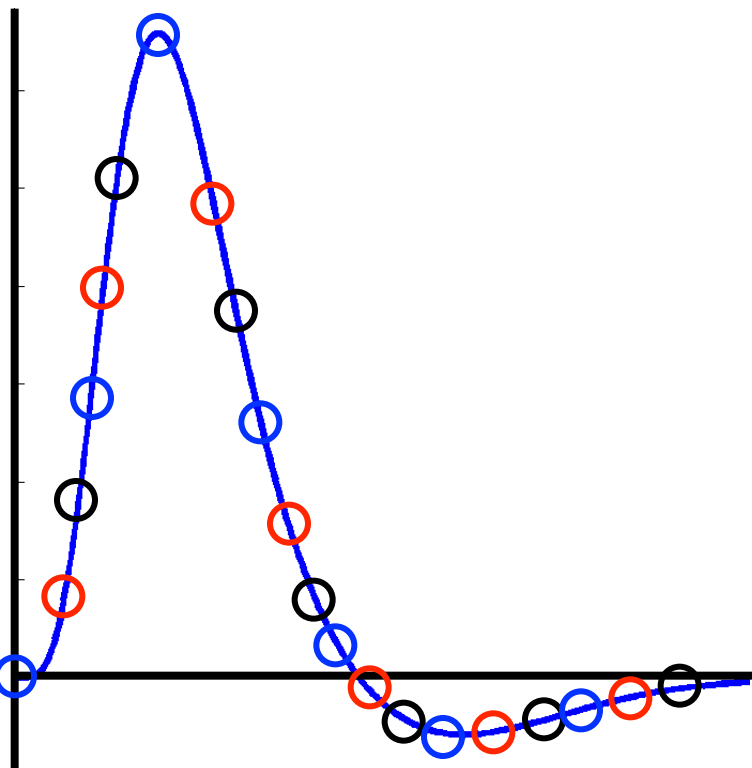


# Oversampling

TR = 3s ISI = 4s



TRs ↑



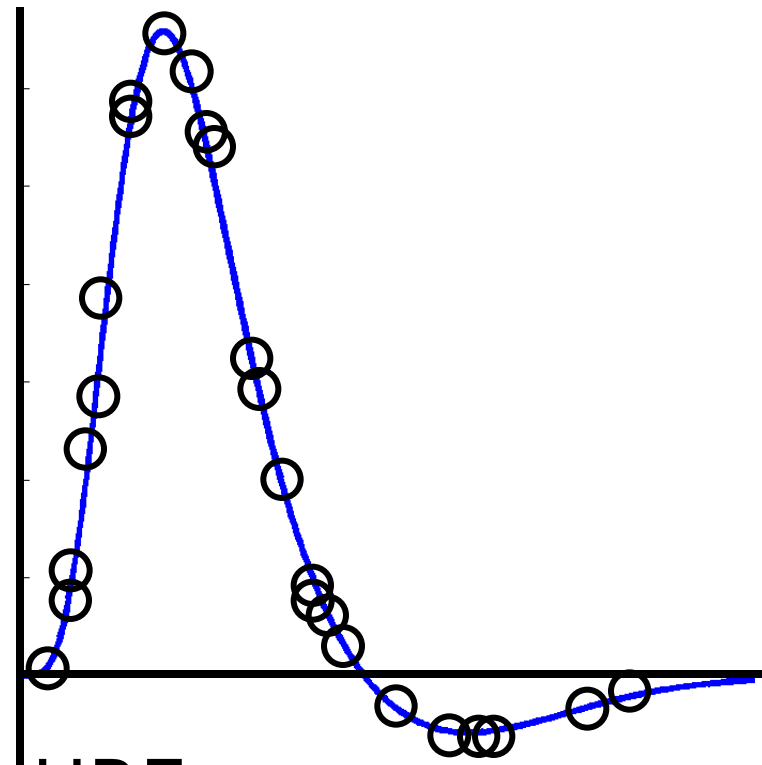
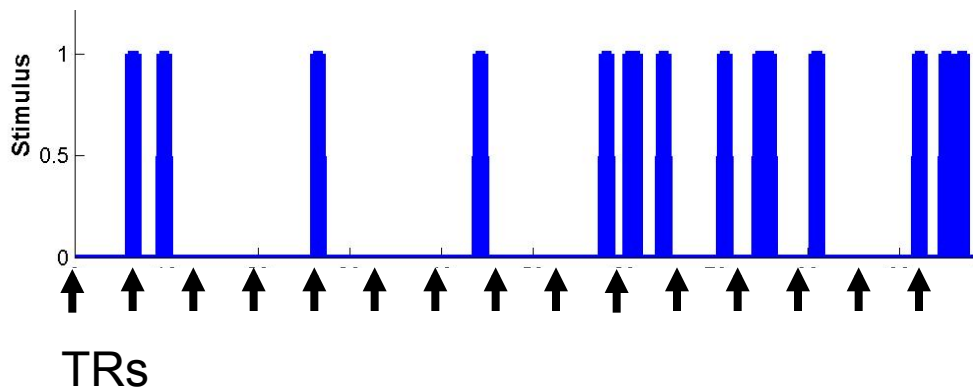
- effective sampling rate of 1s





# Jittering the ISI

Choosing ISI from a random distribution



Random sample of evoked HRF

Inherent benefit of jittered ISI ER designs



# Sampling bias summary

---

- For both blocked and event-related designs:
  - Avoid  $TR = \text{integer multiple of ISI}$
  - Oversampling uses *fixed* ISIs (good for blocked, ER where fixed ISI WITH random trial sequence &/- nulls)
  - Jitter uses *random* ISI (easiest to use with event-related and not mess up either power - ability to differential signal for different trial types - or sampling)



# Mixing blocks and events

---

- Qu: How do response conflict (interference) effects change as a functional of perceptual load?

High perceptual load trial



Low perceptual load trial



Response conflict conditions

**Congruent:** e.g. Target X, Distractor X

**Neutral:** e.g. Target X, Distractor C

**Incongruent:** e.g. Target X, Distractor N





# Mixing blocks and events

---

- Qu: What should we block? What needs to be event-related?
- Issues: attentional set; conflict from Inc trial greater when follows Cong trial than when follows Inc trial – Carter et al. (2000)

High perceptual load trial



Low perceptual load trial



## Response conflict conditions

**Congruent:** e.g. Target X, Distractor X

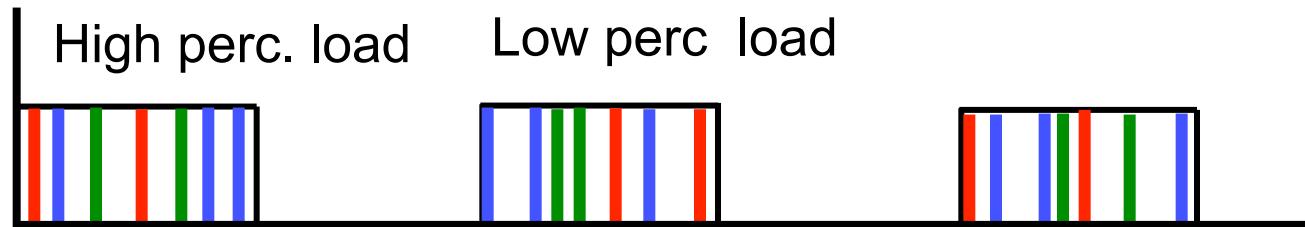
**Neutral:** e.g. Target X, Distractor C

**Incongruent:** e.g. Target X, Distractor N



# Mixing blocks and events

- **Qu:** *How do response conflict (interference) effects change as a functional of perceptual load?*
- **Design:**



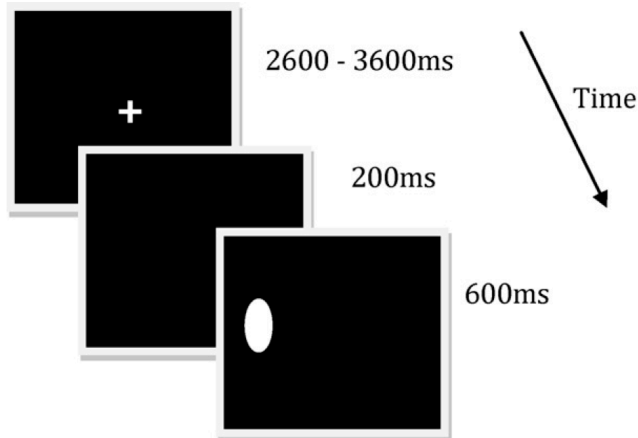
**Congruent:** Target X, Distractor X  
**Neutral:** Target X, Distractor C  
**Incongruent:** Target X, Distractor N

- Block perceptual load manipulation so subjects can get into attentional 'set' and to reduce task switching effects
- Have response conflict vary event-related so avoid response conflict effects washing out due to strings of incongruent trials increasing expectancy of and preparation for high conflict trials

(Bishop Nature Neuroscience, 2009)

# Back to our antisaccade example

- The antisaccade task



**Fig. 1.** Schematic representation of a typical trial: on antisaccade trials, participants were instructed to look AWAY from the oval and on prosaccade trials they were instructed to look TOWARD the oval as soon as it appeared.

- Equal numbers of ‘anti’ and ‘pro’ saccade trials
- 2/3rds of trials are ‘full trials’ (preparation and saccade);
- 1/3<sup>rd</sup> of trials are ‘half’ trials (subjects prepare but instead of the oval the screen remains blank (i.e. only preparation, no saccade))



# Back to our antisaccade example

---

- What type of design is this?
- Factorial: type of saccade (anti vs pro) by trial type (full, half)
- Using *subtraction* logic activity to half antisaccade trial vs half prosaccade trial should reflect preparation for antisaccade **(if you know it is an antisaccade trial)**
- *So should we use blocked or event-related?*
- Solution: block type of saccade (pro or anti) as want subject to know which to prepare for, randomise half and full trials within block so subjects still prepare on half trials = **mixed design.**



# Sparse sampling

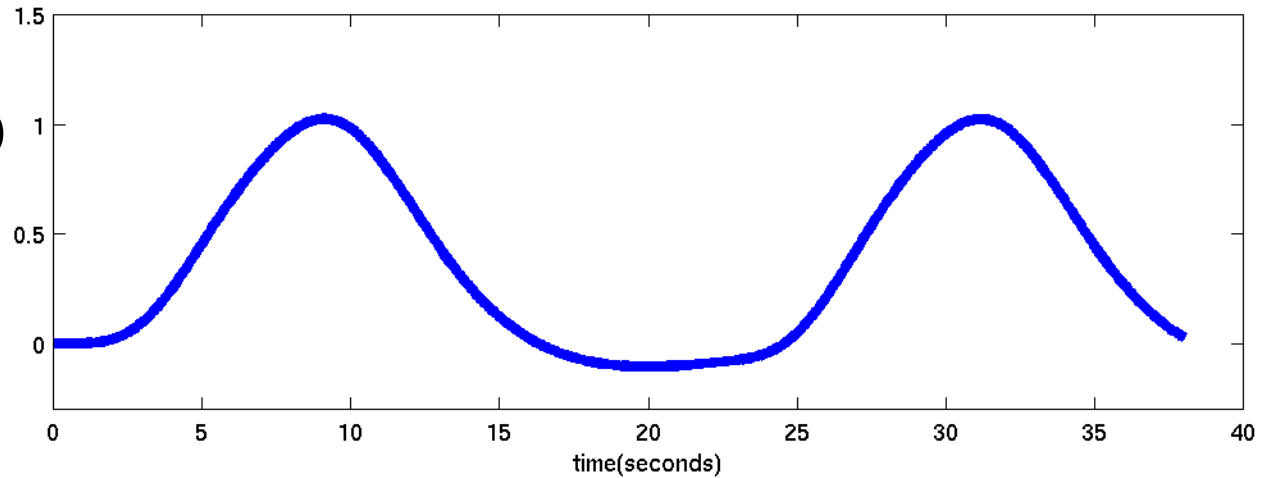
---

- Useful for studying auditory processes without scanner noise by presenting auditory stimuli during silence
- Also for allowing subjects to speak in the scanner without introducing further distortions in the image
- Acquire one volume at peak BOLD response
- Wait until BOLD evoked by scanner noise returns to baseline levels

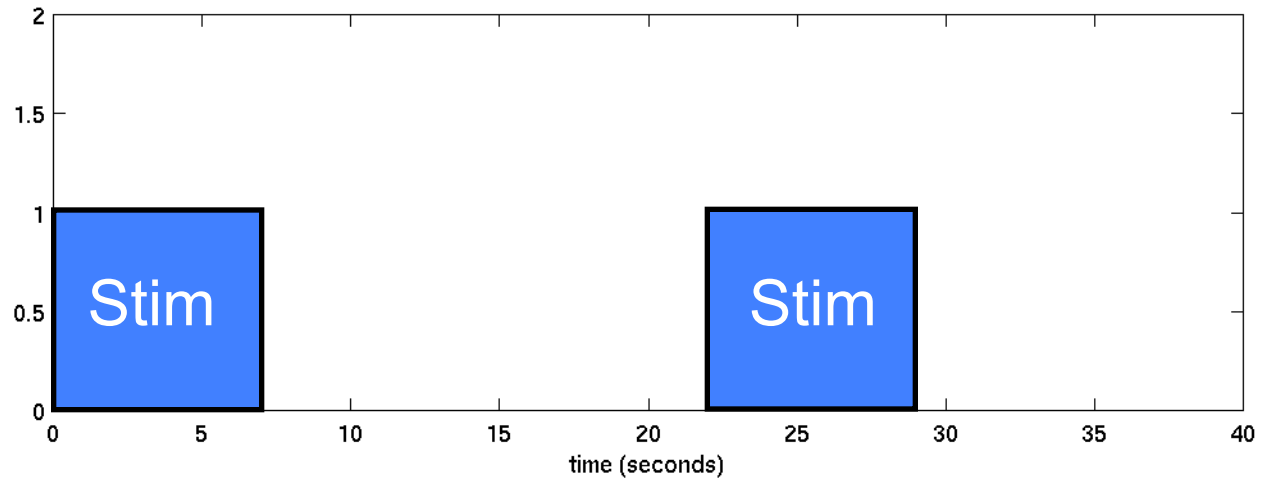


# Sparse sampling

Evoked BOLD response



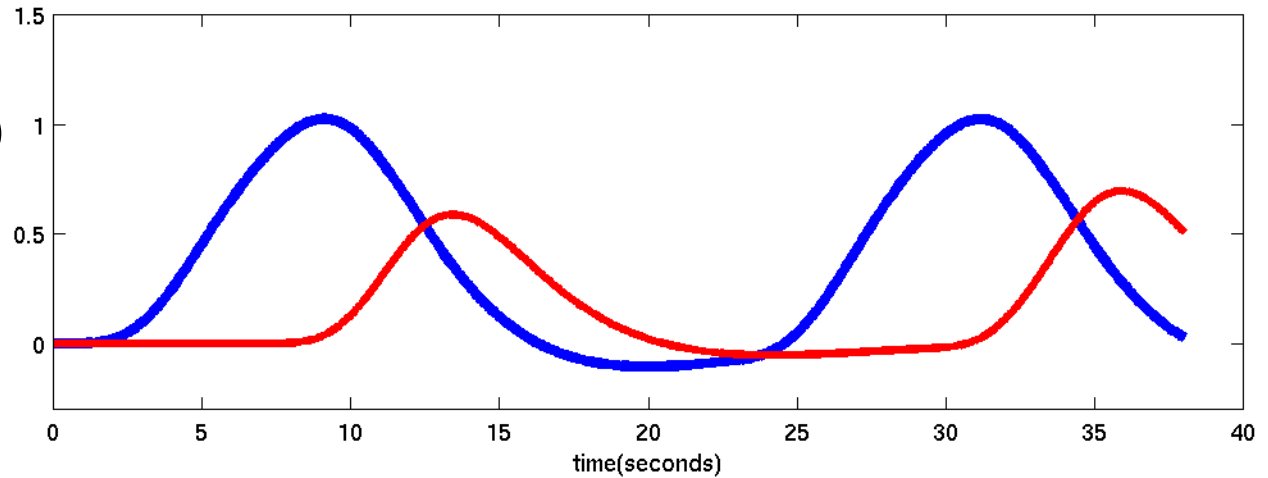
Stimulus paradigm



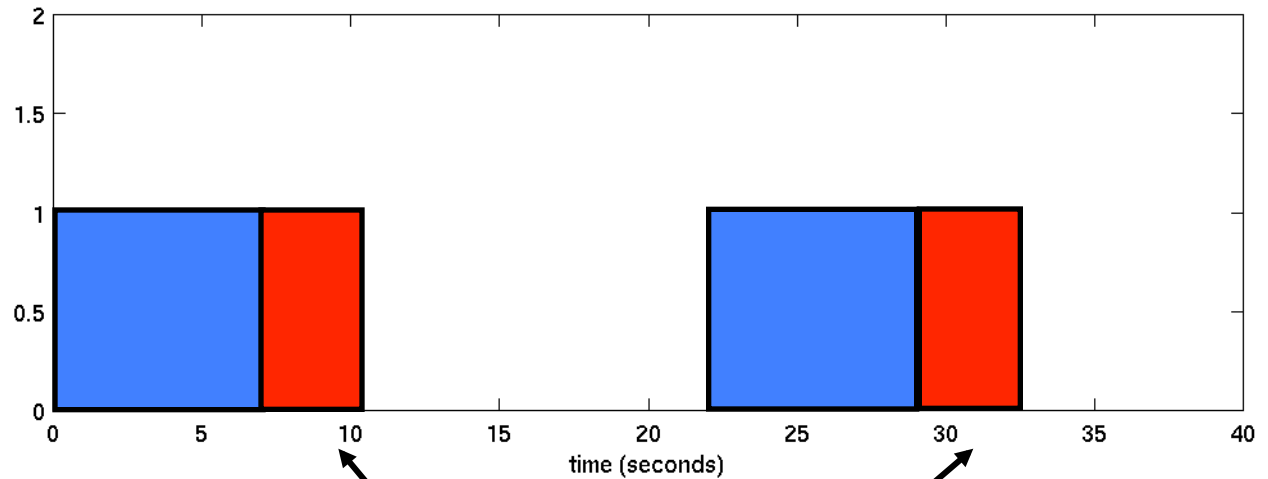


# Sparse sampling

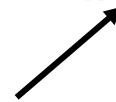
Evoked BOLD response



Stimulus paradigm



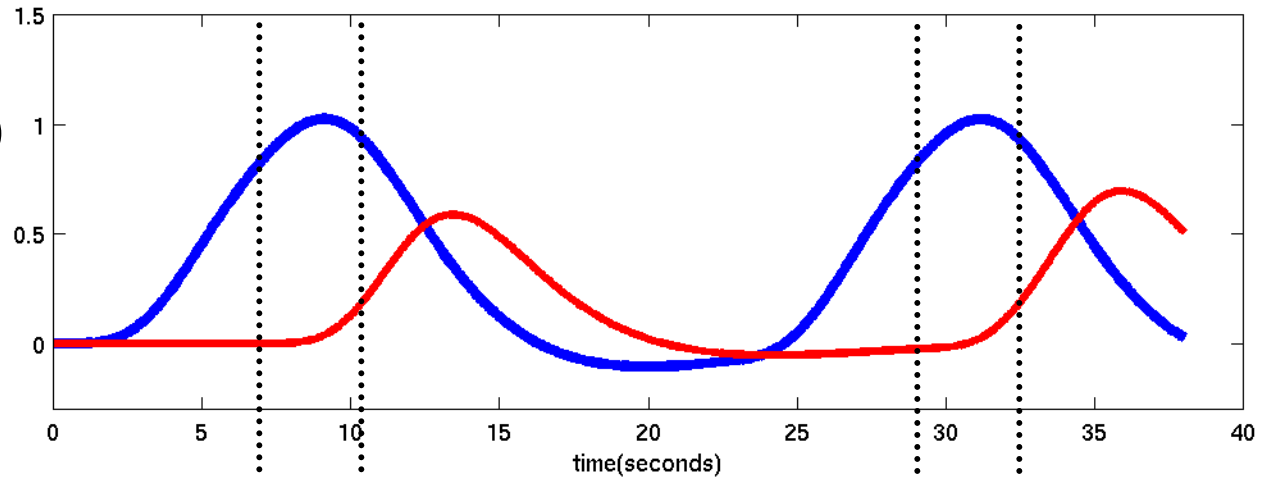
Scanner noise



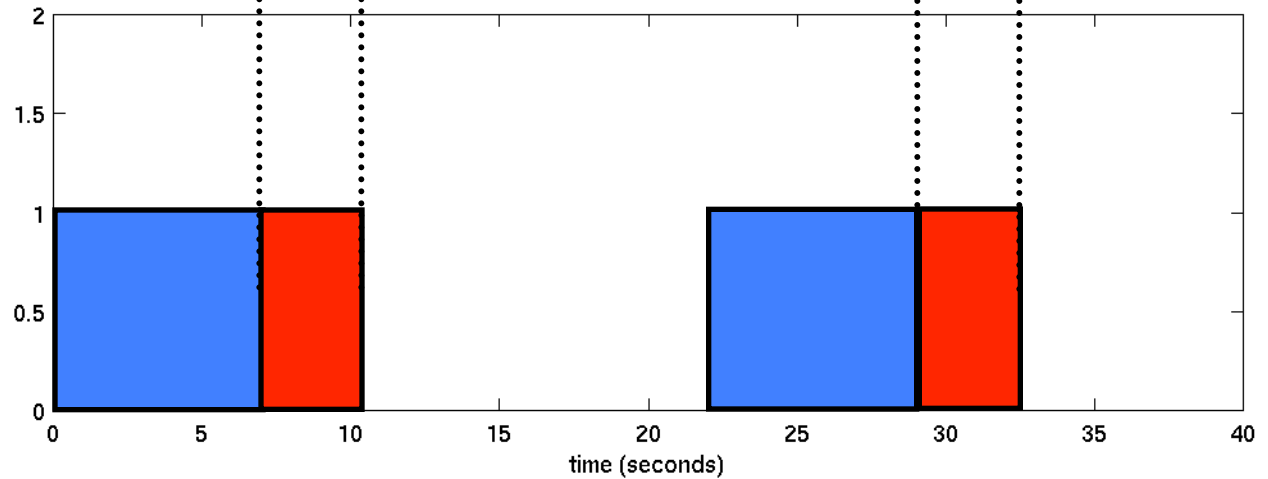


# Sparse sampling

Evoked BOLD response



Stimulus paradigm



Requires knowledge of HRF timing

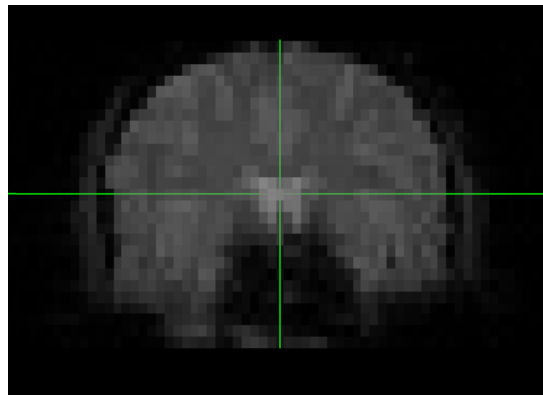




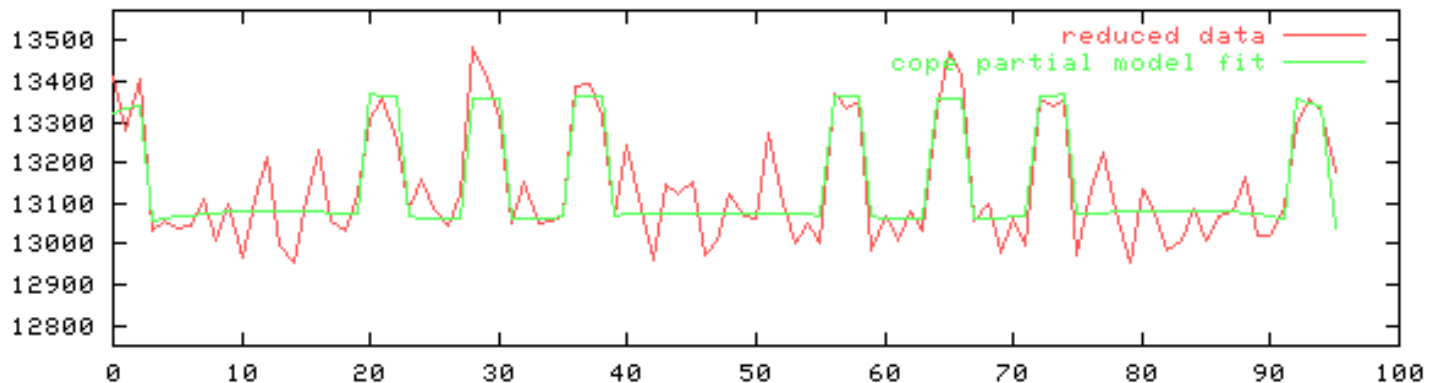
# Sparse sampling: analysis

- a single volume is collected every X seconds

Volume 2



- Not a time series (more like PET data)





# Outline

---

- Other things to think about
  - between subject covariates (e.g. anxiety, age)
  - controlling for (unwanted) effects of arousal, task difficulty etc.
  - counterbalancing
- Recent advances
  - Examining representational similarity
    - Adaptation, MVPA, calculating voxel 'tuning-curves'



# Outline

---

- Other things to think about
  - **between subject covariates (e.g. anxiety, age)**
  - controlling for (unwanted) effects of arousal, task difficulty etc.
  - counterbalancing
- Recent advances
  - Examining representational similarity
    - Adaptation, MVPA, calculating voxel 'tuning-curves'



# Back to our antisaccade example

---

- ***how do we examine whether trait anxiety modulates preparatory activity?***
- Solution: enter trait anxiety as a between subject covariate
- Generally entering such measures as continuous covariates is more powerful than using a median split to create 'low' and 'high' anxious groups where possible
- *Eugene will talk more about how to do this using FEAT tomorrow*



# Important considerations

---

- Is your subject doing what you think they are doing? Can they do the task?
  - Consider practice session
- Match conditions for difficulty, motor demands etc.
- Collect behavioural data
  - Can use for post-hoc sorting of data
  - Correlation with fMRI signal
- Consider collecting physiological data



# Important considerations

---

- Example (Bishop et al., 2008)
- COMT genotype influences on frontal activity during performance of fluid reasoning task.
- Design: 'high 'g' blocks, low 'g' blocks based on prior task design by Duncan and colleagues.
- **Issues:**
- *High 'g' items take longer* -you don't want only 1 high 'g' vs 1 low 'g' item – the latter will be done much faster, so bound to generate less activity ->(do 5 low 'g' for each 1 high 'g')? ->
- *Some subjects may give up* on high 'g' items – (will see activity for 'giving up' /'anxious rest' not for doing high 'g' task)
- **Solution:** have blocks of set length, not set number of items, if get stuck can time out after x sec)
- **Remaining issue:** will be more motor responses for low g items (get through more in a block) than high 'g' items



# Important considerations

---

- Example (Bishop et al., 2008)
- COMT genotype influences on frontal activity during performance of fluid reasoning task.
- Design: 'high 'g' blocks, low 'g' blocks based on prior task design by Duncan and colleagues.
- COMT val carriers showed more frontal activity for high g > low g contrast.

## ***Value of collecting behavioural data:***

Can see if more frontal activity for high g > low g linked to better performance and if it mediates relationship between COMT genotype and performance ...



# Important considerations

---

## ***Value of collecting physiological data:***

*e.g. passive viewing of highly emotional vs low emotional stimuli*

- ***pulse/ respiration***
- *Can help ensure activity differences are not just due to peripheral arousal varying between conditions*
  
- ***Eyetracking***
- *Can make sure subject staying awake*
  
- ***Pupillometry***
- *Pupil dilation tracks changes in stimulus contingencies – very valuable signal for decision making, fear conditioning tasks – also good measure for looking at individual differences*





# Important considerations

---

## Optimal counter-balancing

Consider the letter string task example

- Stimuli: congruent, incongruent , neutral
- Randomise for event-related presentation?
  
- I C N C N I N C N I I C subj 1
- C I N C I N N C I N C I subj 2
- C C N I I I N C N N C I subj 3



# Important considerations

---

- Cohen, Carter and colleagues (Carter et al., 2000) have shown increased conflict effects in ACC from incongruent trials which follow congruent trials and reduced conflict effects from incongruent trials which follow incongruent trials
- I C N C **N I** N C **N I I** C subj 1
- **C I** N **C I** N N **C I** N **C I** subj 2
- C C N **I I I** N C N C **N I** subj 3
- If a single group, may reduce power, need bigger  $n$ .
- If two groups, or covariate (e.g. anxiety) could be disastrous (e.g. if all low anx. happen to get the pattern for subjs 1,3, all high anx. that for subj 2)
- If using group-level covariates (or gps) may want to pre-(pseudo) randomise then keep constant across subjects



# Question checklist before you start

---

- What is your question?
- What is the best way to evaluate it?
  - Subtraction, parametric, factorial, conjunction, adaptation; group level covariate?
  - Blocked, event-related, mixed?
- Movement –is your paradigm extra likely to have this? (e.g. administration of shocks)
  - (*think about training in a mock scanner*)
- How long is your experiment?
  - How many blocks/events needed?
- TR? (oversample, jitter)
- Acquisition: Whole brain? Resolution?
  - Extra few scans to help registration?
- How many subjects?
  - Collect behavioural responses?
  - Collect physiological responses?
  - Counter-balancing
  - Make sure your subjects know what they are doing



# Outline

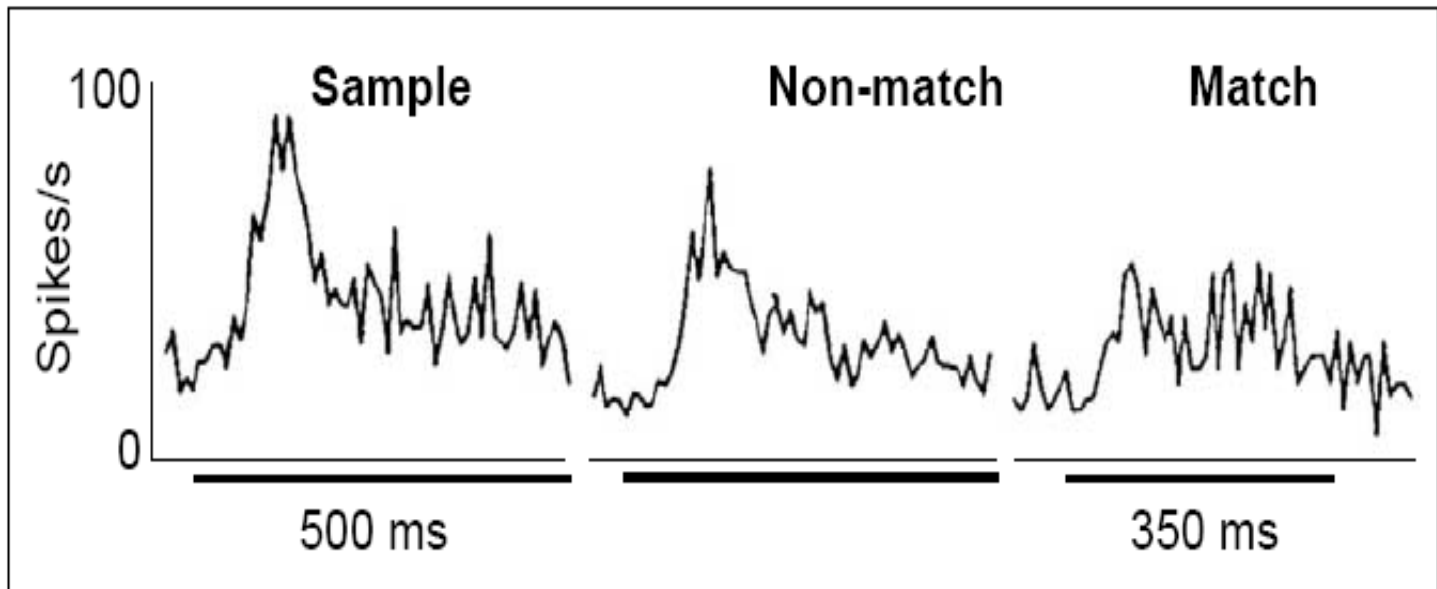
---

- Other things to think about
  - *between subject covariates (e.g. anxiety, age)*
  - controlling for (unwanted) effects of arousal, task difficulty etc.
  - counterbalancing
- Recent advances
  - Examining representational similarity
    - Adaptation, MVPA, calculating voxel 'tuning-curves'



# Recent(ish) advances: fMRI adaptation

- Based on neural repetition suppression



Averaged neuronal response (spike/s)

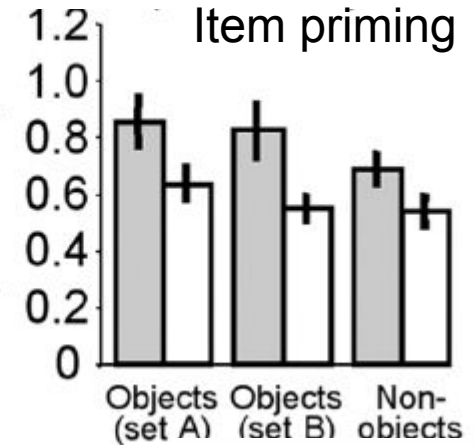
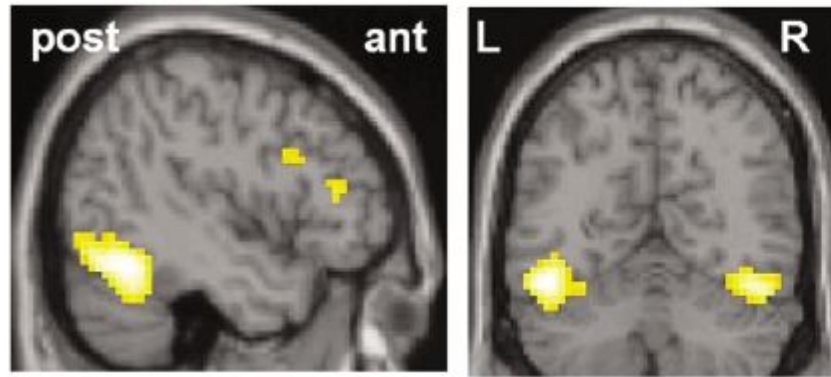
Desimone (1996)

# Repetition Priming

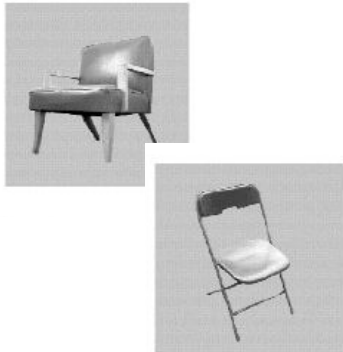
Same picture



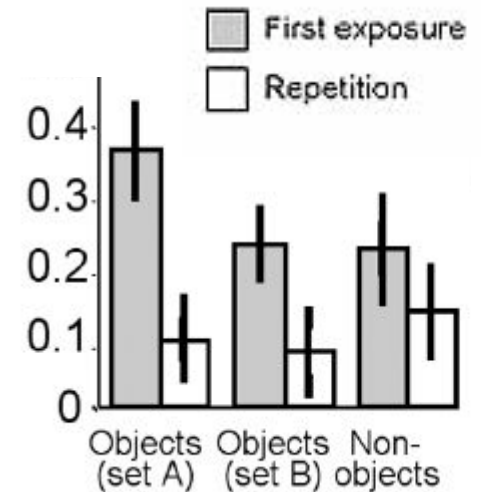
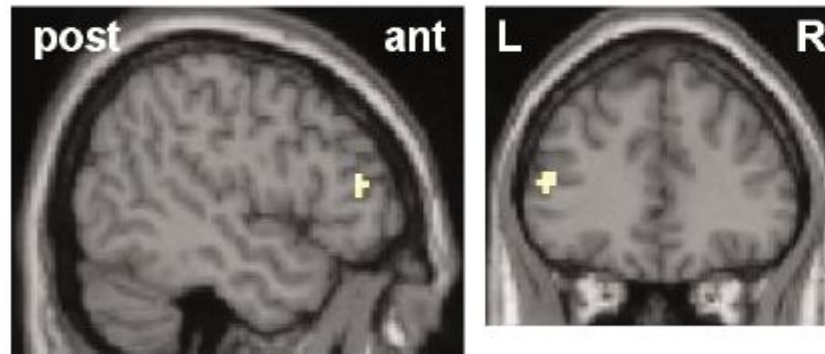
Posterior fusiform gyrus



Same category  
Different picture

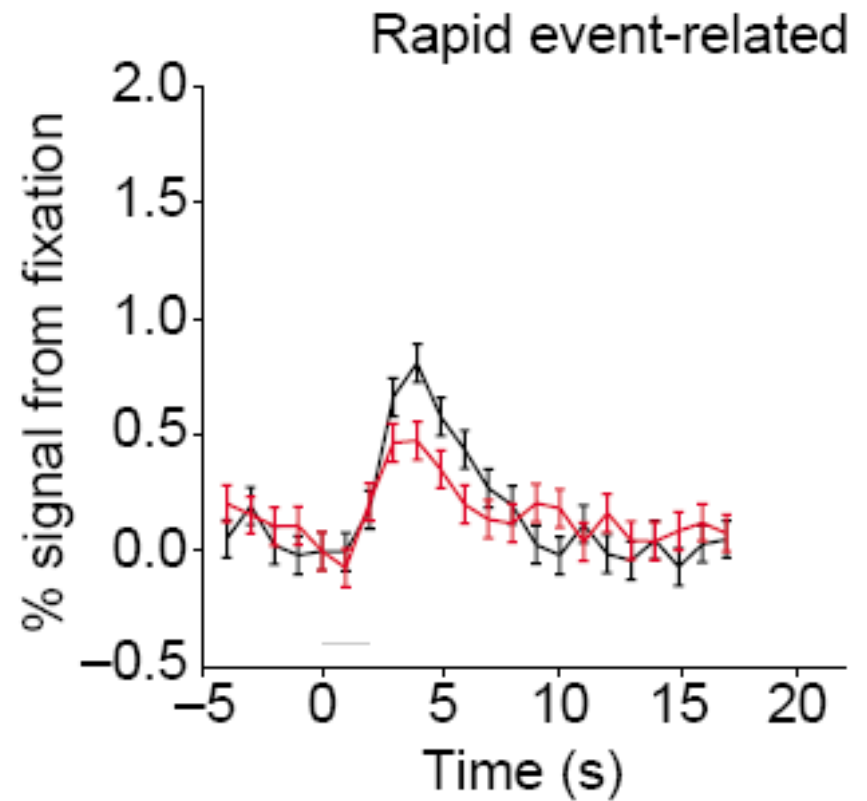
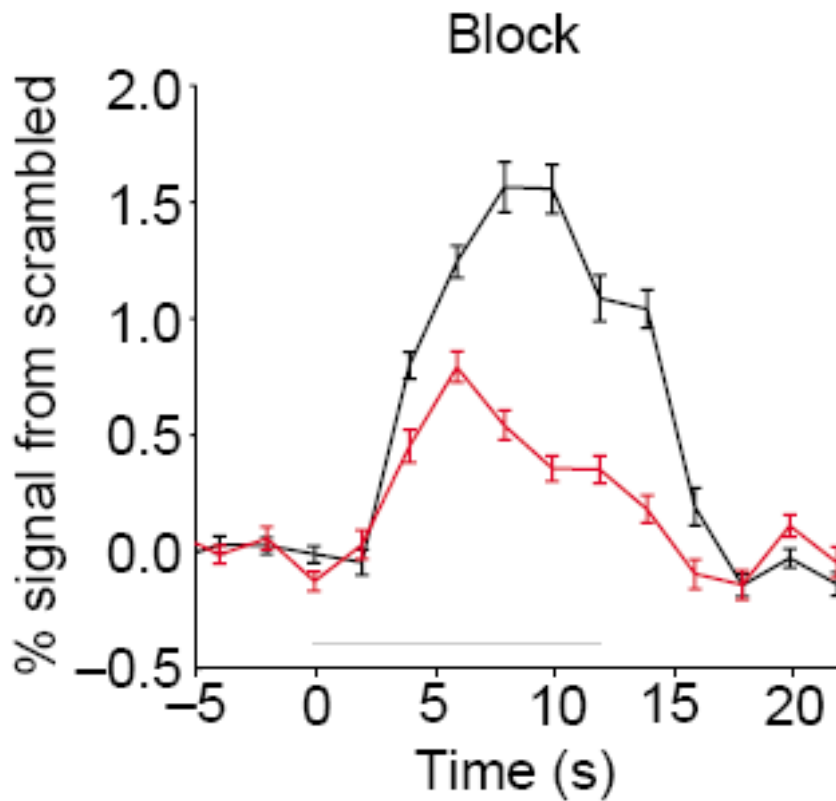


Left inferior frontal cortex



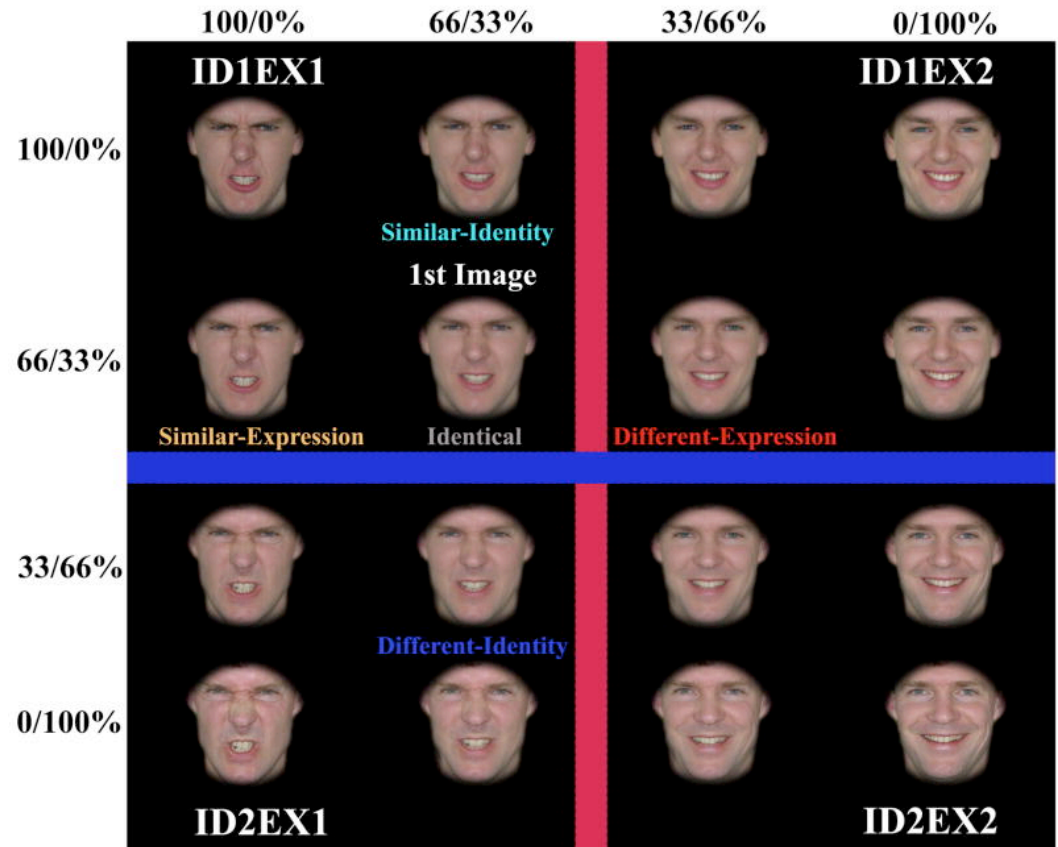


# BOLD decrease for repetition



Each trial = pair of faces

2<sup>nd</sup> face can involve change that does/ does not cross 50/50 boundary for either (i) expression or (ii) identity



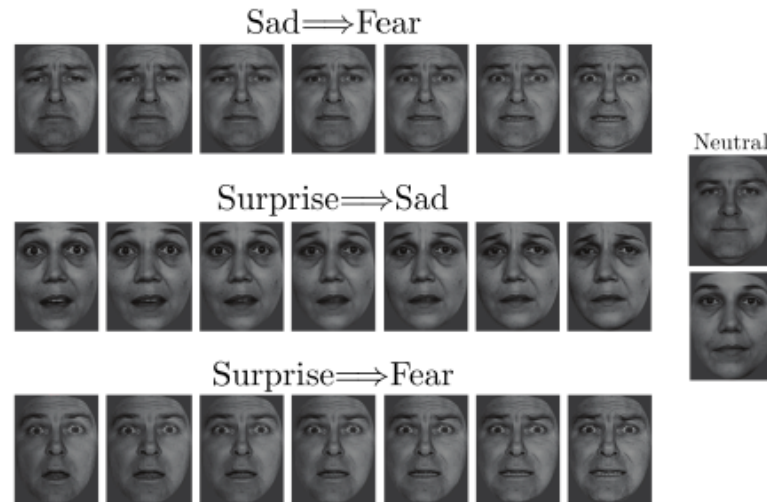


As described by Aguirre (2007)

Stimuli vary along given dimension

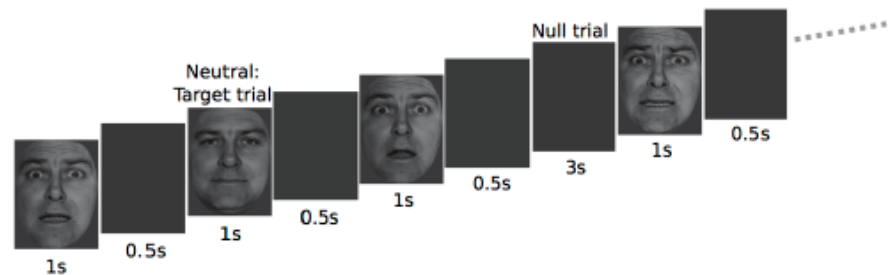
Can look at adaptation between sequential stimuli as function of steps along dimension

## A Stimuli



## B fMRI task:

Press when you see a neutral face



unpub data.

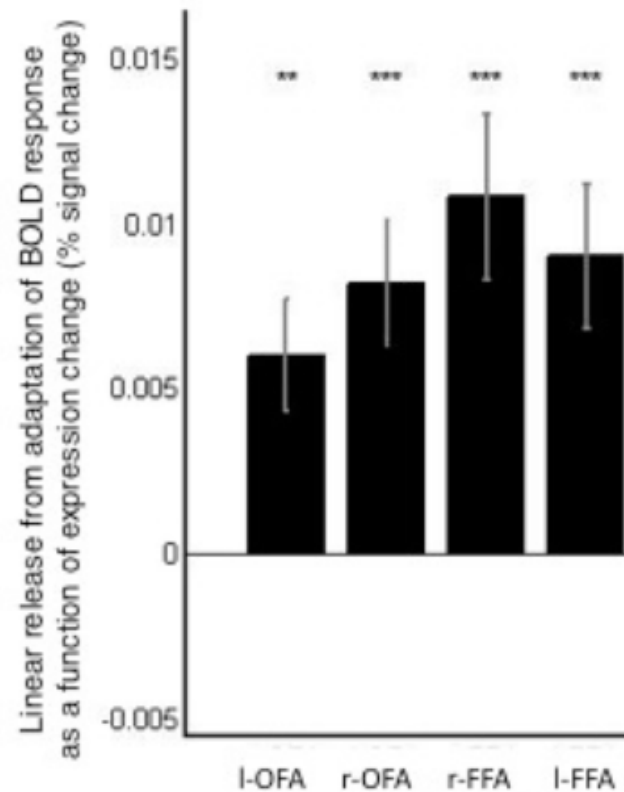


# Continuous carry-over adaptation design

As described by  
Aguirre (2007)

Stimuli vary along  
given dimension

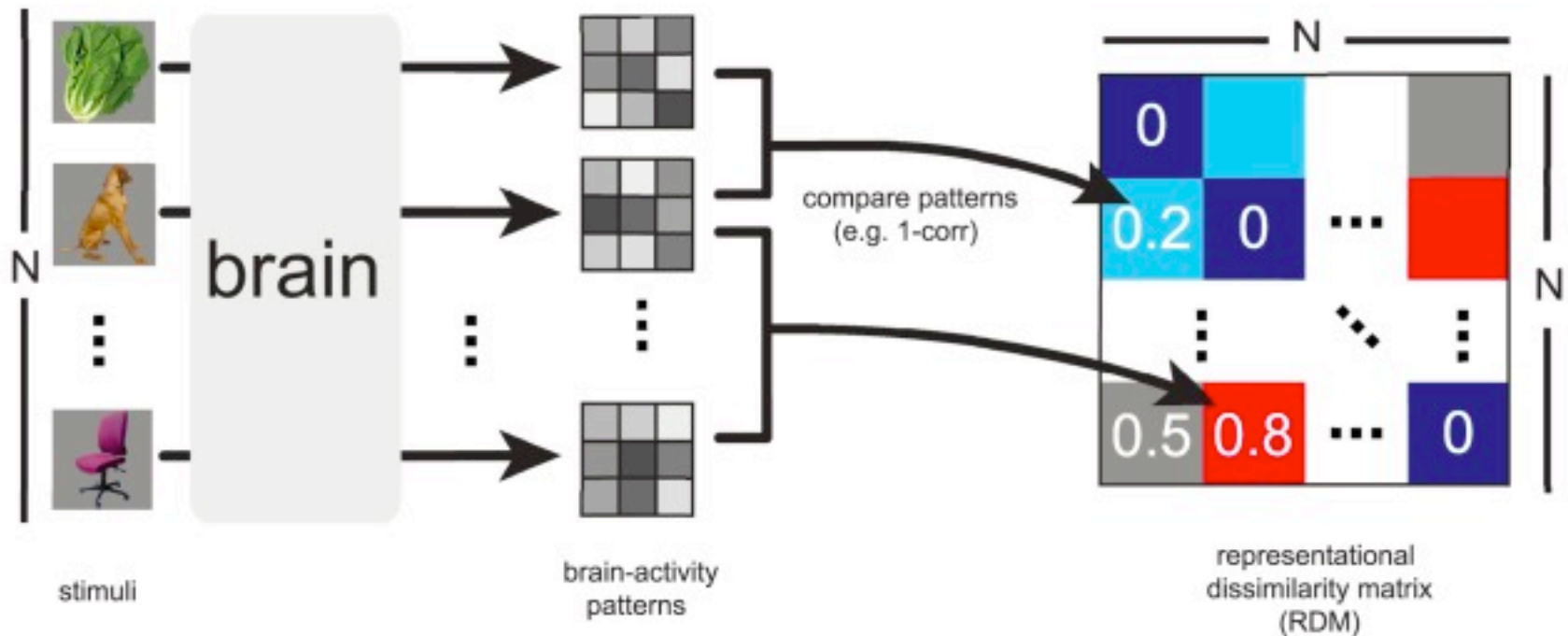
Can look at adaptation  
between sequential  
stimuli as function of  
steps along dimension



unpub data.

# Multi voxel pattern analysis

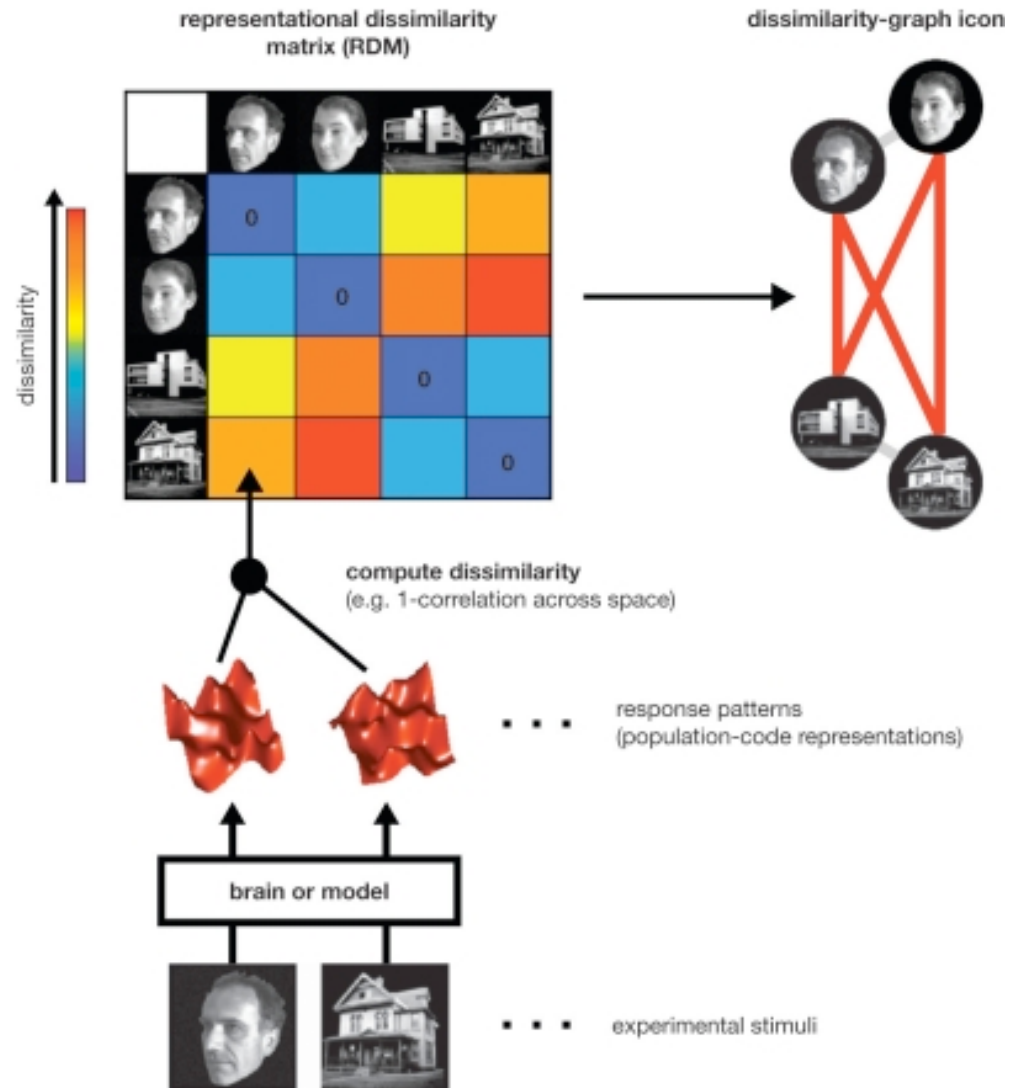
Allows examination of whether distributed pattern of activity differs between different classes of stimuli, even if univariate response across region is the same





# Multi voxel pattern analysis

Can use it to get measure of 'representational similarity' of pattern across voxels for different stimuli

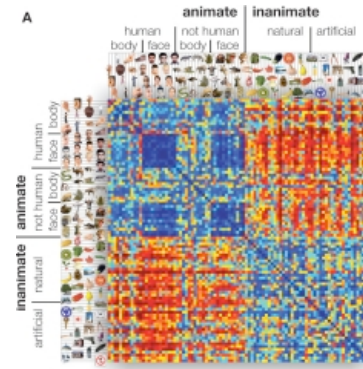


# Multi voxel pattern analysis

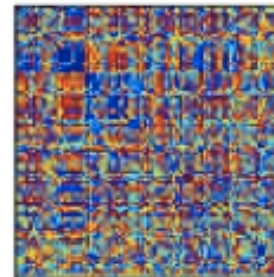
Can use it to get measure of 'representational similarity' of pattern across voxels for different stimuli

*and see if this tracks differences in low or high level physical features of the stimuli*

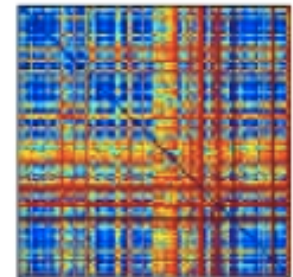
brain data



**luminance image (low-pass)**

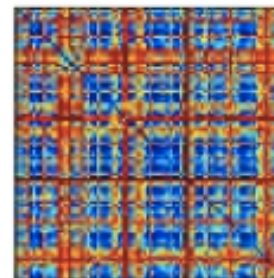


**stimulus image (Lab)**

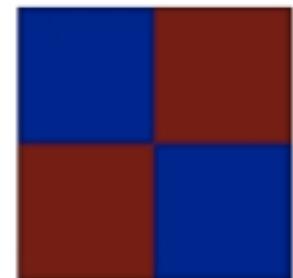


models

**RADON (smoothed)**



**animate-inanimate**

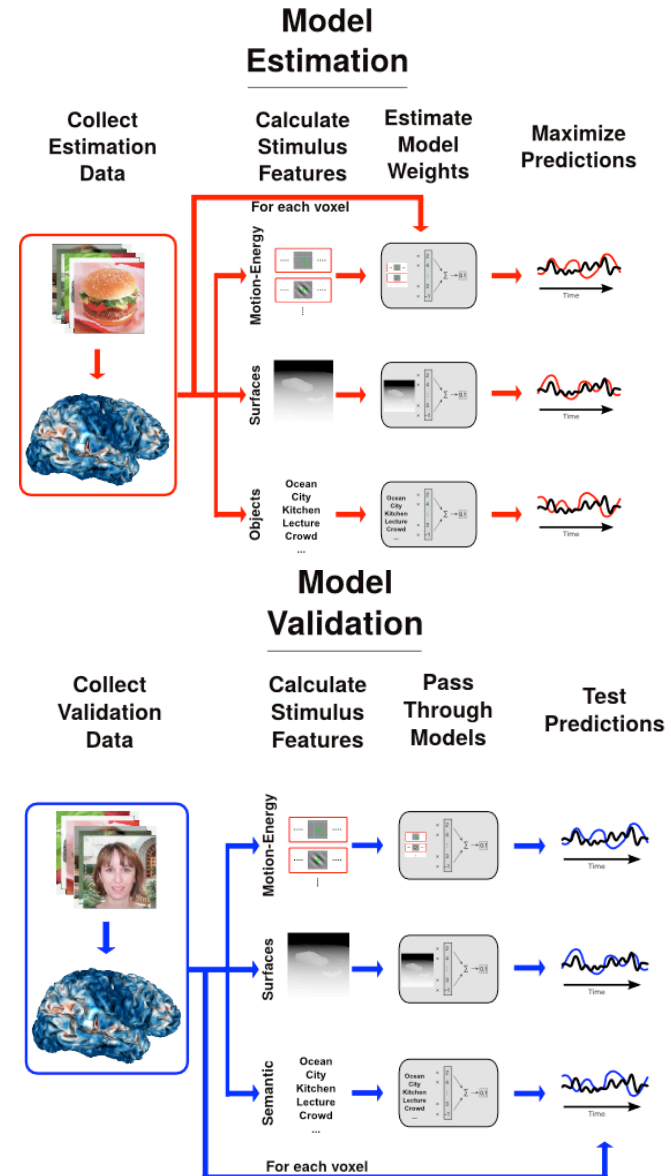




# Voxel 'tuning' curves

But .. univariate models do not need to be limited to simple contrasts using smoothed data

can model voxel 'tuning' curves -> compare different models to see what types of 'features' individual voxels are most responsive to.



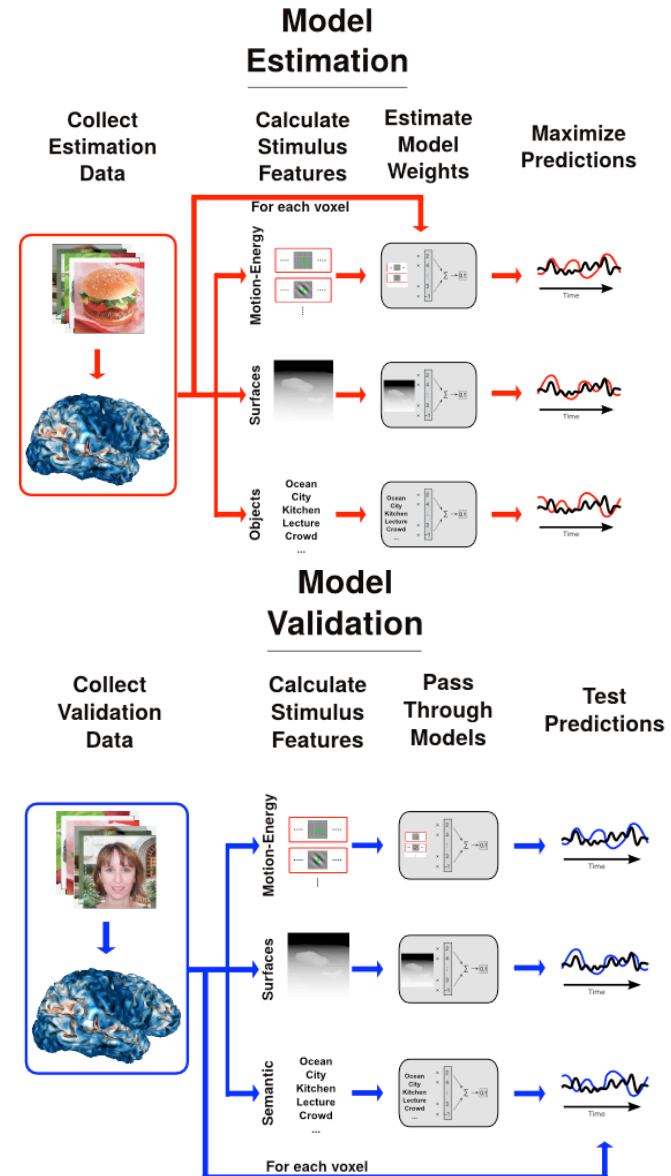


# Voxel 'tuning' curves

Fits for different models are compared using permutation testing.

Each model will have different beta weights for each 'feature' e.g. can see the category to which a voxel sensitive to semantics responds the most

<http://gallantlab.org>







# Experimental Design

---

*That's all folks ...*

With thanks to:  
Kate Watkins  
Heidi Johansen-Berg  
Joe Devlin