Donders Institute
for Brain, Cognition and Behaviour

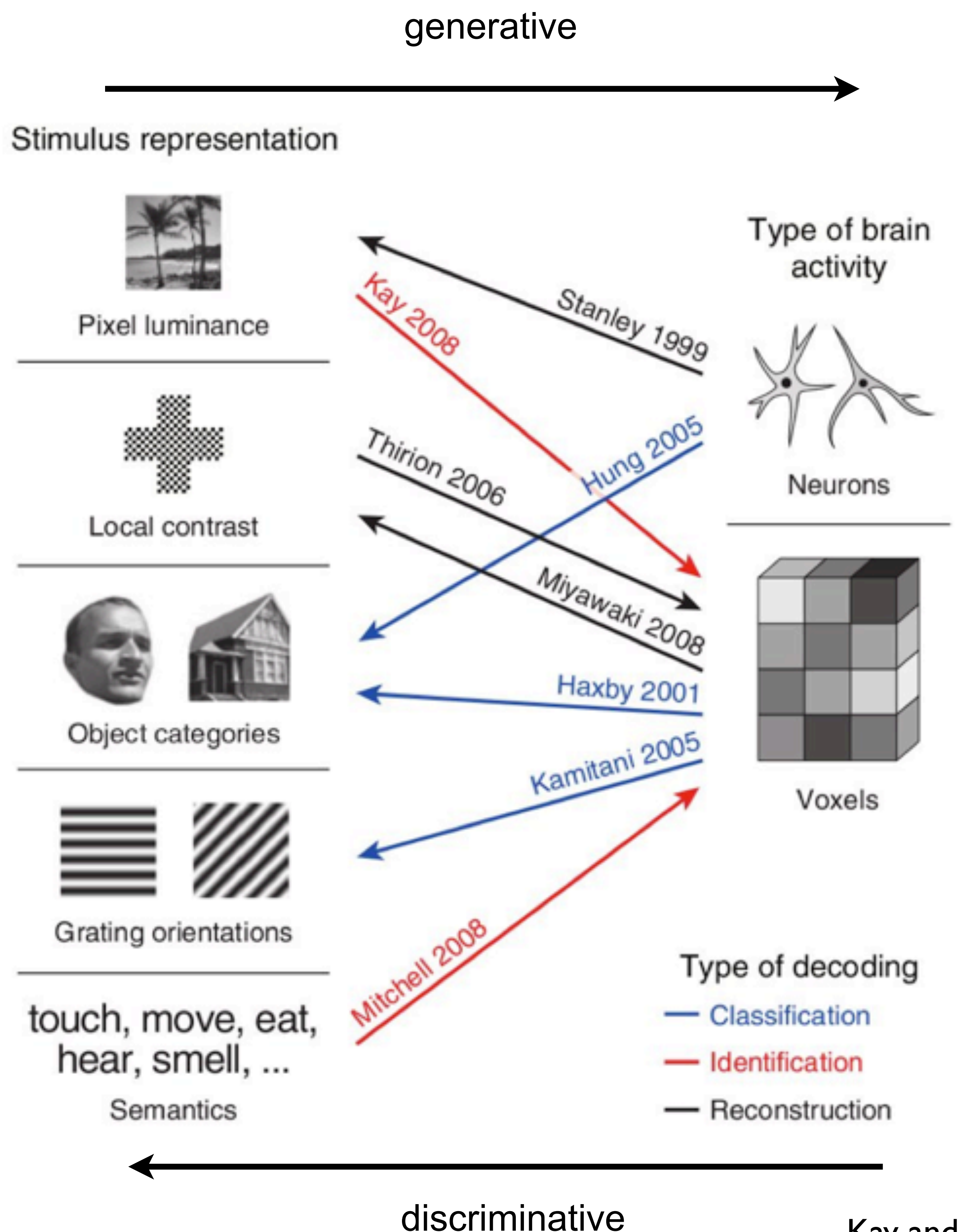
Brain Reading (MKI43)

Lecture 2: Discriminative decoding

Marcel van Gerven
Assistant Professor
Distributed Representations Group
Donders Centre for Cognition

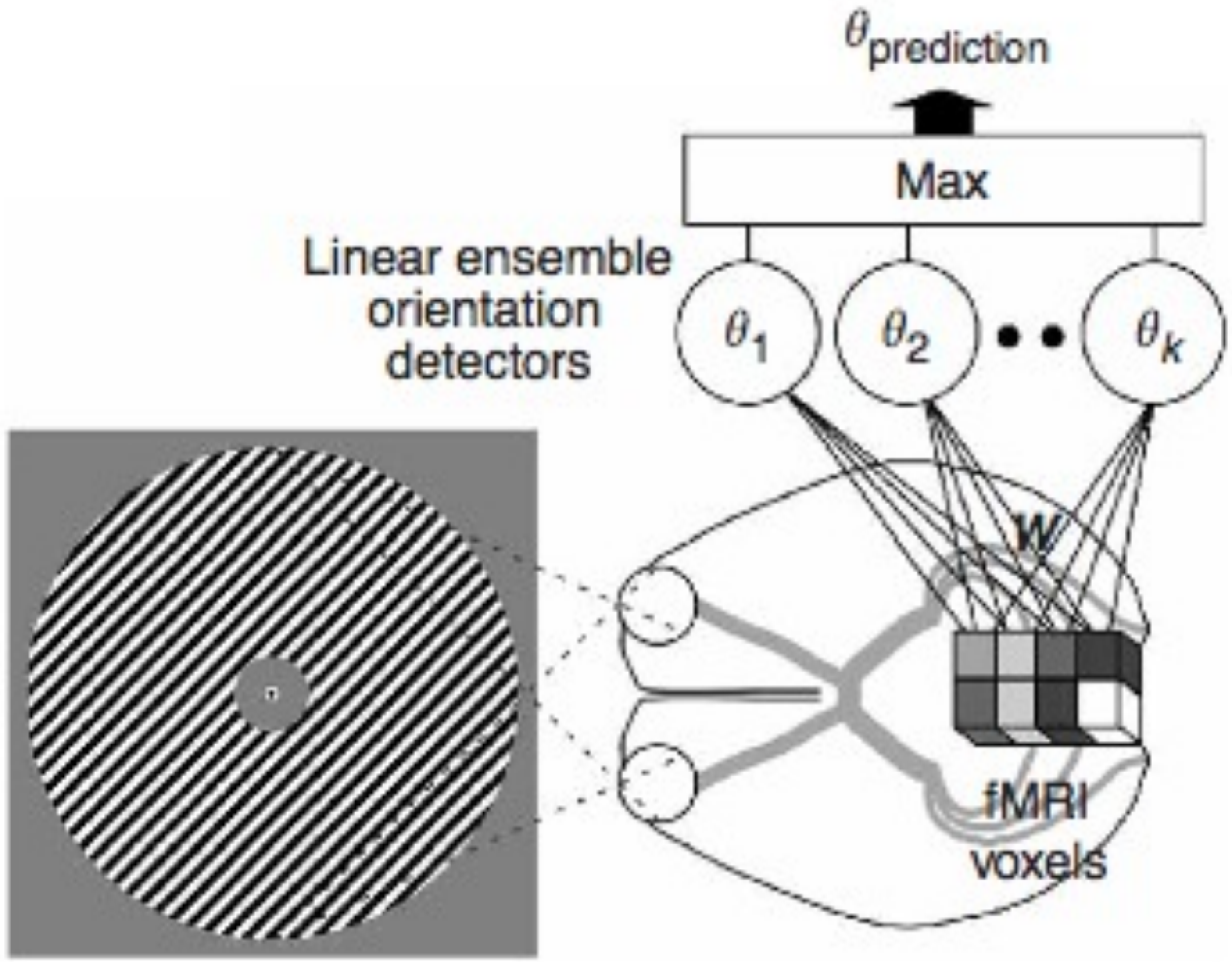
Radboud University Nijmegen

Decoding flavors



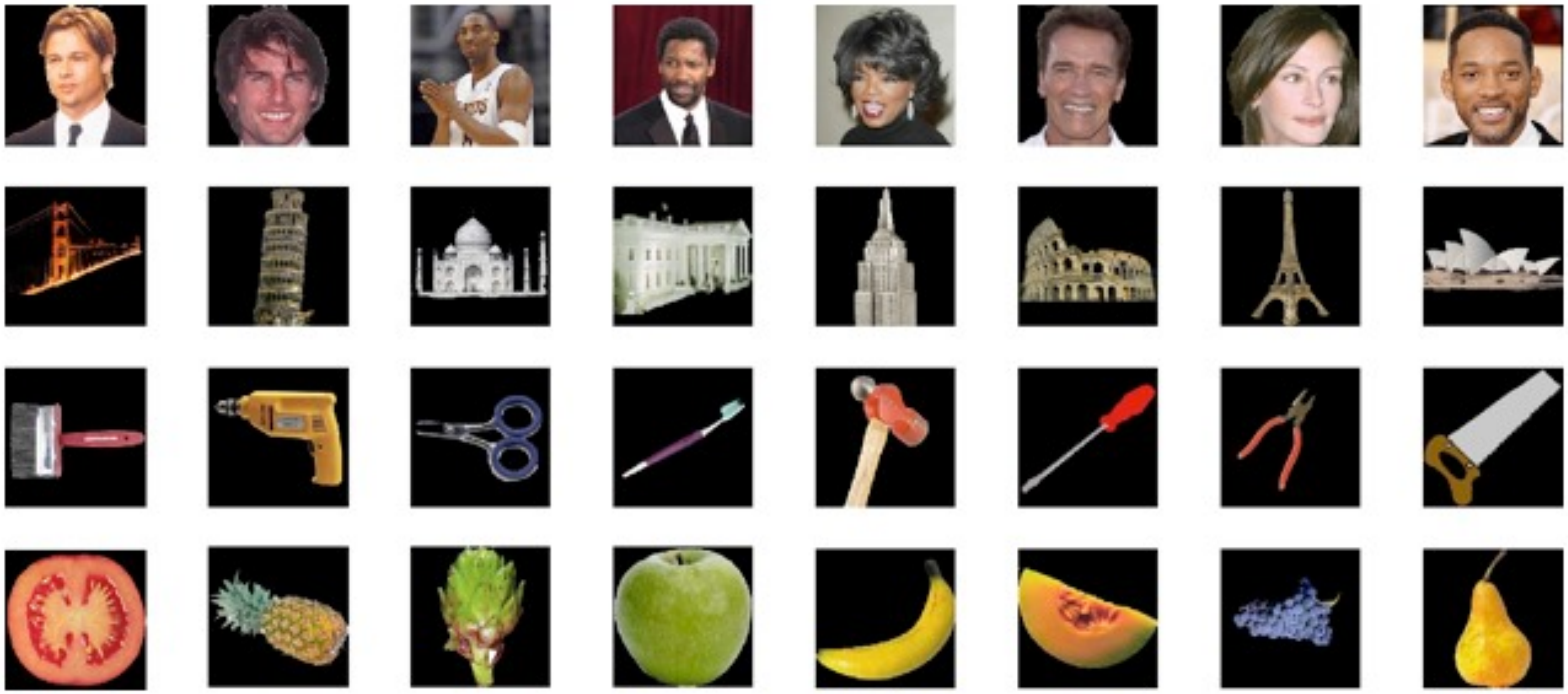
Kay and Gallant, Nature, 2009

Orientation decoding



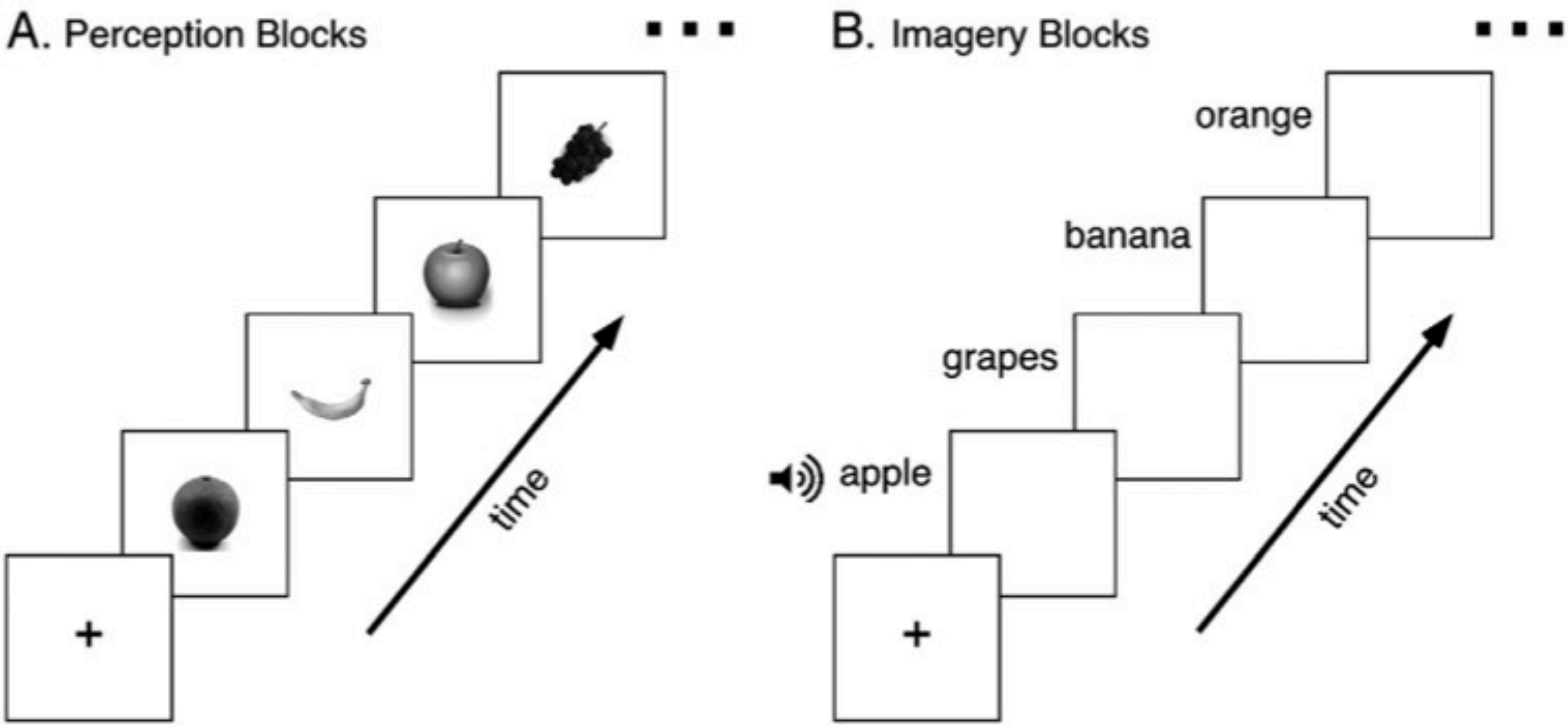
Kamitani and Tong, Nature Neuroscience, 2005

Imagined category decoding

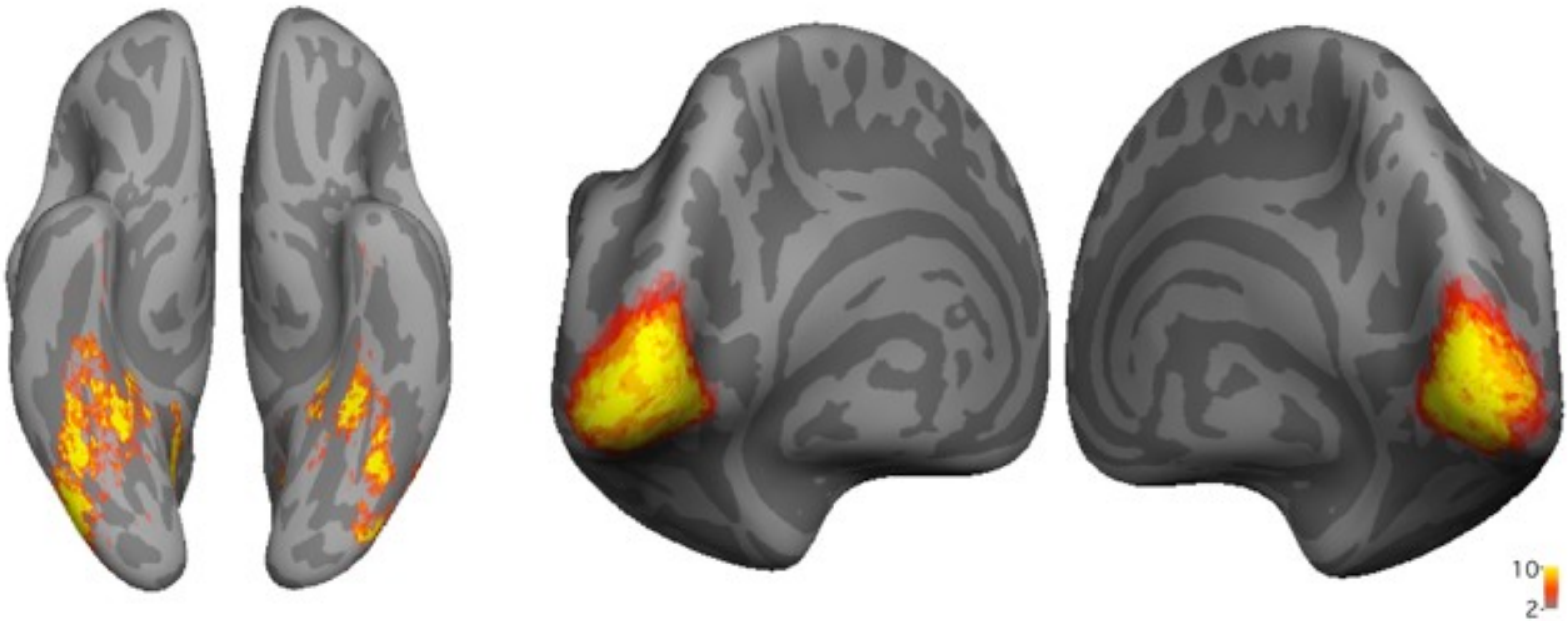


Reddy L, Tsuchiya N, Serre T. Reading the mind's eye: decoding category information during mental imagery. Neuroimage. 2010;50(2):818–825.

Imagined category decoding



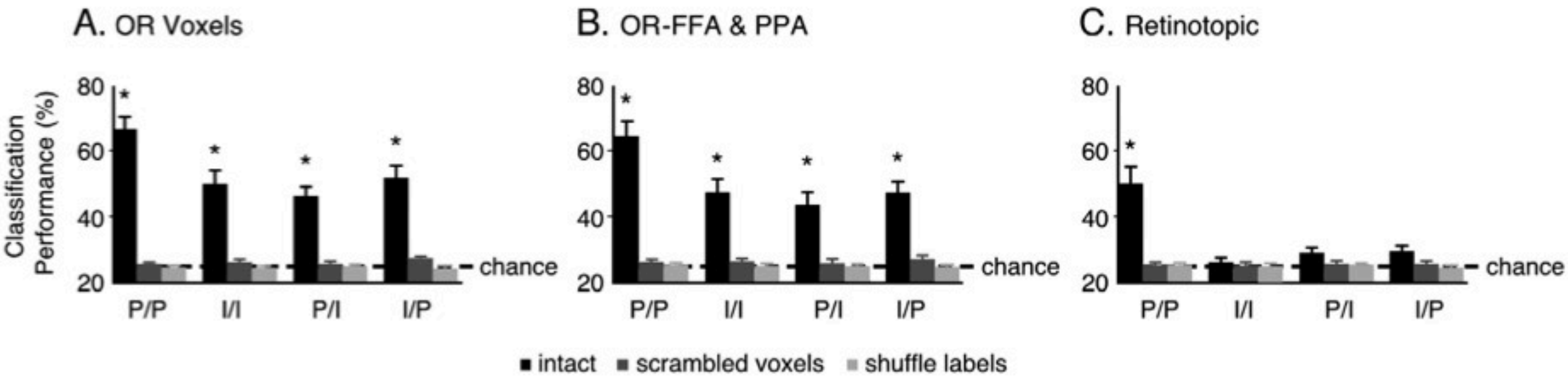
Imagined category decoding

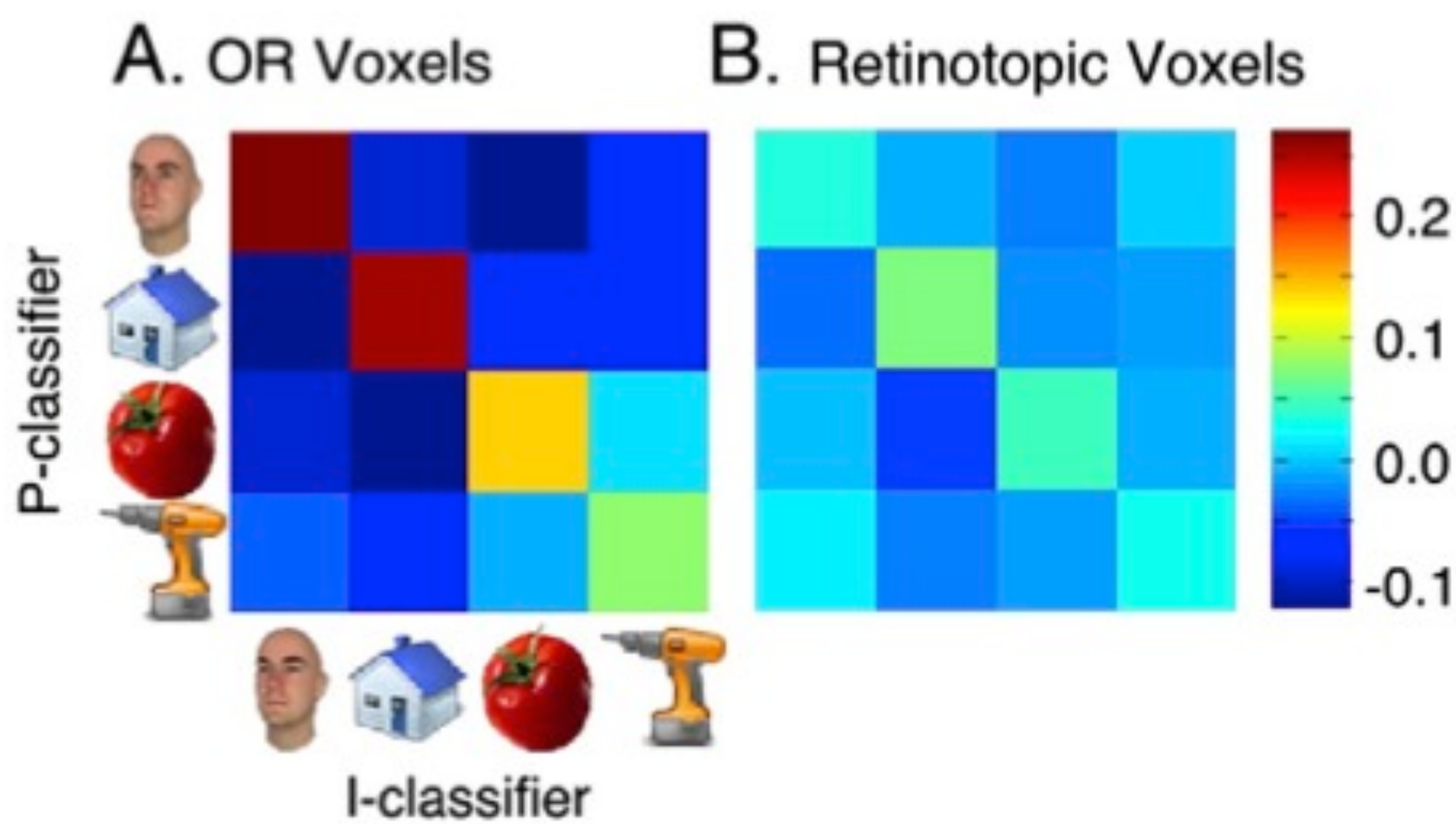


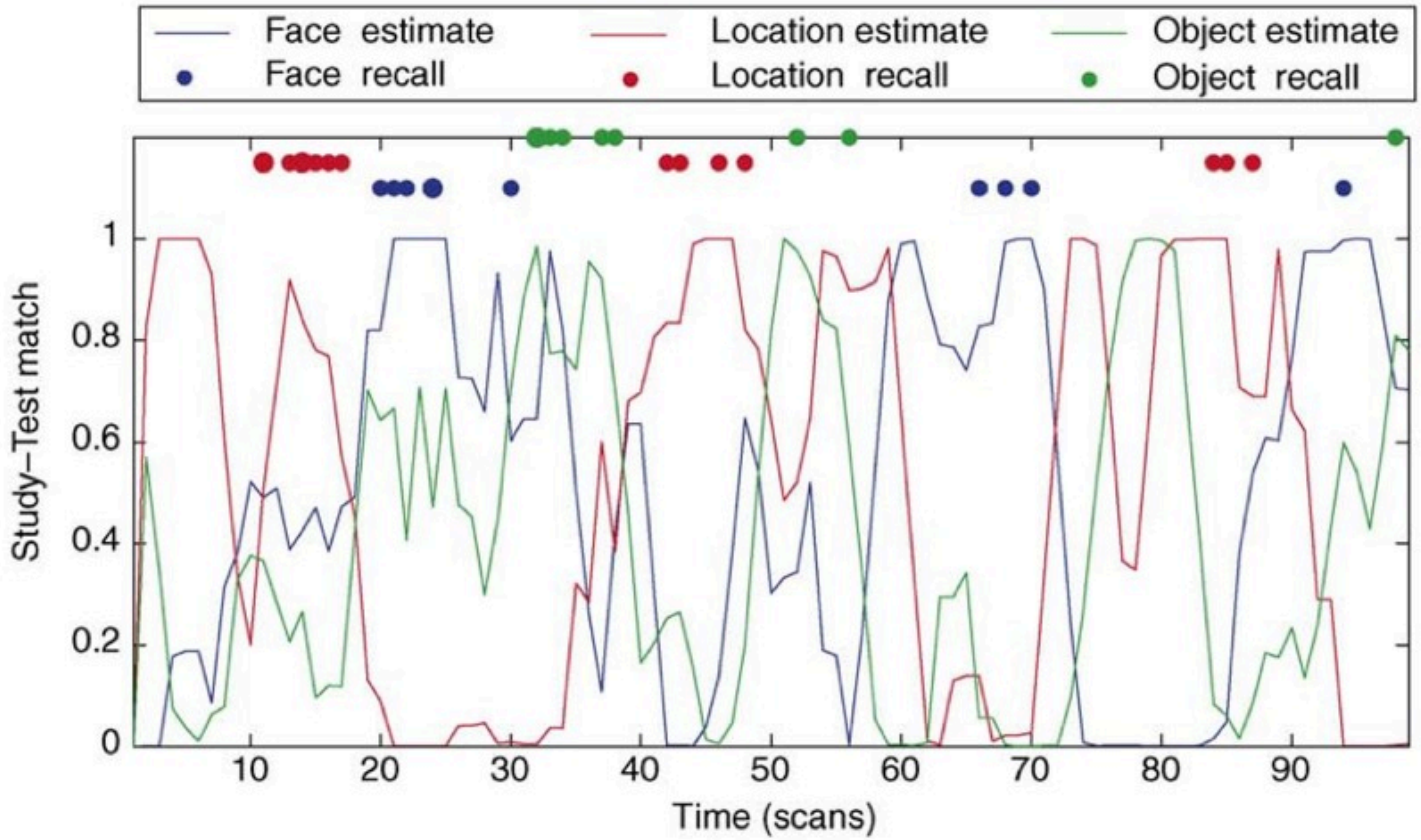
A. Average OR ROI

B. Average Retinotopic ROI

Imagined category decoding

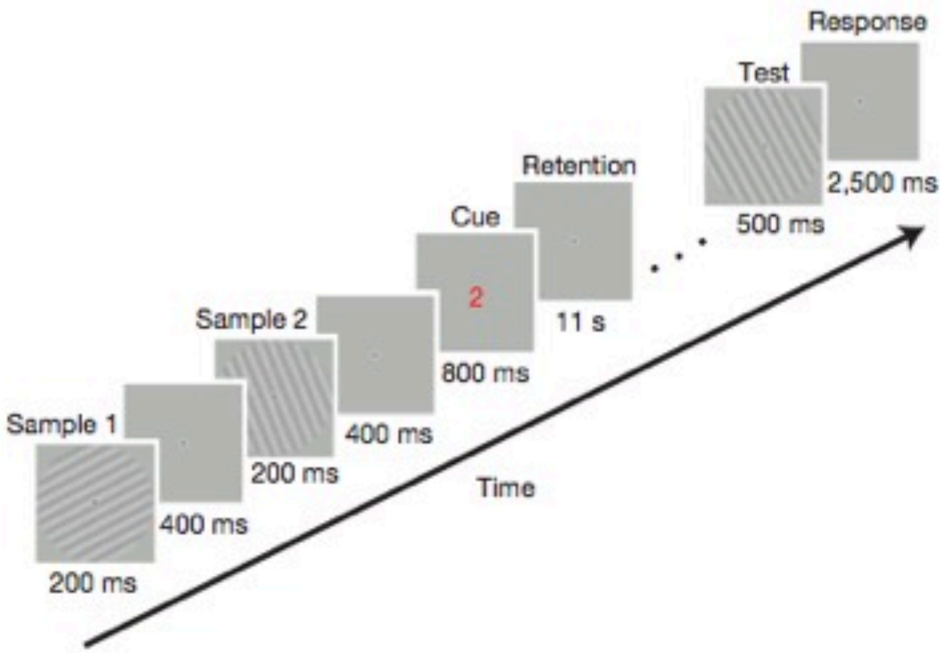


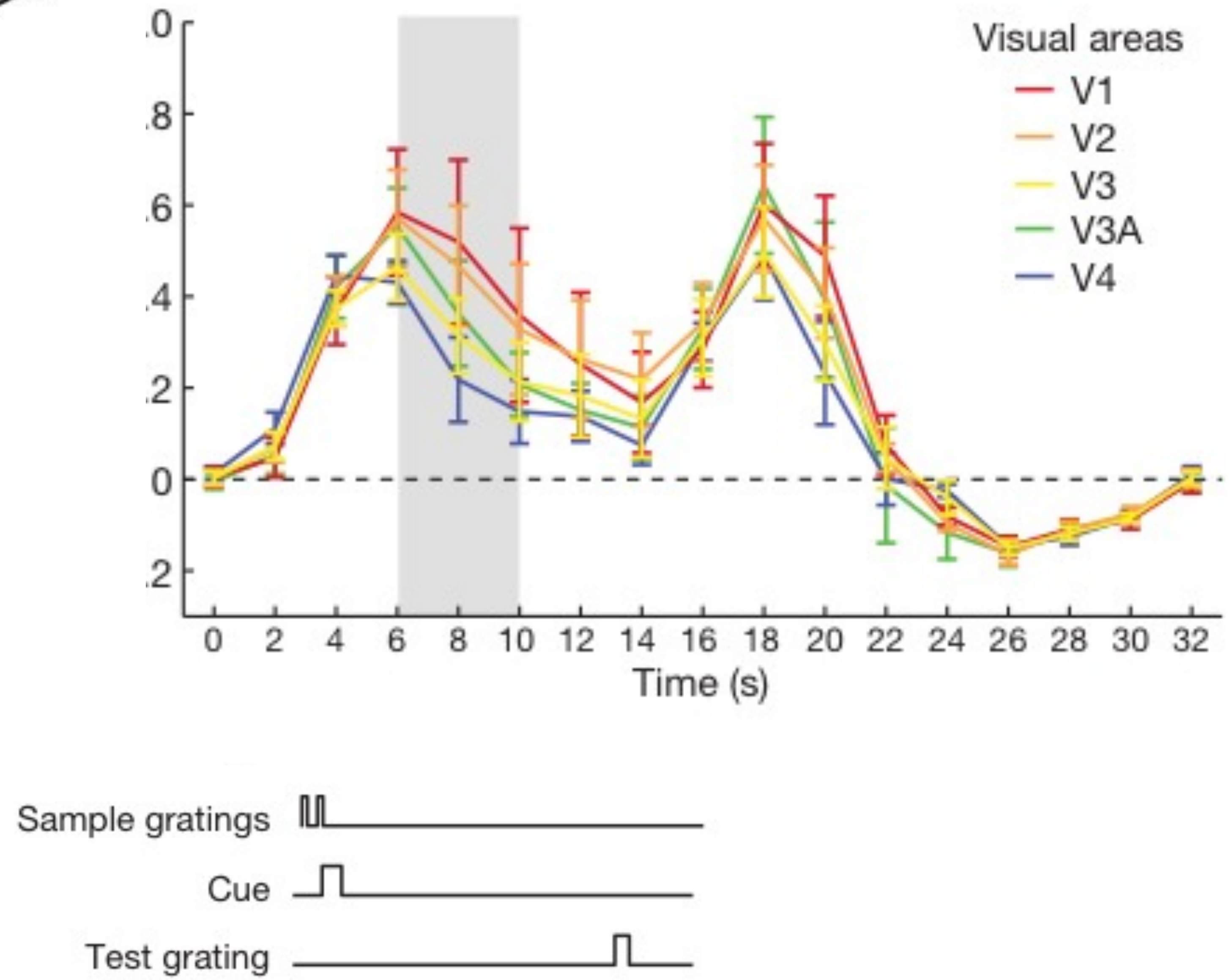




Norman et al. TICS, 2006

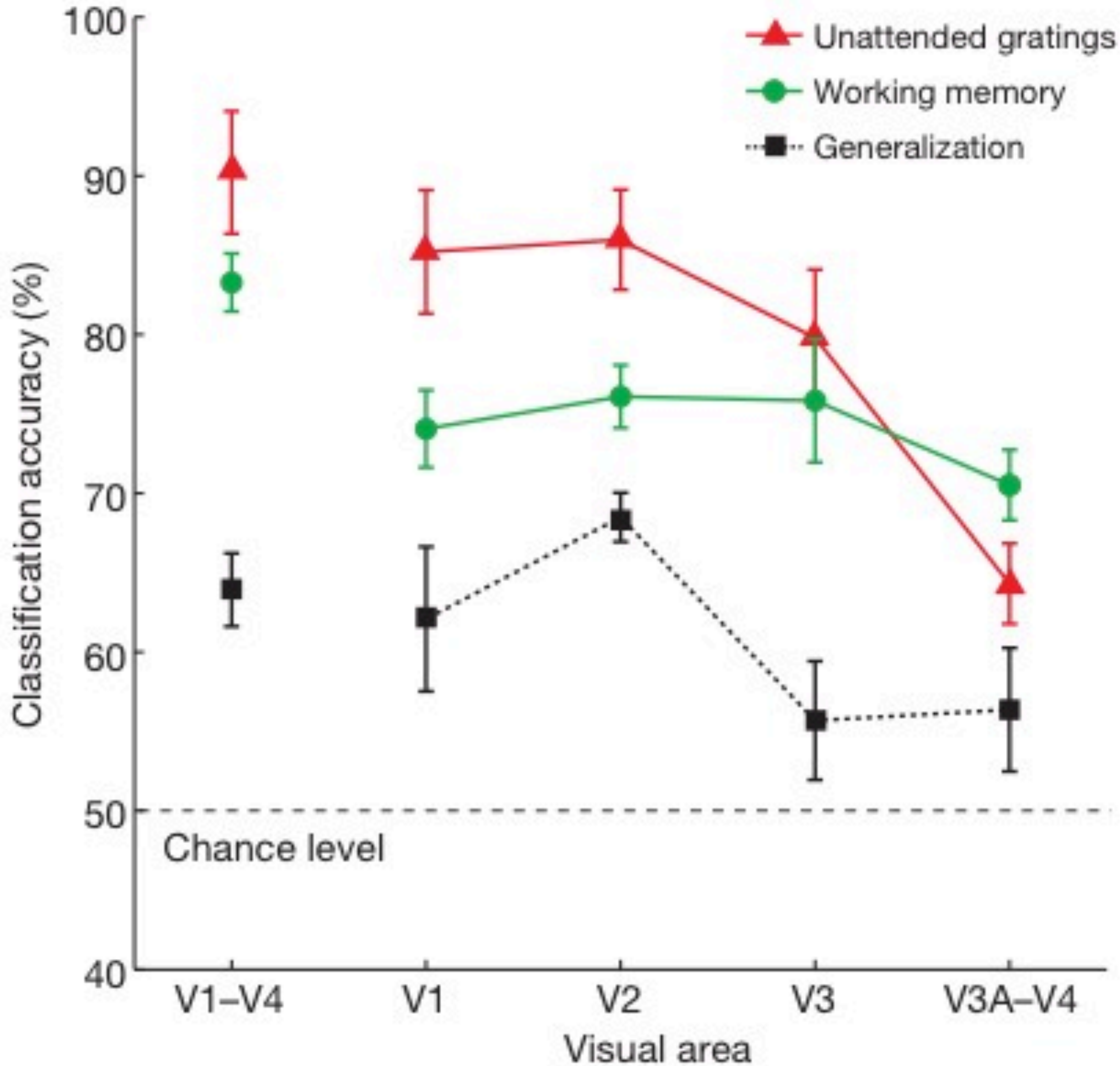
Decoding working memory



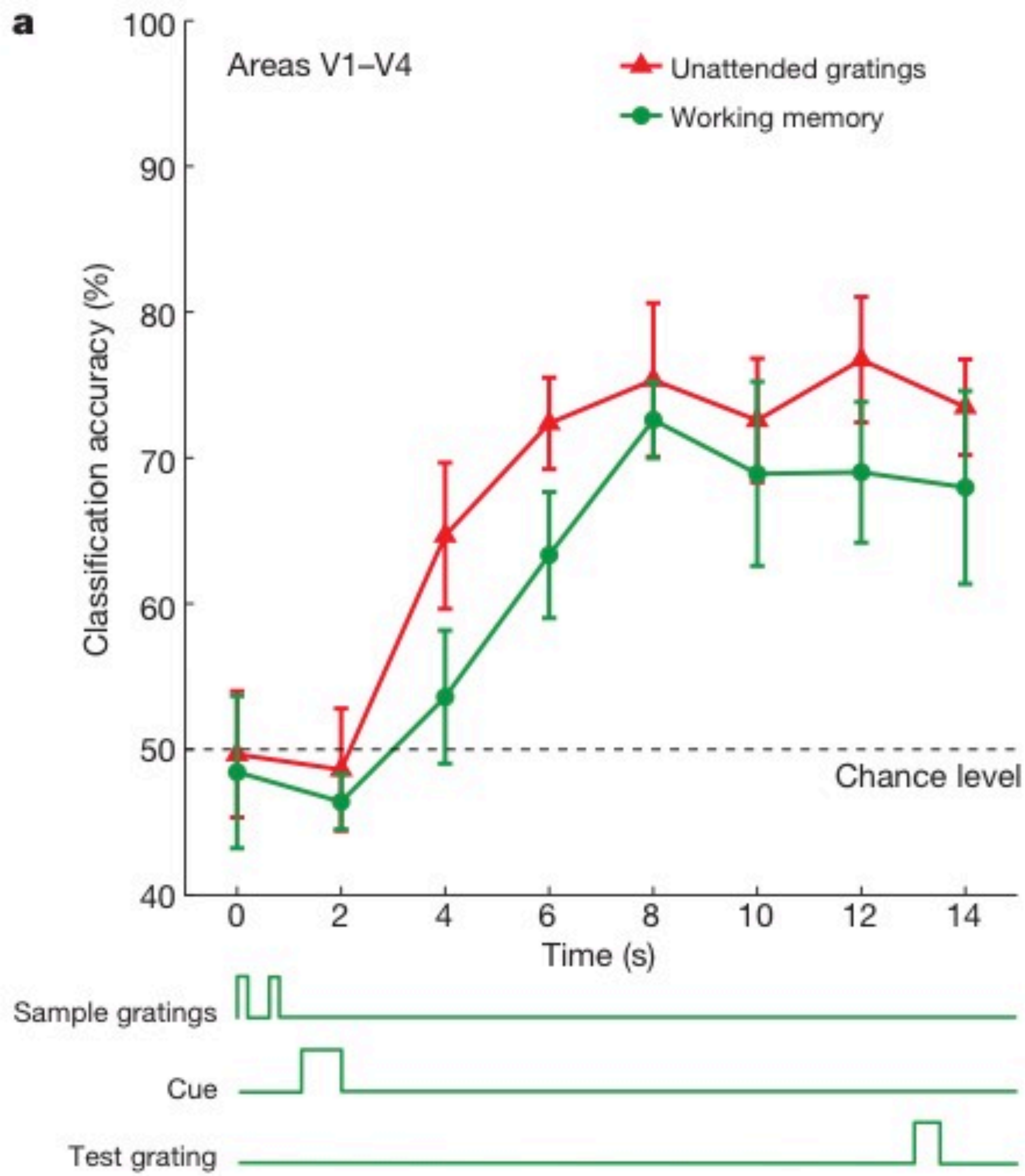


S. A. Harrison and F. Tong. Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458:632–635, 2009.

Decoding working memory



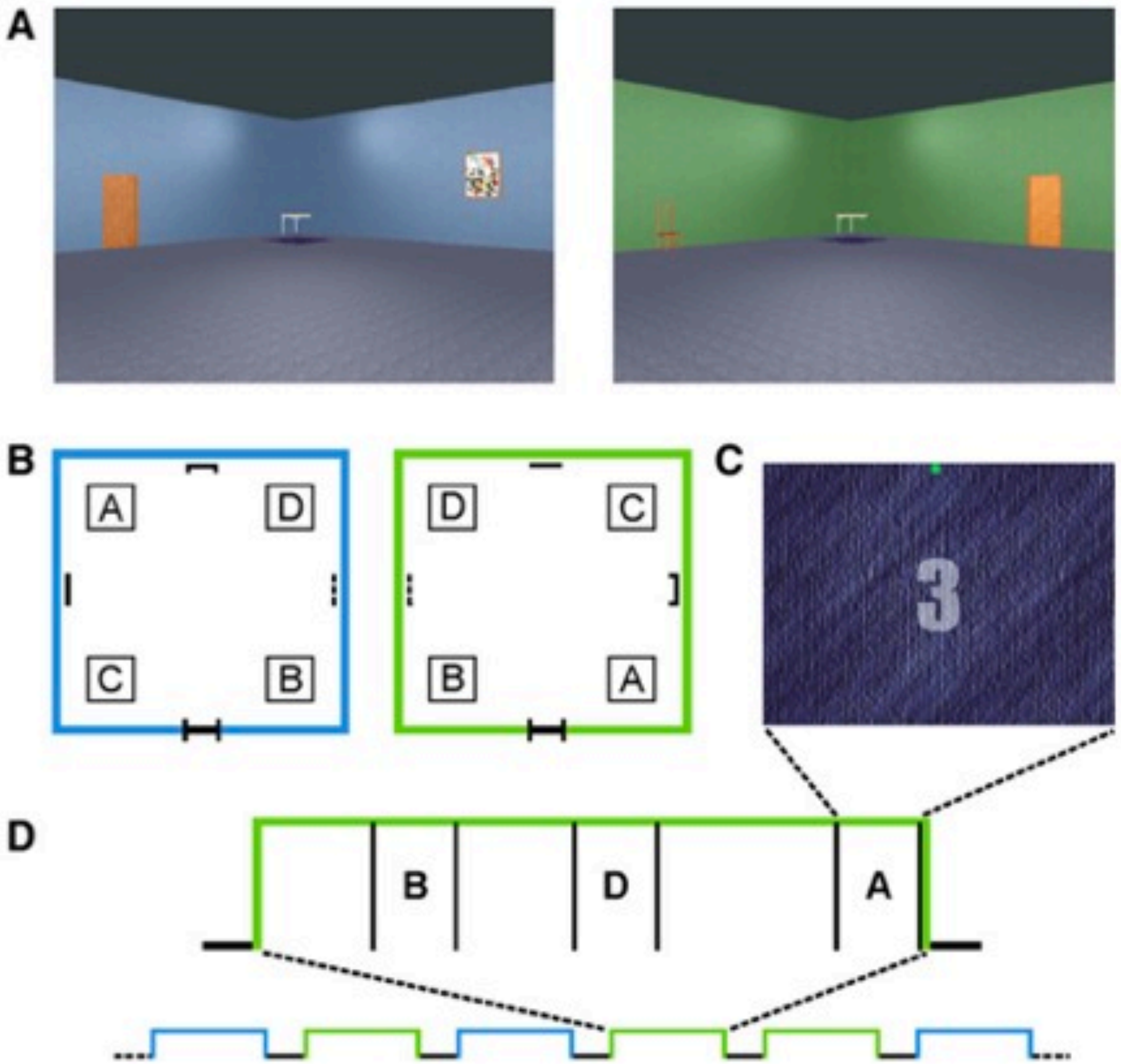
Decoding working memory

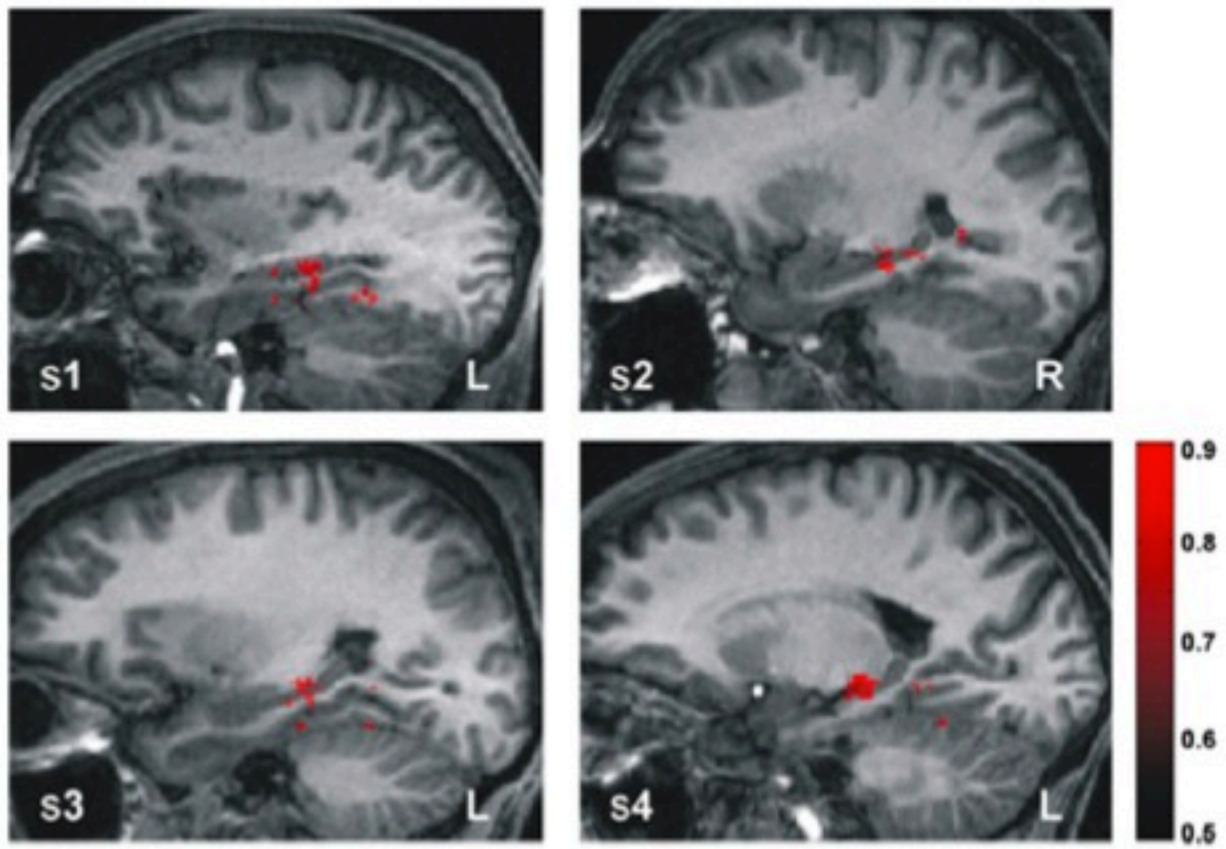


Place decoding

Decoding Neuronal Ensembles in the Human Hippocampus

Demis Hassabis,[1,*] Carlton Chu,[1] Geraint Rees,[1,2]
Nikolaus Weiskopf,[1] Peter D. Molyneux,[3]
and Eleanor A. Maguire[1,*]





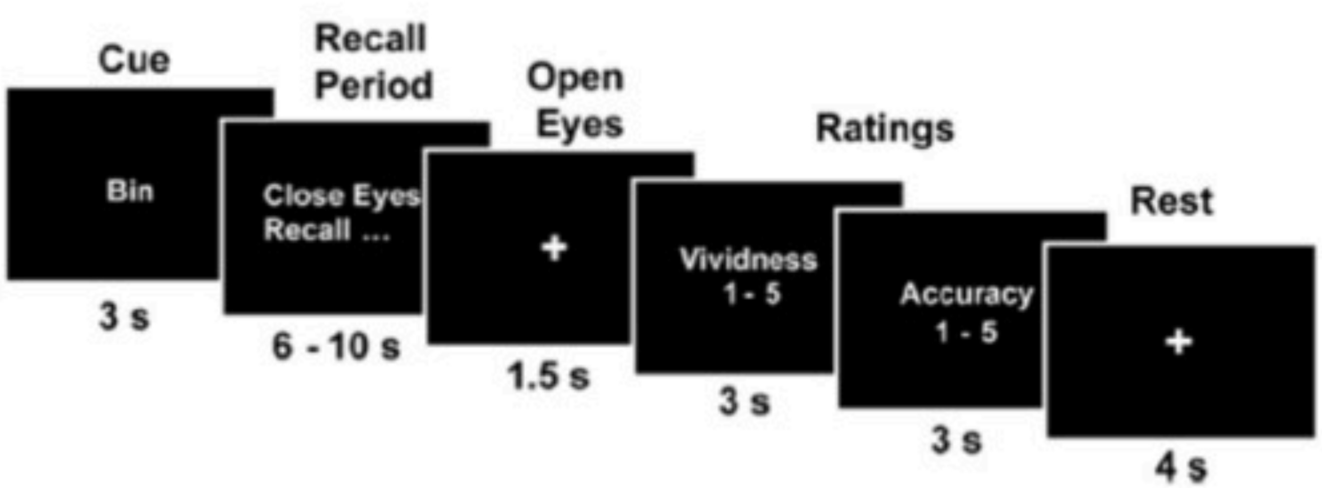
Current Biology, 2009

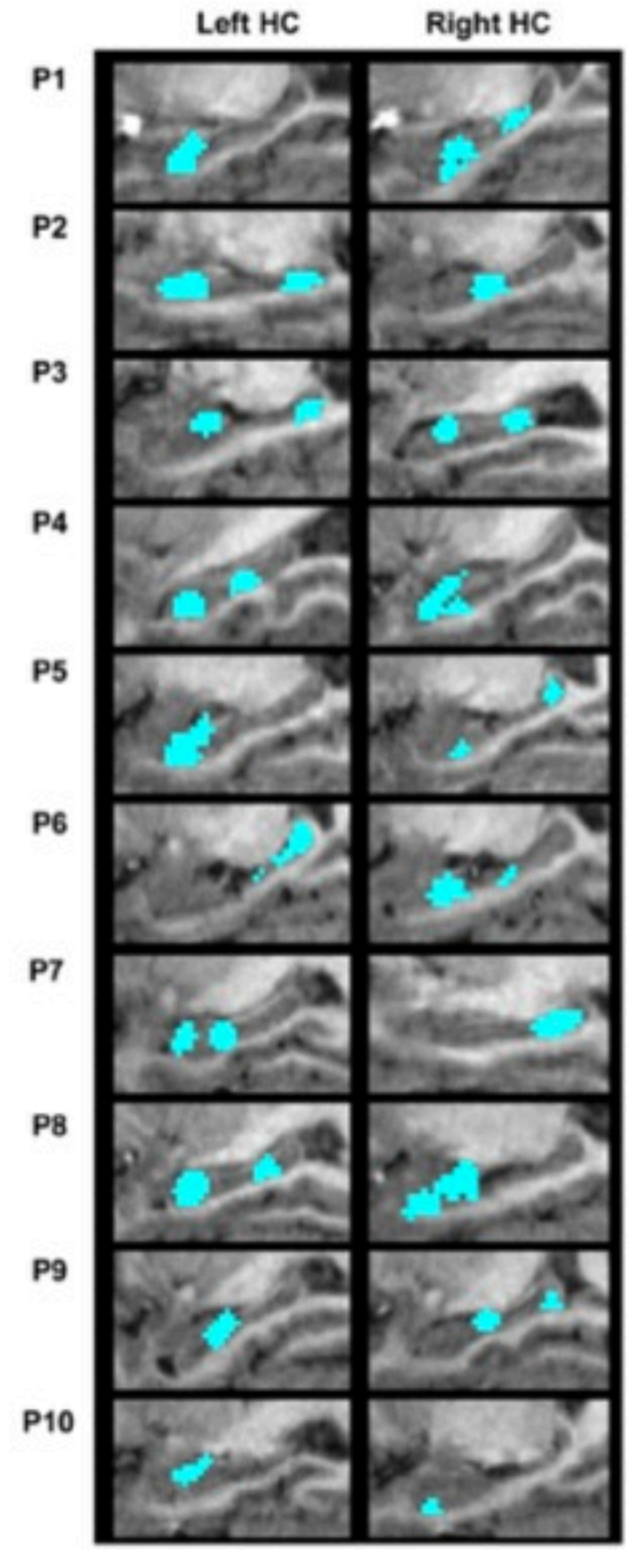
Decoding Individual Episodic Memory Traces in the Human Hippocampus

Martin J. Chadwick,[1,2] Demis Hassabis,[1,2]
Nikolaus Weiskopf,[1] and Eleanor A. Maguire[1,*]

A

B

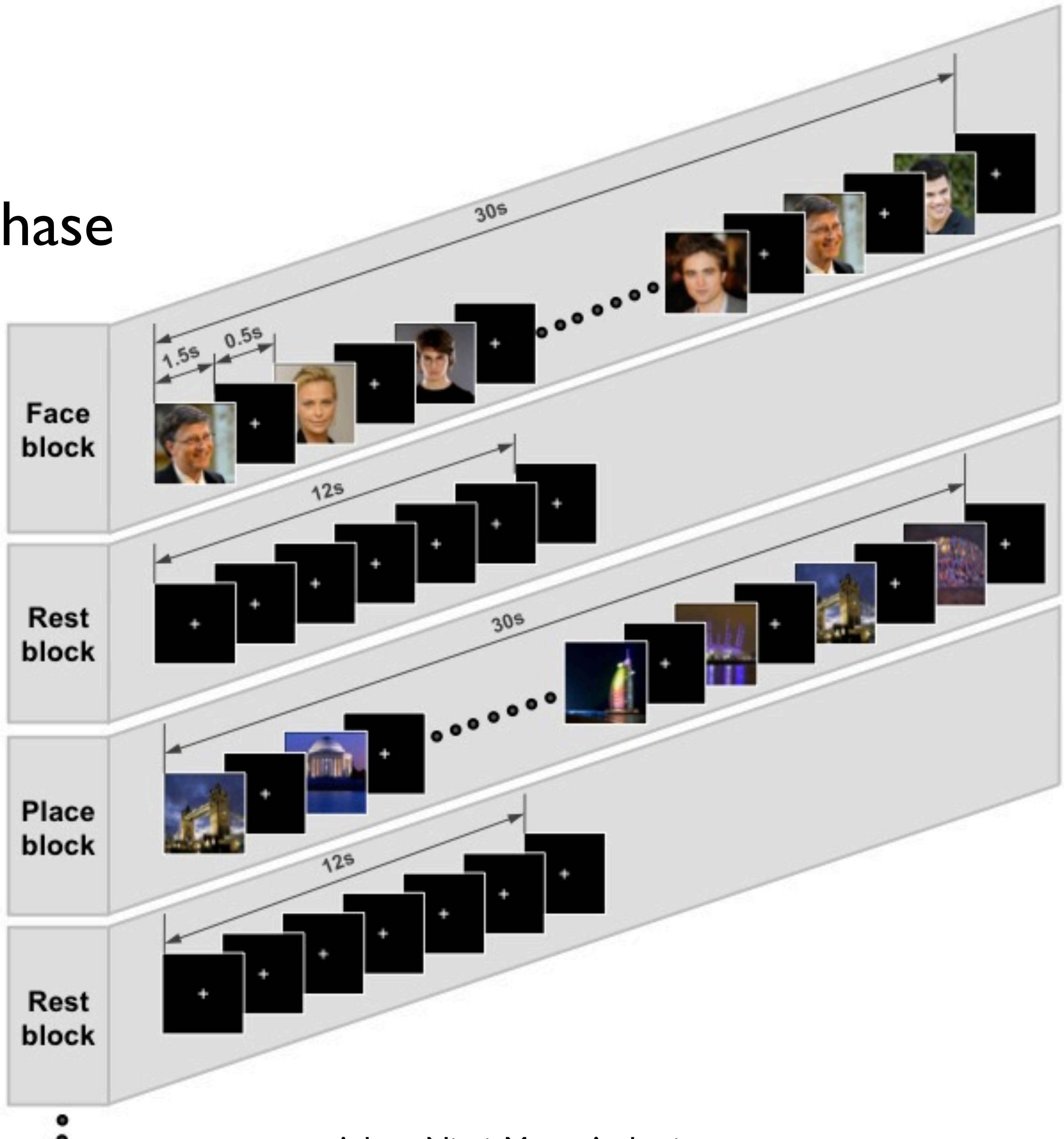




Current Biology, 2010

Real-time decoding of attended object categories

training phase



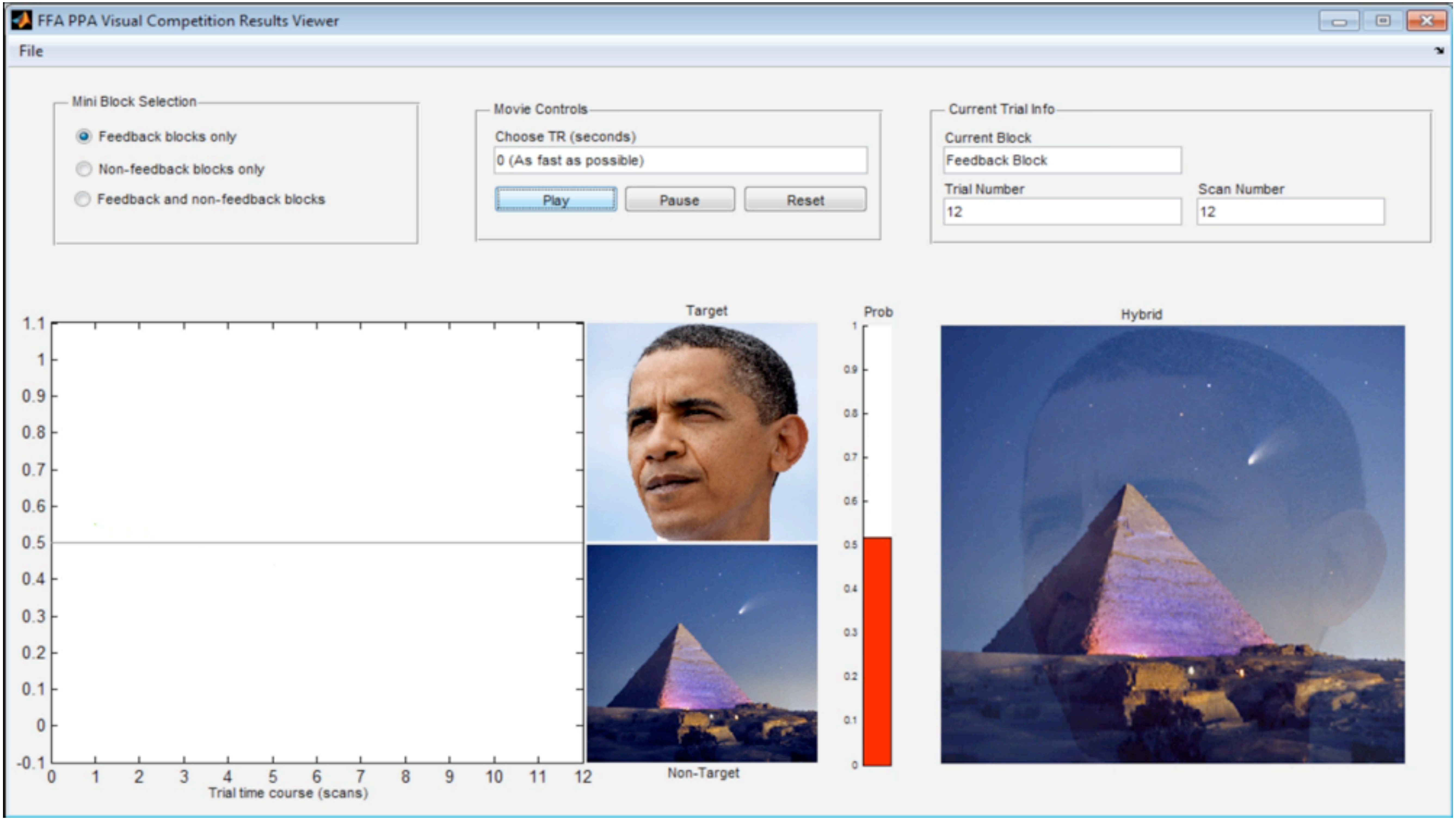
Adnan Niazi. Master's thesis

Real-time decoding of attended object categories

accuracy ~ 75%

Adnan Niazi. Master's thesis

Real-time decoding of attended object categories

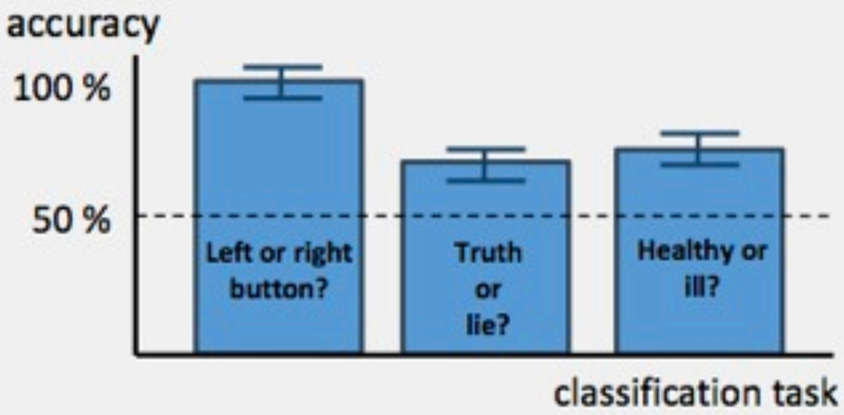


accuracy ~ 75%

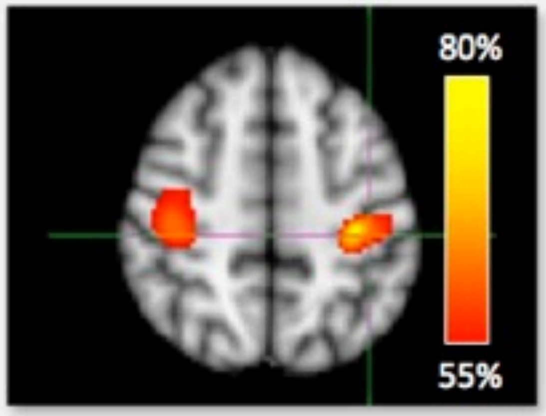
Adnan Niazi. Master's thesis

Research questions for classification

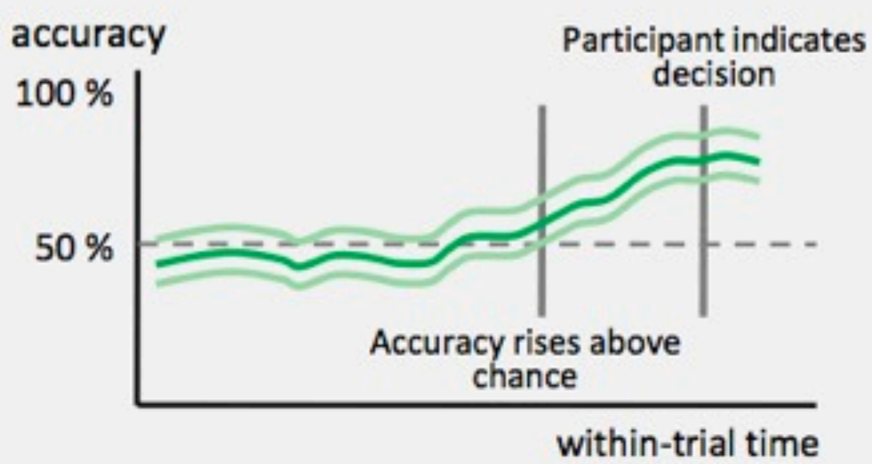
Overall classification accuracy



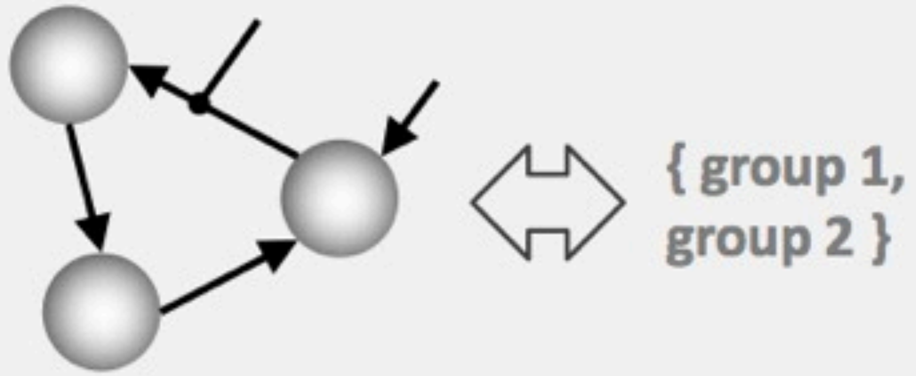
Spatial deployment of discriminative regions



Temporal evolution of discriminability



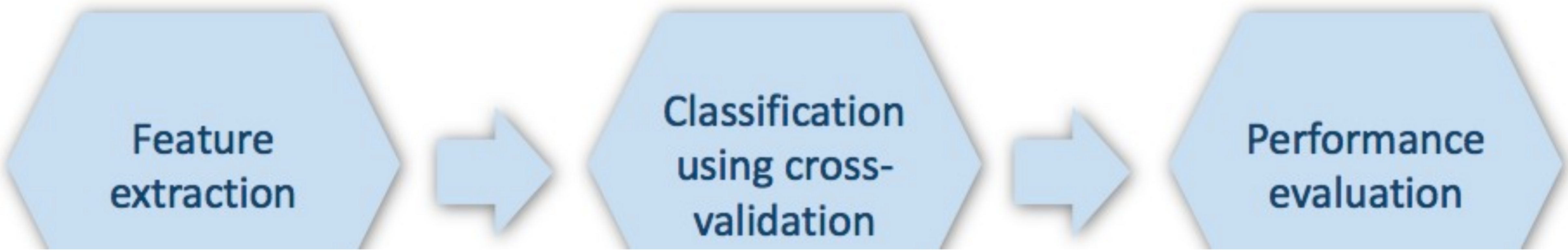
Model-based classification



Slide courtesy Kai Brodersen

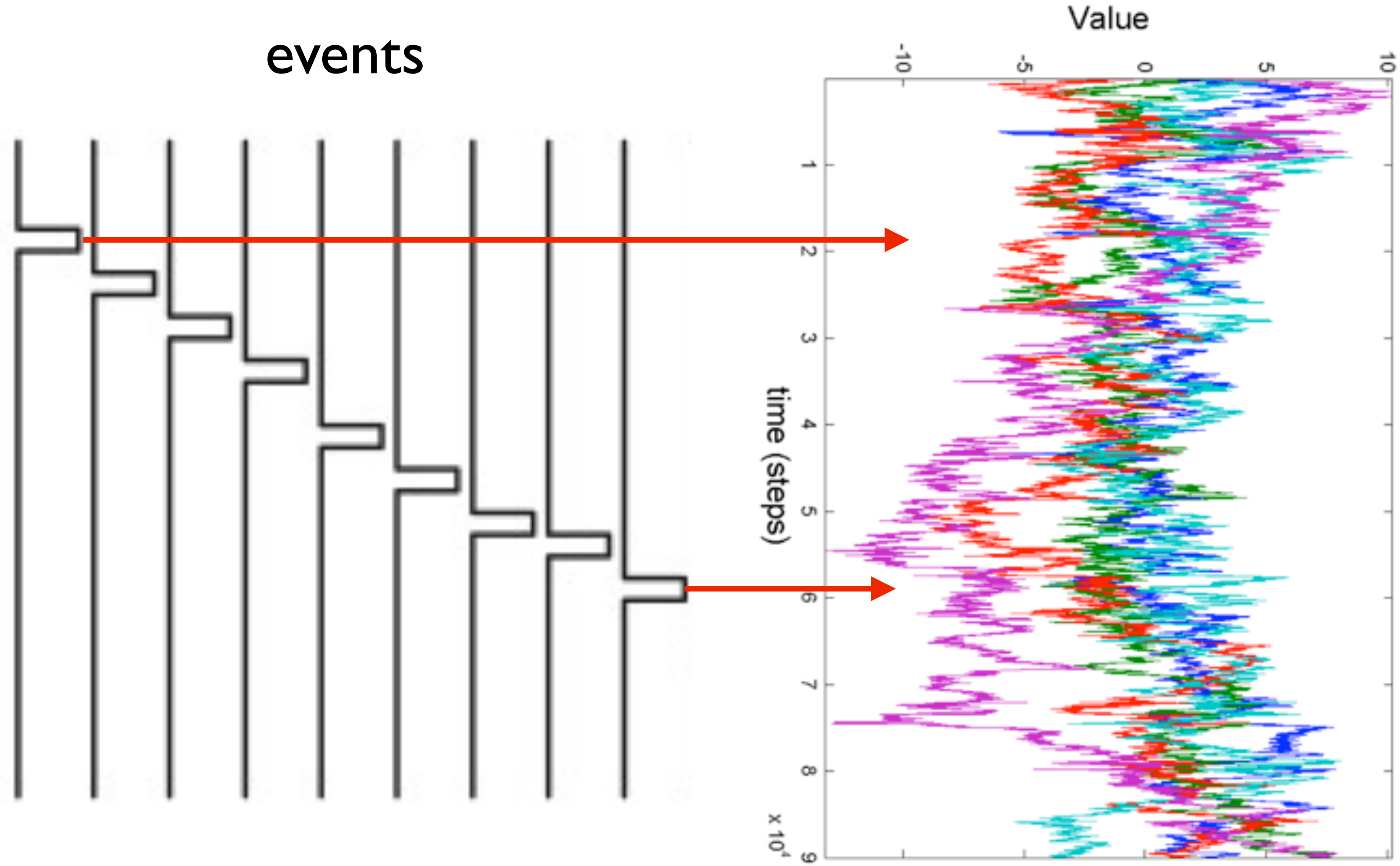
Stages in a classification analysis

predicting one from a discrete set of cognitive states from measured data (e.g. perceived animal versus tool, patient versus control)



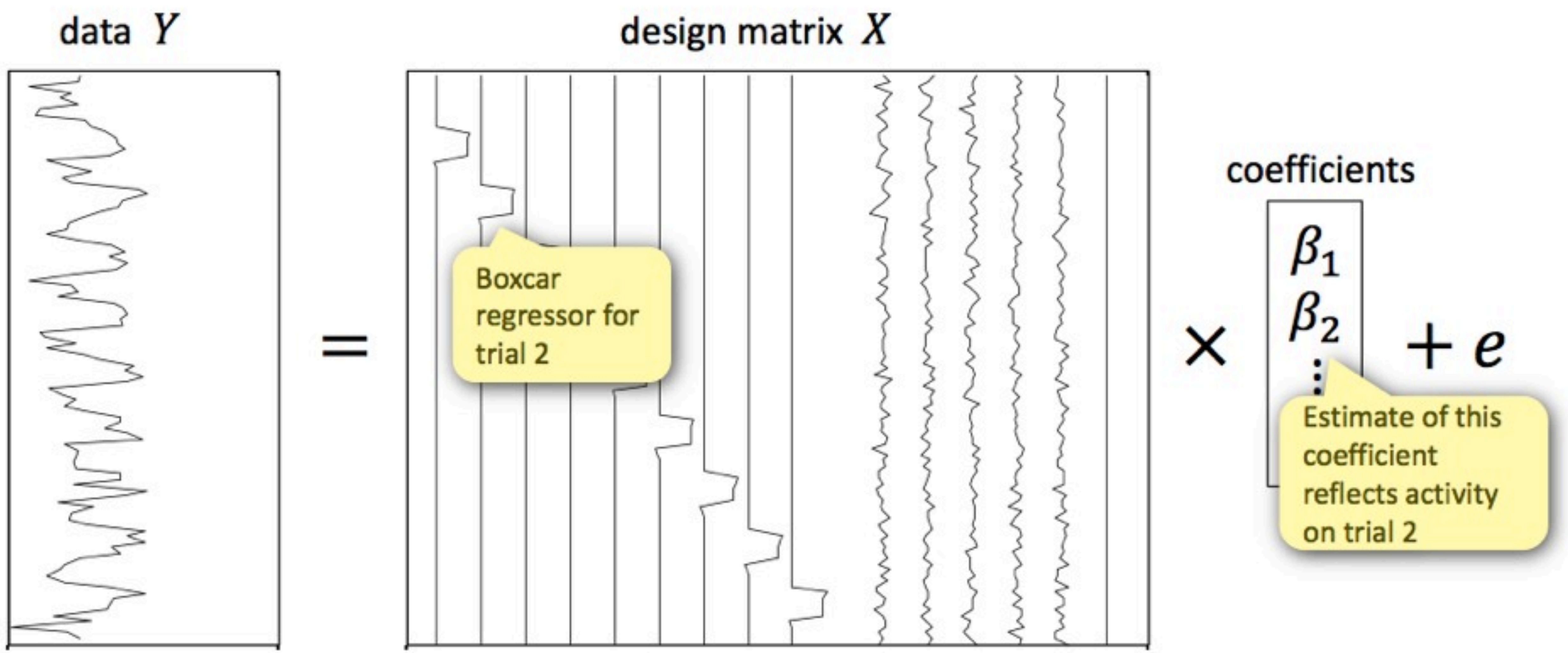
Note:
if the cognitive state is a continous quantity then we are dealing with a regression problem (e.g. level of arousal)

What’s the input?



Better: single-trial beta estimates

We can obtain trial-wise estimates of neural activity by filtering the data with a GLM.



Slide courtesy Kai Brodersen

Classifiers

A principled way of designing a classifier would be to adopt a probabilistic approach:

$Y_t \rightarrow \boxed{f} \rightarrow$ that k which maximizes $p(X_t = k | Y_t, X, Y)$

In practice, classifiers differ in terms of how strictly they implement this principle.

Generative classifiers

use Bayes' rule to estimate $p(X_t|Y_t) \propto p(Y_t|X_t)p(X_t)$

- *Gaussian Naïve Bayes*
- *Linear Discriminant Analysis*

Discriminative classifiers

estimate $p(X_t|Y_t)$ directly without Bayes' theorem

- *Logistic regression*
- *Relevance Vector Machine*

Discriminant classifiers

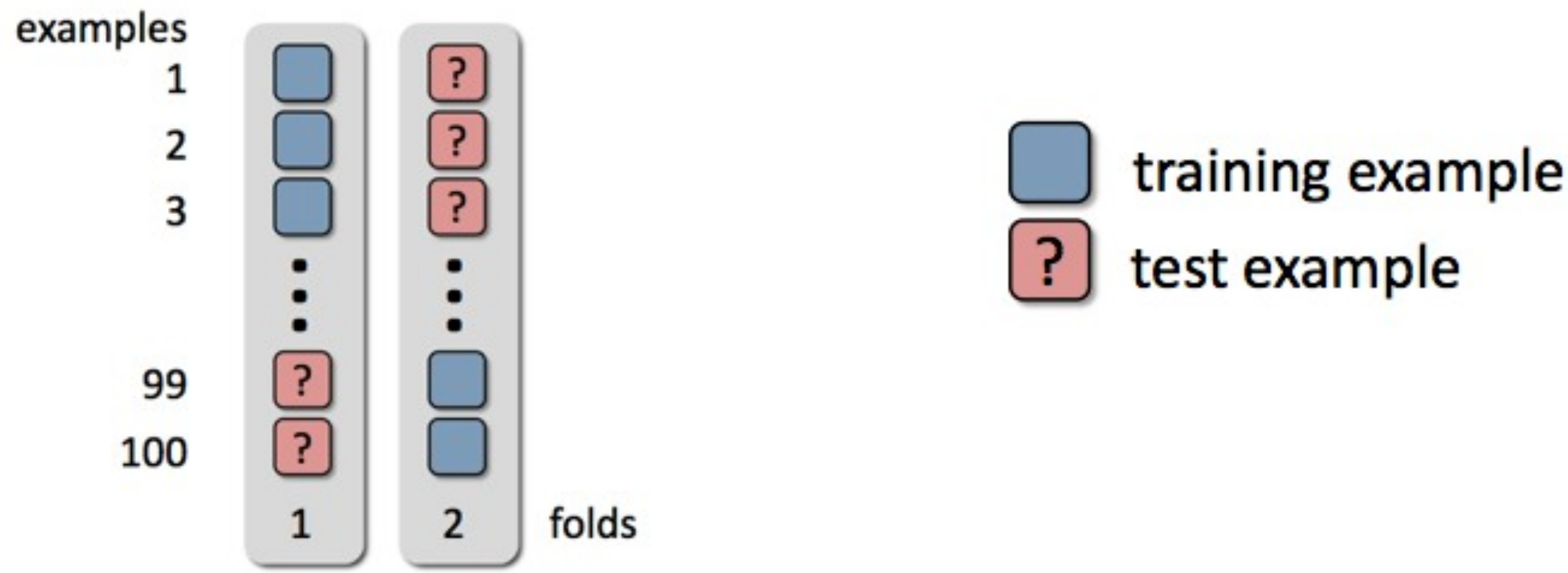
estimate $f(Y_t)$ directly

- *Fisher's Linear Discriminant*
- *Support Vector Machine*

Slide courtesy Kai Brodersen

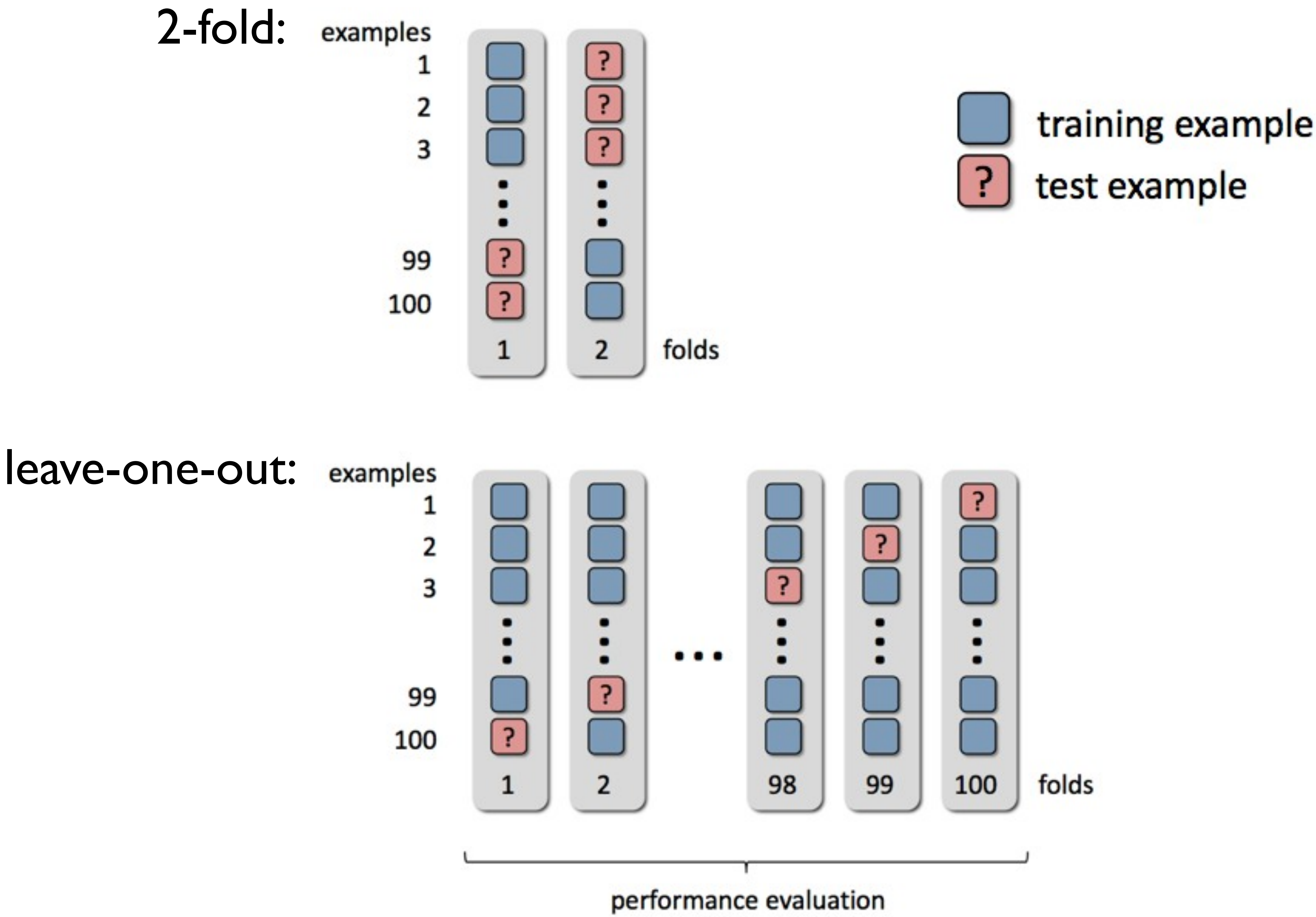
Cross-validation

2-fold:



Slide courtesy Kai Brodersen

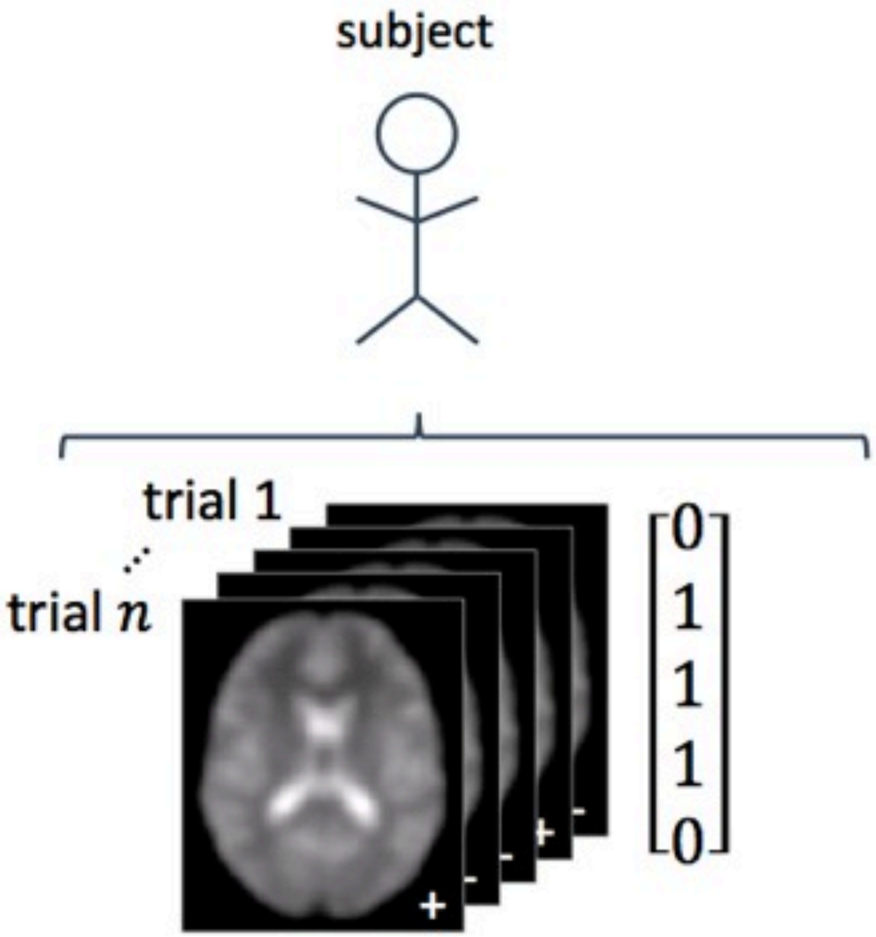
Cross-validation



Slide courtesy Kai Brodersen

Single-subject study with n trials

The most common approach is to assess how likely the obtained number of correctly classified trials could have occurred by chance.



Binomial test

$$p = P(X \geq k | H_0) = 1 - \Pr(X \leq x - 1)$$

In MATLAB:

```
p = 1 - binocdf(k-1,n,pi_0)
```

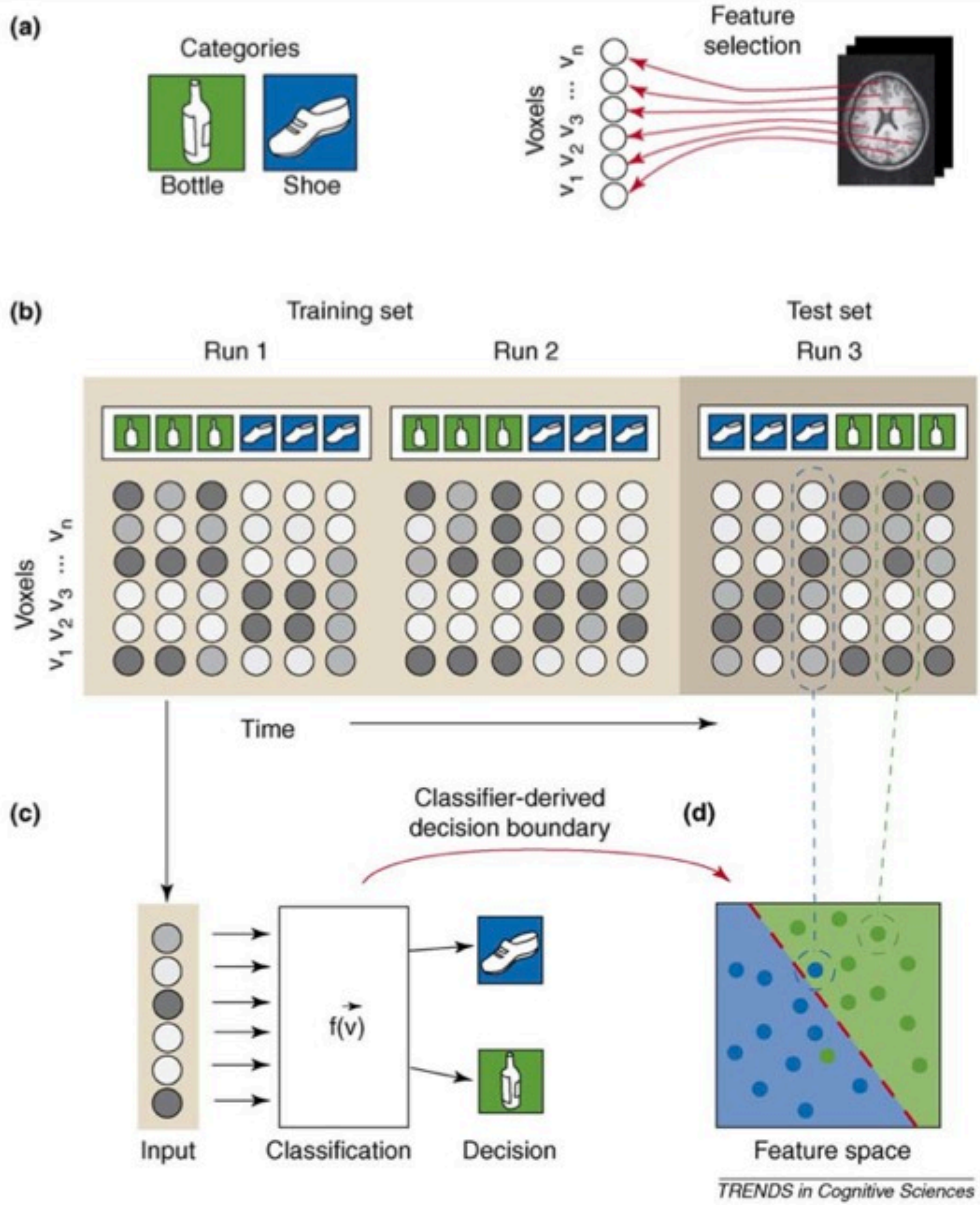
k	number of correctly classified trials
n	total number of trials
π_0	chance level (typically 0.5)
B	binomial cumulative density function

- compute number of subjects s from population of n that are significant under a certain p value

- compute probability of observing such a result, again under a binomial distribution:

```
pvalue = 1 - binocdf(s-1,n,p);
```

Classification of neuroimaging data



Norman et al. TICS, 2006

Discriminant classifiers

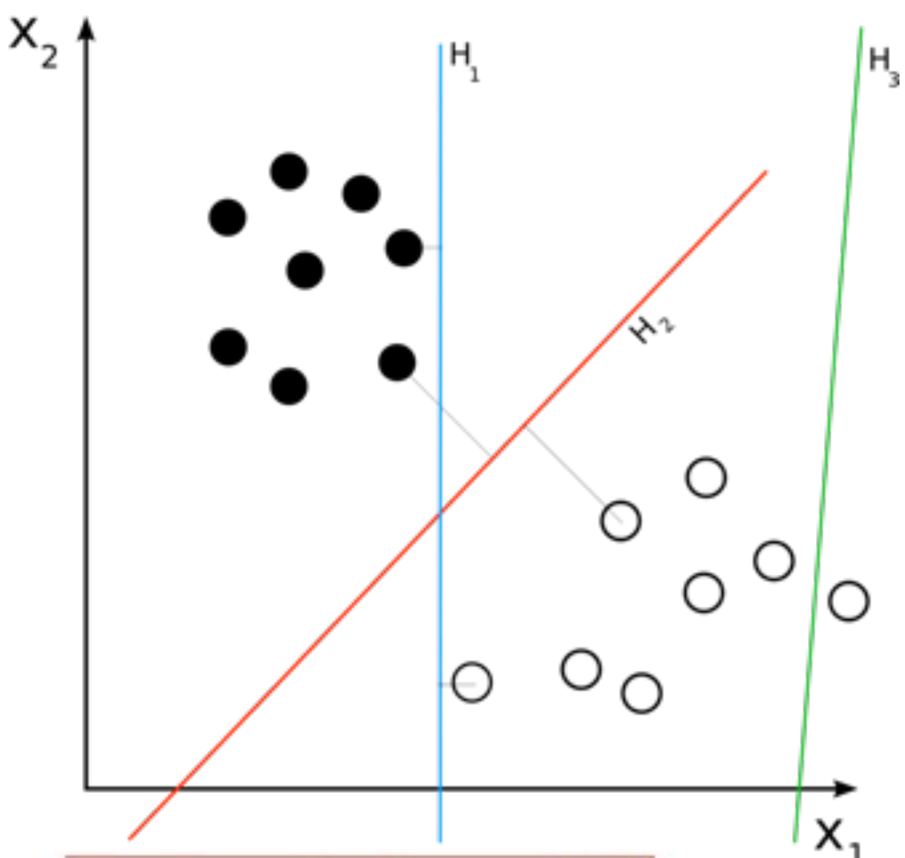
condition model data

$$x = f(\mathbf{y}; D)$$

measurements

E.g. Linear support vector machine

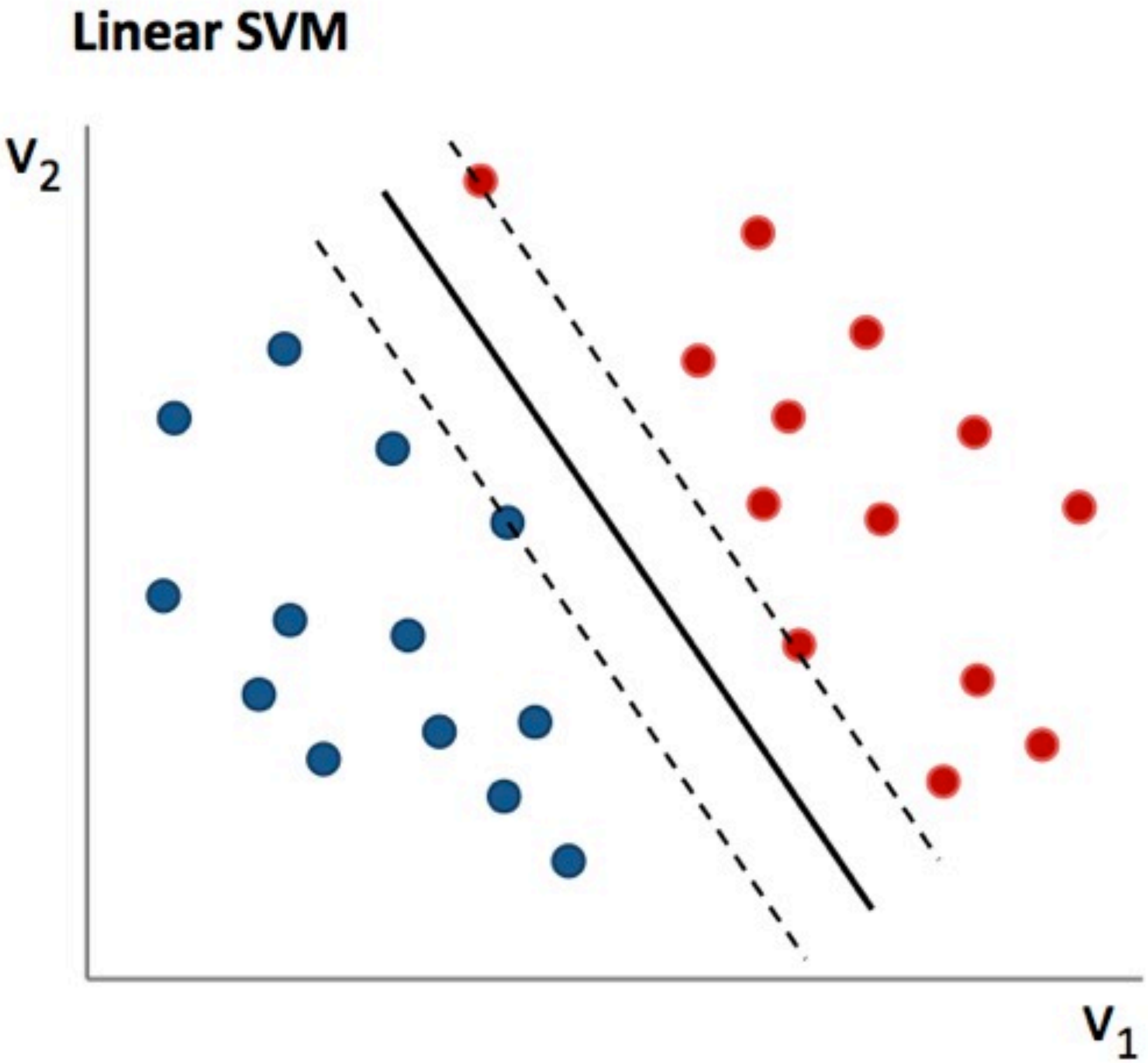
hyperplane maximizes distance between classes

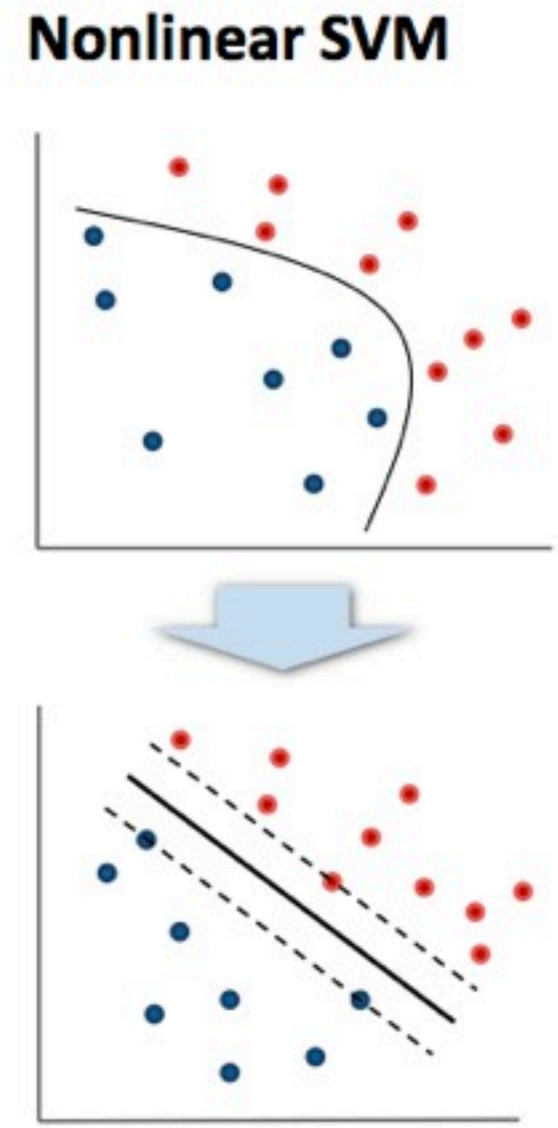


very popular approach in neuroimaging:

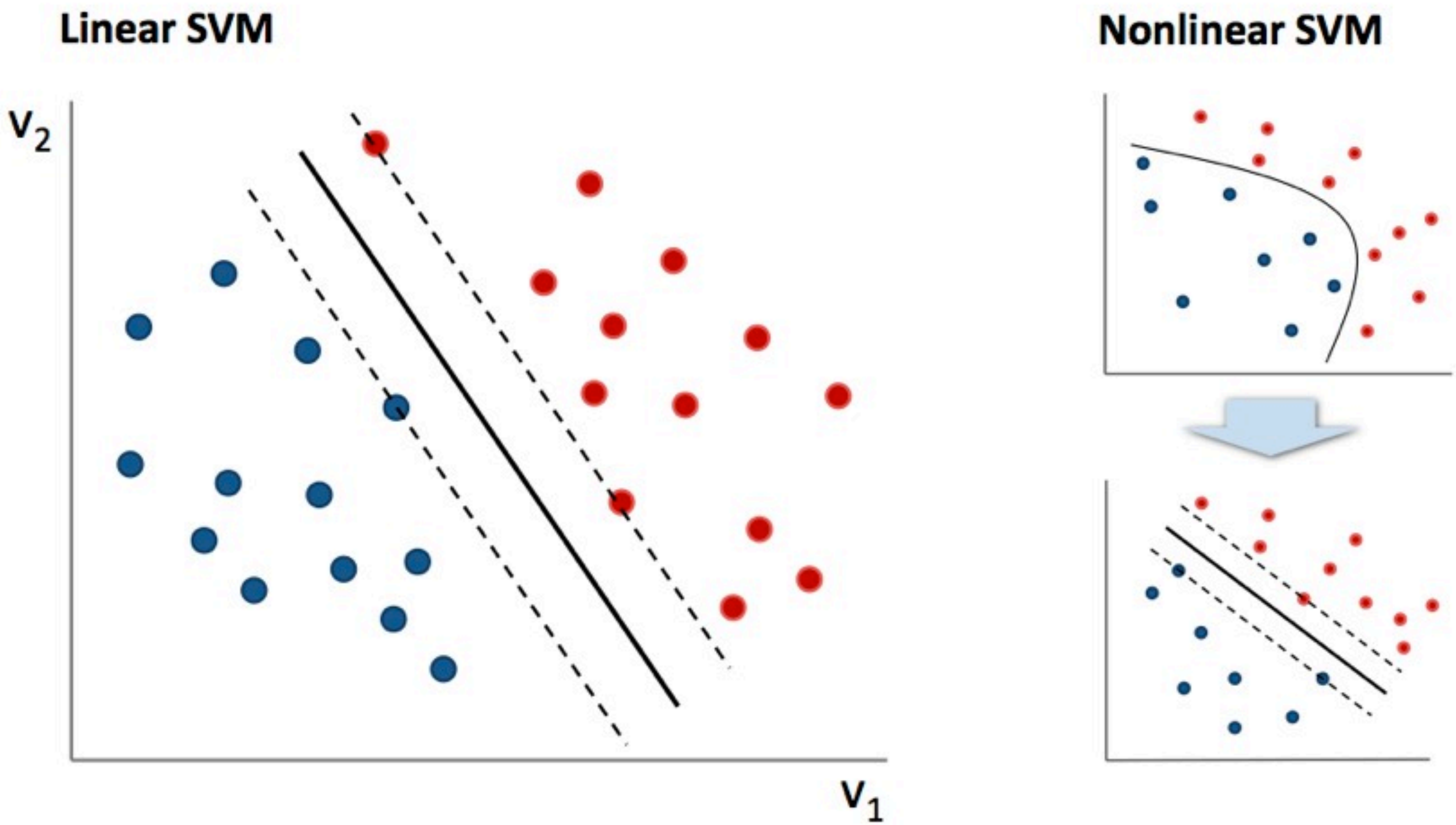
- fast
- robust
- easy-to-use packages

Linear versus non-linear classifiers





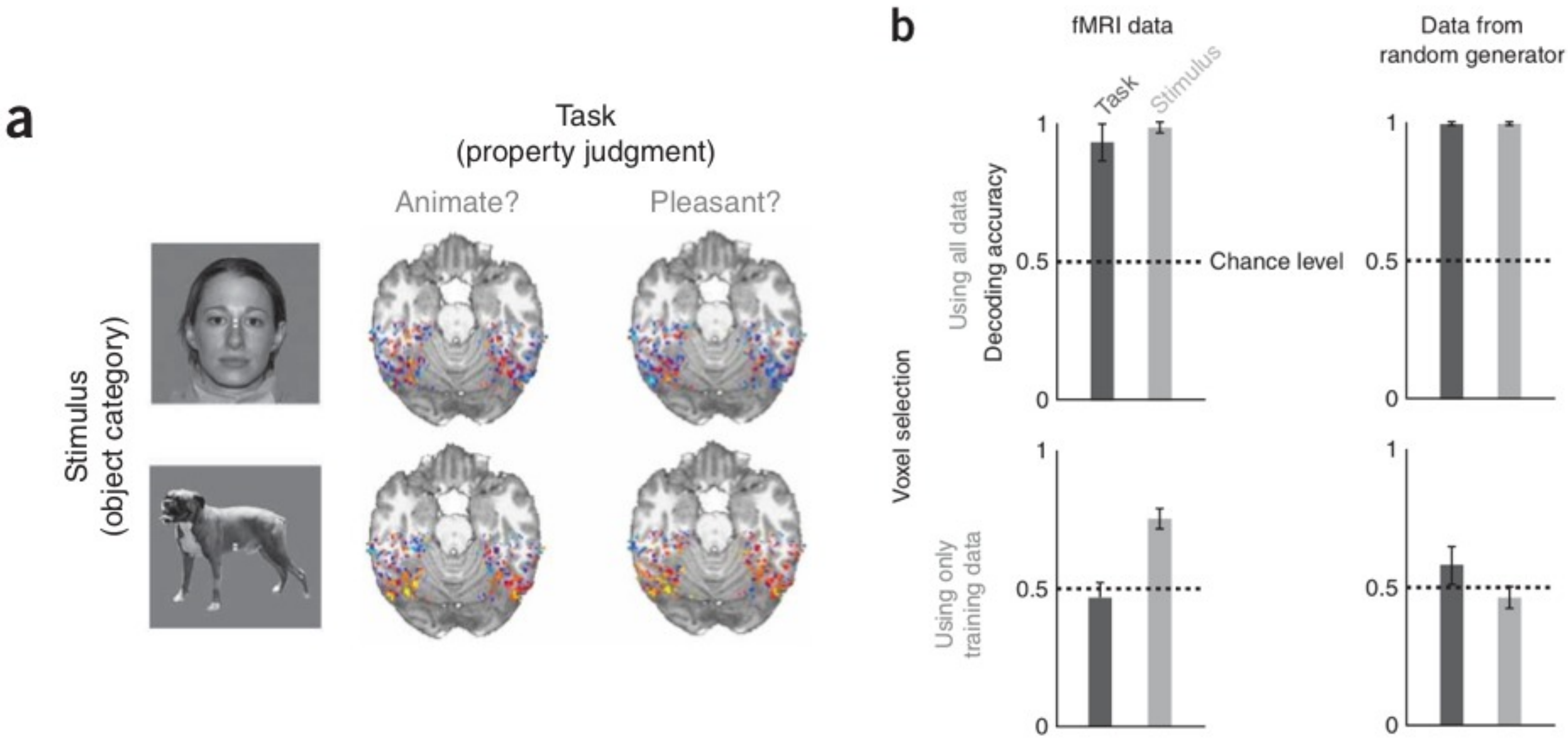
Vapnik (1999) Springer; Schölkopf et al. (2002) MIT Press



Vapnik (1999) Springer; Schölkopf et al. (2002) MIT Press

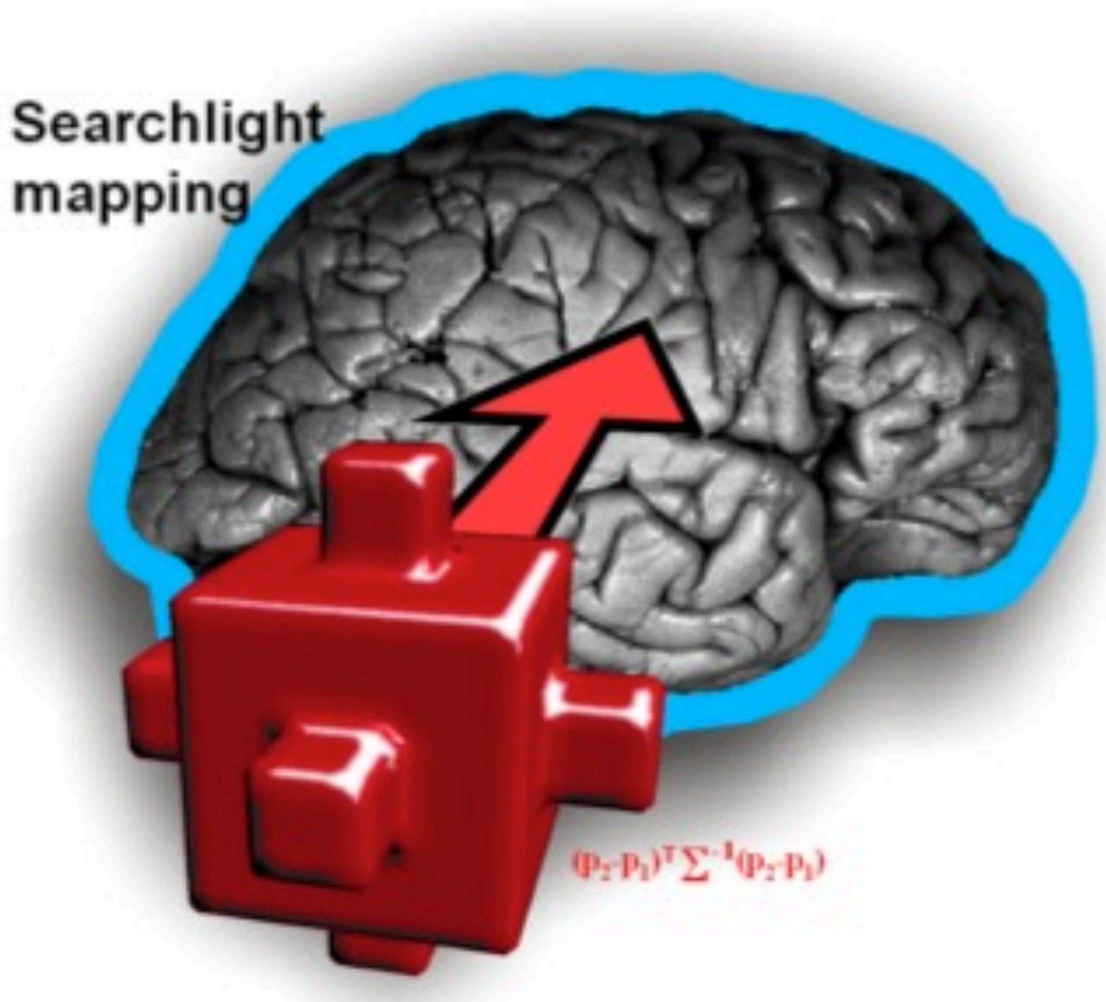
Linear classifiers are more often used in neuroimaging:
Robust in case of small sample sizes

How to cheat with classifiers



Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci. 2009;12(5):535–540.

How to use classifiers to find which regions are functionally involved?



In each "sphere":

Support Vector Machine classifier

- classification accuracy

- classification significance

local classification performance ~ voxel relevance

1. select voxels within radius
2. train
3. test
4. set estimated significance to center voxel
5. correct for multiple comparisons

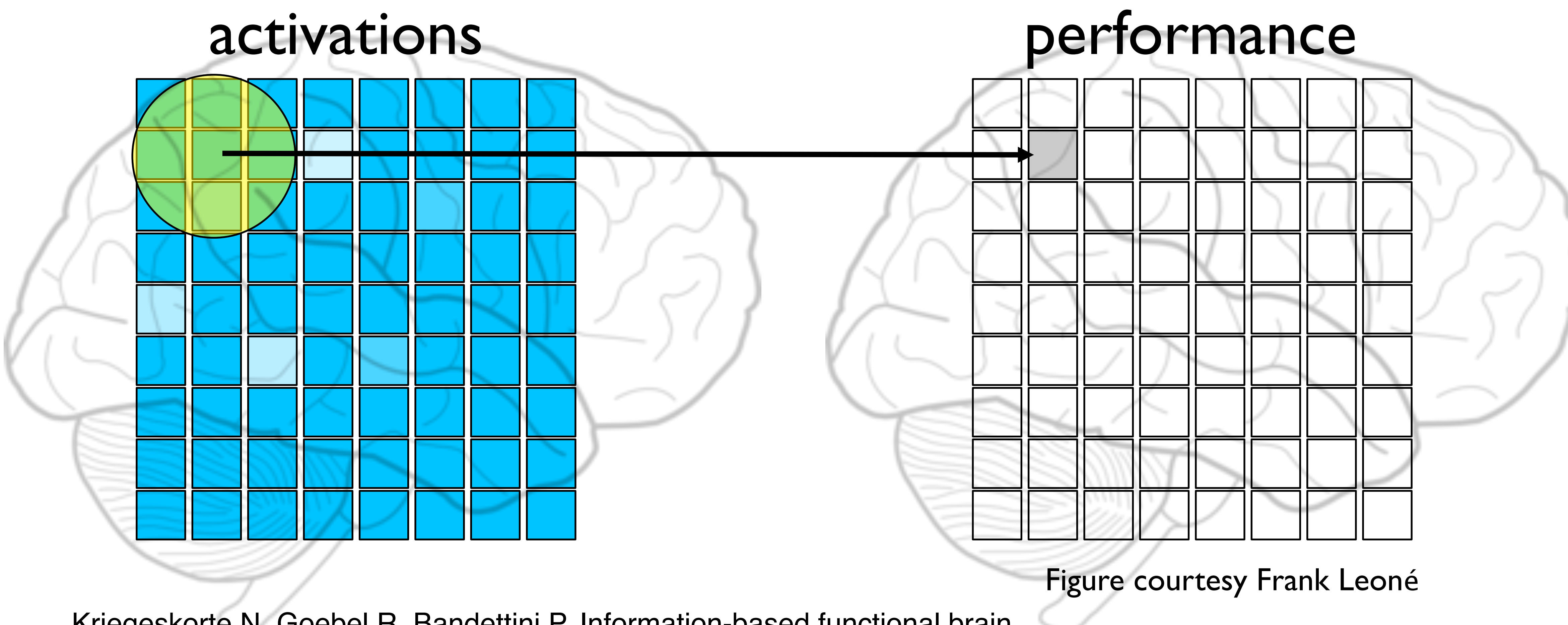


Figure courtesy Frank Leoné

Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 2006 Mar.;103(10):3863–3868.

local classification performance ~ voxel relevance

1. select voxels within radius
2. train
3. test
4. set estimated significance to center voxel
5. correct for multiple comparisons

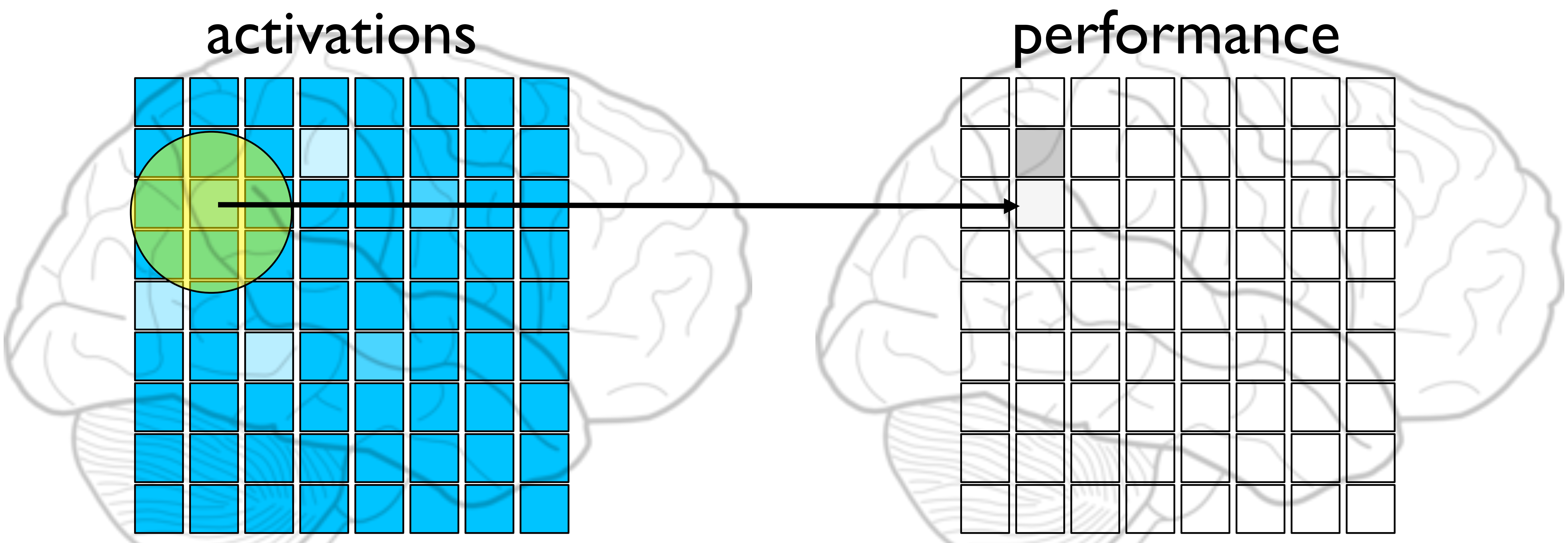


Figure courtesy Frank Leoné

Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 2006 Mar.;103(10):3863–3868.

local classification performance ~ voxel relevance

1. select voxels within radius
2. train
3. test
4. set estimated significance to center voxel
5. correct for multiple comparisons

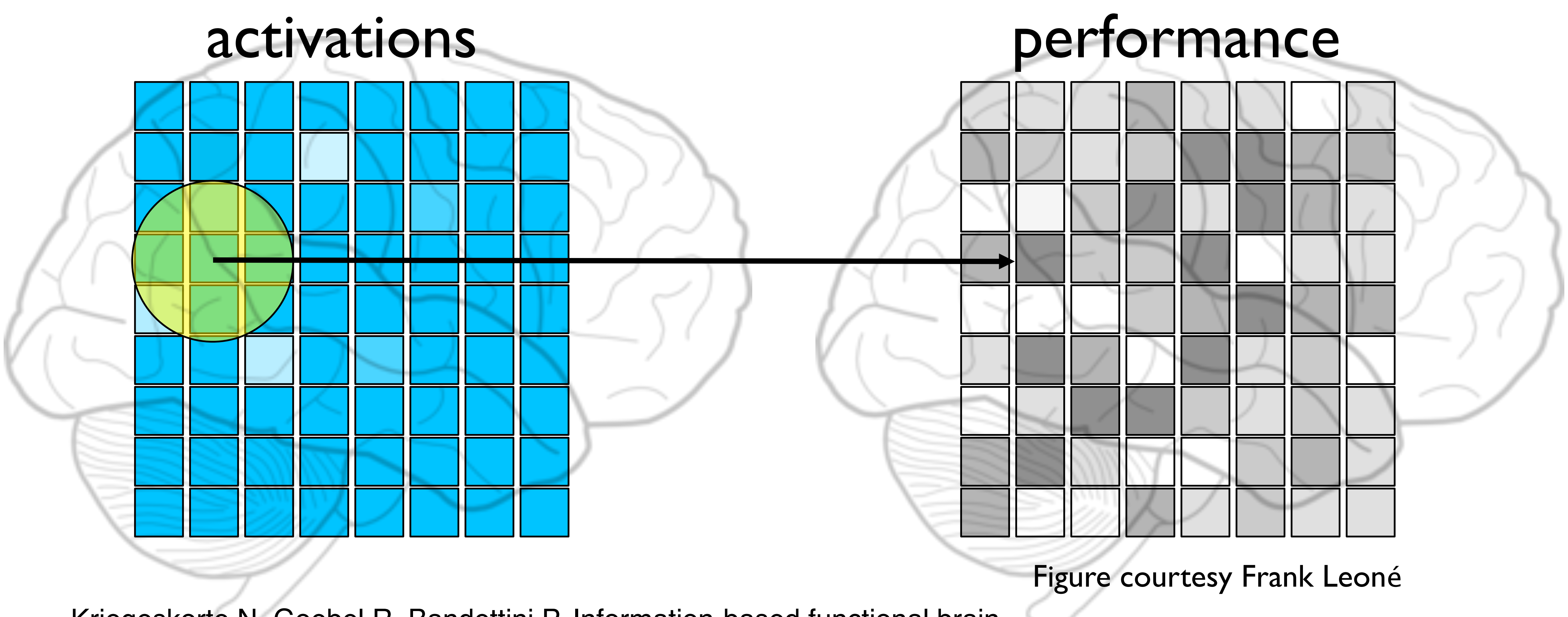
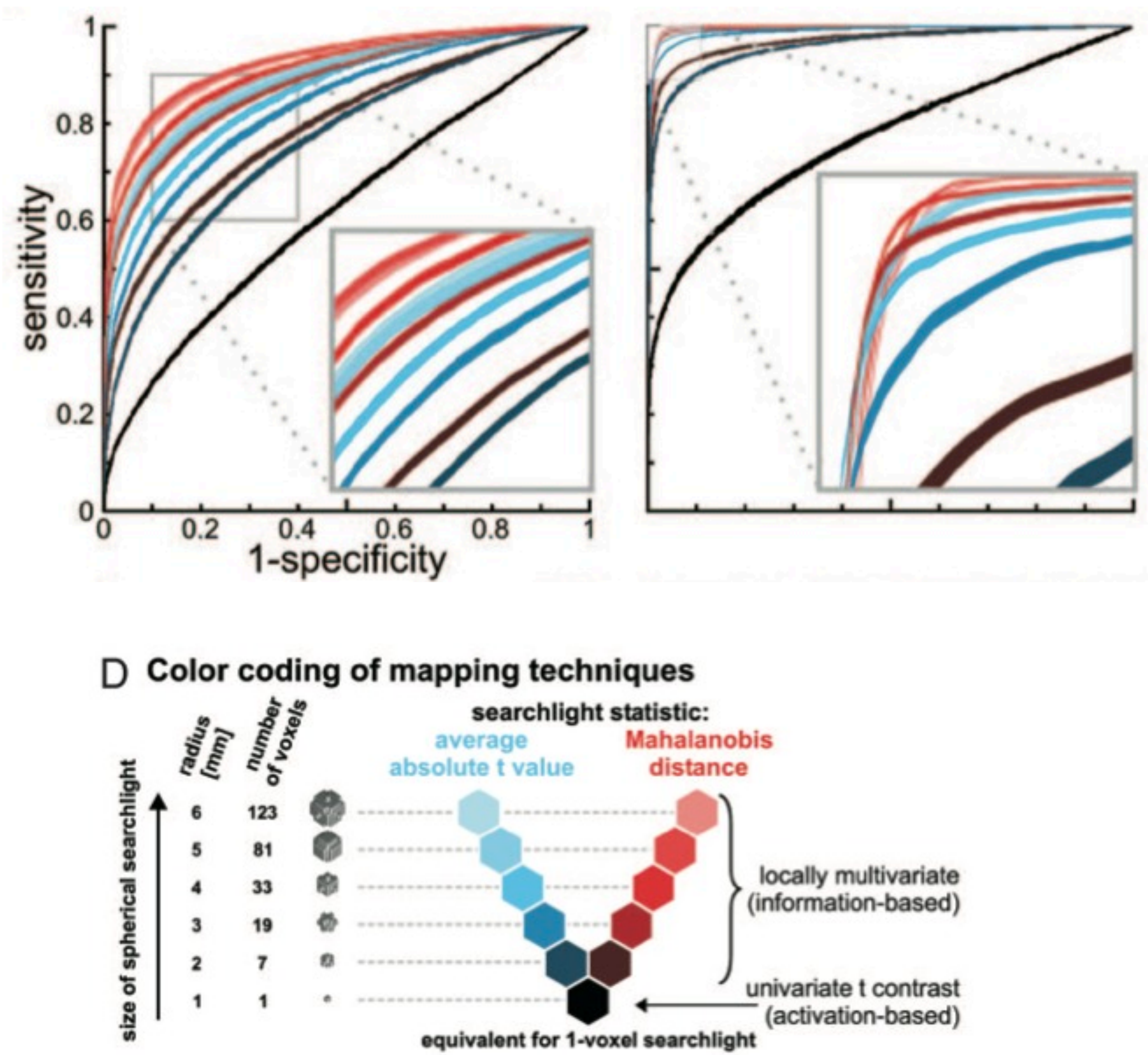


Figure courtesy Frank Leoné

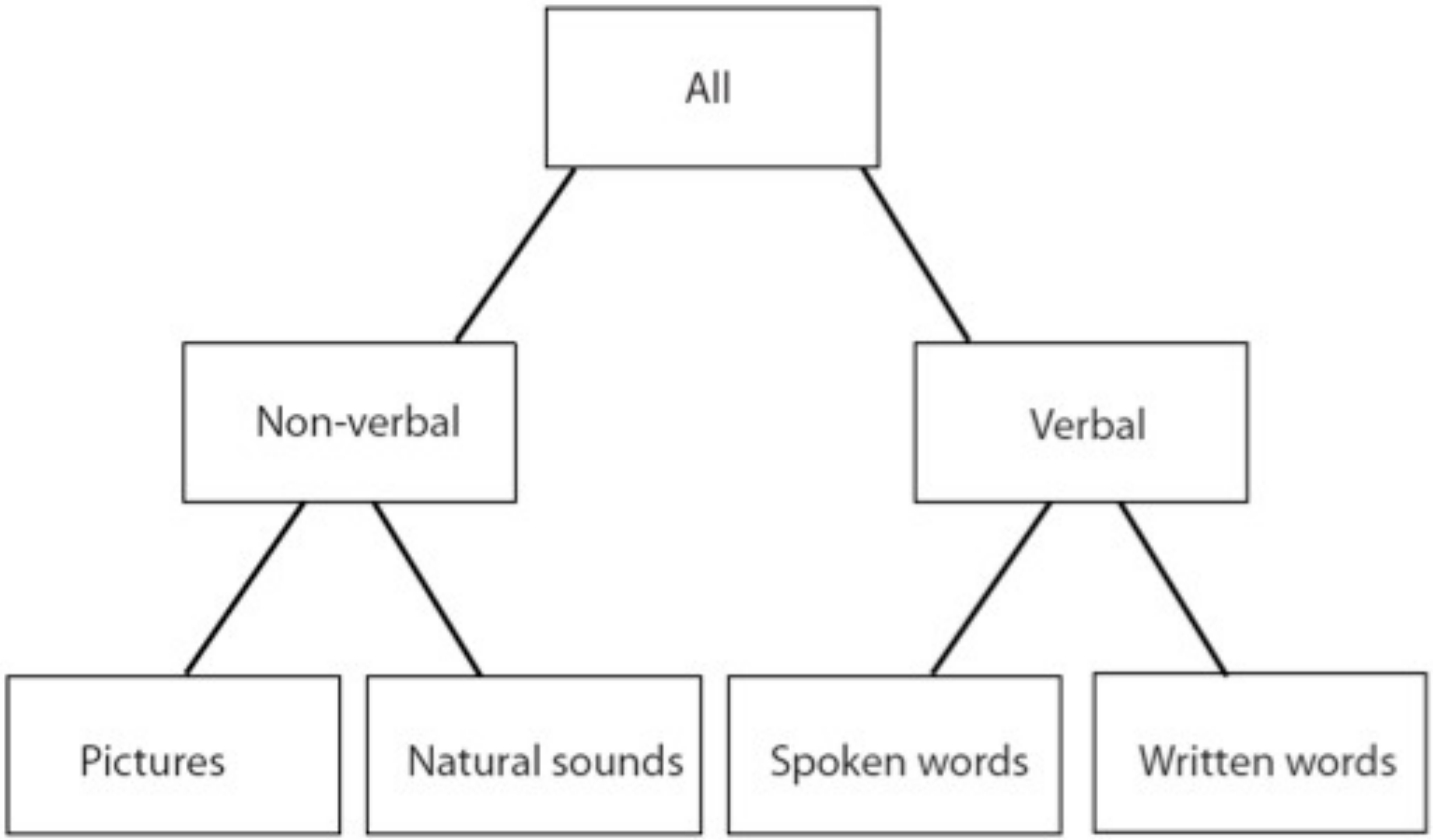
Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 2006 Mar.;103(10):3863–3868.

Searchlight analysis

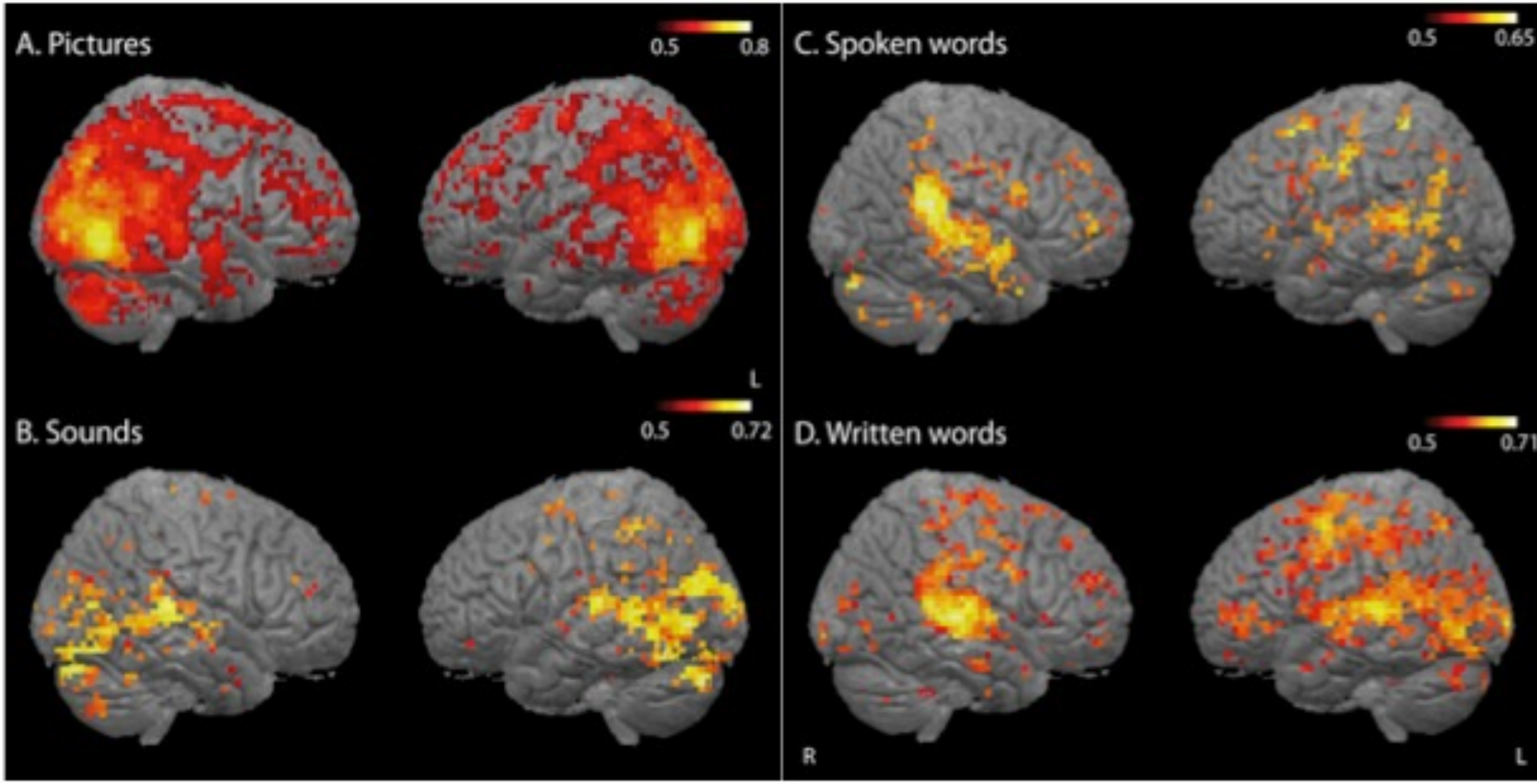


Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 2006Mar.;103(10):3863–8.

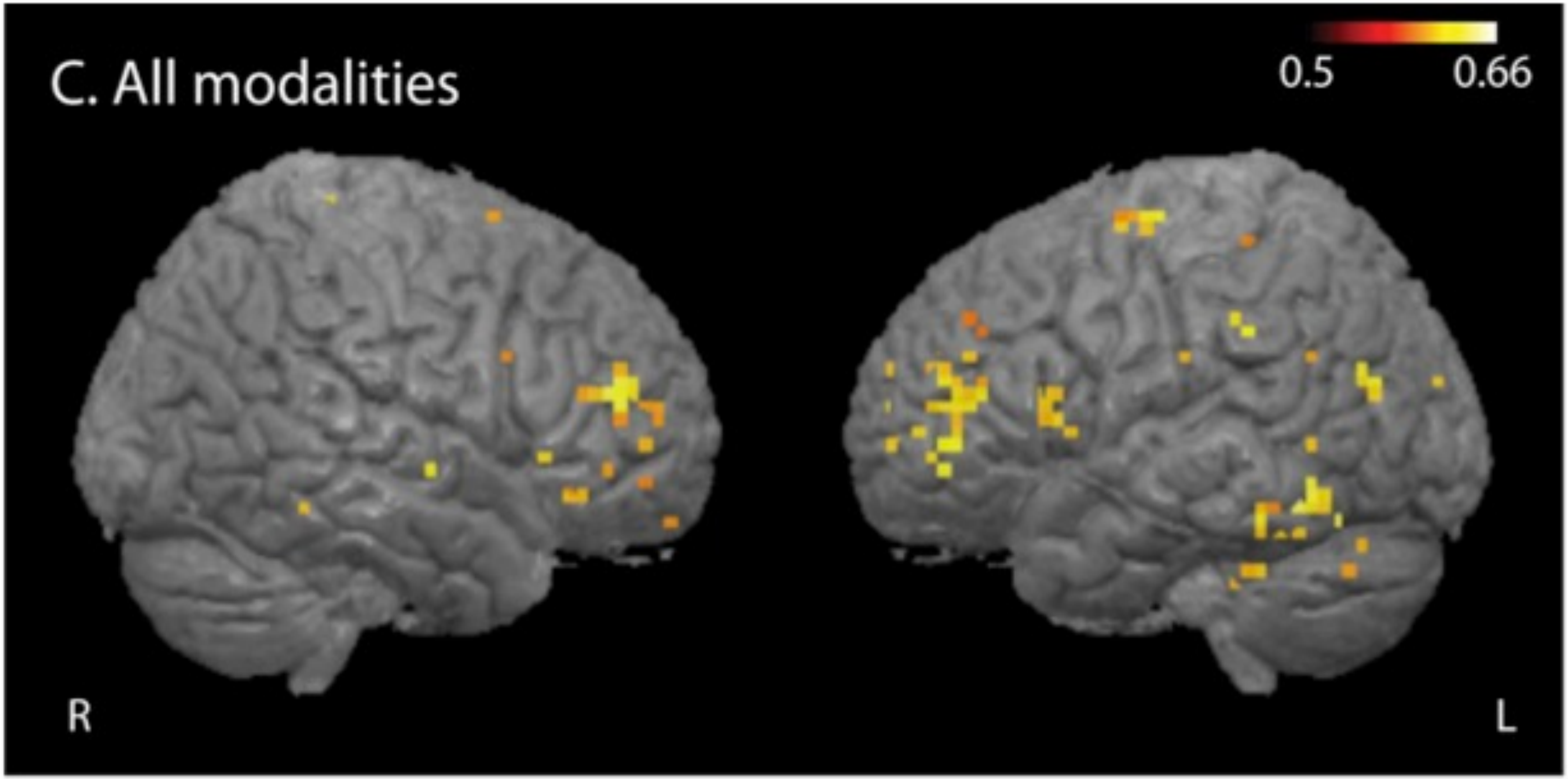
Searchlight example: cross-modal decoding of perceived animals and tools



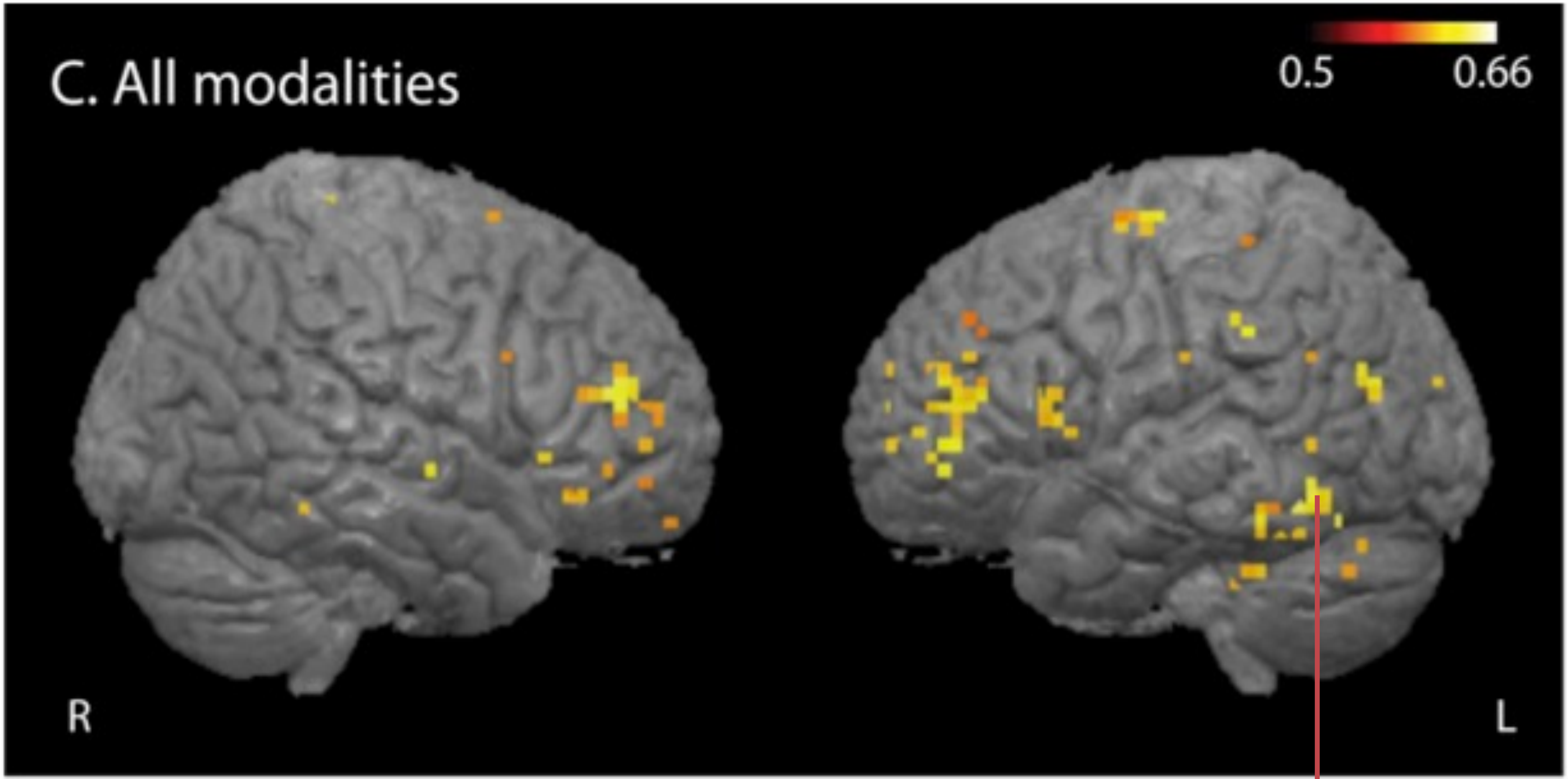
Single modalities



Cross-modal training



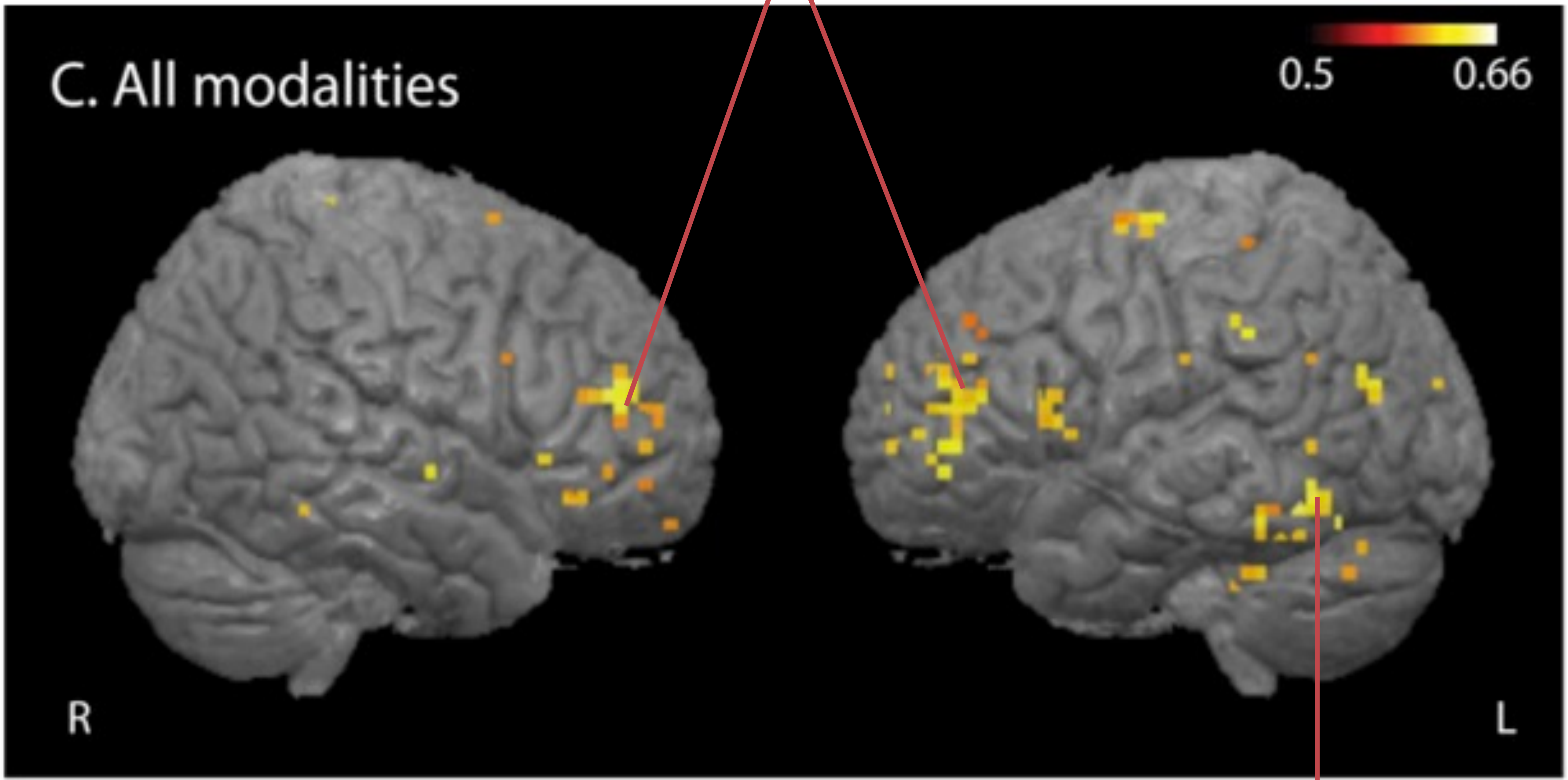
Cross-modal training



left posterior MTG

Cross-modal training

bilateral MFG



left posterior MTG

- correct p-value for m comparisons!
- use Bonferroni or FDR correction on single or group p-values

Bonferroni:
- pnew = p/m
- conservative

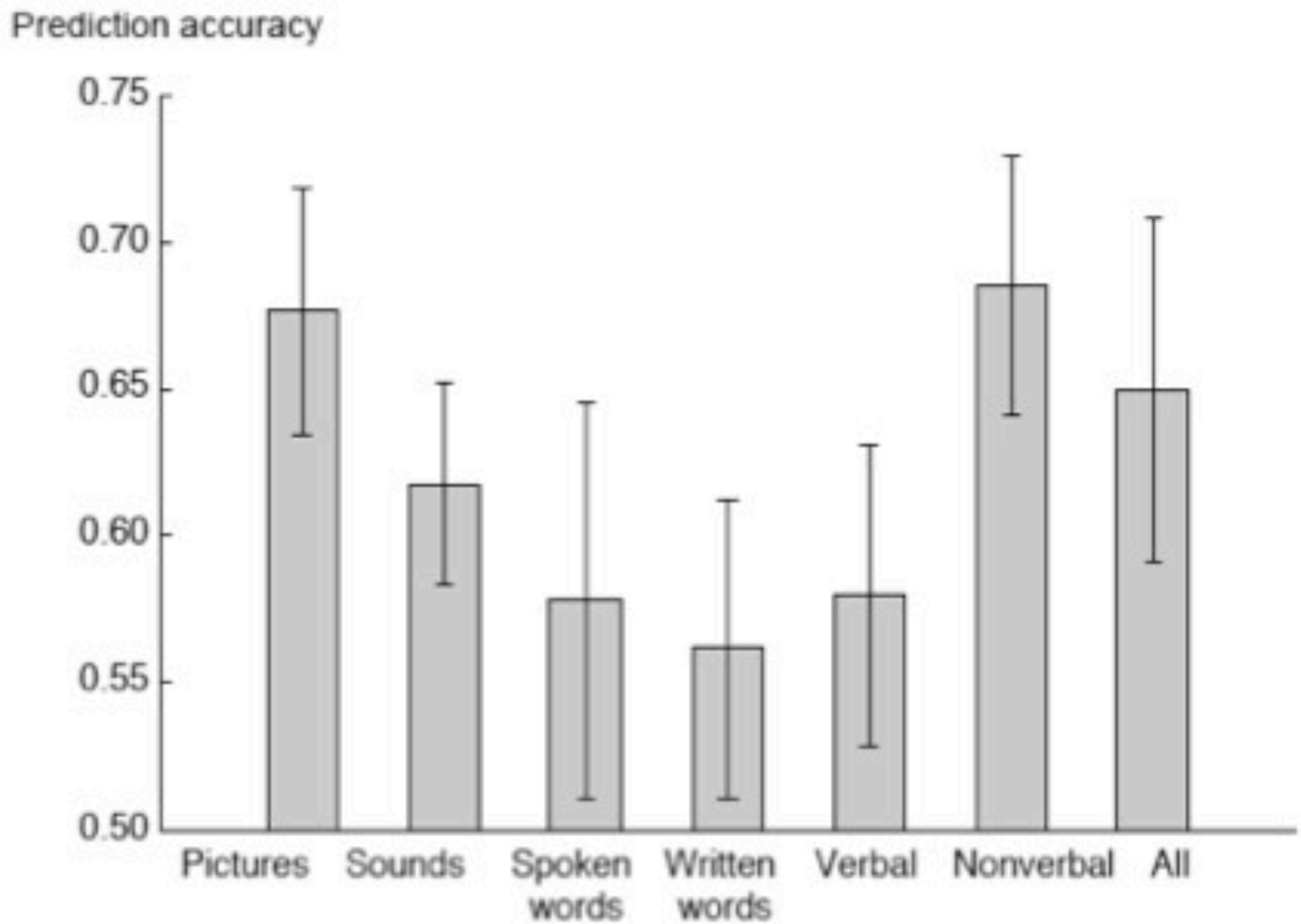
False discovery rate:

1. choose alpha
2. $H_1 \ldots H_m$ hypotheses with $P_{(1)} \ldots P_{(m)}$ p-values ordered in increasing order
3. find largest k such that $P_{(k)} \leq \frac{k}{m}\alpha$
4. reject null hypotheses $H_1 \ldots H_k$

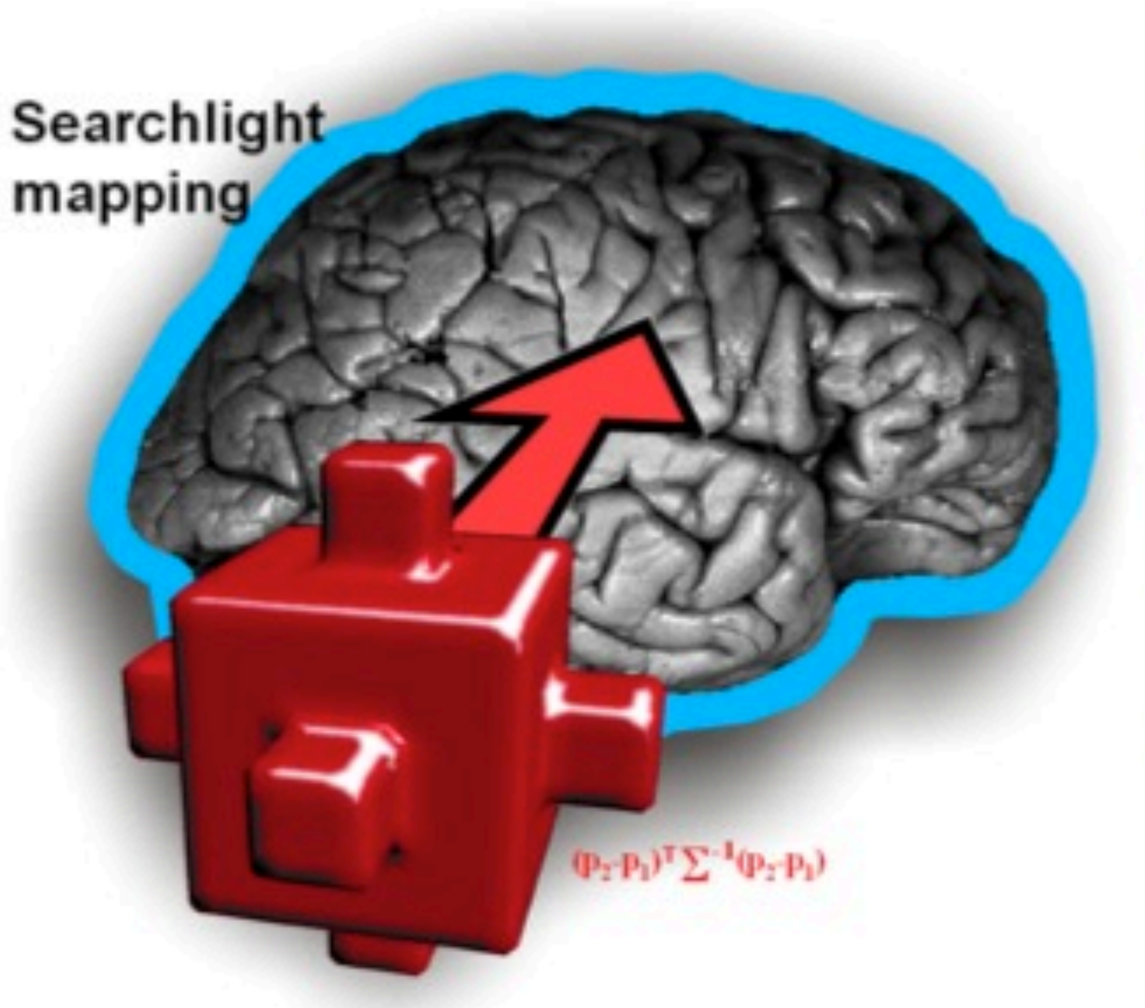
alpha gives expected proportion of false positives among all significant hypotheses

Free-recall test

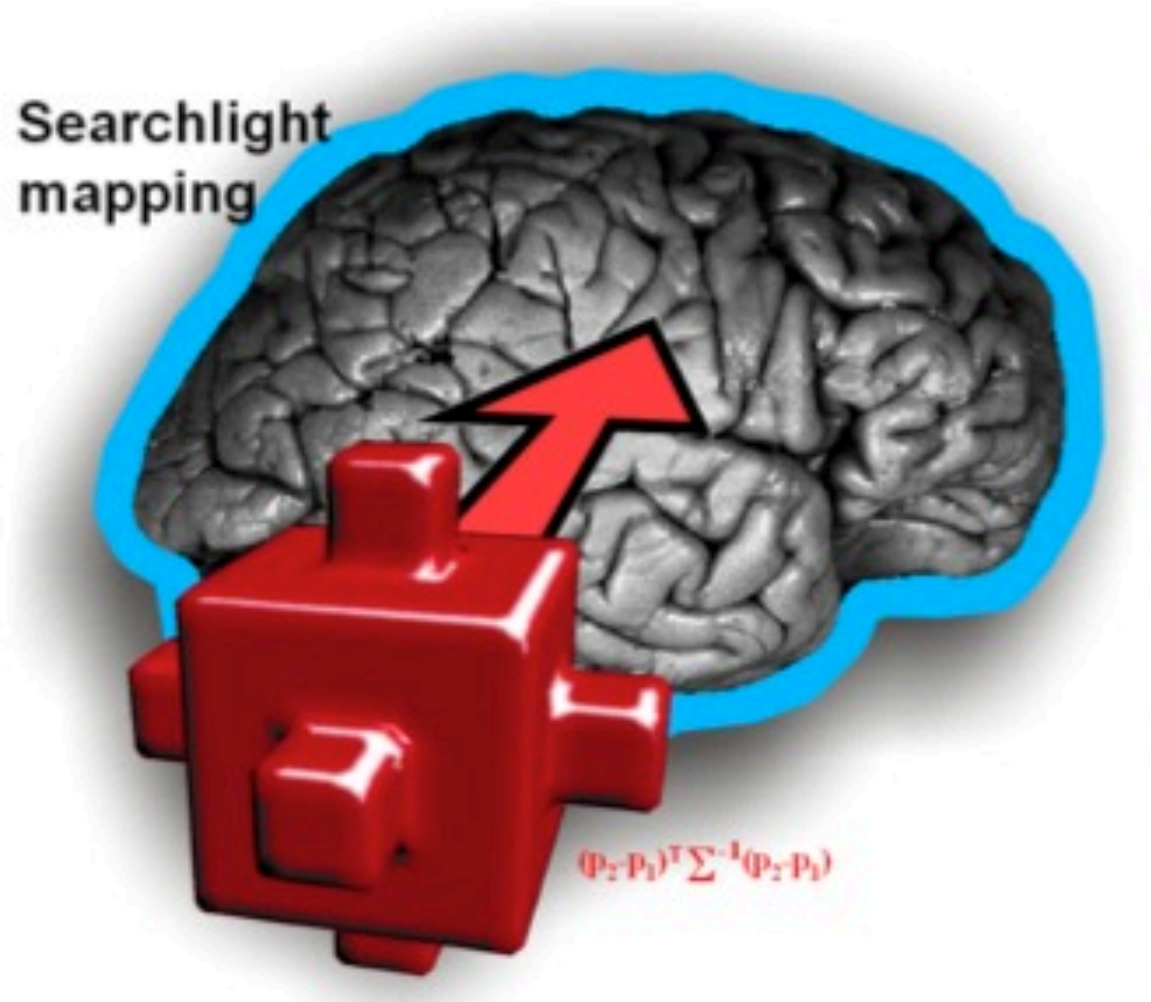
Think about all animals that you have seen during the experiment...
OR
Think about all tools that you have seen....



How to use classifiers to find which regions are functionally involved?



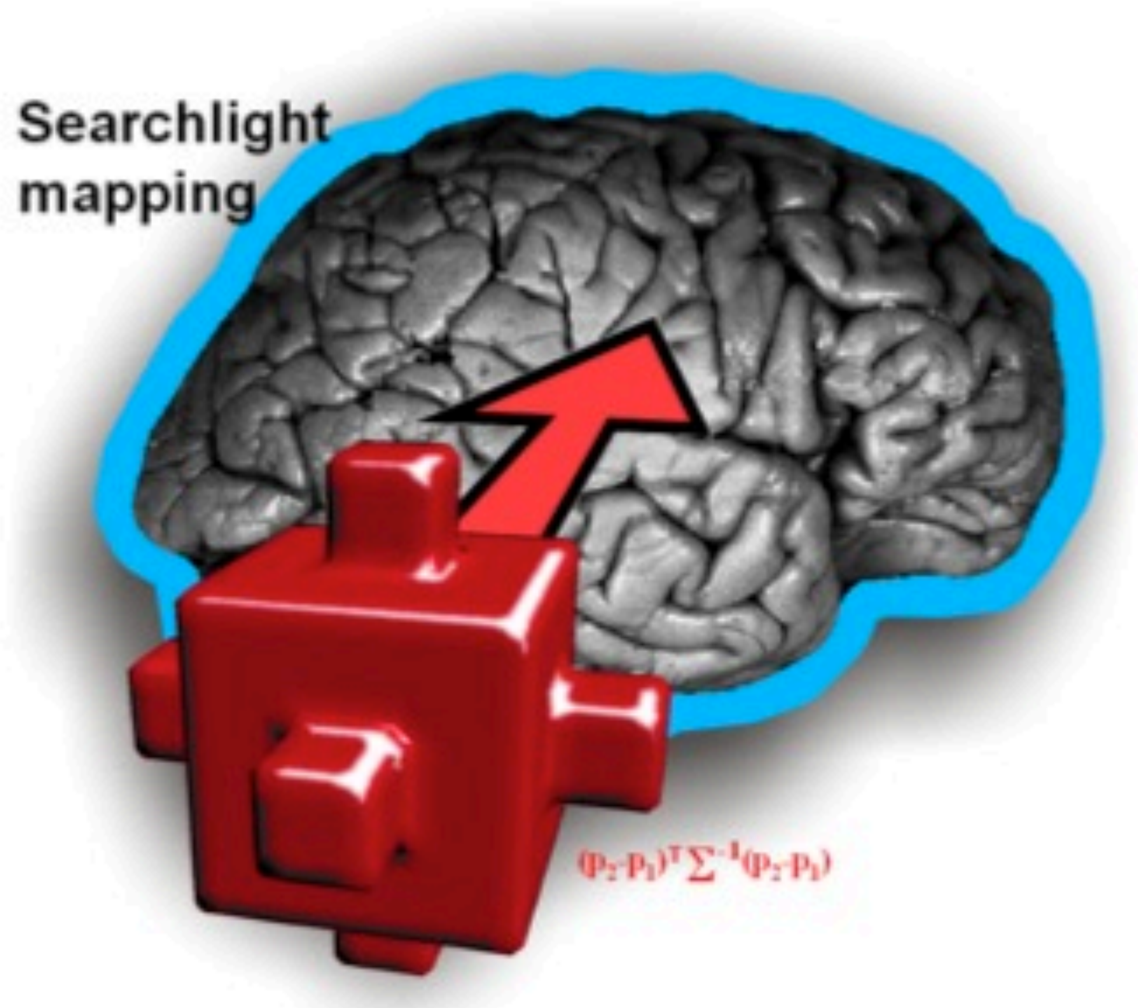
How to use classifiers to find which regions are functionally involved?



- nice estimate of local decoding performance

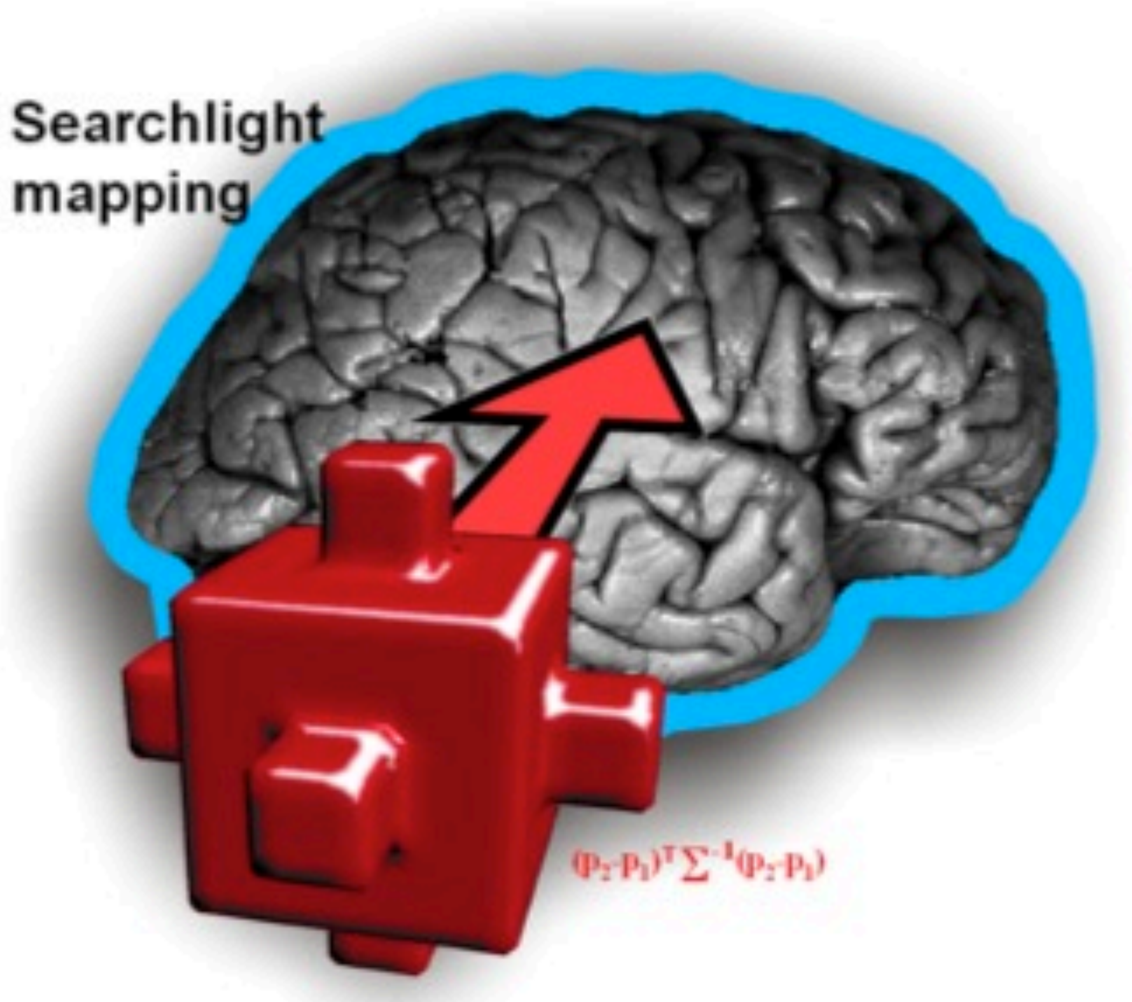
Multivariate pattern analysis

How to use classifiers to find which regions are functionally involved?



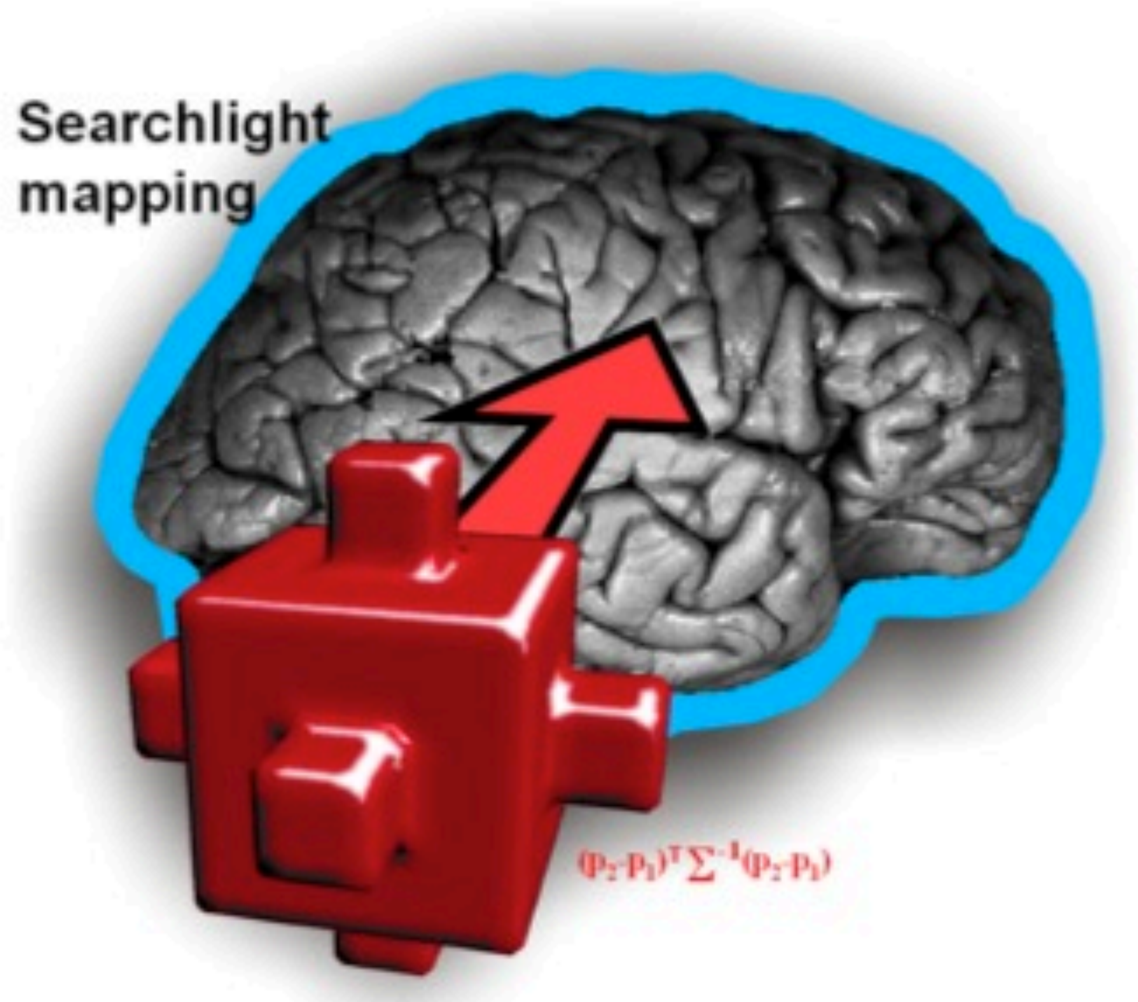
- nice estimate of local decoding performance
- naive approach to feature selection

How to use classifiers to find which regions are functionally involved?



- nice estimate of local decoding performance
- naive approach to feature selection
- huge multiple comparison problem

How to use classifiers to find which regions are functionally involved?



- nice estimate of local decoding performance
- naive approach to feature selection
- huge multiple comparison problem
- closely related to mass-univariate approach