Donders Institute
for Brain, Cognition and Behaviour

Brain Reading (MKI43)

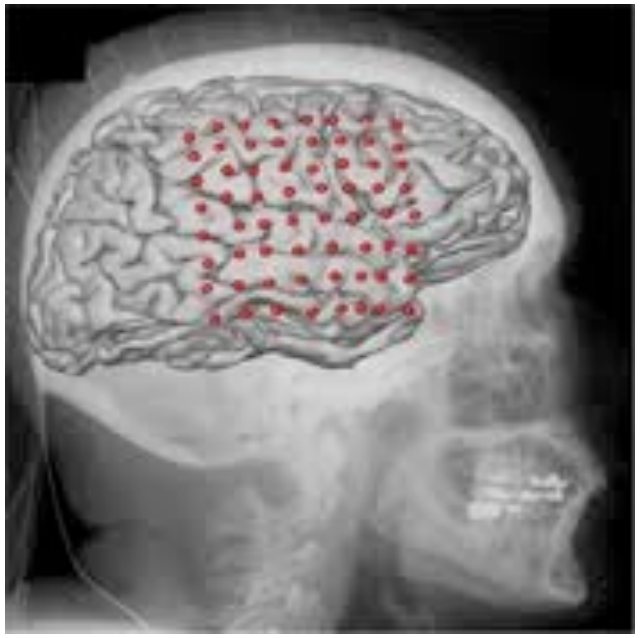
Lecture 1: Introduction

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

Radboud University Nijmegen

Who am I?

- Marcel van Gerven
- Assistant professor in Artificial Intelligence
- Principal investigator at the Donders Institute
- Research group: www.distrep.org
- Understanding brain function using machine learning techniques:
  - Decoding cognitive states from neural data
  - Inferring connectivity patterns from neural data



- Course outline
- Introduction to brain reading:

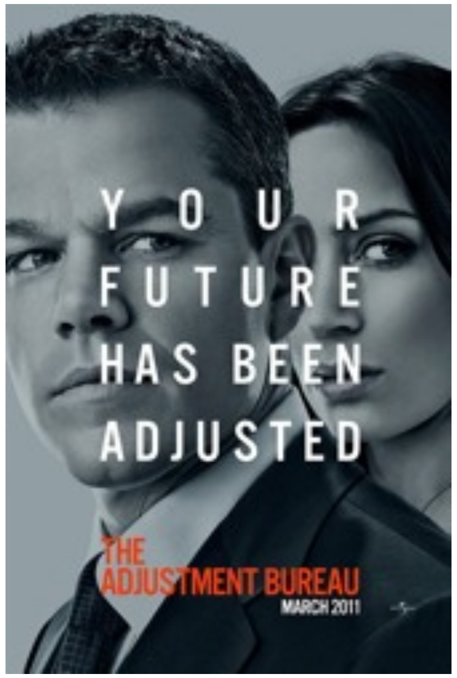
  Working definition:

  Understanding cognitive processing based on brain activity as analysed using sophisticated statistical machine learning techniques

Presented clip

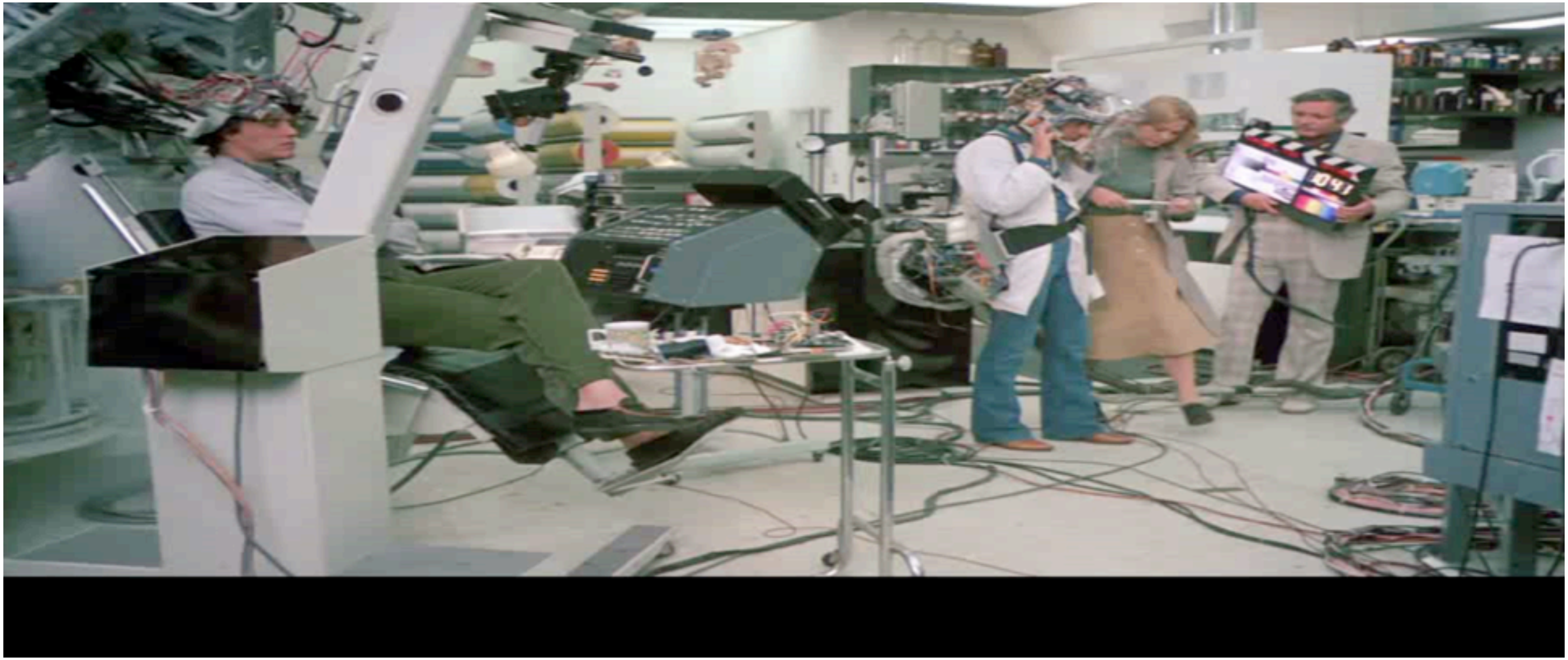
Clip reconstructed from brain activity

Decoding speaks to the imagination

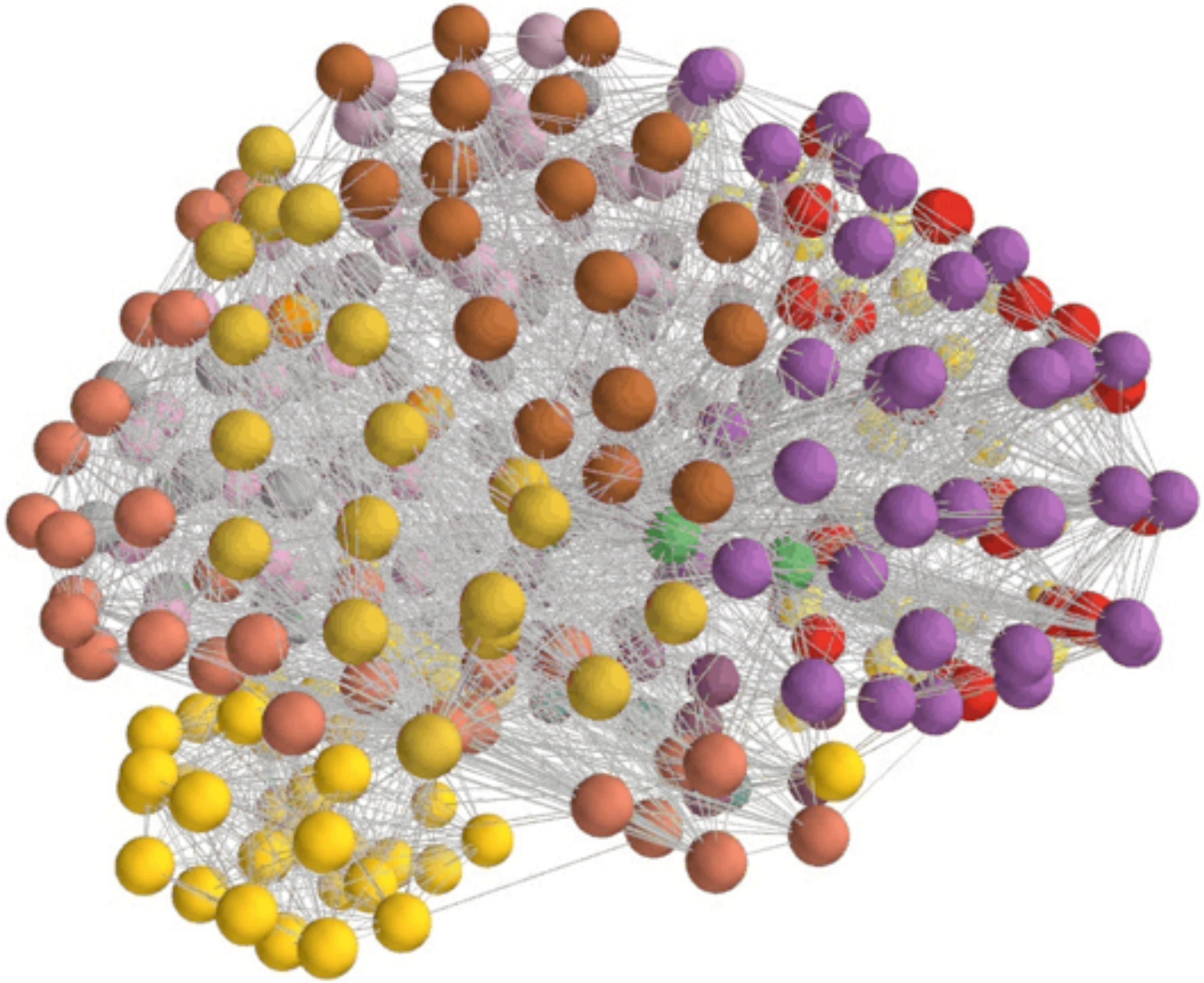


Brainstorm (1983)

Brainstorm (1983)



Connectivity



Dynamic changes

Study goals:

Study goals:

- Understand how to decode cognitive states from neuroimaging data

Study goals:

- Understand how to decode cognitive states from neuroimaging data
- Get acquainted with the mathematics underlying multivariate analysis

Study goals:

- Understand how to decode cognitive states from neuroimaging data
- Get acquainted with the mathematics underlying multivariate analysis
- Analyse (preprocessed) fMRI data using multivariate analysis methods

Study goals:

- Understand how to decode cognitive states from neuroimaging data
- Get acquainted with the mathematics underlying multivariate analysis
- Analyse (preprocessed) fMRI data using multivariate analysis methods
- Read and understand the literature on this topic

Study goals:

- Understand how to decode cognitive states from neuroimaging data
- Get acquainted with the mathematics underlying multivariate analysis
- Analyse (preprocessed) fMRI data using multivariate analysis methods
- Read and understand the literature on this topic
- Being able to place this research in a broader context

The course consists of the following components:

1. Lecture
2. Practical sessions
3. Essay
4. Final exam

For each lecture, required reading (journal papers or book chapter) and slide handouts will be made available in Blackboard.

A second component of the course is the practical. In the practical sessions, you will learn how to perform multivariate analysis of fMRI data. The practical assignments must be handed in before the exam date. Consult the practical manual for further details.

A third component of the course is writing an essay of about five to seven pages (2000-3000 words) on multivariate decoding.

The essay should address a topic that is of relevance in the context of multivariate decoding and will be evaluated on originality, scientific relevance and scientific soundness.

The final mark will be based on three parts:

i. Practical assigments (20%)

ii. Essay on multivariate decoding (30%)

iii. Final written, closed book exam (50%)

The exam consist of a series of open questions.

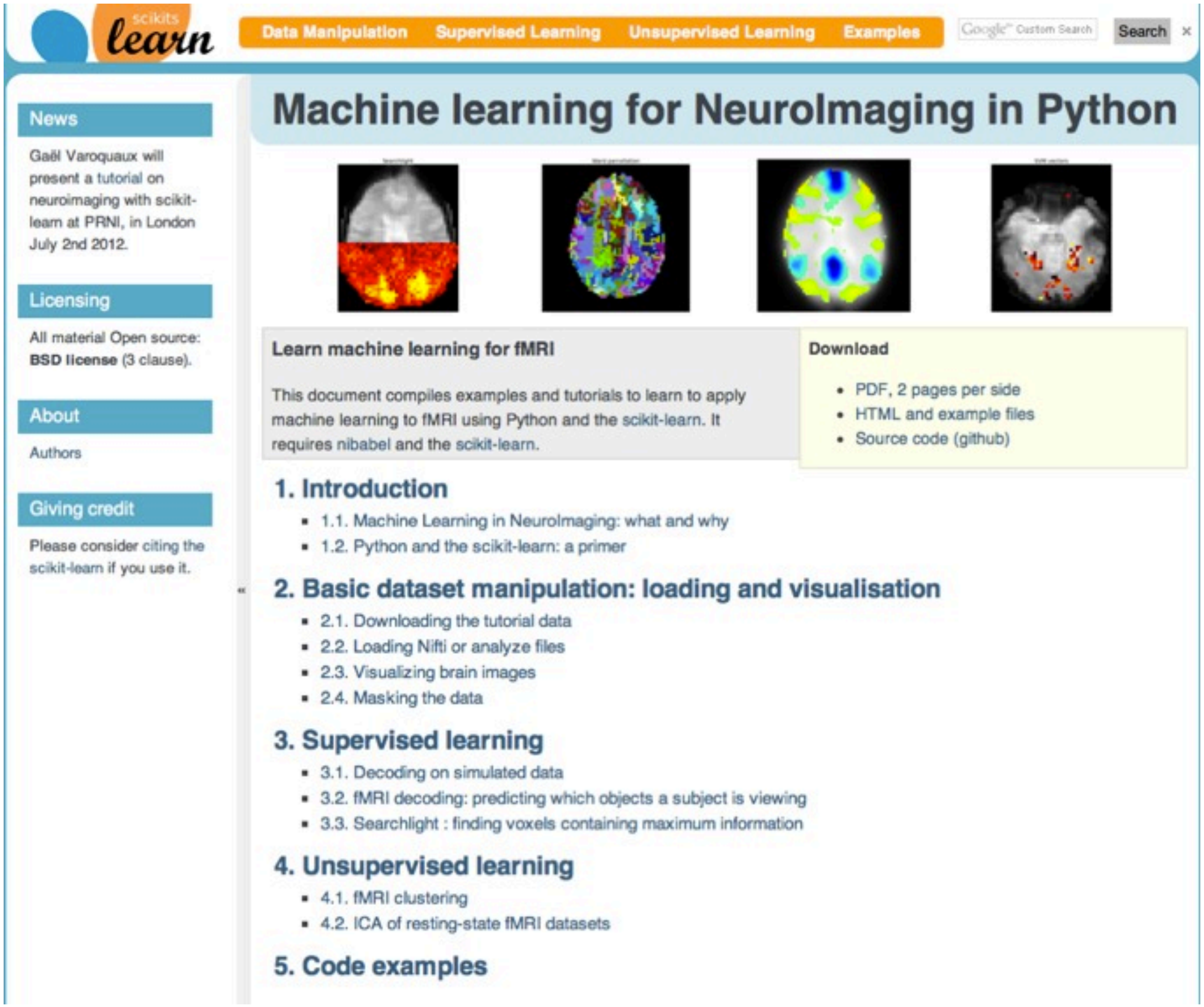
Examinable is:

- ♣ All that has been discussed during lectures
- ♣ All required background reading material

Prerequisites:

- Hand in of practical assignments
- Hand in of essay

http://nisl.github.com


scikits learn

Data Manipulation | Supervised Learning | Unsupervised Learning | Examples

Search

News

Gaël Varoquaux will present a tutorial on neuroimaging with scikit-learn at PRNI, in London July 2nd 2012.

Licensing

All material Open source: **BSD license** (3 clause).

About

Authors

Giving credit

Please consider citing the scikit-learn if you use it.

Machine learning for NeuroImaging in Python

Learn machine learning for fMRI

This document compiles examples and tutorials to learn to apply machine learning to fMRI using Python and the scikit-learn. It requires nibabel and the scikit-learn.

Download

- PDF, 2 pages per side
- HTML and example files
- Source code (github)

1. Introduction

- 1.1. Machine Learning in NeuroImaging: what and why
- 1.2. Python and the scikit-learn: a primer

2. Basic dataset manipulation: loading and visualisation

- 2.1. Downloading the tutorial data
- 2.2. Loading Nifti or analyze files
- 2.3. Visualizing brain images
- 2.4. Masking the data

3. Supervised learning

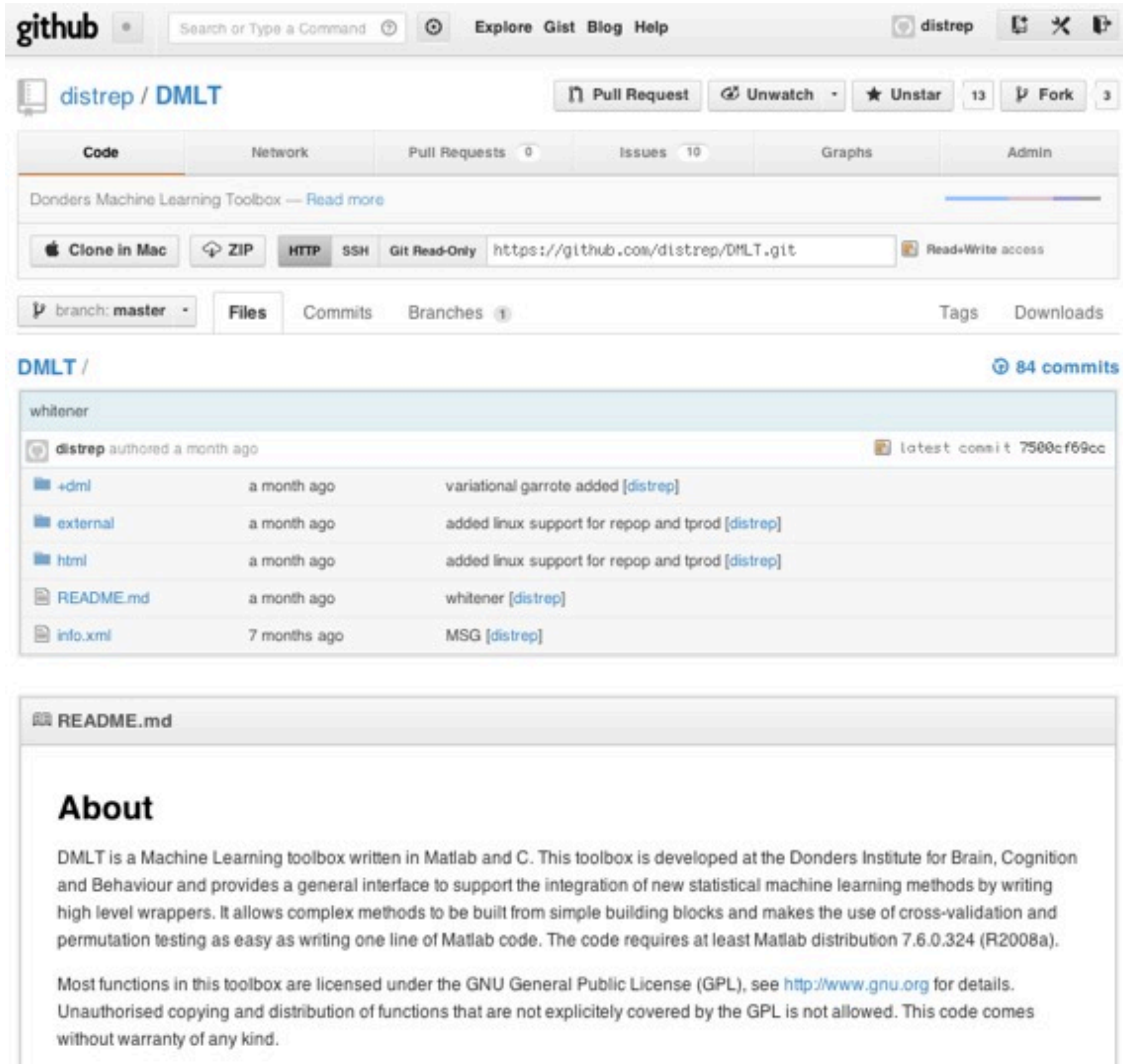
- 3.1. Decoding on simulated data
- 3.2. fMRI decoding: predicting which objects a subject is viewing
- 3.3. Searchlight : finding voxels containing maximum information

4. Unsupervised learning

- 4.1. fMRI clustering
- 4.2. ICA of resting-state fMRI datasets

5. Code examples

also see: https://github.com/distrep/DMLT



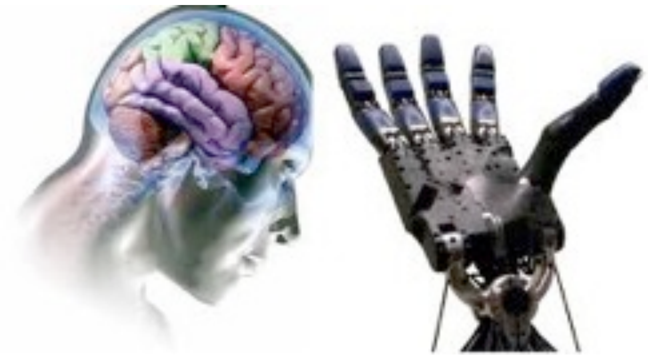
Finally

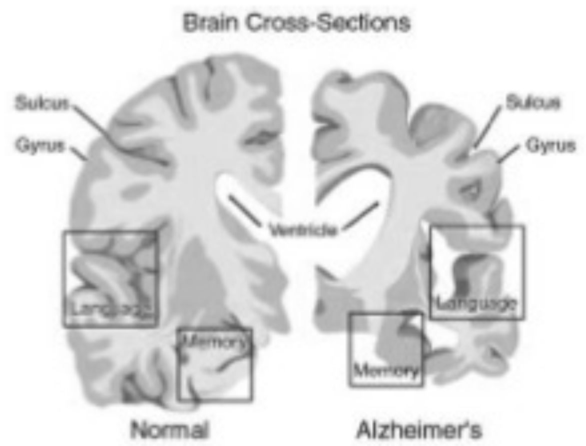
- Check Blackboard for announcements/deadlines/course material
- Check out the Reader and Practical Manual
- Study background literature belonging to each lecture (exam material)
- Interactive lectures (don't be afraid to ask questions or give input!)

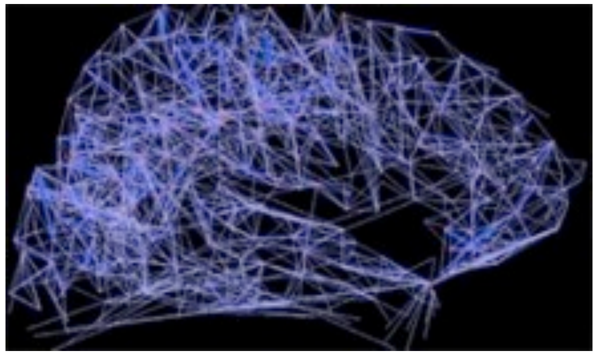
Introduction to brain reading

statistical machine learning

prediction ← → inference







Going beyond the standard GLM analysis

Why Statistical machine learning

Probabilistic models as a lingua franca

Probabilistic models as a lingua franca

Statistics as inverse probability theory:

$$p(\theta \mid D) = \frac{p(D \mid \theta)p(\theta)}{P(D)}$$

"Getting from the data to the models"

Probabilistic models as a lingua franca

Statistics as inverse probability theory:

$$p(\theta \mid D) = \frac{p(D \mid \theta)p(\theta)}{P(D)}$$

"Getting from the data to the models"

Bayesian statistics as the theoretically optimal way to
- analyze neural data
- think about brain function

Exploratory analysis based on large-scale data-sharing initiatives

The fMRI Data Center
fMRIDC

ADNI Alzheimer's Disease Neuroimaging Initiative

HUMAN Connectome PROJECT

NEUROTYCHO Beta

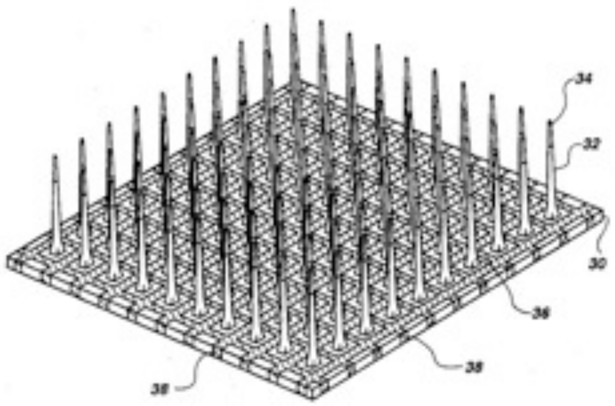
ALLEN INSTITUTE for BRAIN SCIENCE
Fueling Discovery

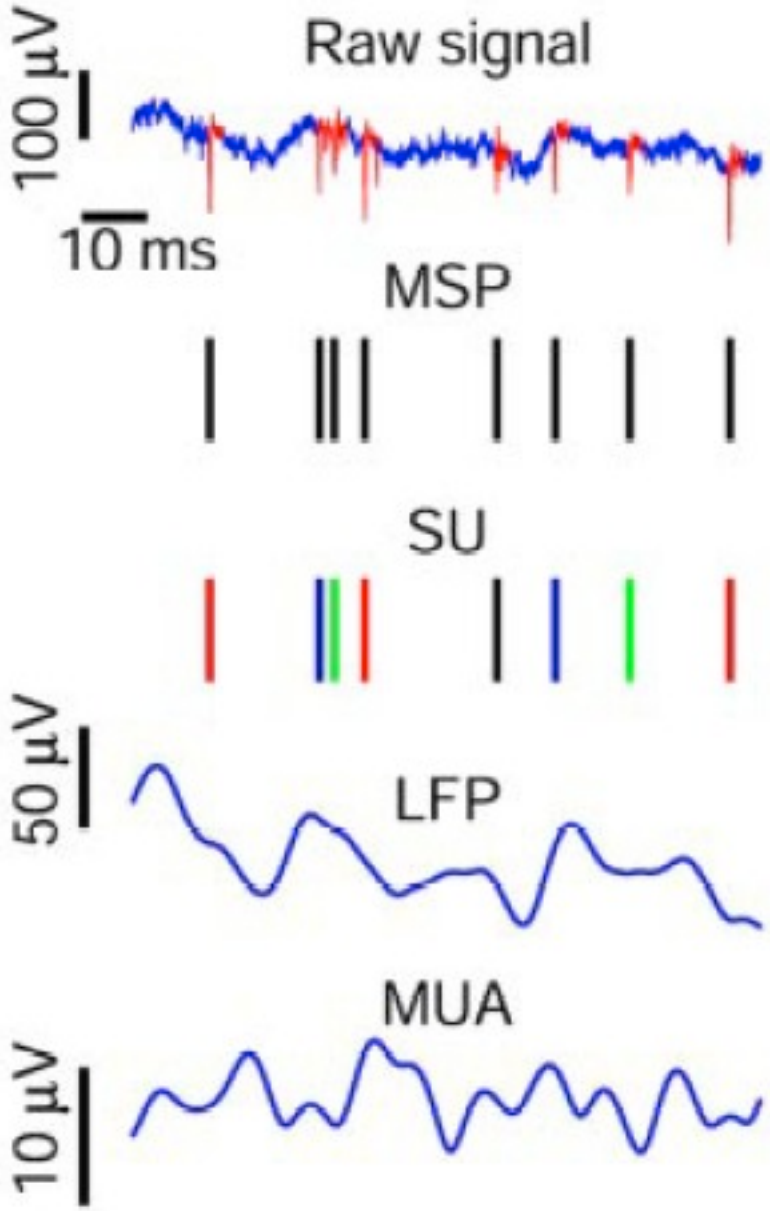
Challenges and Opportunities in Mining Neuroscience Data

Huda Akil,[1*] Maryann E. Martone,[2] David C. Van Essen[3]

Science, 2011

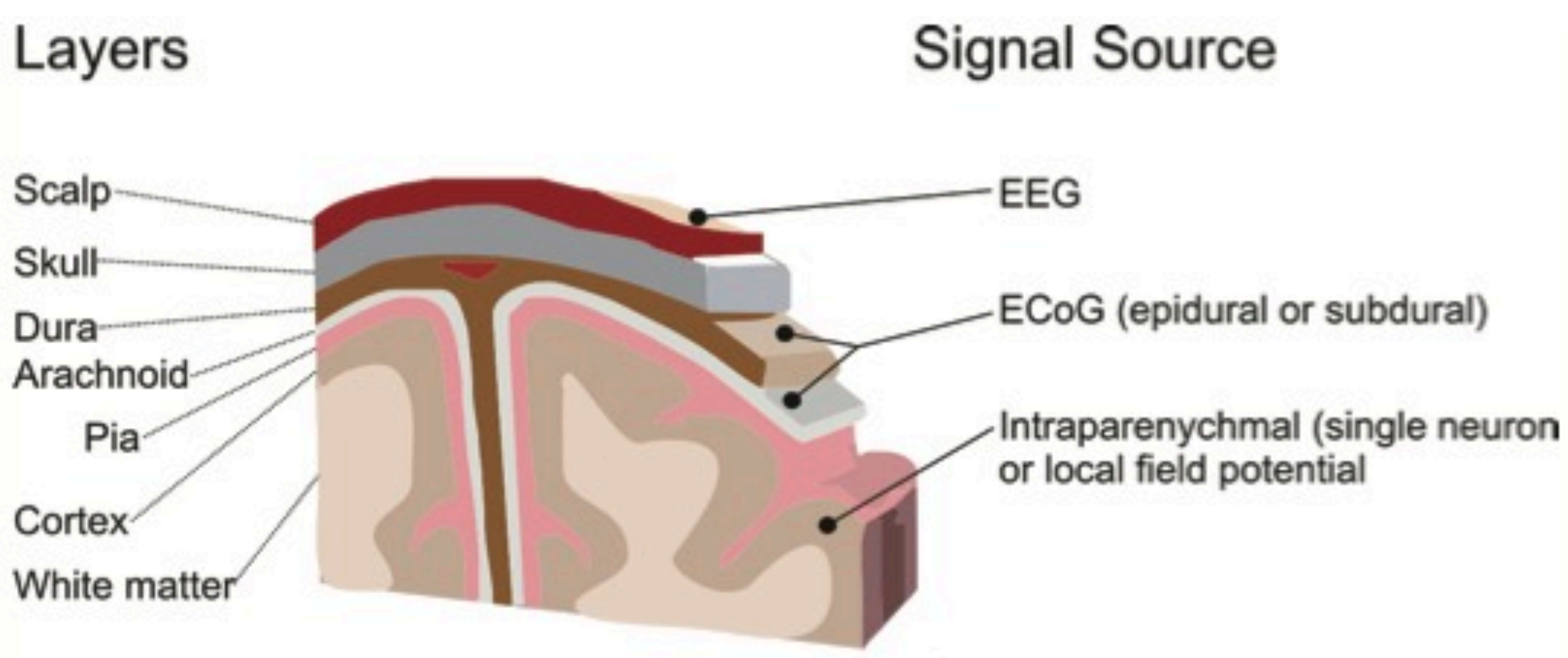
Neuroimaging data: electrophysiology





Neuroimaging data: electrophysiology

Neuroimaging data: electrophysiology

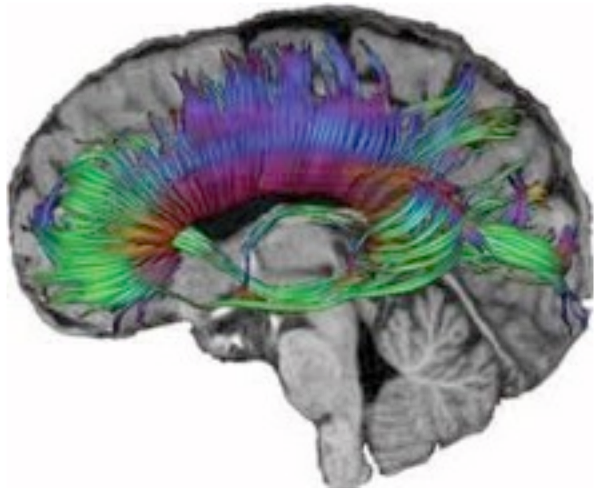


Leuthardt, E. C., Schalk, G., Roland, J., Rouse, A., & Moran, D. W. (2009). Evolution of brain-computer interfaces: going beyond classic motor physiology Neurosurgical focus, 27(1)

Neuroimaging data: MRI

functional

structural

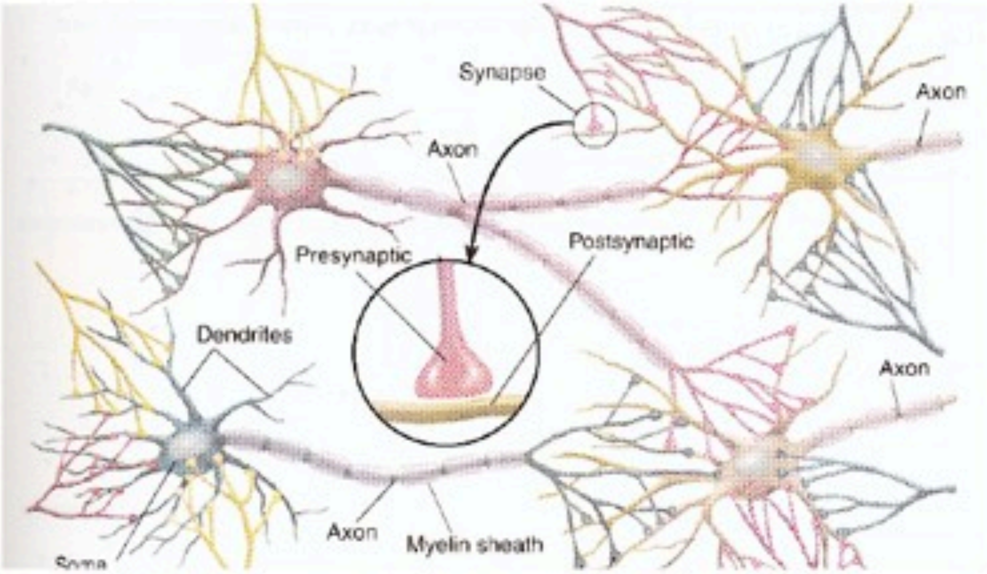


diffusion

The BOLD Contrast

BOLD (Blood Oxygenation Level Dependent) contrast = measures inhomogeneities in the magnetic field due to changes in the level of O_2 in the blood

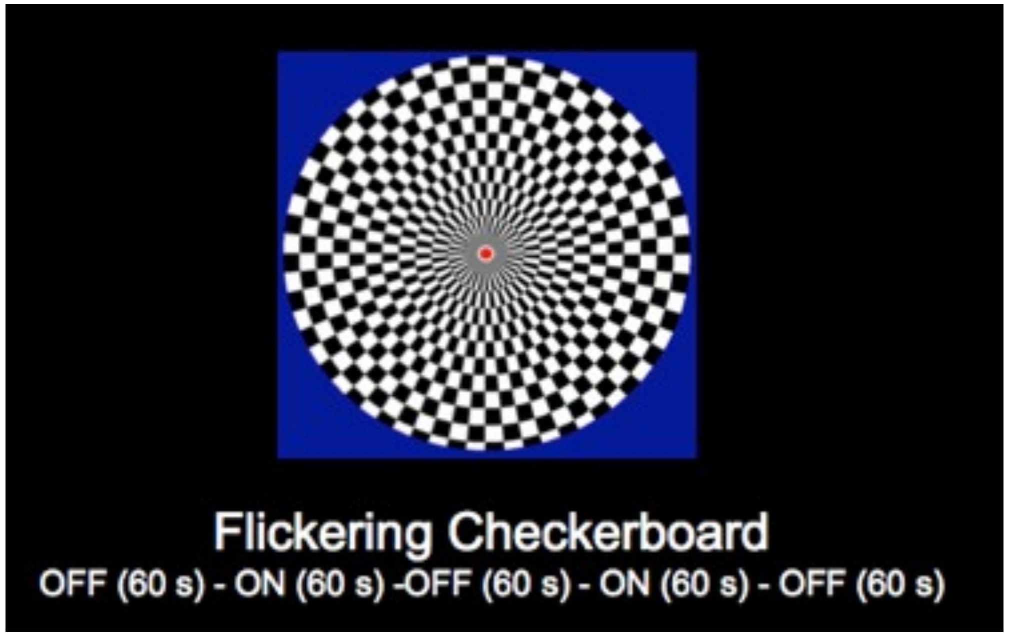
Current conclusion: BOLD signal seems to be more strongly correlated to postsynaptic activity

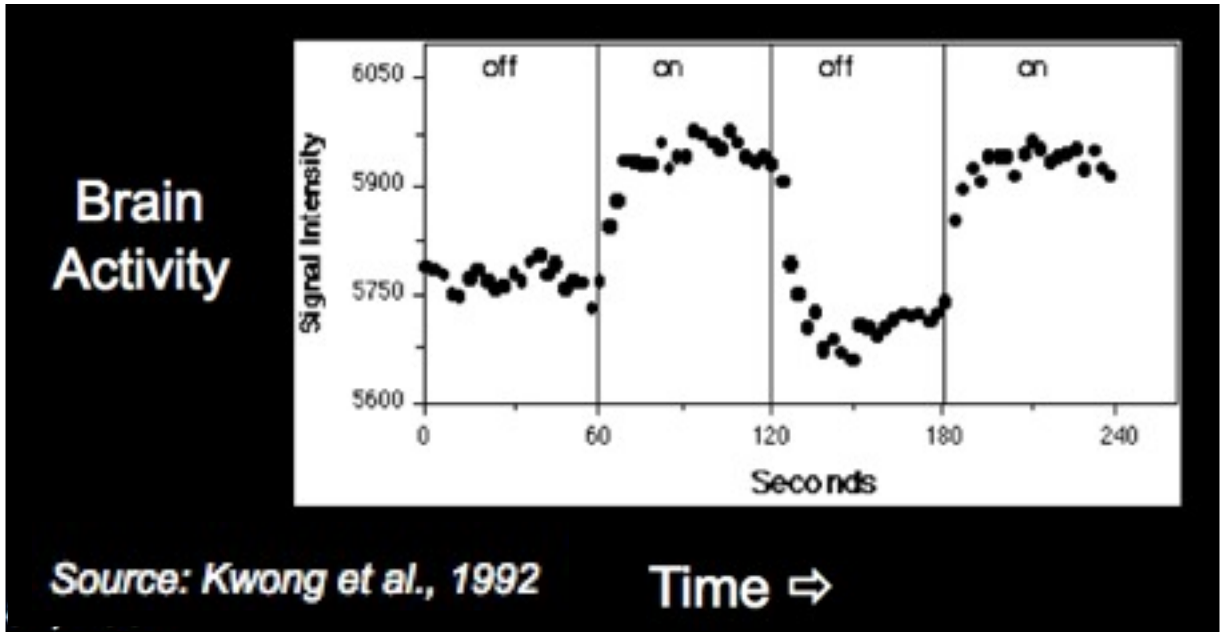


BOLD seems to reflect the input to a neuronal population as well as its intrinsic processing.

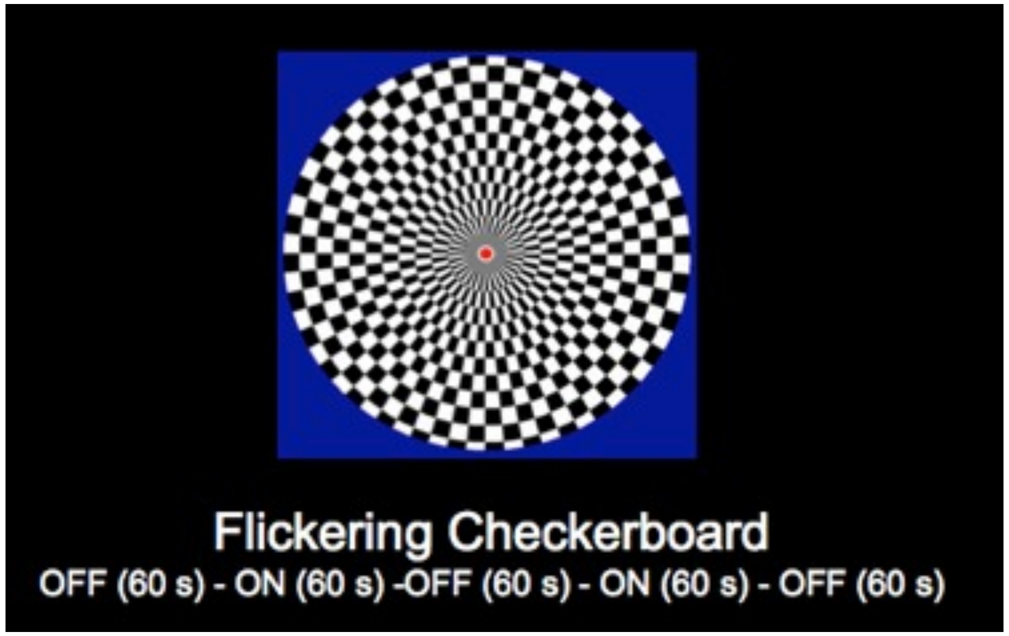
Lauritzen 2005, *Nat. Neurosci. Rev.*

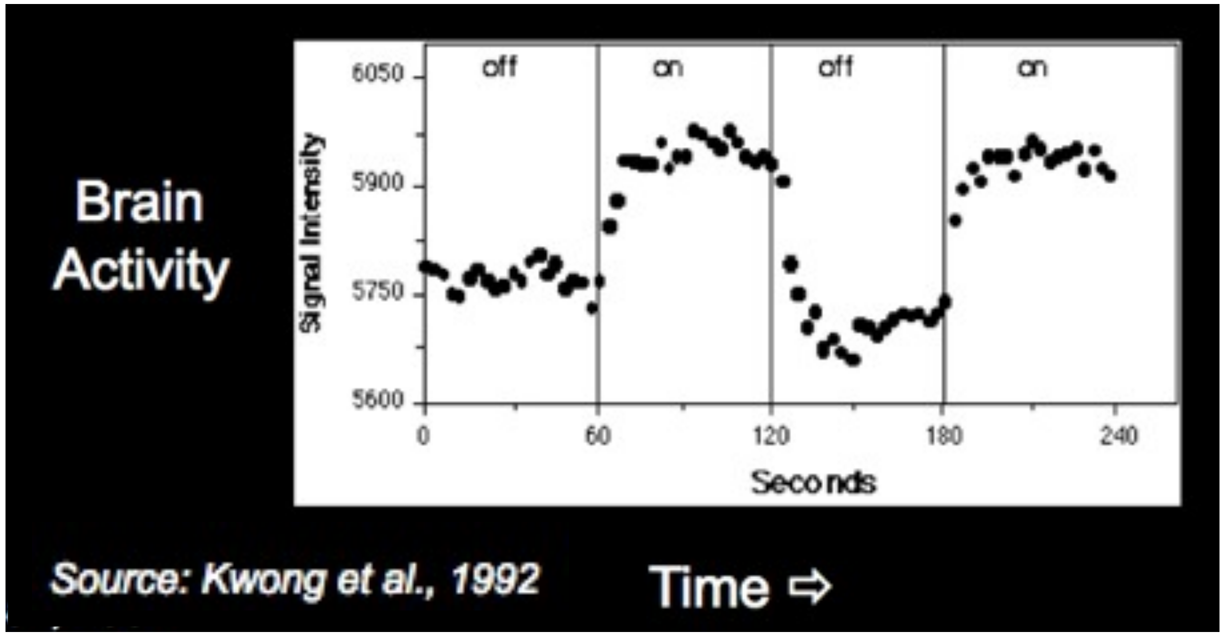
Hemodynamic response

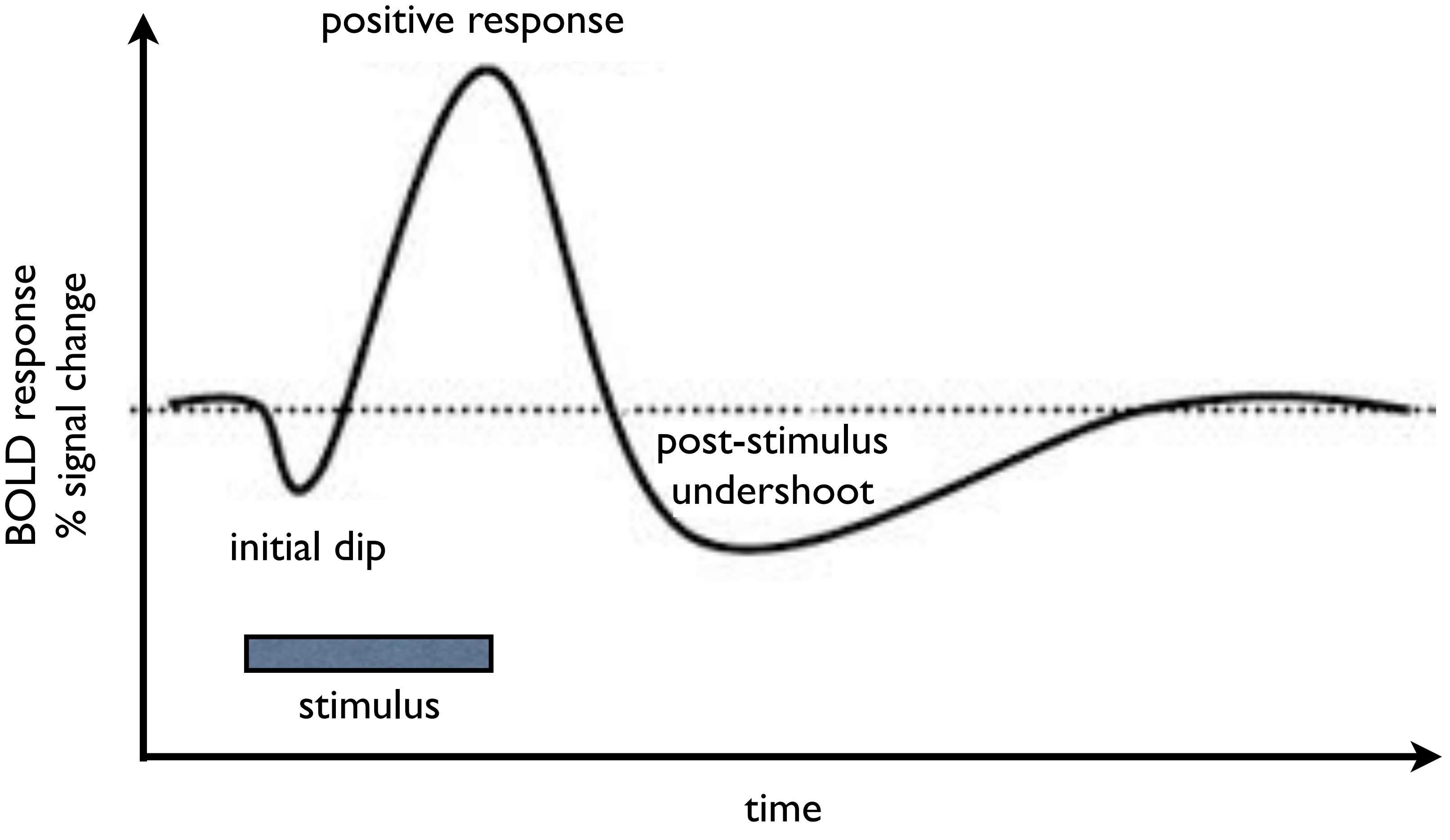


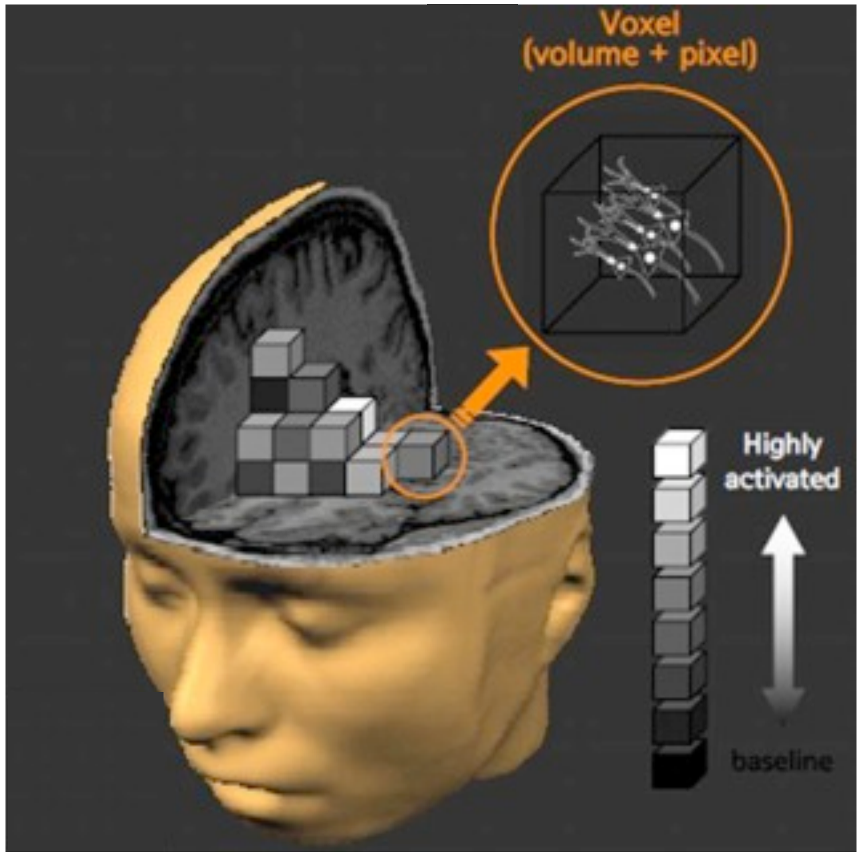


Hemodynamic response



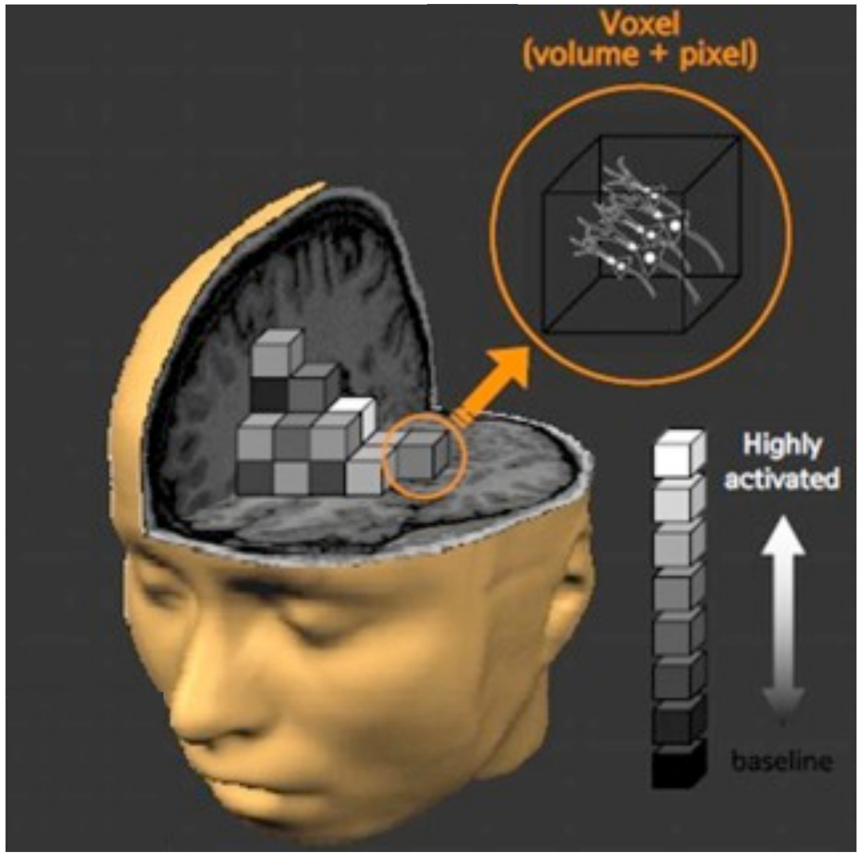




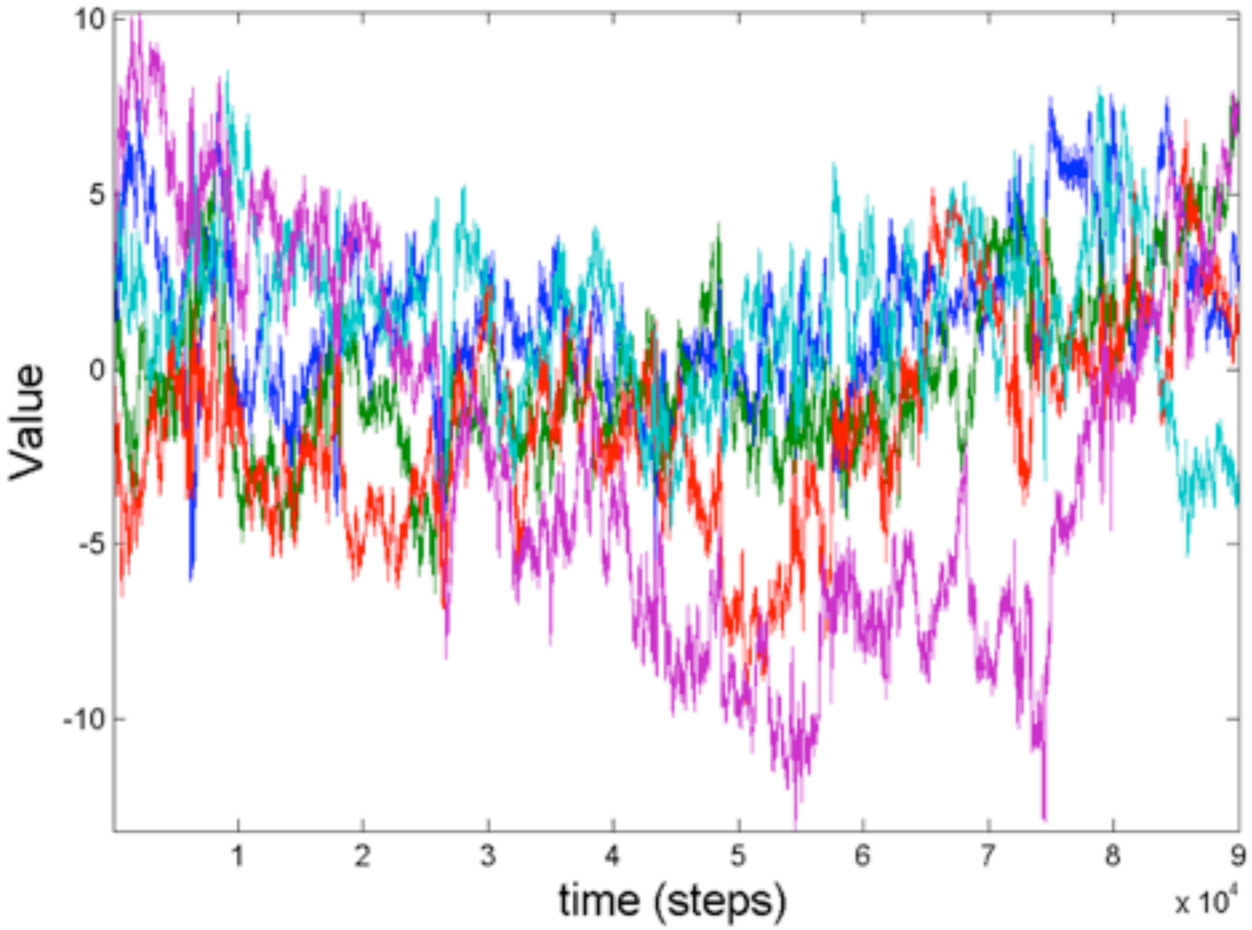


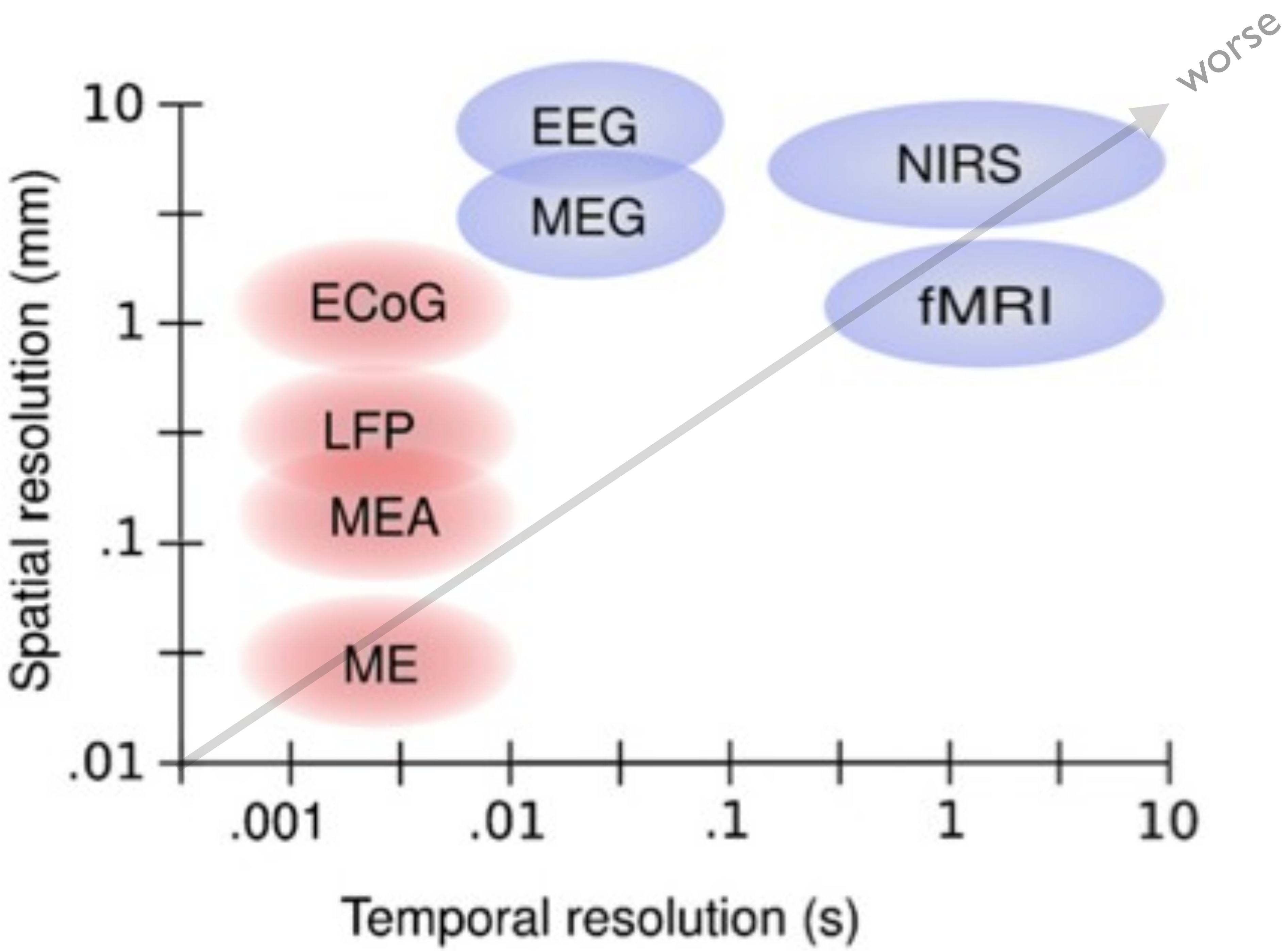
Acquisition of about 50000 2x2x2 mm voxels every 2 seconds

Acquired data



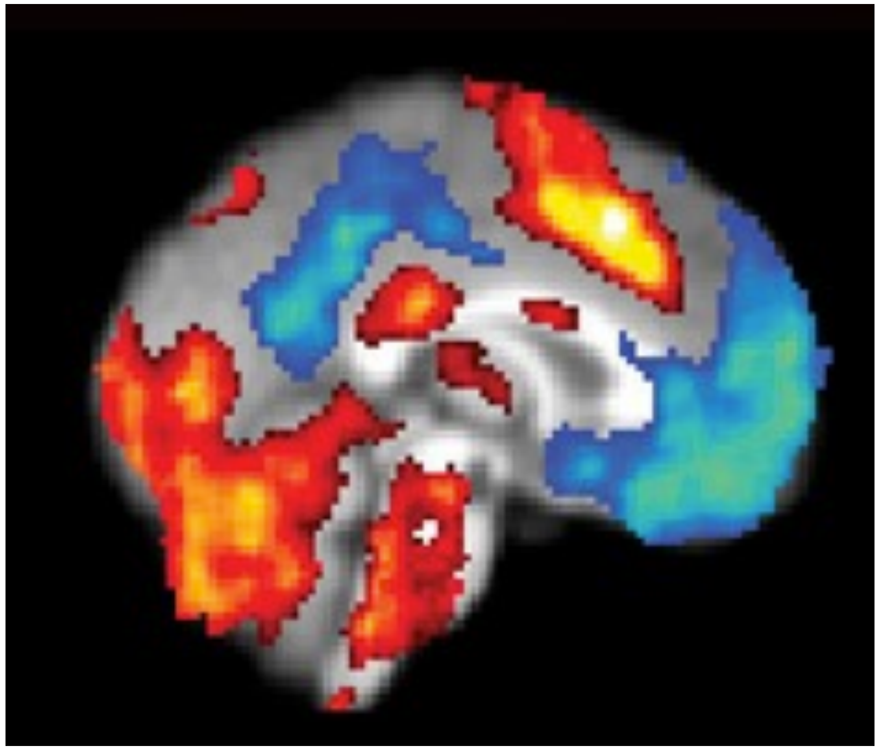
Acquisition of about 50000 2x2x2 mm voxels every 2 seconds




10
1
.1
.01
Spatial resolution (mm)
.001
.01
.1
1
10
Temporal resolution (s)
EEG
MEG
NIRS
fMRI
ECoG
LFP
MEA
ME
worse

Local analyses of regionally-specific effects

Mass-univariate analysis:
Conventional analysis based on the general linear model

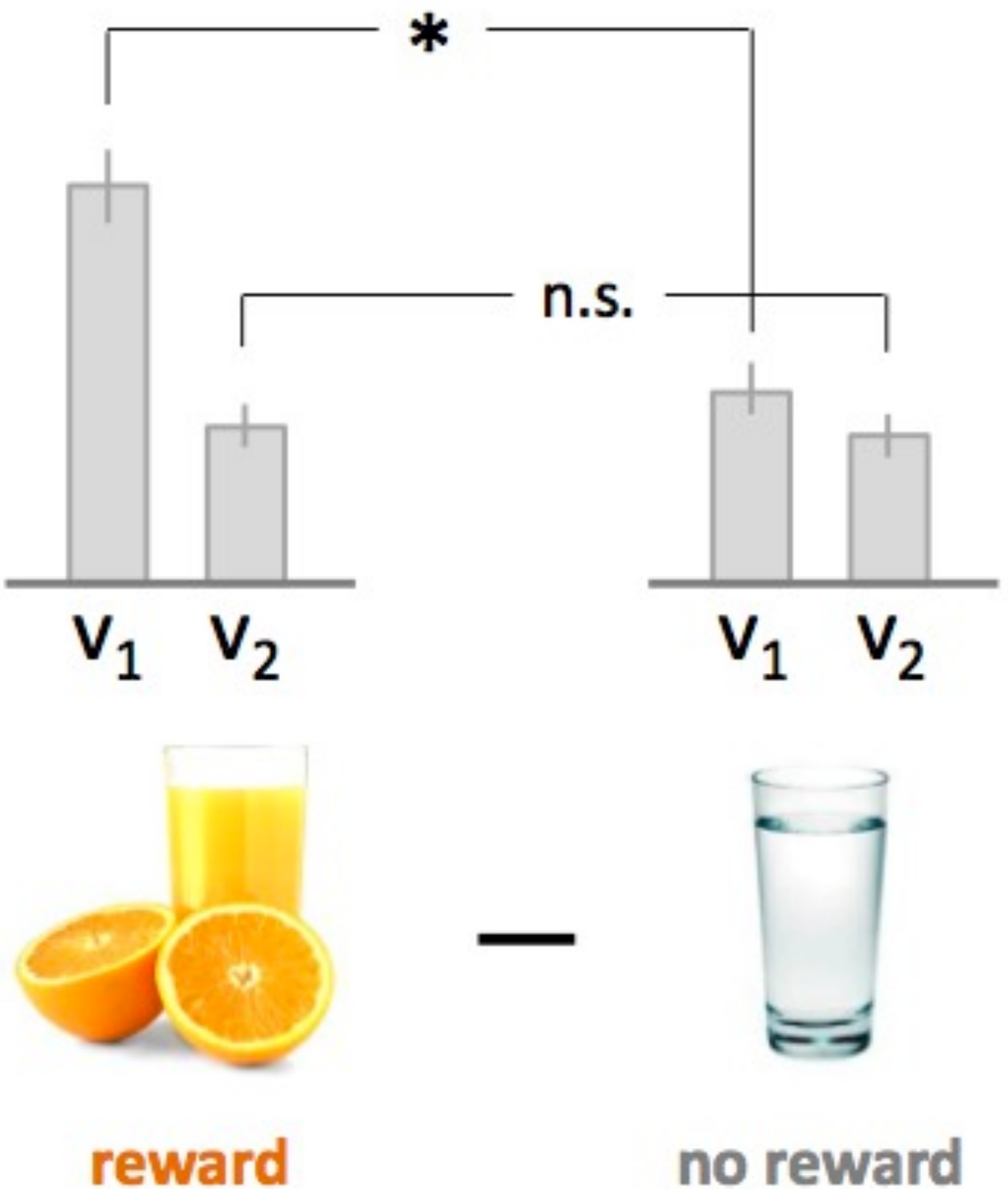


Multivariate analysis:
Multivariate analysis which takes distributed patterns of activity into account

Which regions are activated during a cognitive task?

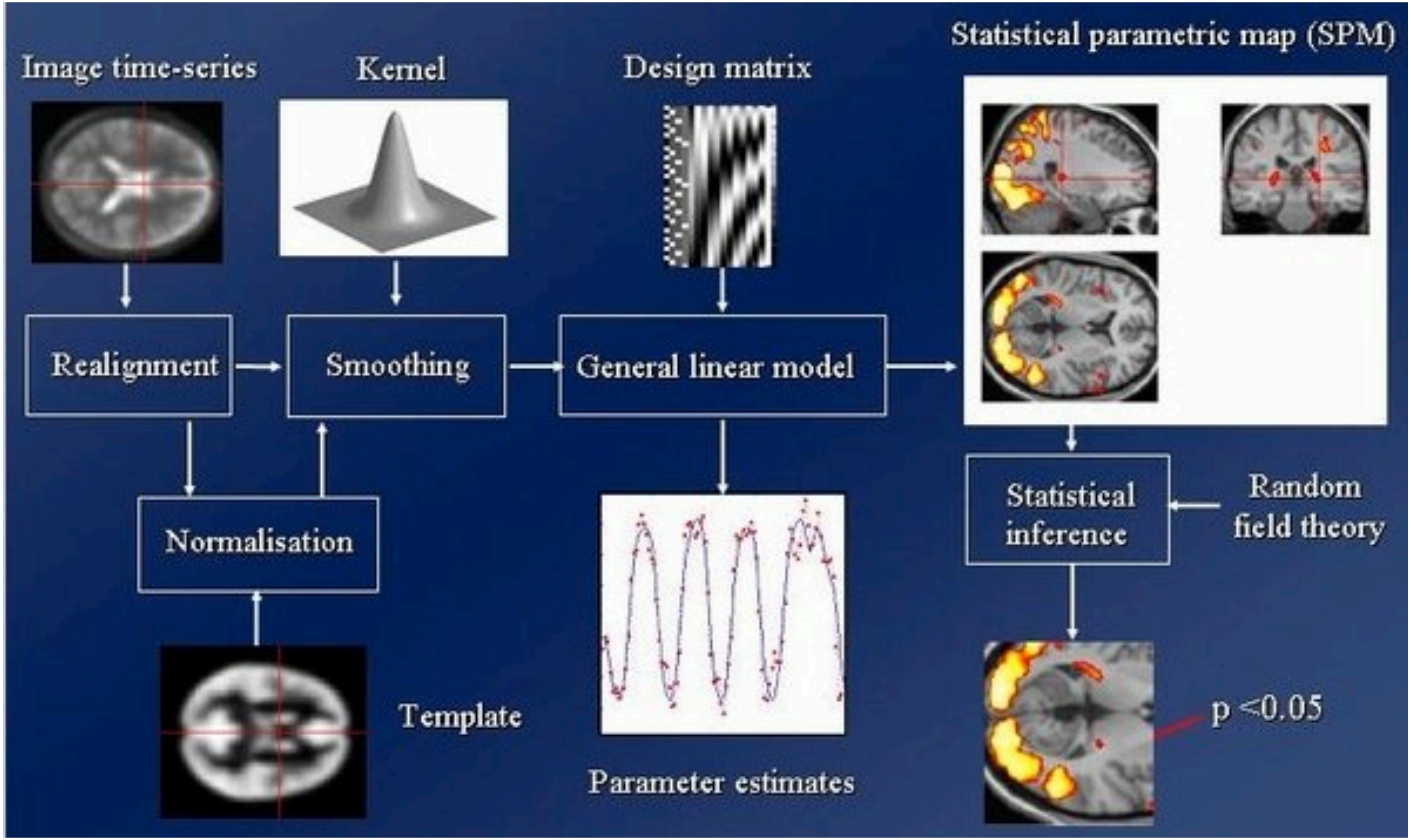
Functional segregation: Mass-univariate analysis

- analyzes each voxel independently
- excellent for localizing robust activations in individual voxels



Slide courtesy Kai Brodersen

Mass-univariate analysis processing pipeline



[www.fil.ion.ucl.ac.uk/**spm**](http://www.fil.ion.ucl.ac.uk/spm)

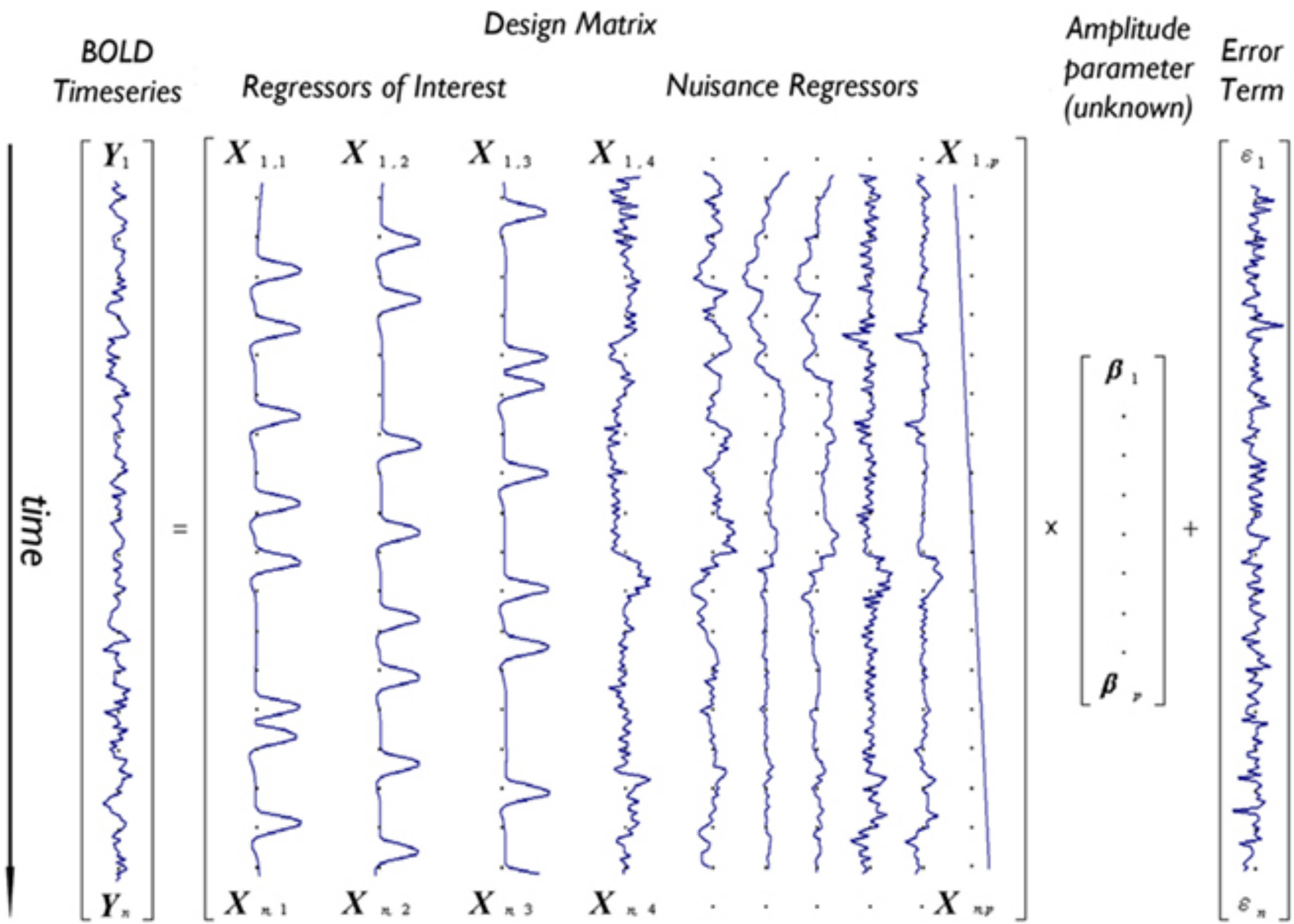
Univariate analysis of regionally-specific effects

$$\underset{N \times K}{\mathbf{Y}} = \underset{N \times P}{\mathbf{X}}\underset{P \times K}{\mathbf{B}} + \underset{N \times K}{\mathbf{U}}$$

voxel time-series | design matrix | regression coefficients | residual

N = number of time points
K = number of voxels
P = number of regressors

Example for one voxel



Estimation of the regression coefficients via ordinary least squares:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Regression coefficients inform about which regressors predict a response

Interest is normally in some of the regressors; modeled using a contrast vector:

$c^T = [1\ 0\ 0 \ldots]$ contribution of first regressor to the data

$c^T = [1\ {-1}\ 0 \ldots]$ whether the first regressor is more active than the second

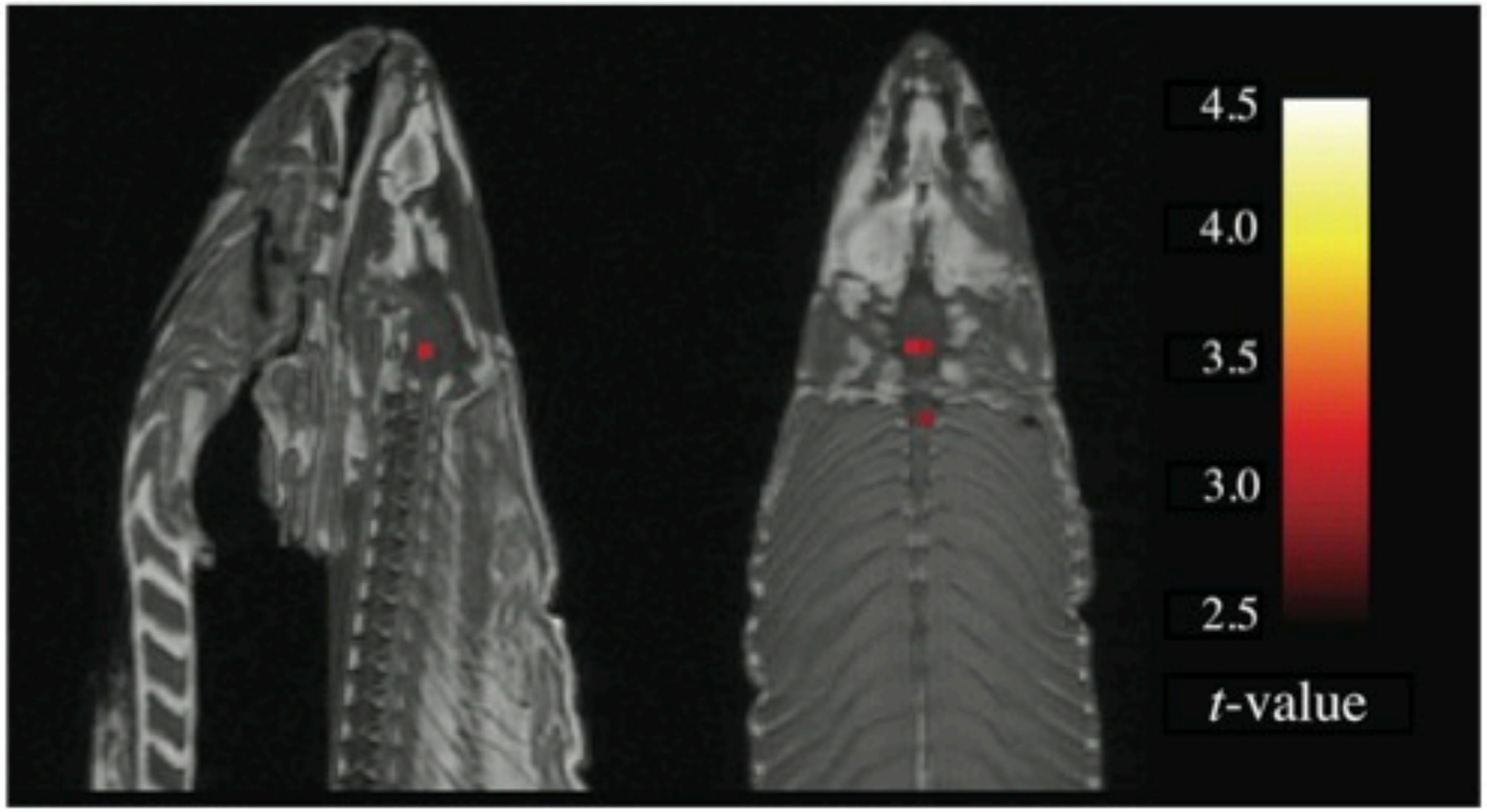
Whether a voxel is active under a contrast can (for example) be estimated using a t-statistic:

$$t_{\mathrm{df}} = \frac{c^T \hat{\beta}}{\sqrt{\mathrm{var}(c^T \hat{\beta})}}$$

For Gaussian errors, t_{df} approximately follows a Student's t-distribution with *df* degrees of freedom.

The p-value of t_{df} is then computed by testing against the null-distribution.

A voxel is declared significantly activated if the p-value is below some fixed alpha level.



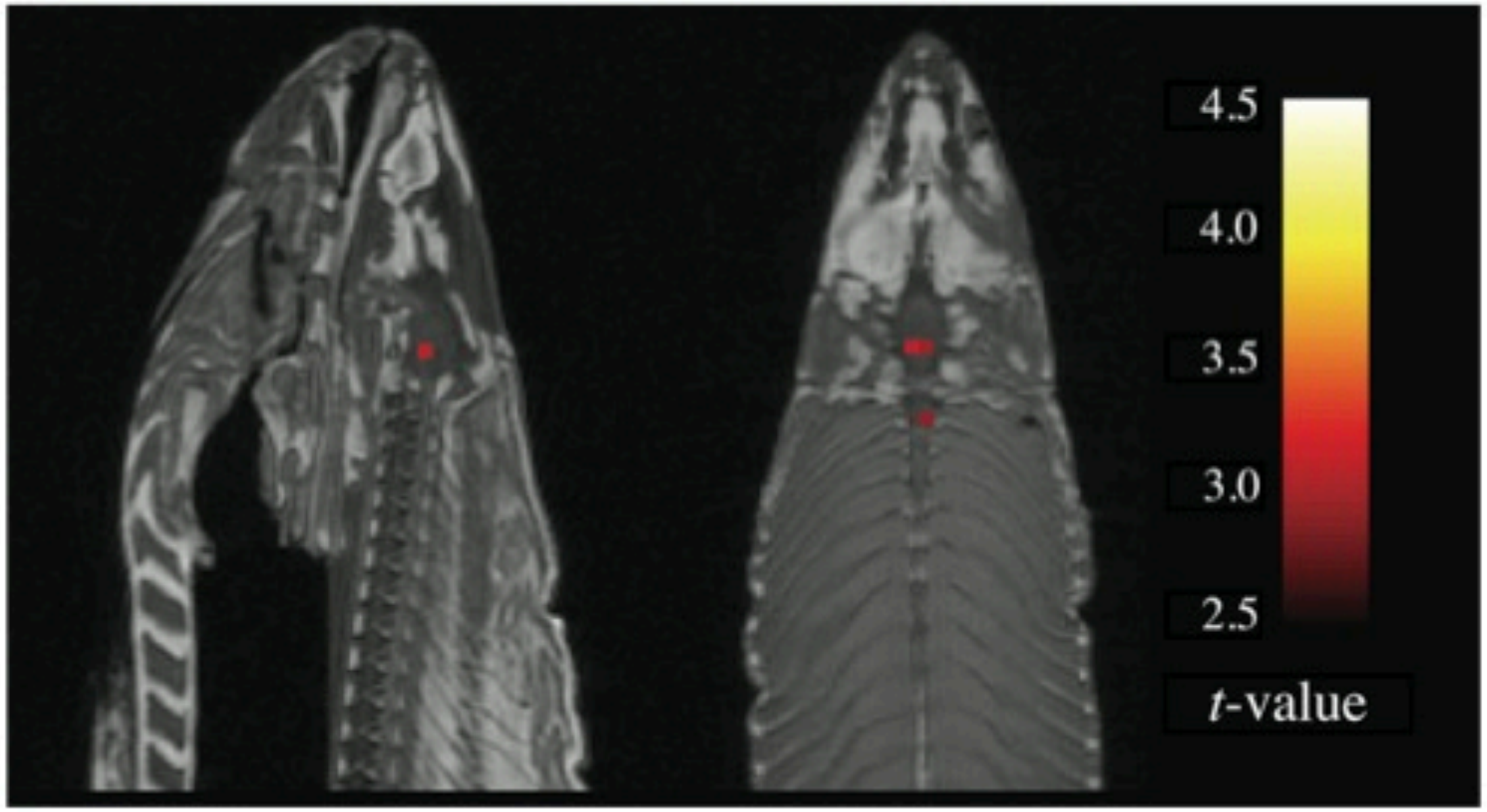
A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, p(uncorrected) < 0.001, 3 voxel extent threshold.

Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction

by: Craig M. Bennett, Abigail A. Baird, Michael B. Miller, George L. Wolford

Journal of Serendipitous and Unexpected Results, Vol. 1 (2011), pp. 1-5 Key: citeulike:9170380

Mass-univariate analysis: the general linear model



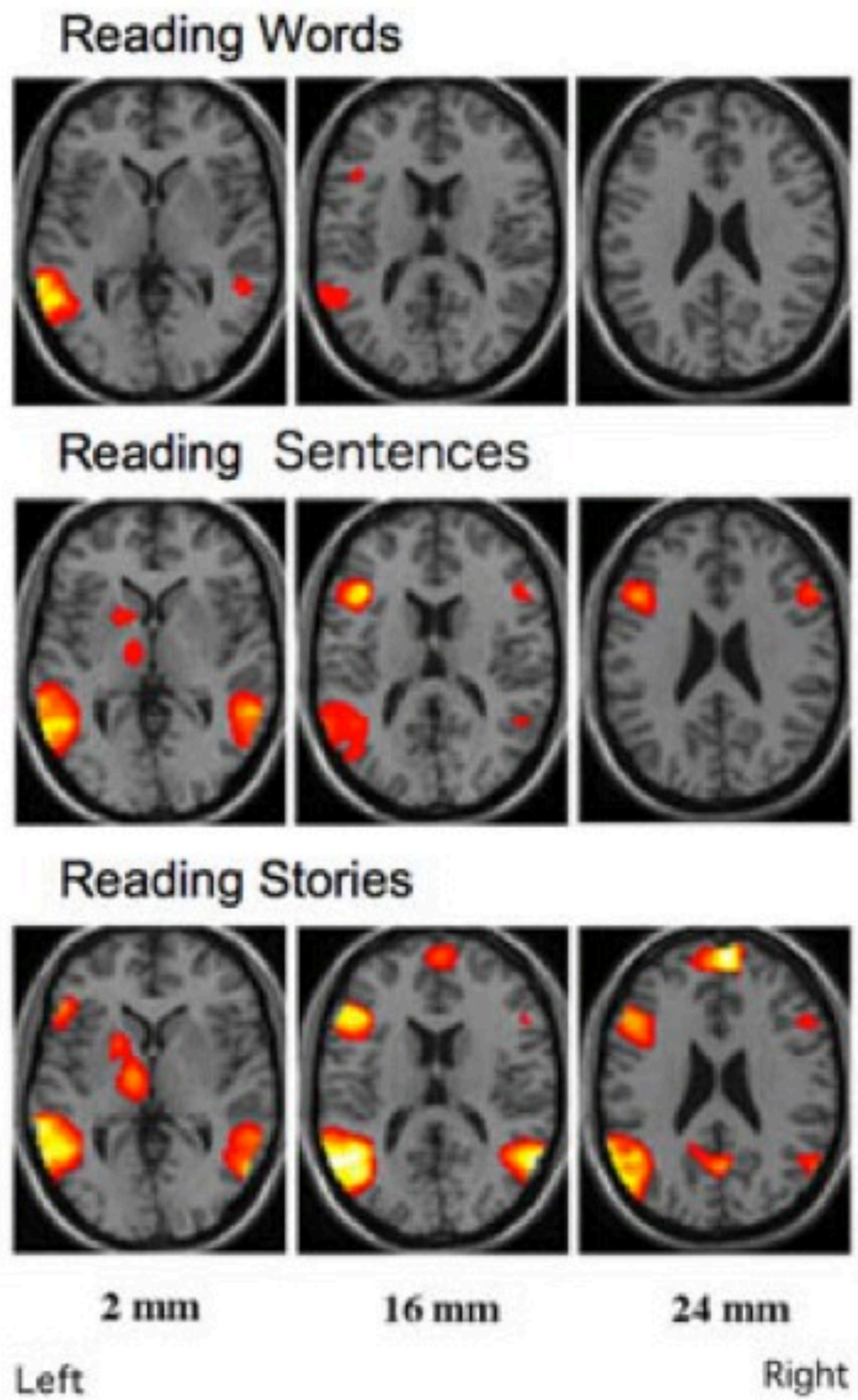
A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, $p(\text{uncorrected}) < 0.001$, 3 voxel extent threshold.

Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction
by: Craig M. Bennett, Abigail A. Baird, Michael B. Miller, George L. Wolford

Journal of Serendipitous and Unexpected Results, Vol. 1 (2011), pp. 1-5 Key: citeulike:9170380

A voxel is declared significantly activated if the p-value is below some fixed alpha level, <u>corrected for multiple comparisons</u>.

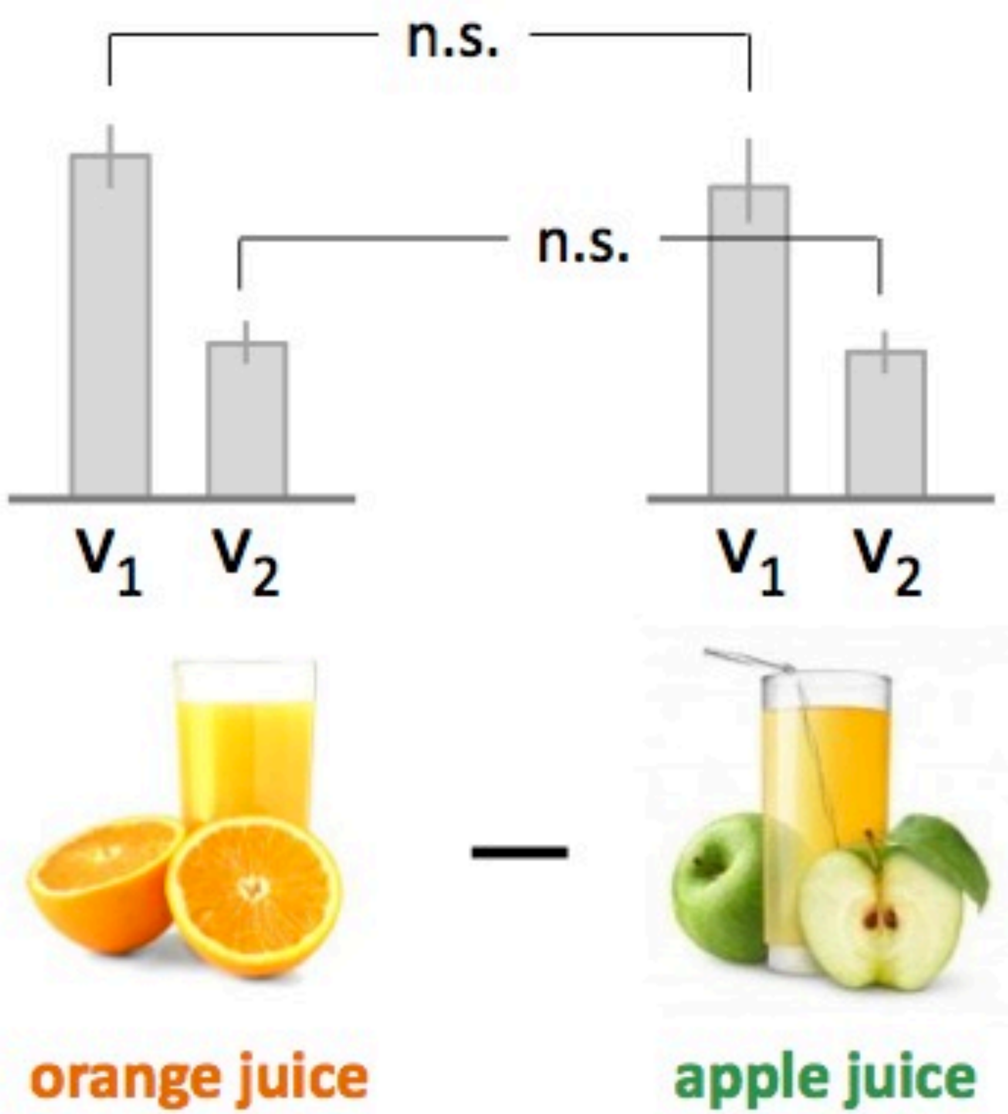
Example

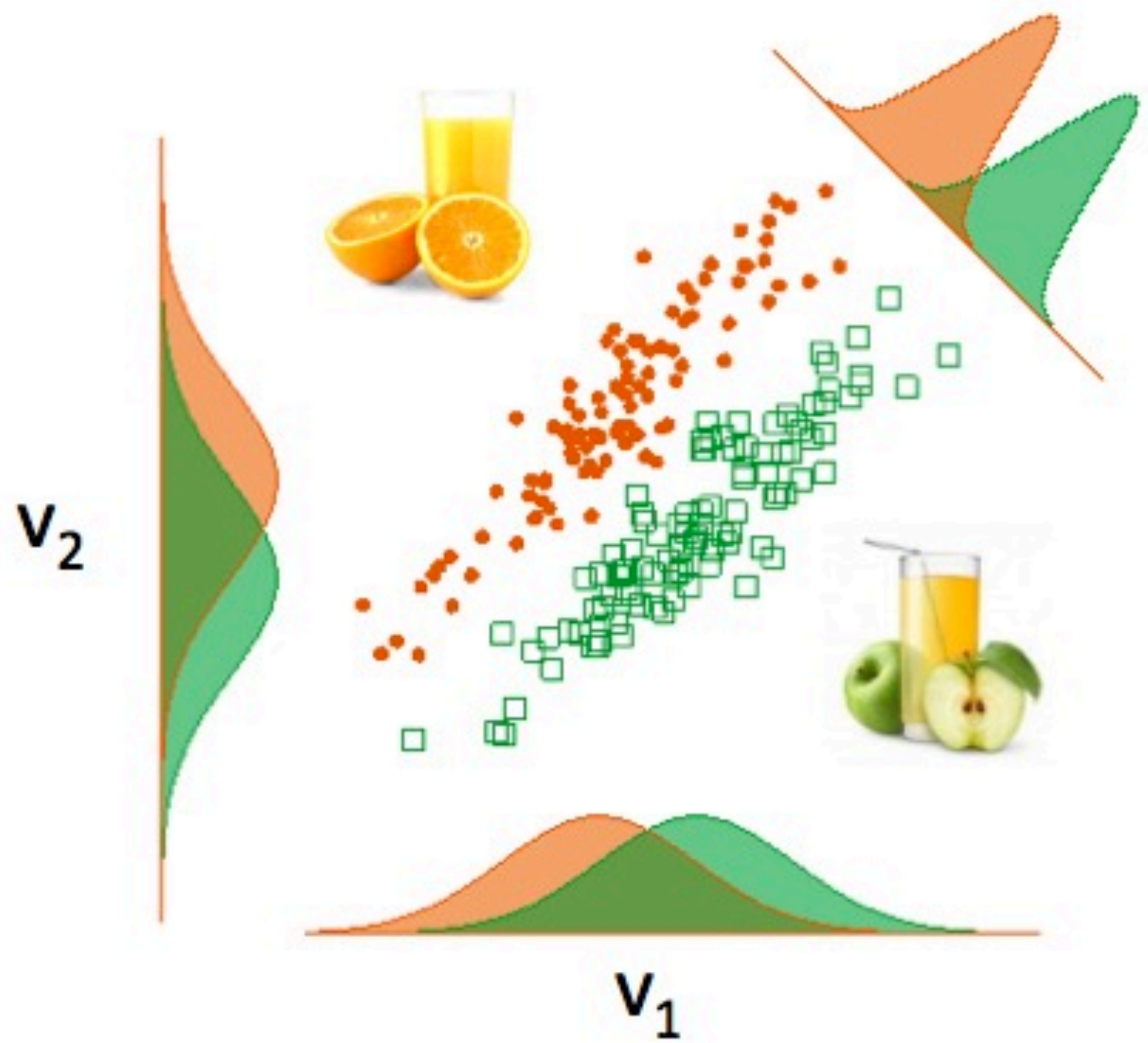


See Poline J-B, Brett M. The general linear model and fMRI: Does love last forever? Neuroimage. 2012 for a gentle overview.

Multivariate analysis

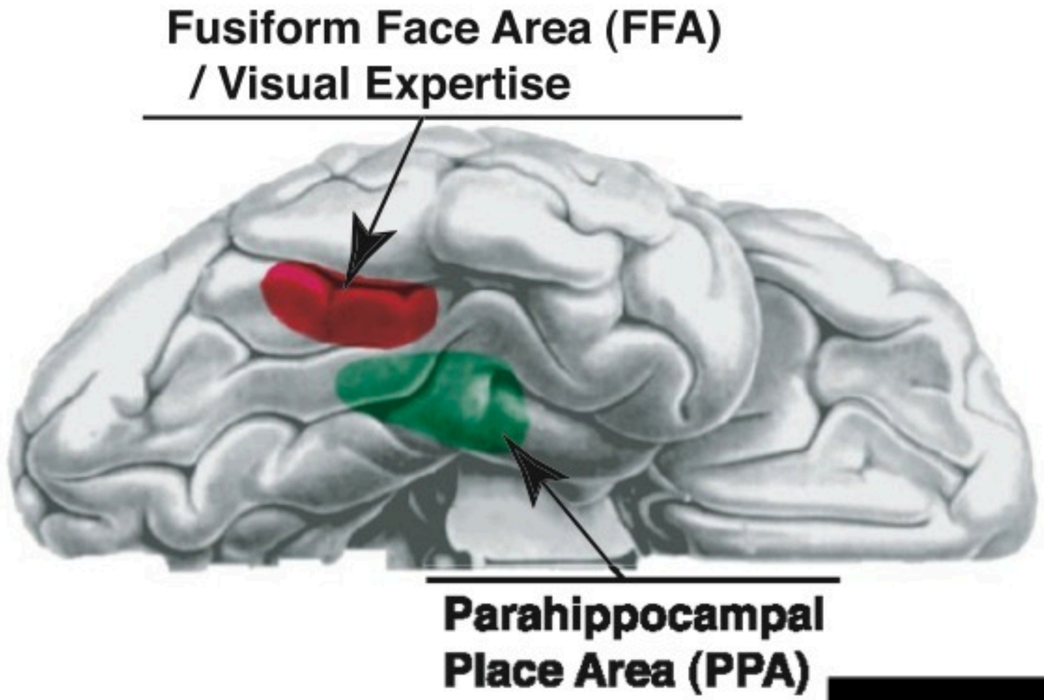
- examine responses that are jointly encoded in multiple voxels

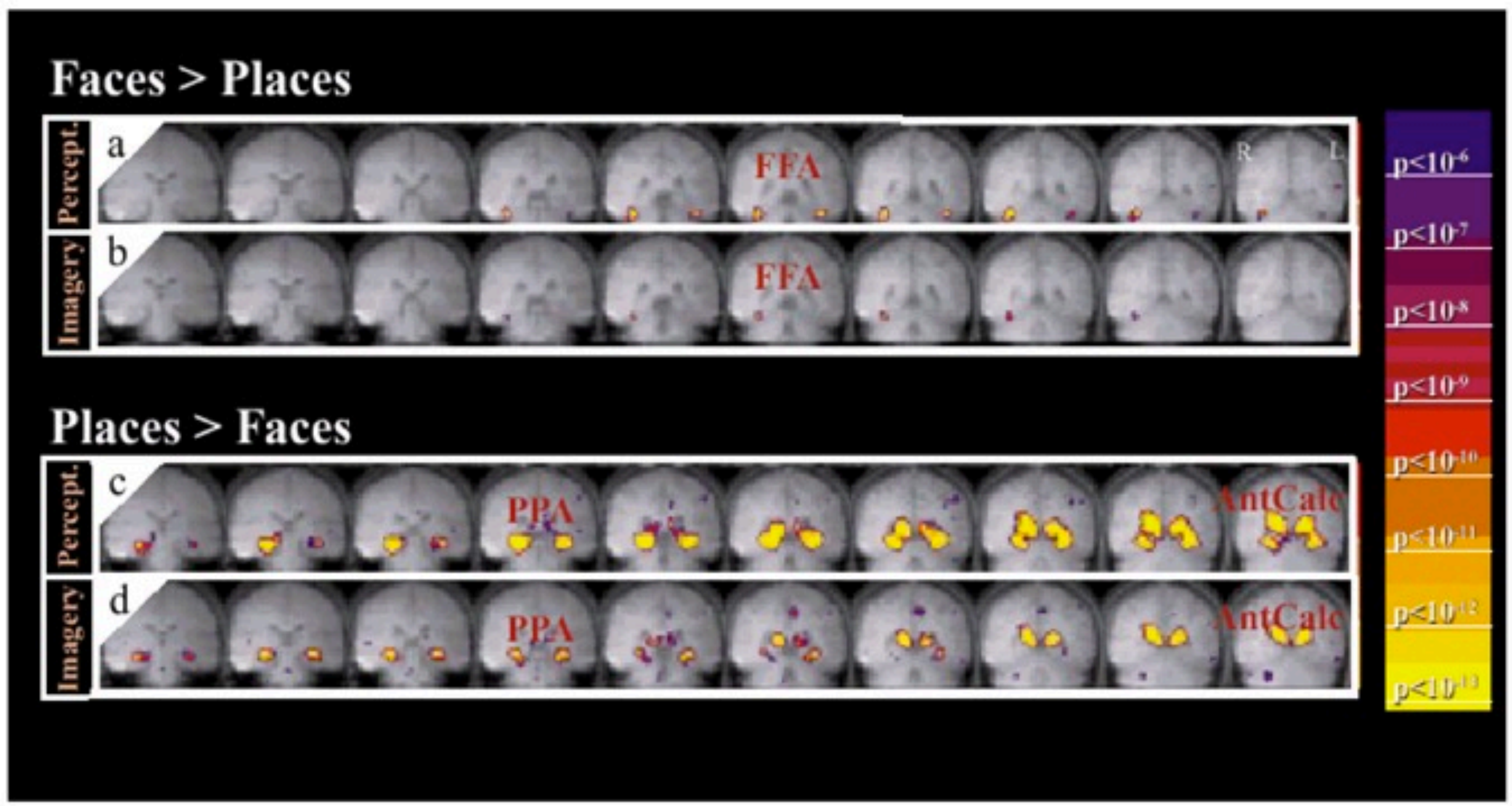




Slide courtesy Kai Brodersen

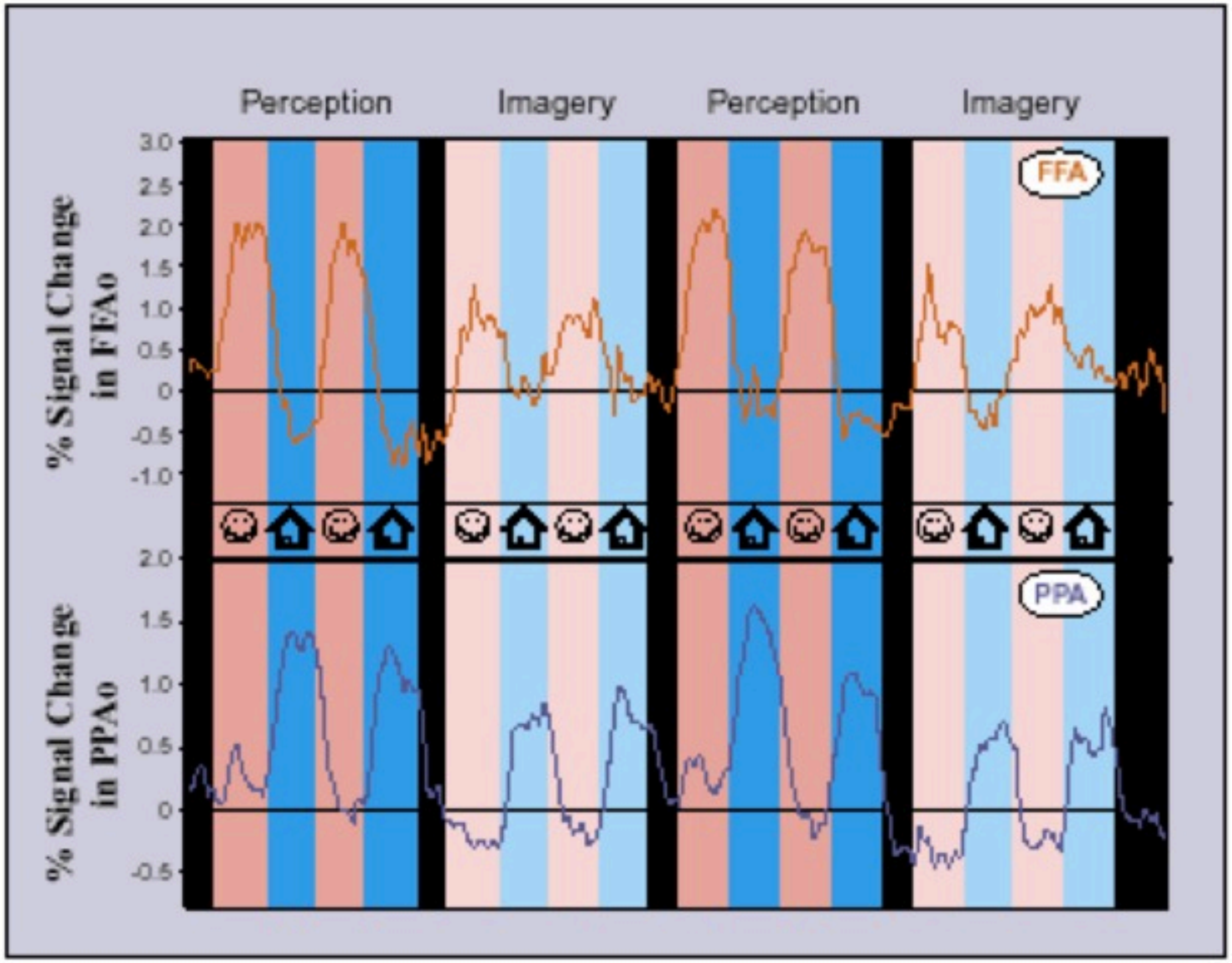
Faces versus places





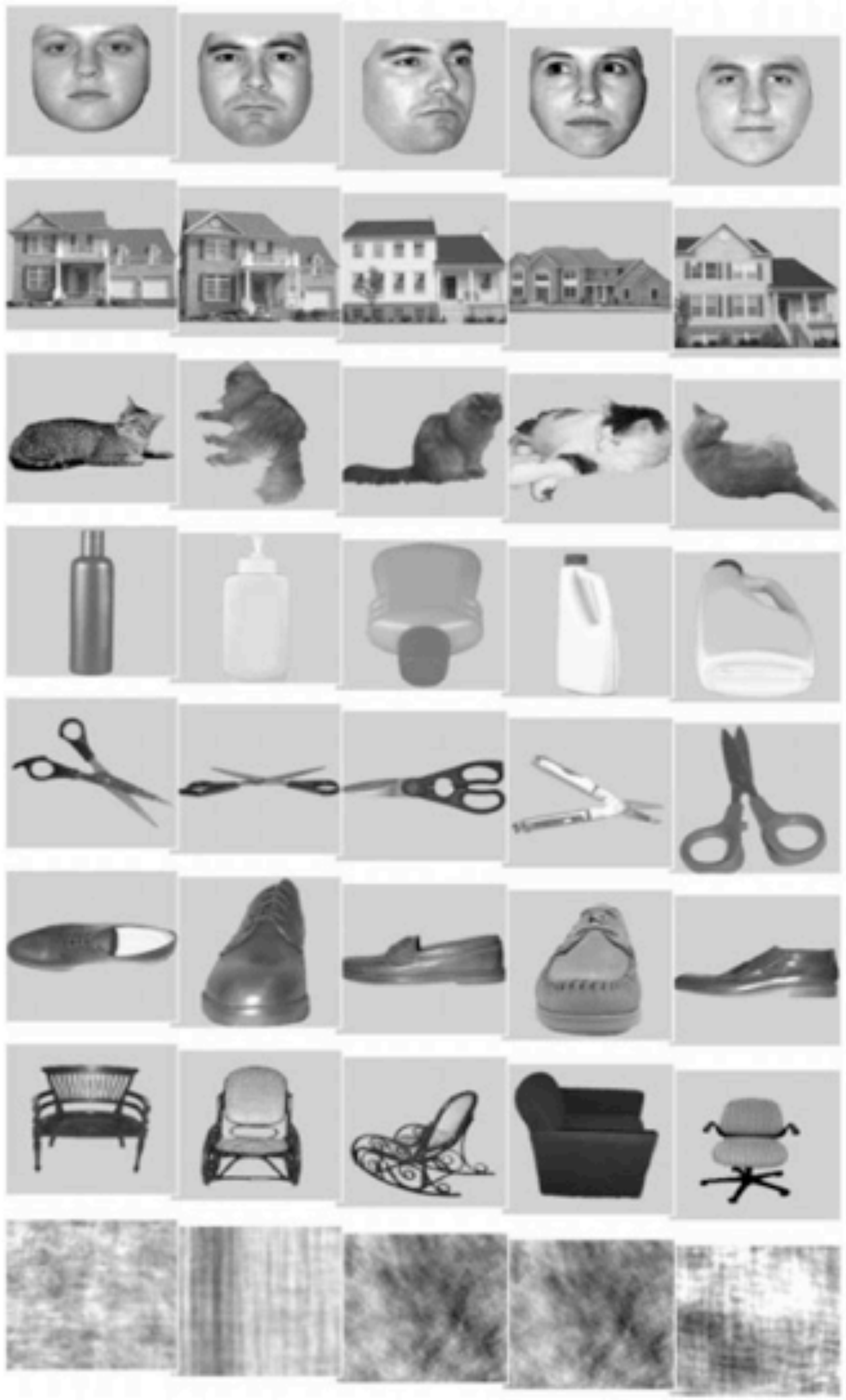
O’Craven and Kanwisher, J Cog Neu, 2000

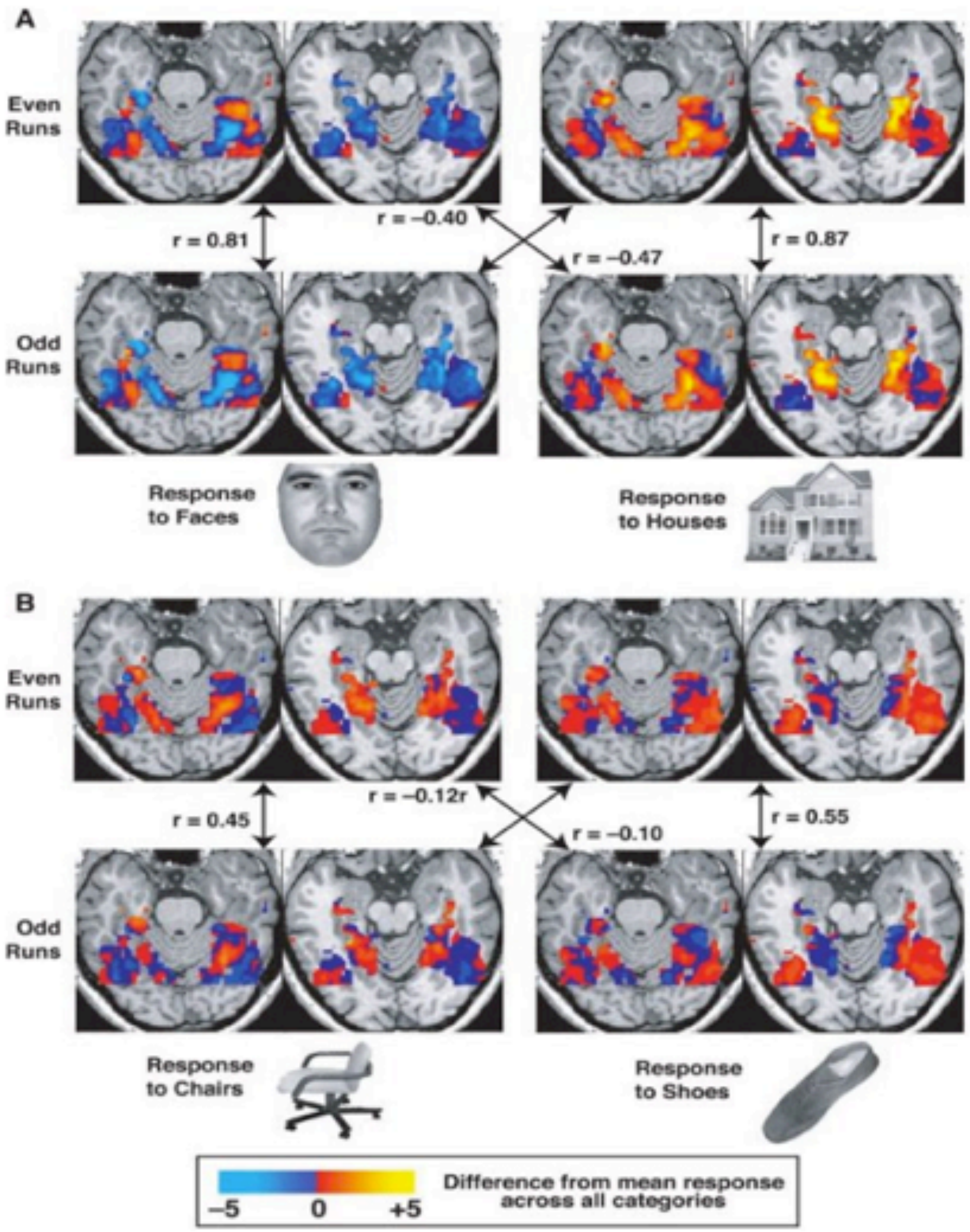
Imagery



David Hume saw the relationship between percepts and images as follows: "*The difference betwixt these consists in the degrees of force and liveliness, with which they strike upon the mind . . . [Perceptions] enter with most force and violence . . . By ideas I mean the faint images of these in thinking and reasoning*" (Hume, 1739).

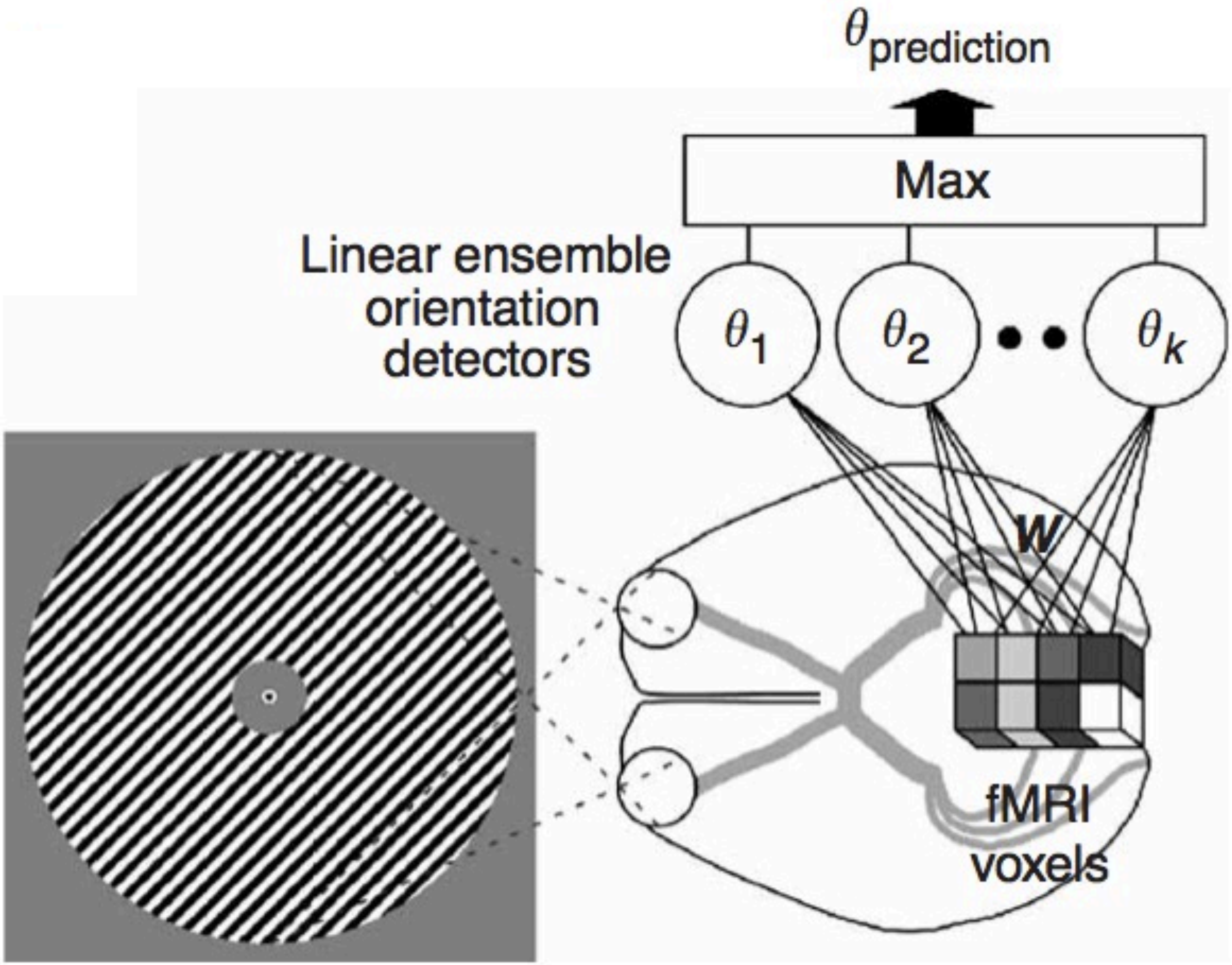
The Haxby study



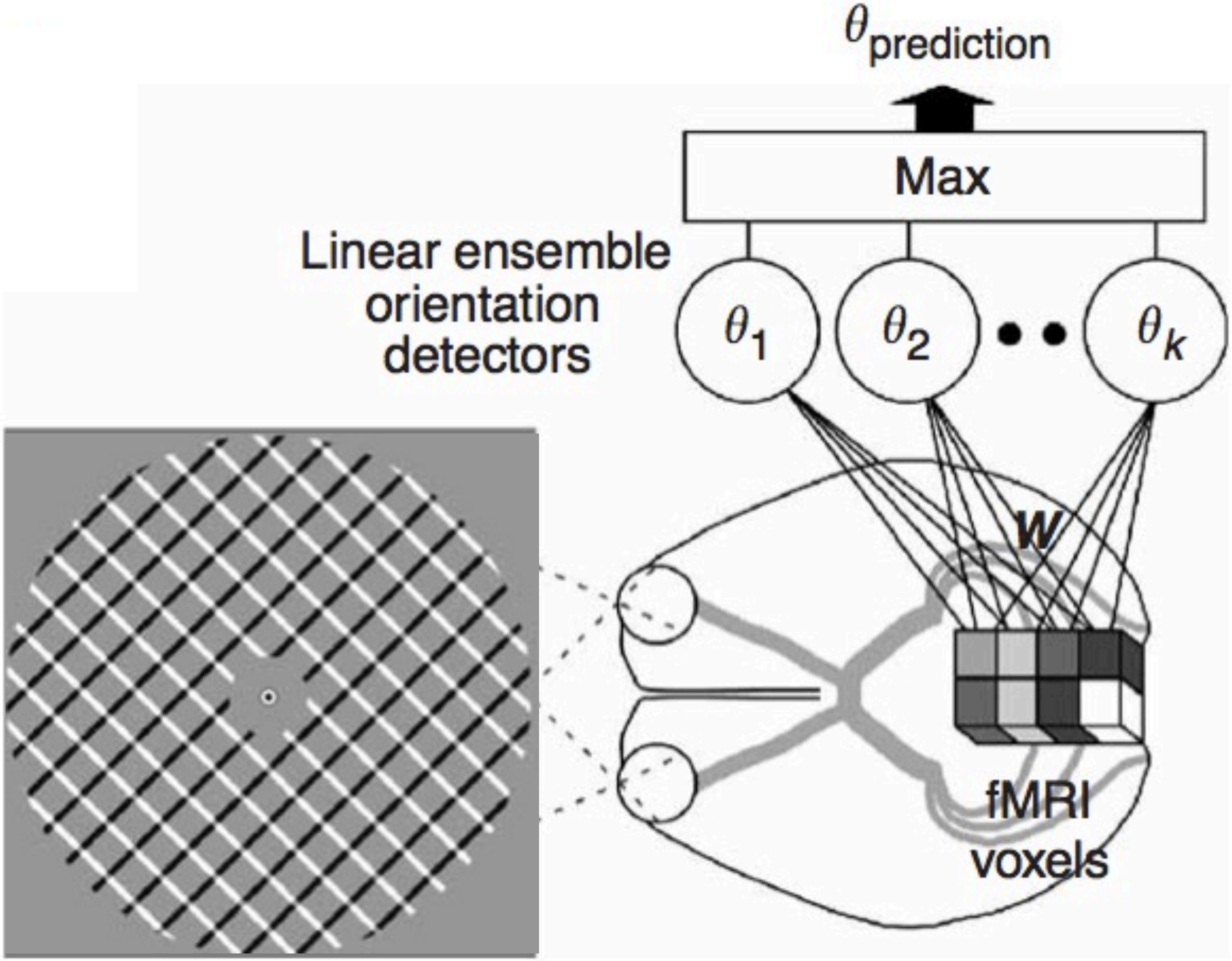


Haxby et al. Science, 2001.

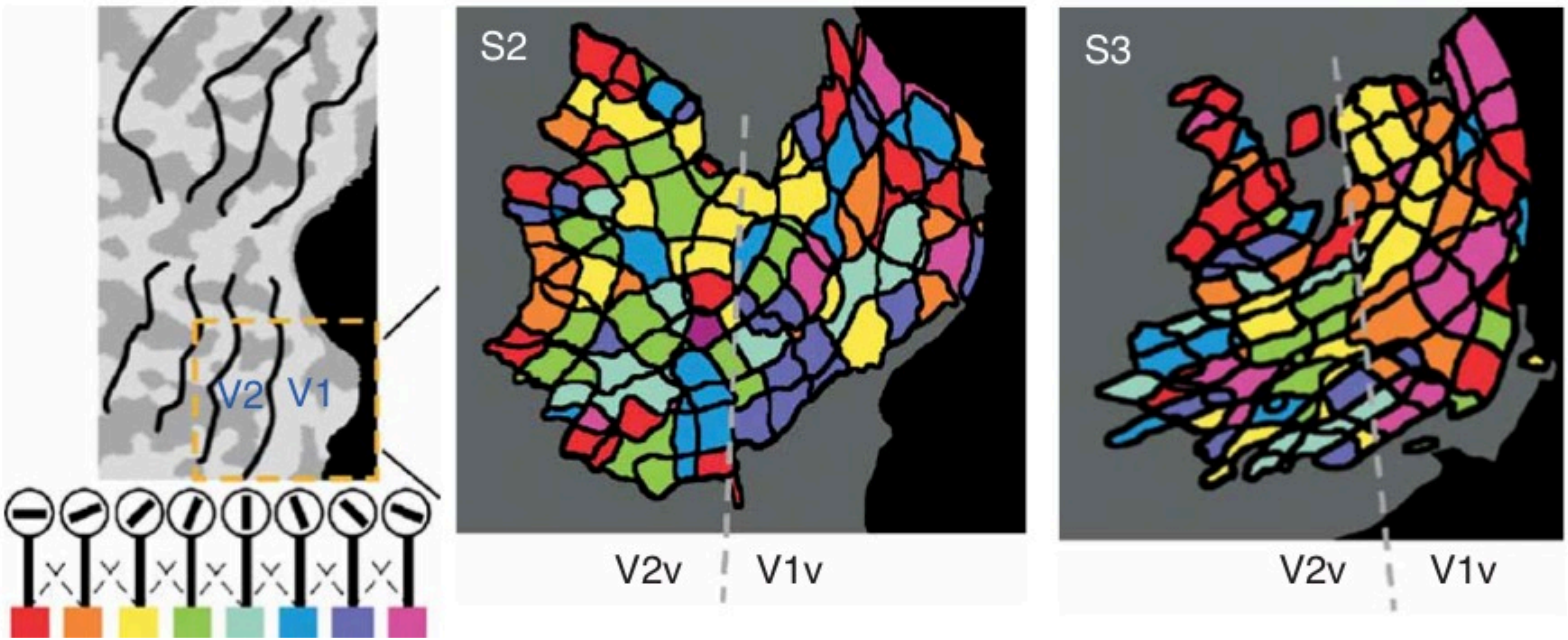
Application of machine learning



Kamitani and Tong, Nature Neuroscience, 2005

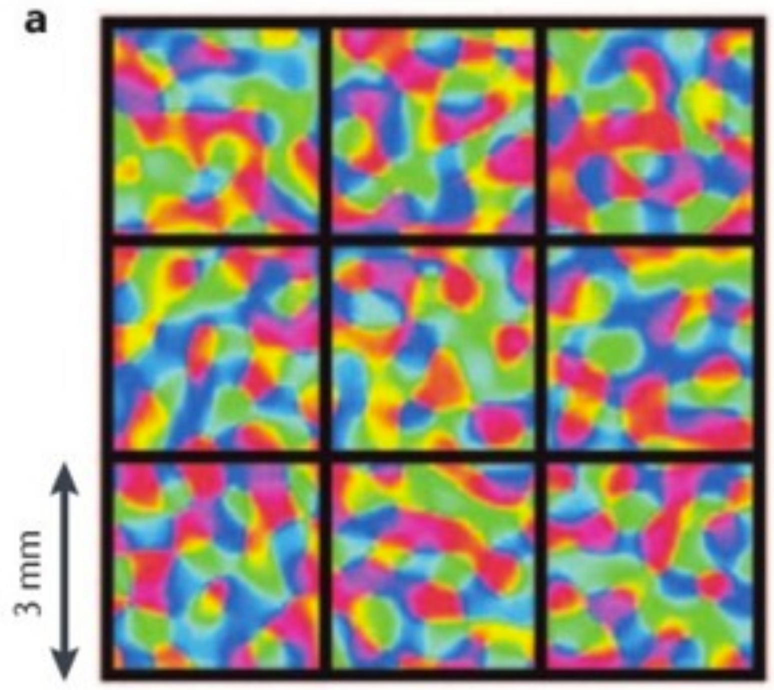


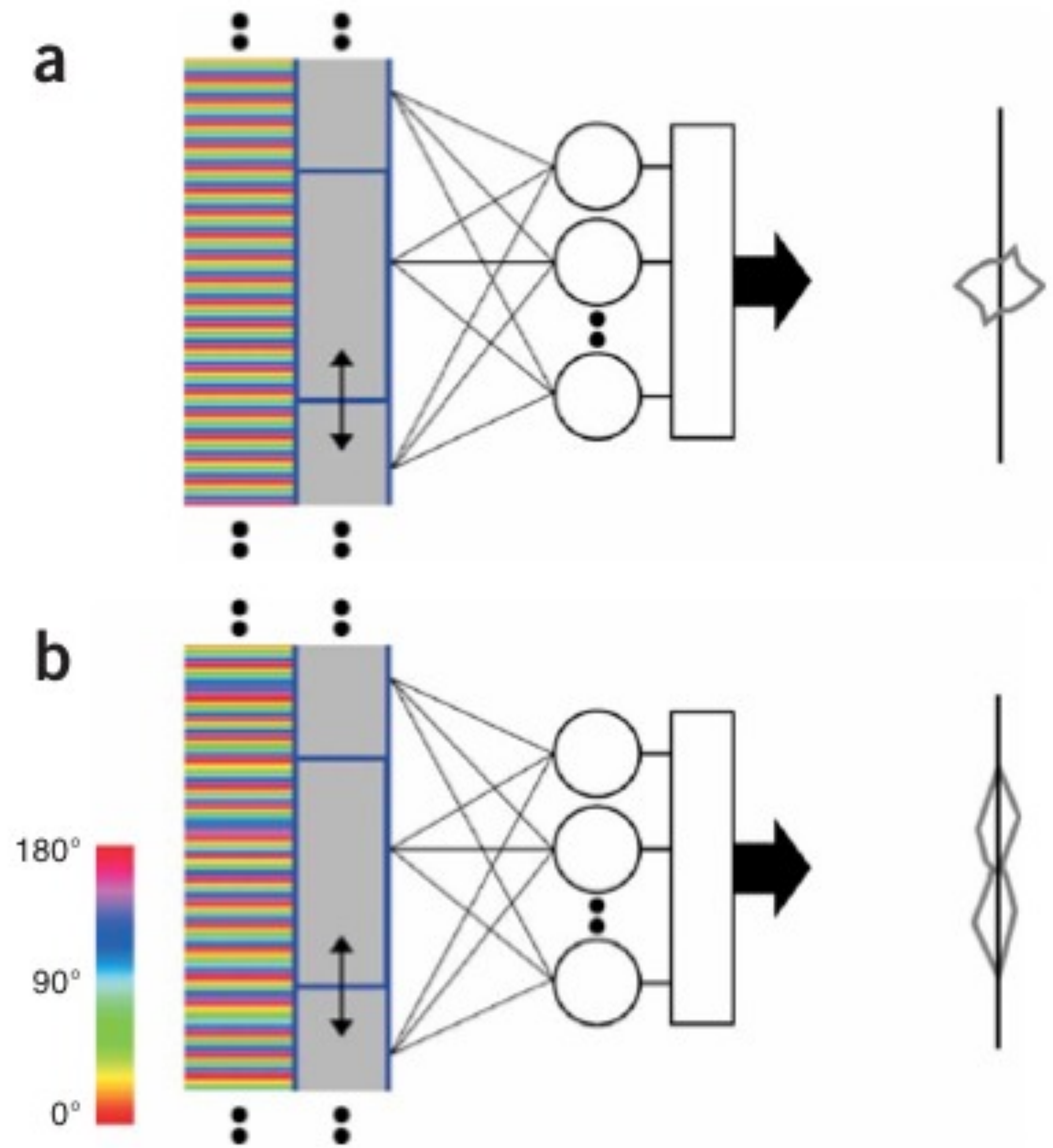
Kamitani and Tong, Nature Neuroscience, 2005



Kamitani and Tong, Nature Neuroscience, 2005

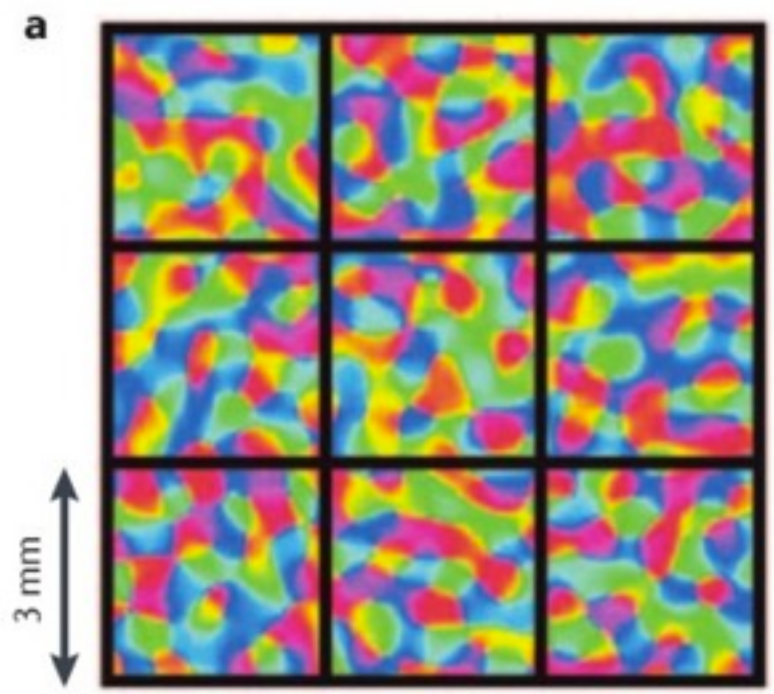
Where does the information come from: biased sampling?

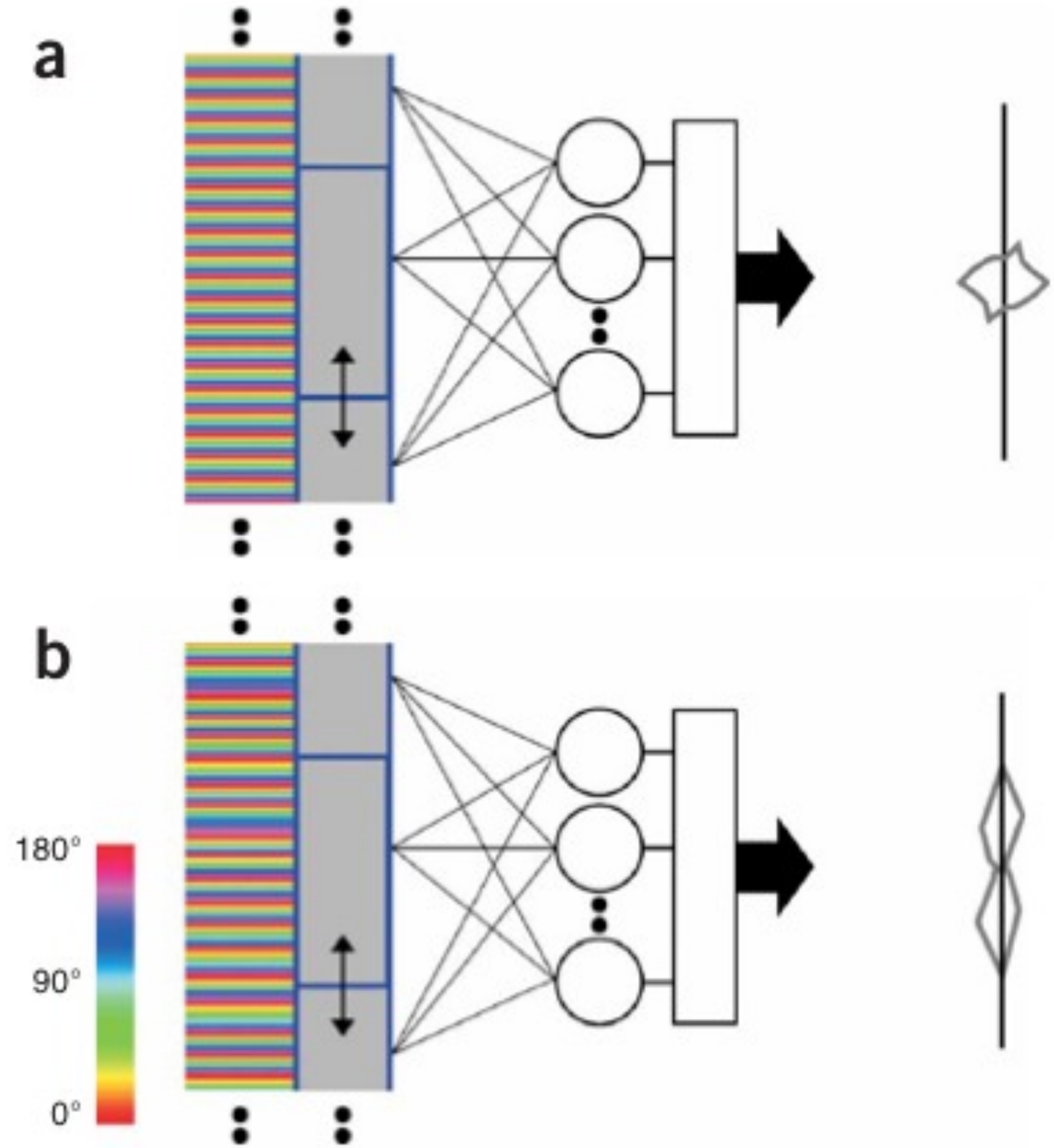




Kamitani and Tong, Nature Neuroscience, 2005

Where does the information come from: biased sampling?



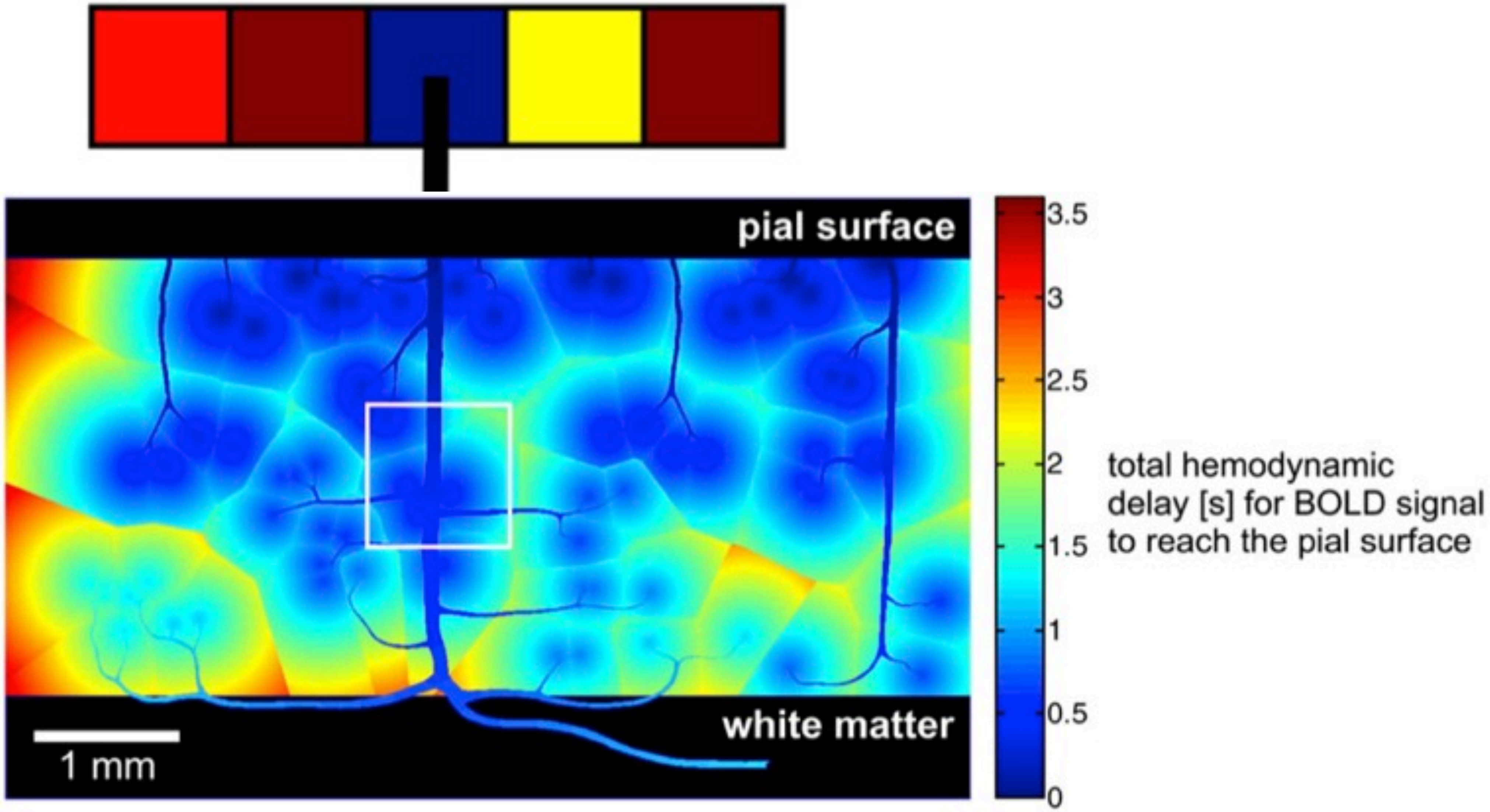


unlikely:

- sensitivity to draining veins
- motion artifacts

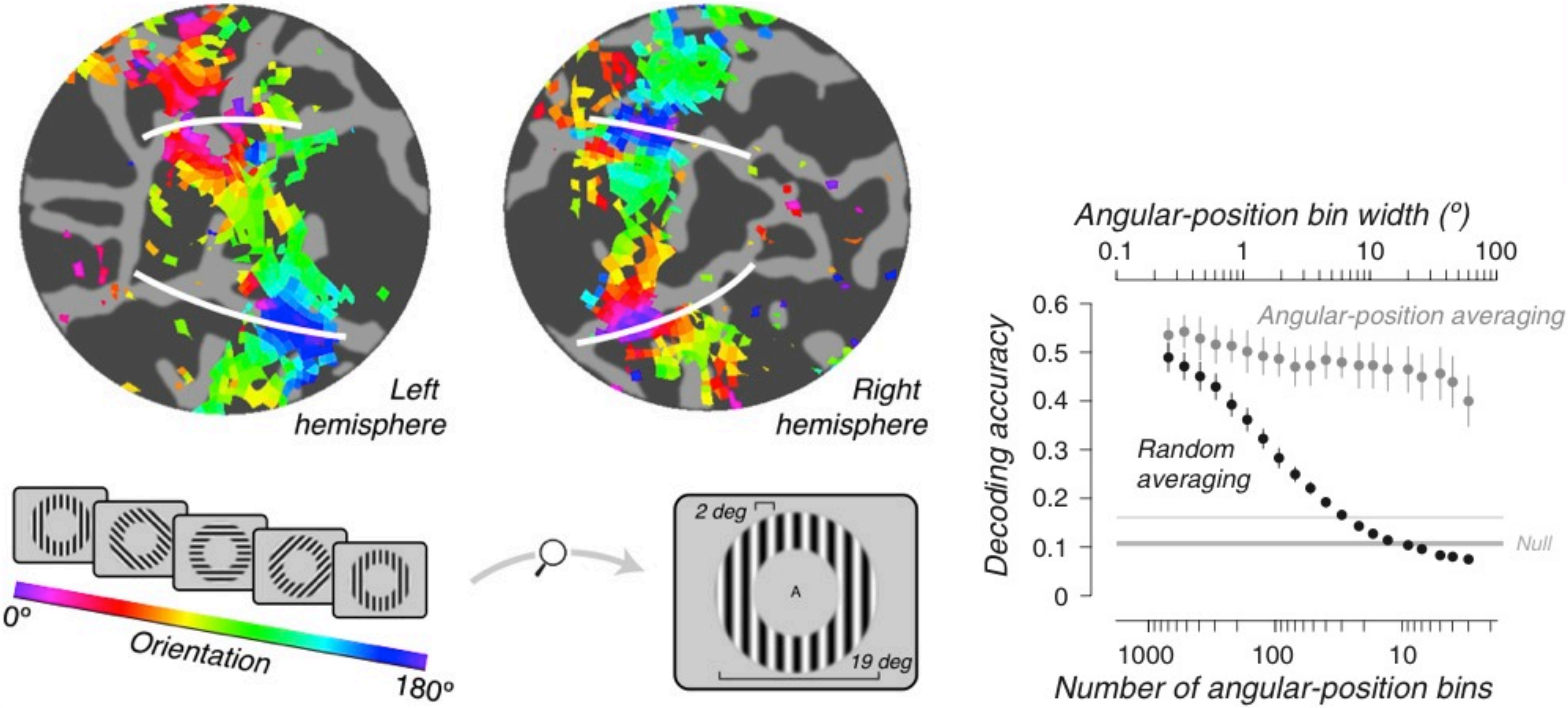
Kamitani and Tong, Nature Neuroscience, 2005

Where does it come from: complex spatiotemporal filter?



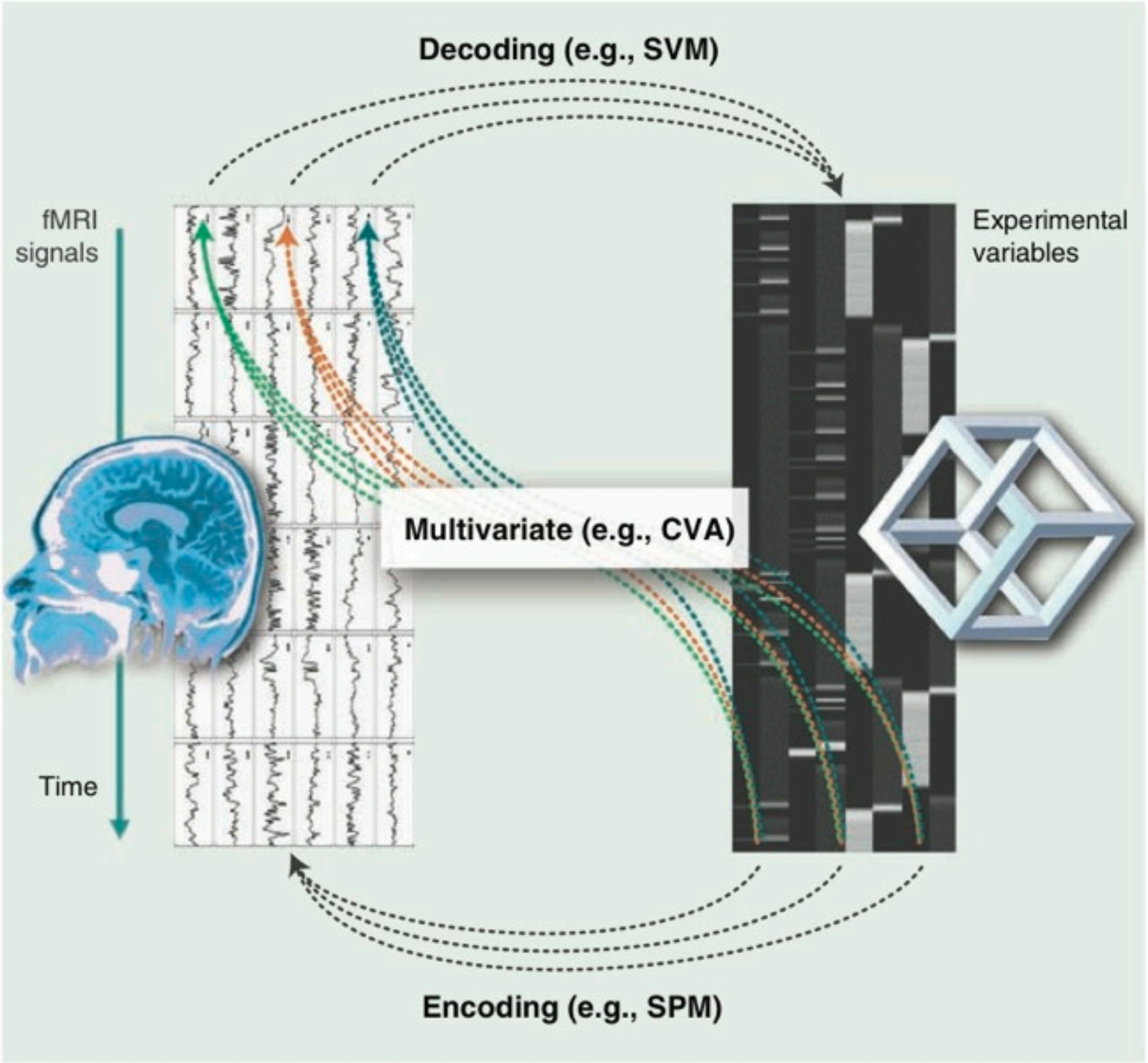
Kriegeskorte et al., Neuroimage, 2010

Where does it come from: coarse maps?



Freeman et al., J. Neurosci, 2011

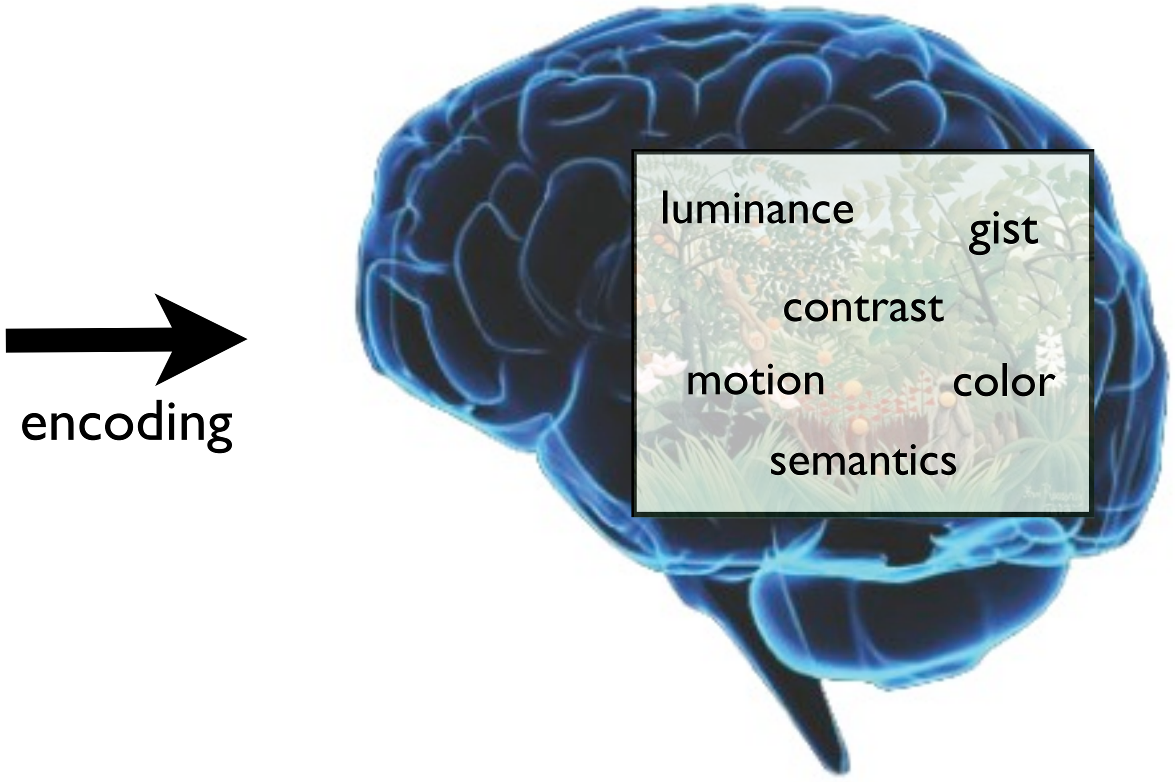
Encoding and decoding



Friston, Science, 2009

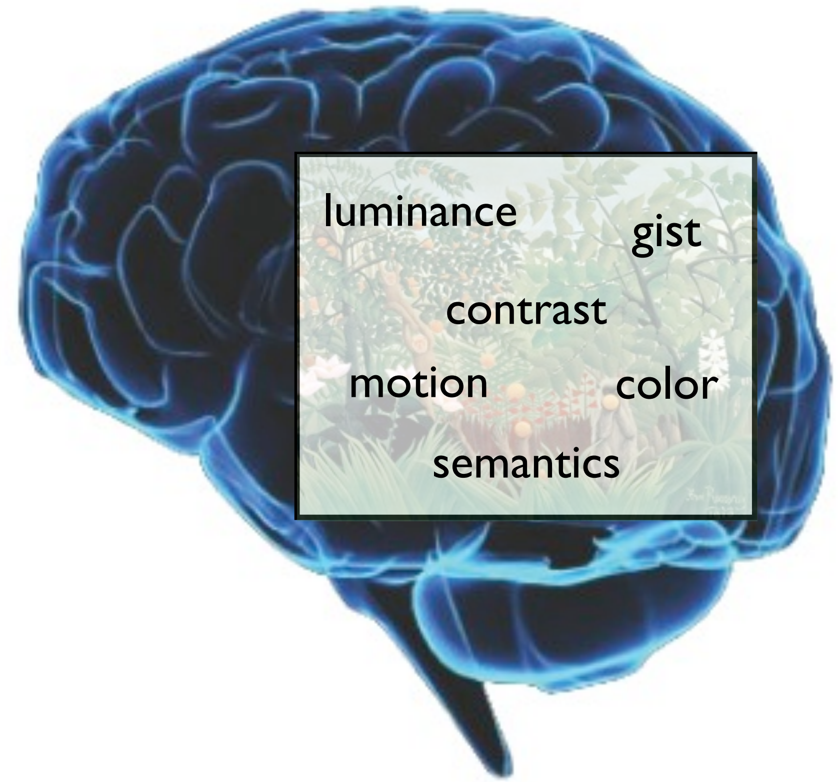
Encoding and decoding

Encoding and decoding



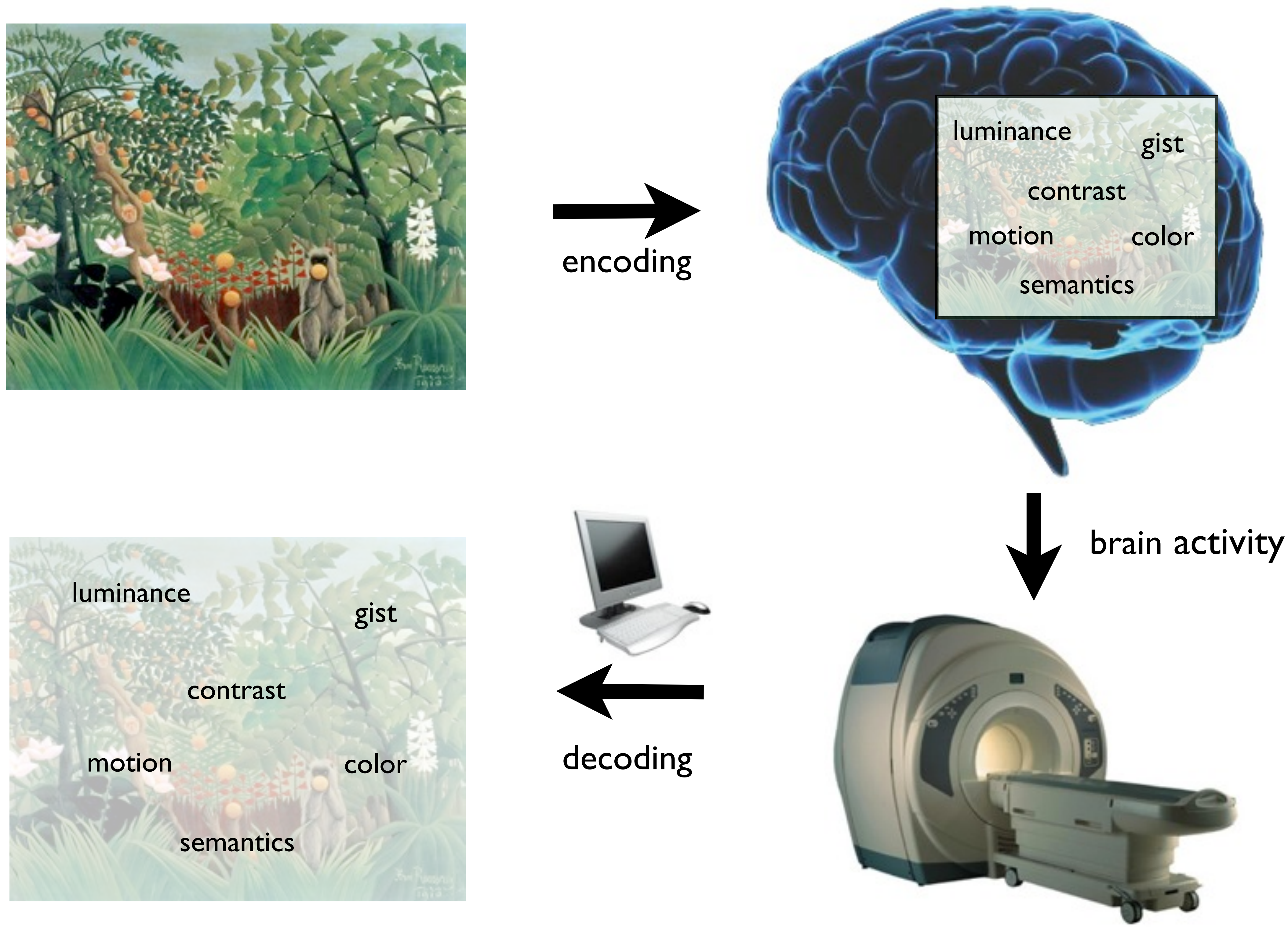
Encoding and decoding

encoding

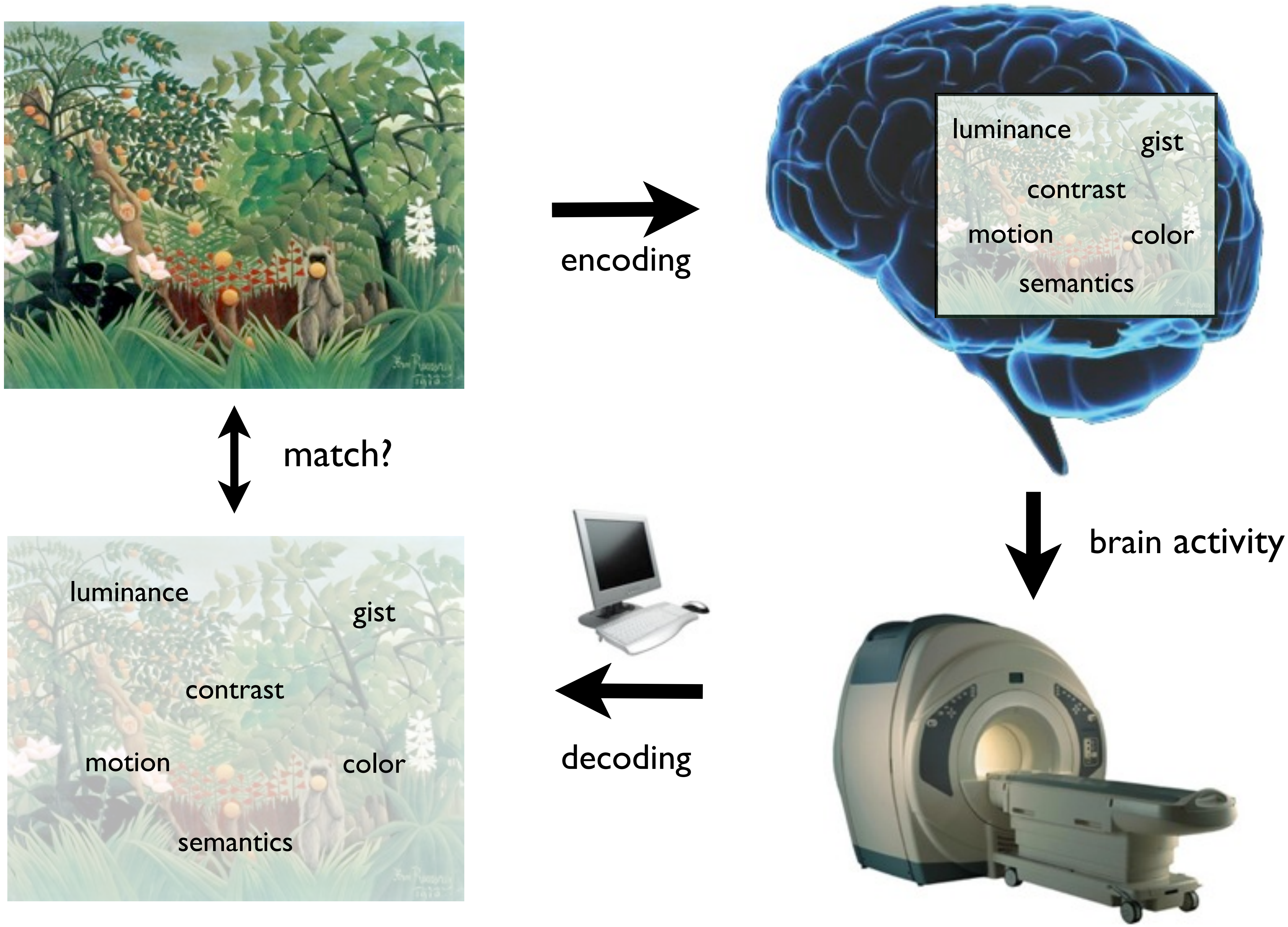


brain activity

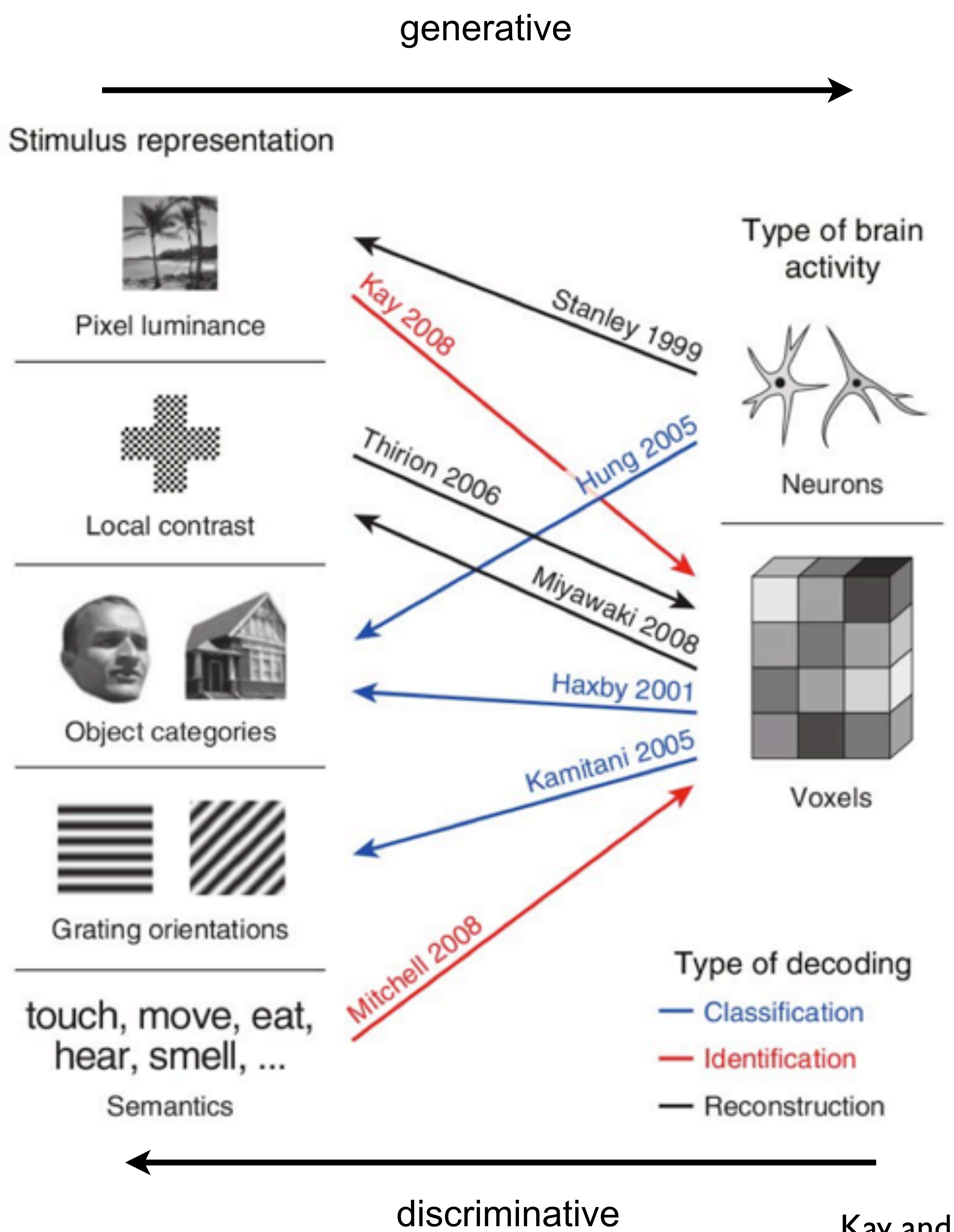
Encoding and decoding



Encoding and decoding



Decoding flavors



Kay and Gallant, Nature, 2009

Univariate versus multivariate analysis

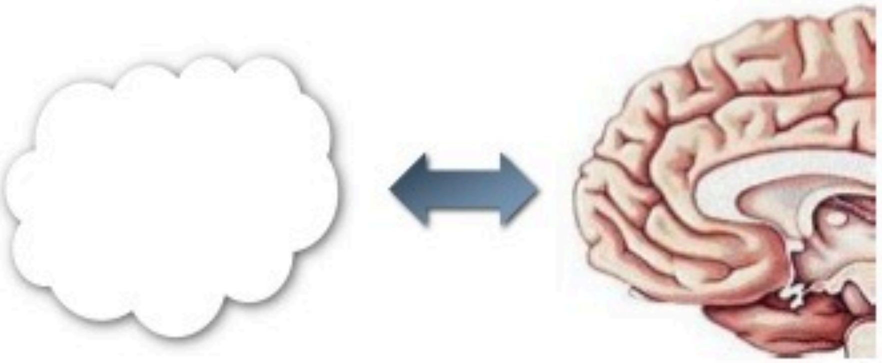
A univariate model considers a single voxel at a time.

context $X_t \in \mathbb{R}^d$

BOLD signal $Y_t \in \mathbb{R}$

Spatial dependencies between voxels are only introduced afterwards, through random field theory.

A multivariate model considers many voxels at once.



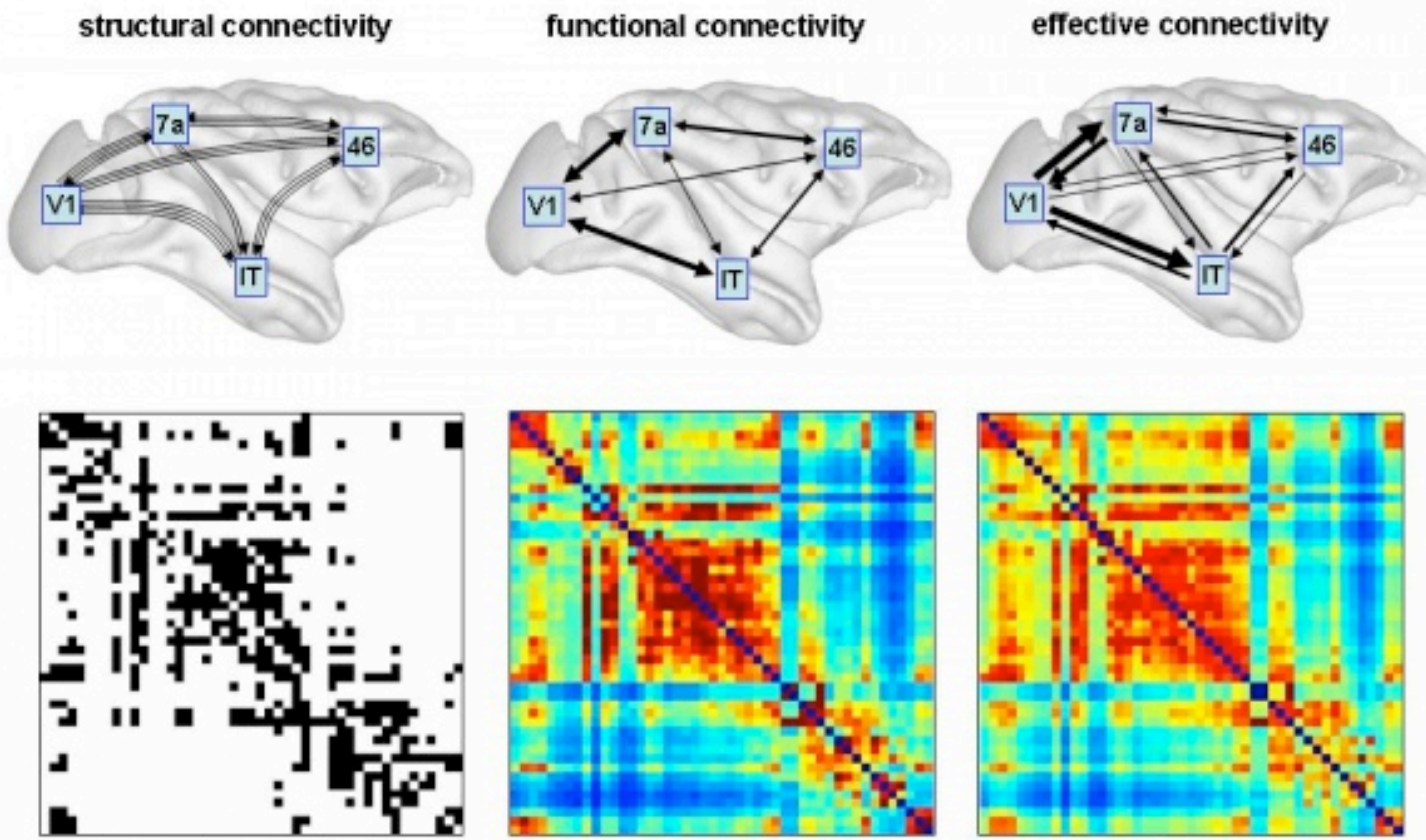
context $X_t \in \mathbb{R}^d$

BOLD signal $Y_t \in \mathbb{R}^v$, $v \gg 1$

Multivariate models enable inferences on distributed responses without requiring focal activations.

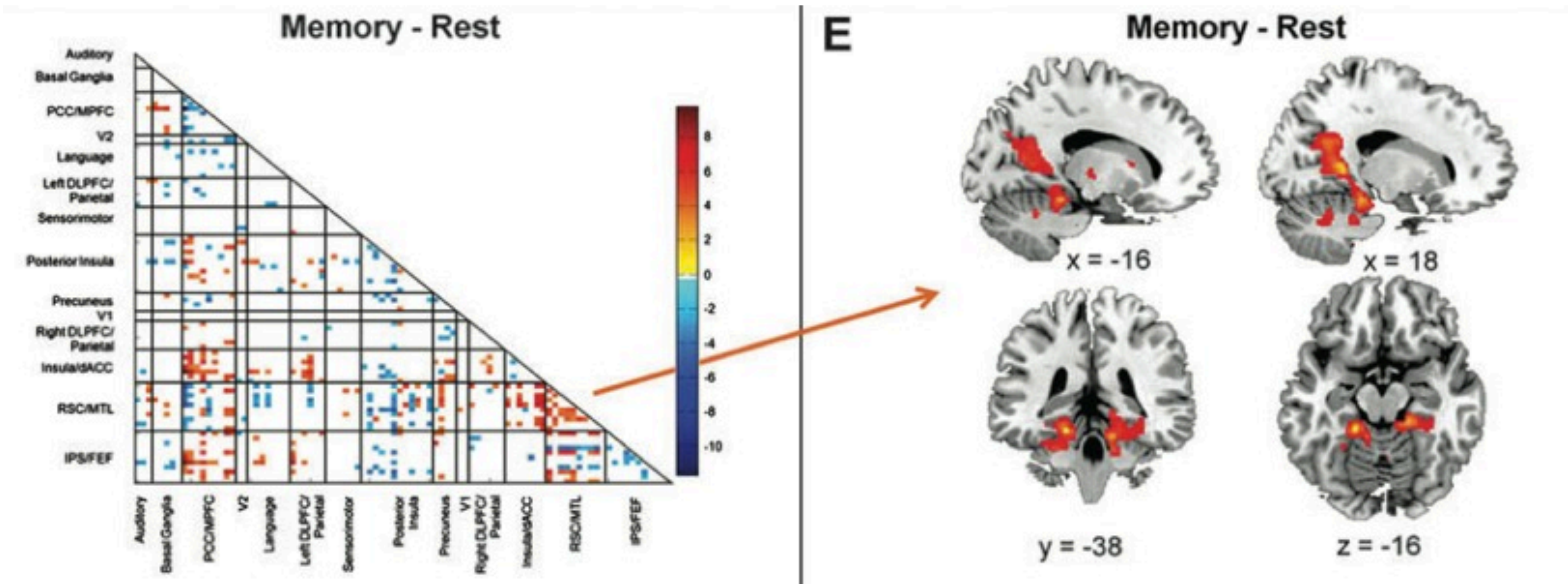
Slide courtesy Kai Brodersen

Multivariate analyses of regional interactions



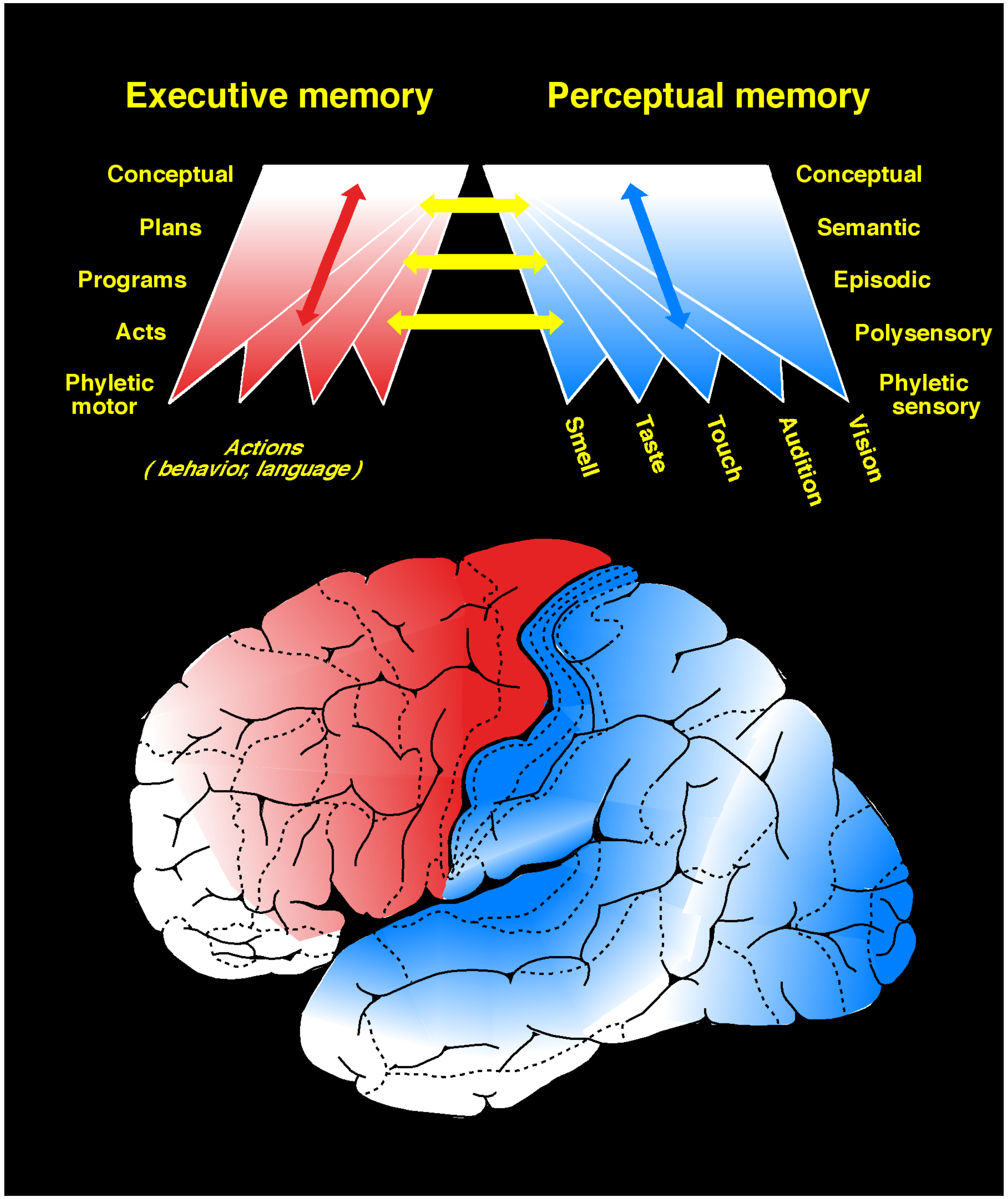
How do we estimate and describe connectivity?

Graph-based decoding



Can we predict cognitive states based on differences in connectivity?

Representations are hierarchically organized in terms of specific activity patterns within and across perceptual and executive networks.

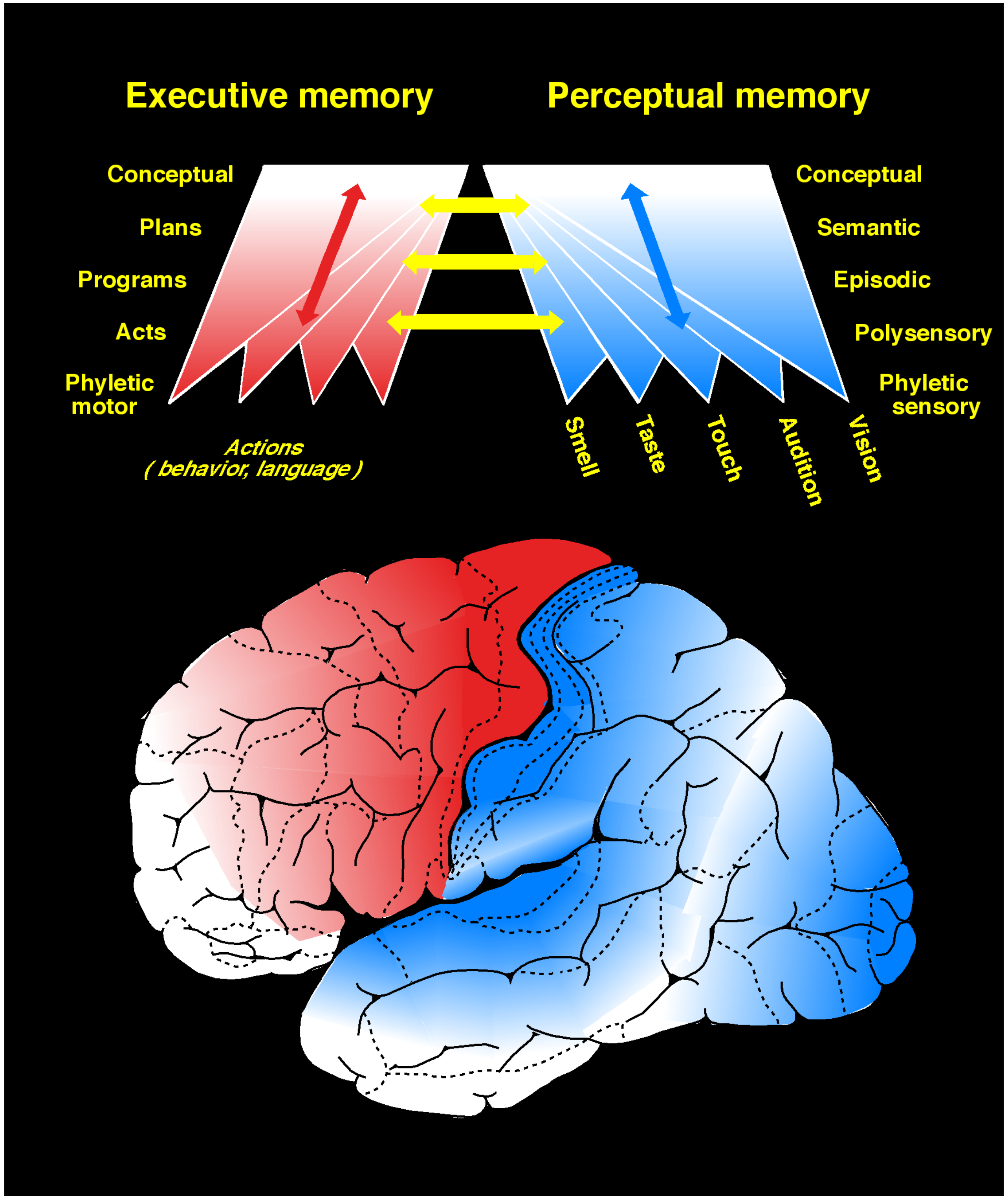


Fuster, TICS, 2004

Distributed representations

Representations are hierarchically organized in terms of specific activity patterns within and across perceptual and executive networks.

(Re-)activation of these representations involved in perception, WM maintenance, imagery, recall, decision making.



Fuster, TICS, 2004

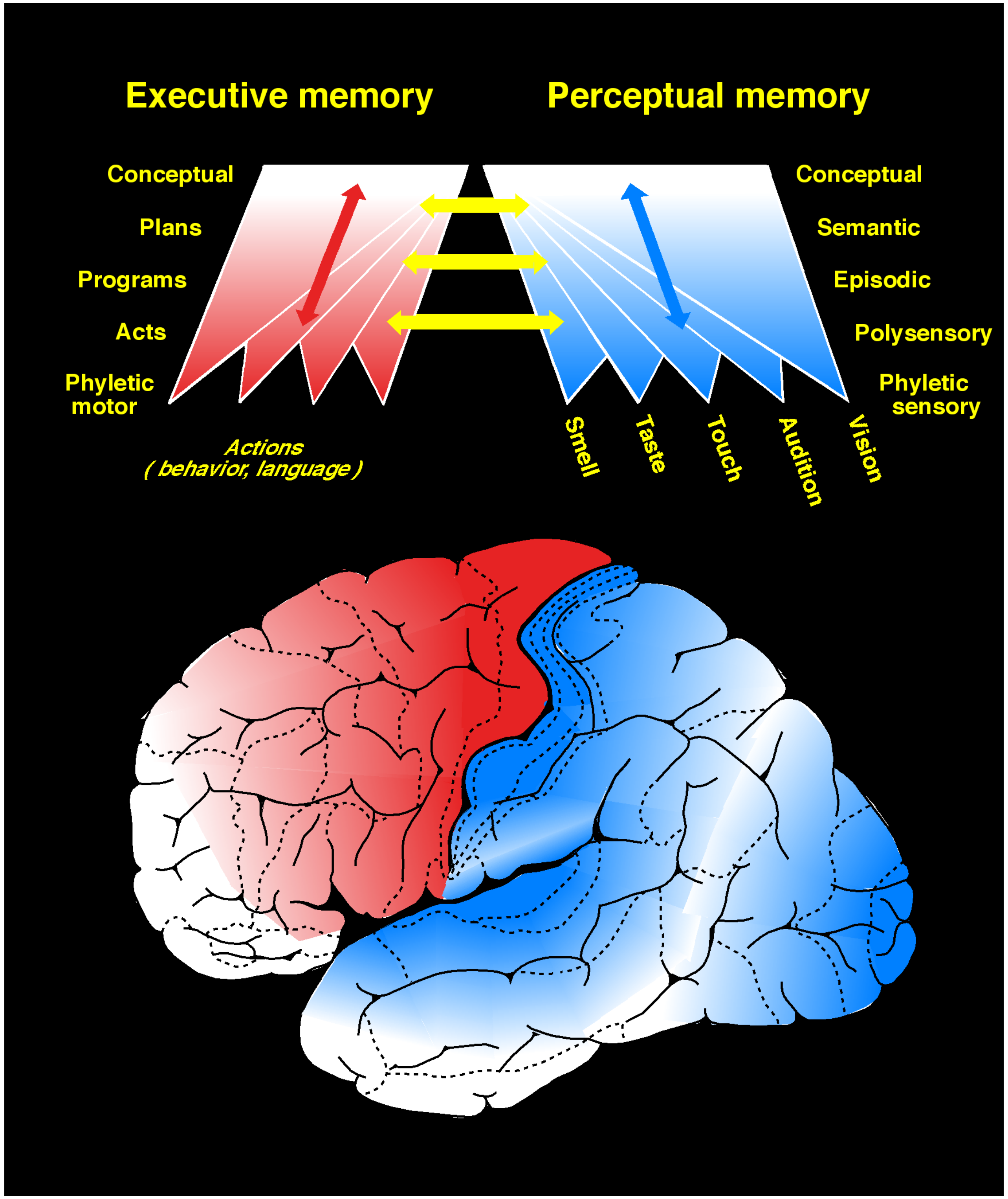
Distributed representations

Representations are hierarchically organized in terms of specific activity patterns within and across perceptual and executive networks.

(Re-)activation of these representations involved in perception, WM maintenance, imagery, recall, decision making.

New machine learning techniques offer new insights into the brain:

- decoding in cognitive/clinical neuroscience
- more sophisticated encoding models
- new approaches to connectivity analysis



Fuster, TICS, 2004

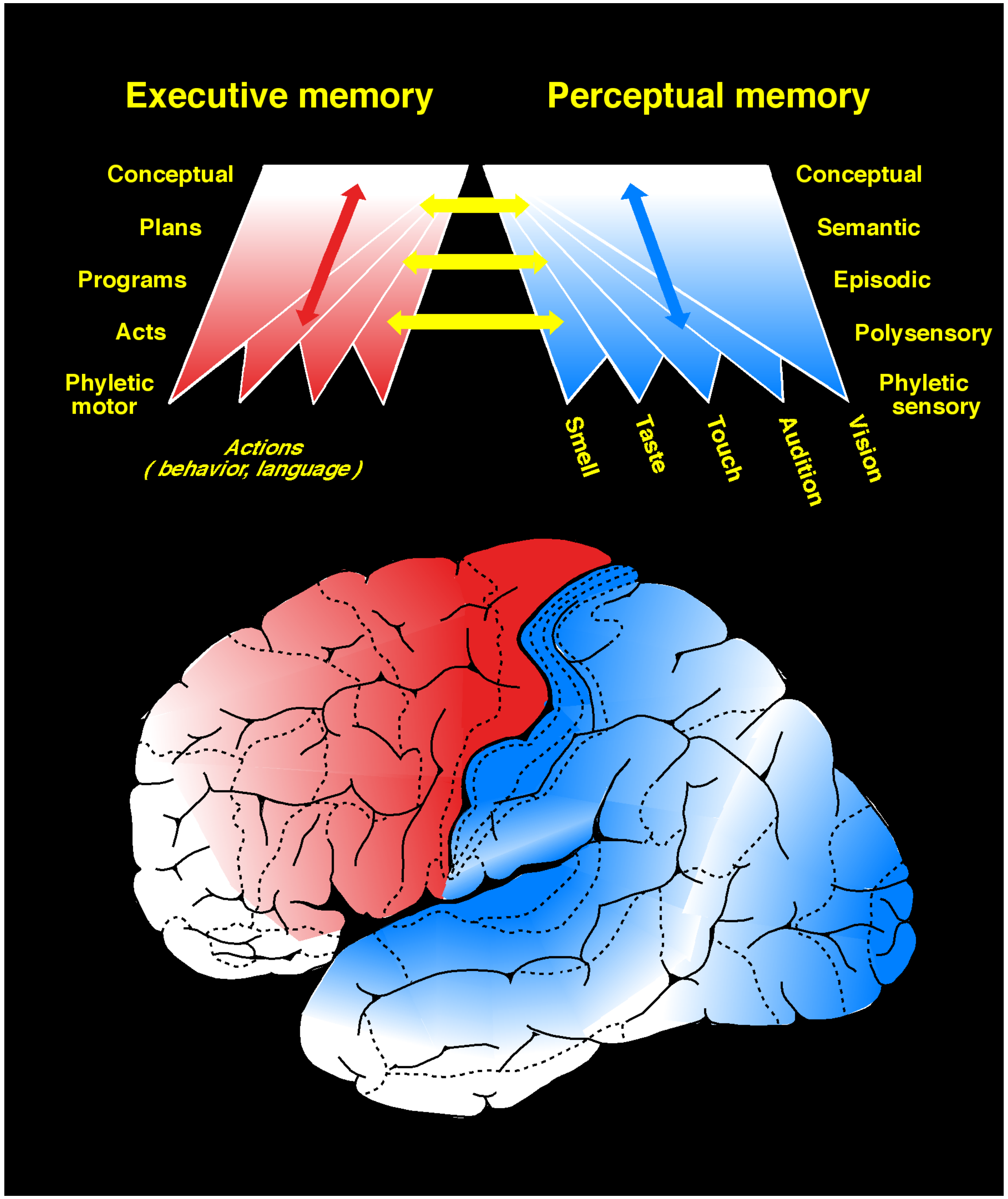
Representations are hierarchically organized in terms of specific activity patterns within and across perceptual and executive networks.

(Re-)activation of these representations involved in perception, WM maintenance, imagery, recall, decision making.

New machine learning techniques offer new insights into the brain:

• decoding in cognitive/clinical neuroscience
• more sophisticated encoding models
• new approaches to connectivity analysis

Bayesian statistics as a lingua franca



Fuster, TICS, 2004

Course overview

1. **Introduction**

1. **Introduction**
2. **Discriminative decoding**

1. **Introduction**
2. **Discriminative decoding**
3. **Advanced discriminative approaches**

1. **Introduction**
2. **Discriminative decoding**
3. **Advanced discriminative approaches**
4. **Generative decoding**

1. Introduction
2. Discriminative decoding
3. Advanced discriminative approaches
4. Generative decoding
5. Advanced generative approaches

1. Introduction
2. Discriminative decoding
3. Advanced discriminative approaches
4. Generative decoding
5. Advanced generative approaches
6. Connectivity analysis

1. Introduction
2. Discriminative decoding
3. Advanced discriminative approaches
4. Generative decoding
5. Advanced generative approaches
6. Connectivity analysis
7. Bayesian connectomics

1. Introduction
2. Discriminative decoding
3. Advanced discriminative approaches
4. Generative decoding
5. Advanced generative approaches
6. Connectivity analysis
7. Bayesian connectomics
8. Guest lectures

Course overview

1. Introduction
2. Discriminative decoding
3. Advanced discriminative approaches
4. Generative decoding
5. Advanced generative approaches
6. Connectivity analysis
7. Bayesian connectomics
8. Guest lectures
9. Bayesian brain / Course conclusion

Required reading

- Paper: Friston KJ. Modalities, Modes, and Models in Functional Neuroimaging. Science. 2009; 326(5951):399–403.

REVIEW

Modalities, Modes, and Models in Functional Neuroimaging

Karl J. Friston

In this, the 21st century, human-brain mapping celebrates 21 years of cognitive activation studies. This review looks at imaging neuroscience and key ideas it has pursued; some ideas portend exciting developments, and others have failed gloriously. In terms of achievements, there is much to celebrate, in the sense that it is difficult to imagine modern neuroscience without brain imaging. I will look at recent advances from the perspectives of functional segregation and integration in the brain, paying special attention to approaches that deal with the distributed and integrated nature of neuronal processing and the questions they address.

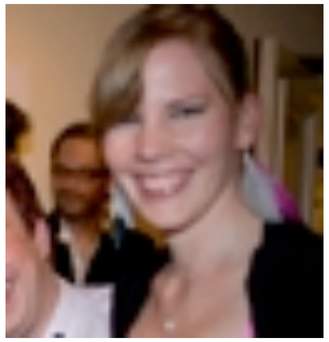
Learning goals

- Advantages of new statistical machine learning methods in cognitive neuroscience
- Prediction vs inference
- Mathematical form of the general linear model
- Difference between mass-univariate and multivariate analysis
- Encoding vs decoding, generative vs discriminative
- Classification, identification, reconstruction
- Functional segregation vs functional integration
- Structural, functional, effective connectivity

Coming week

Practical

Excursion to the Donders Magnetic Resonance Imaging facility!



Lecture

Discriminative decoding
Room TvA 4.00.19