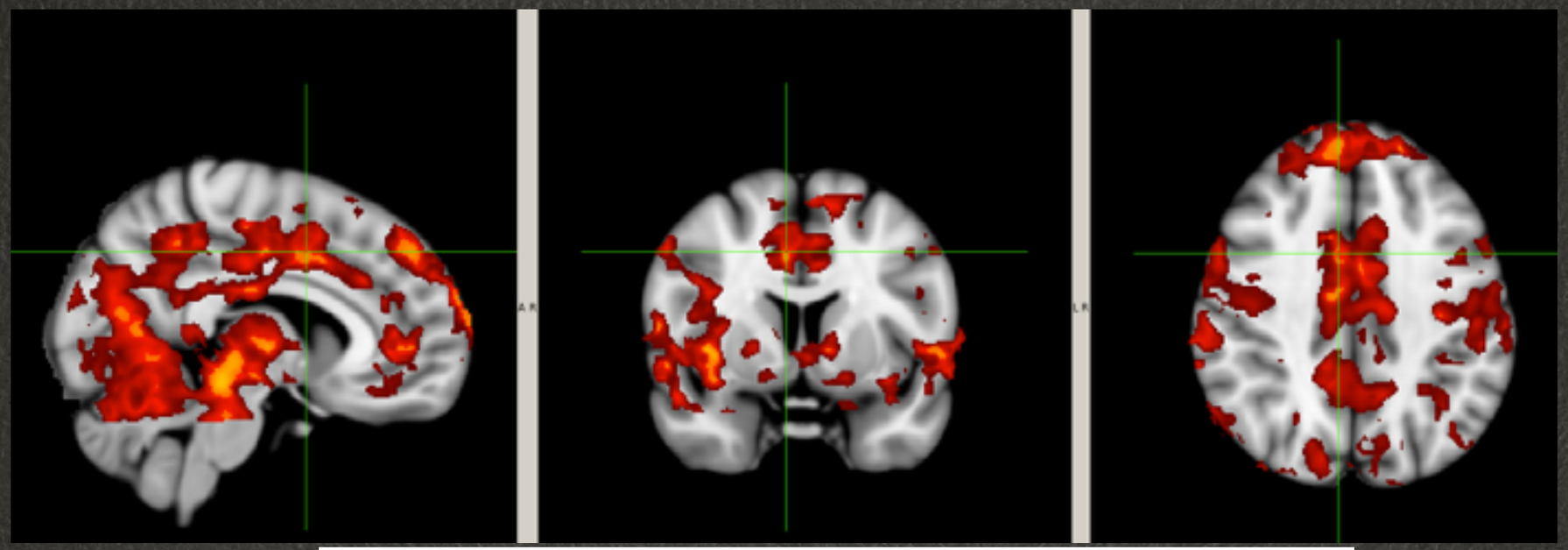
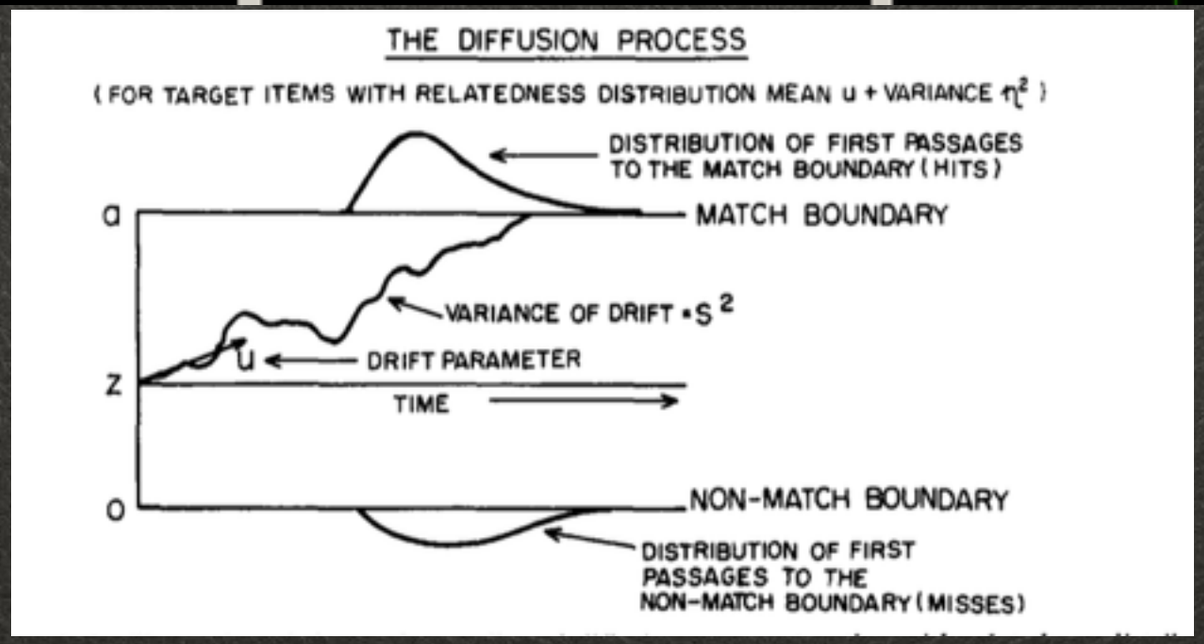
INTRODUCTION TO MODEL-BASED FMRI





Gilles de Hollander

Overview

- Functional magnetic resonance imaging (fMRI)
- Multivariate techniques
- “How to understand the brain”: linking computational cognitive models to neural data
- Practical: Python for model-based cognitive neuroscience

Overview

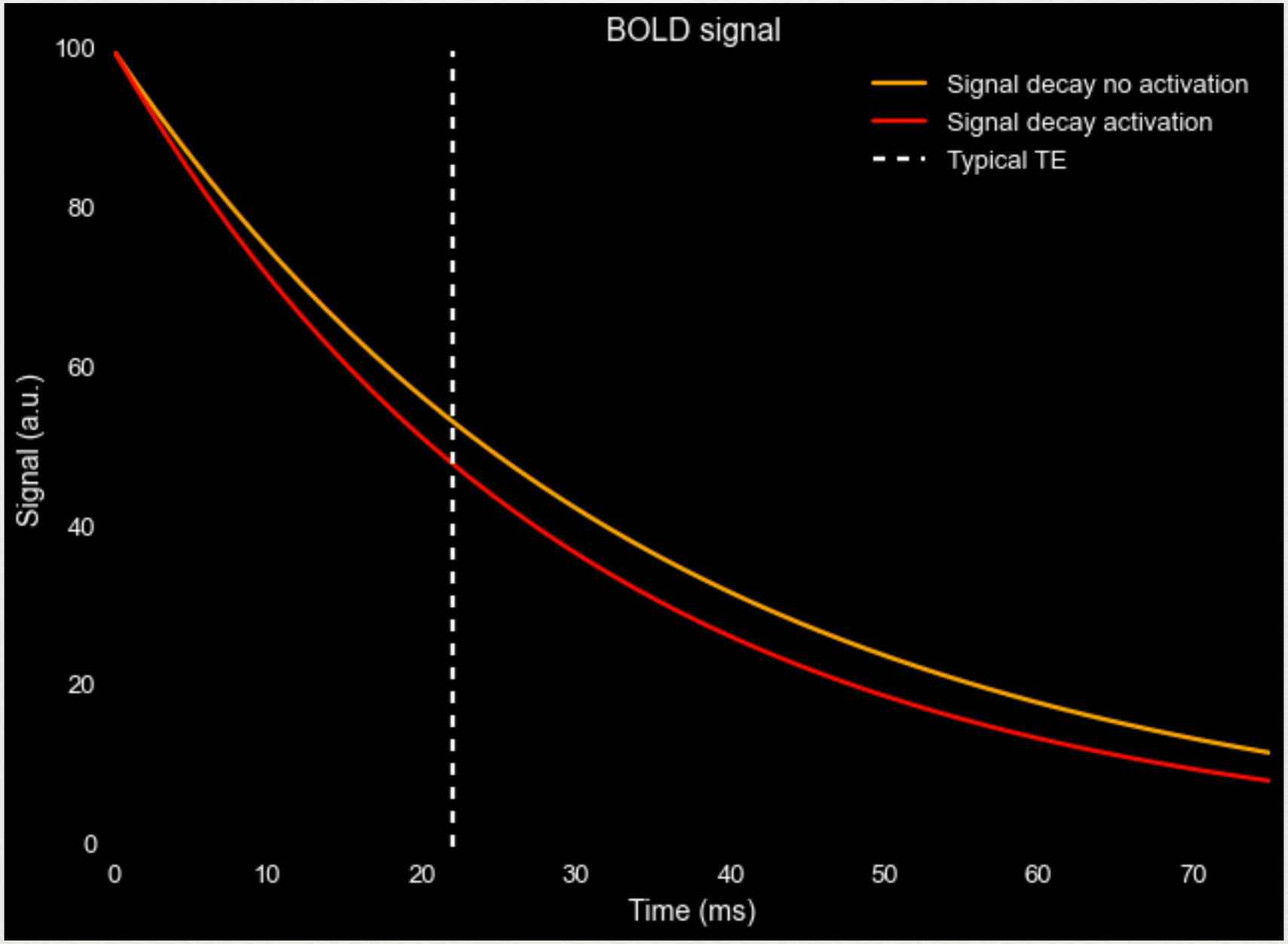
- **Functional magnetic resonance imaging (fMRI)**
- Multivariate techniques
- “How to understand the brain”: linking computational cognitive models to neural data
- Practical: Python for model-based cognitive neuroscience

Magnetic Resonance Imaging

- Explore magnetic properties of tissue using very strong magnetic field
    - 1.5, 3, and 7 Tesla
    - electromagnetic pulses
- Dynamical 3D images

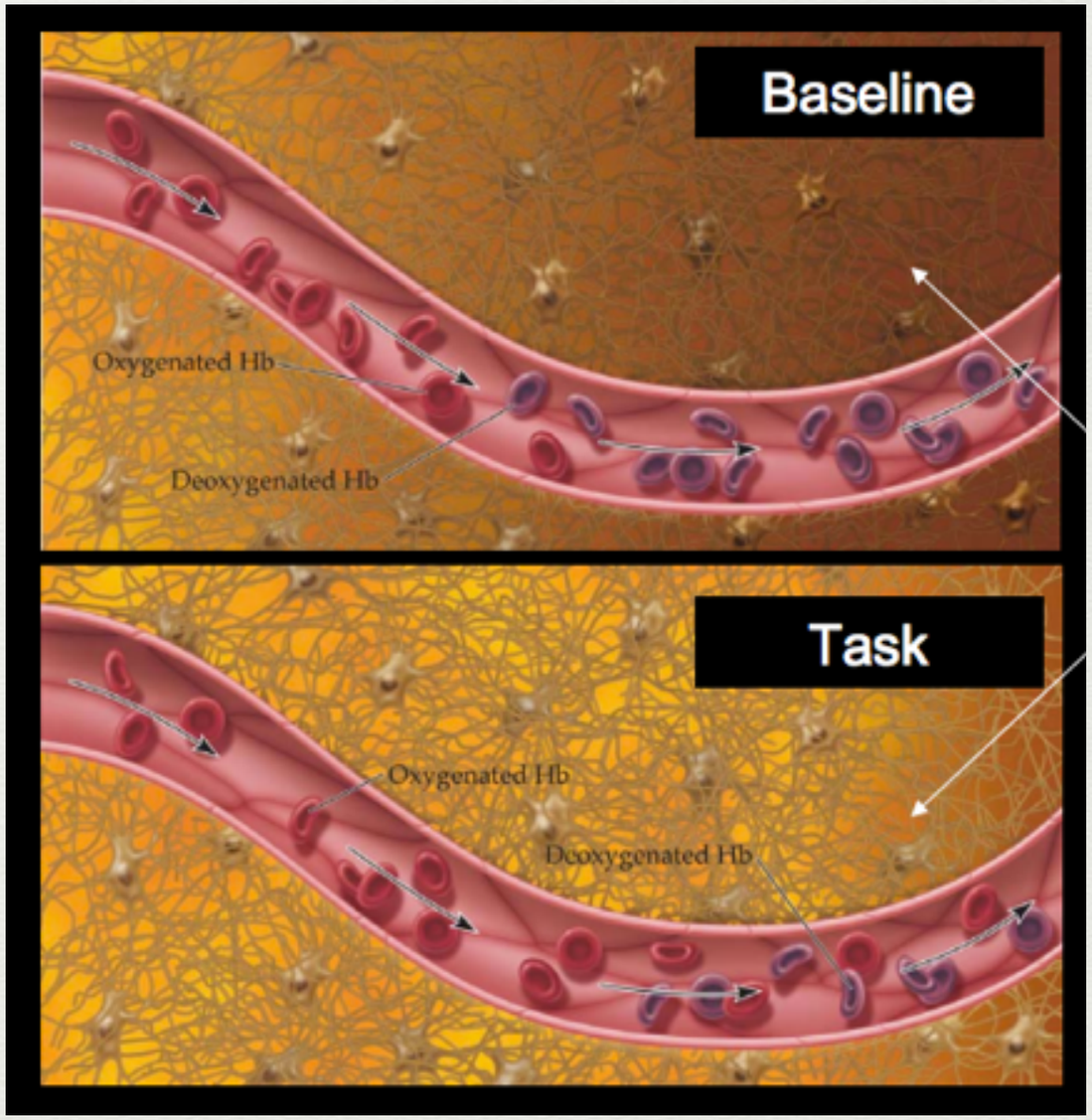
Functional magnetic resonance imaging

- Dynamic scanning (TR ~= 2s)
- T2*-weighted
- Sensitive to BOLD


BOLD signal
Signal decay no activation
Signal decay activation
Typical TE
Signal (a.u.)
100
80
60
40
20
0
0
10
20
30
40
50
60
70
Time (ms)

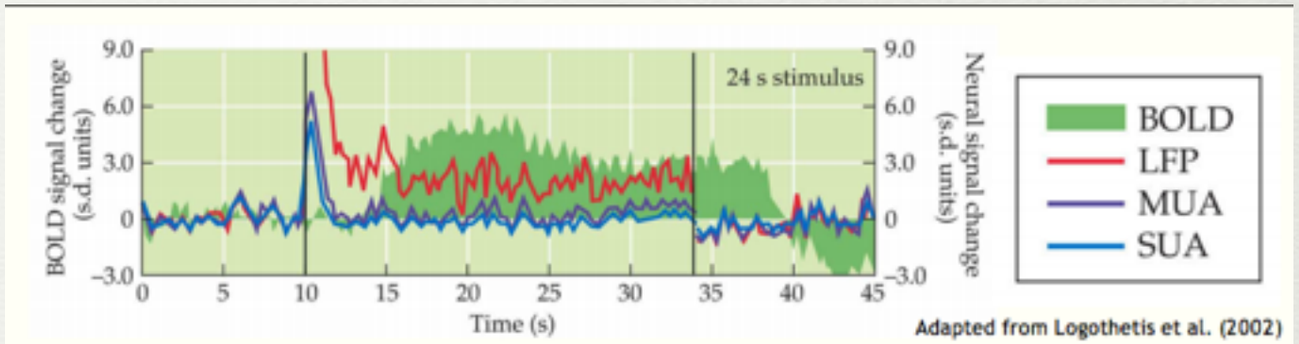
BOLD effect

- Combination of factors
  - ratio deoxygenated blood/oxygenated blood
  - CBV
  - Blood flow

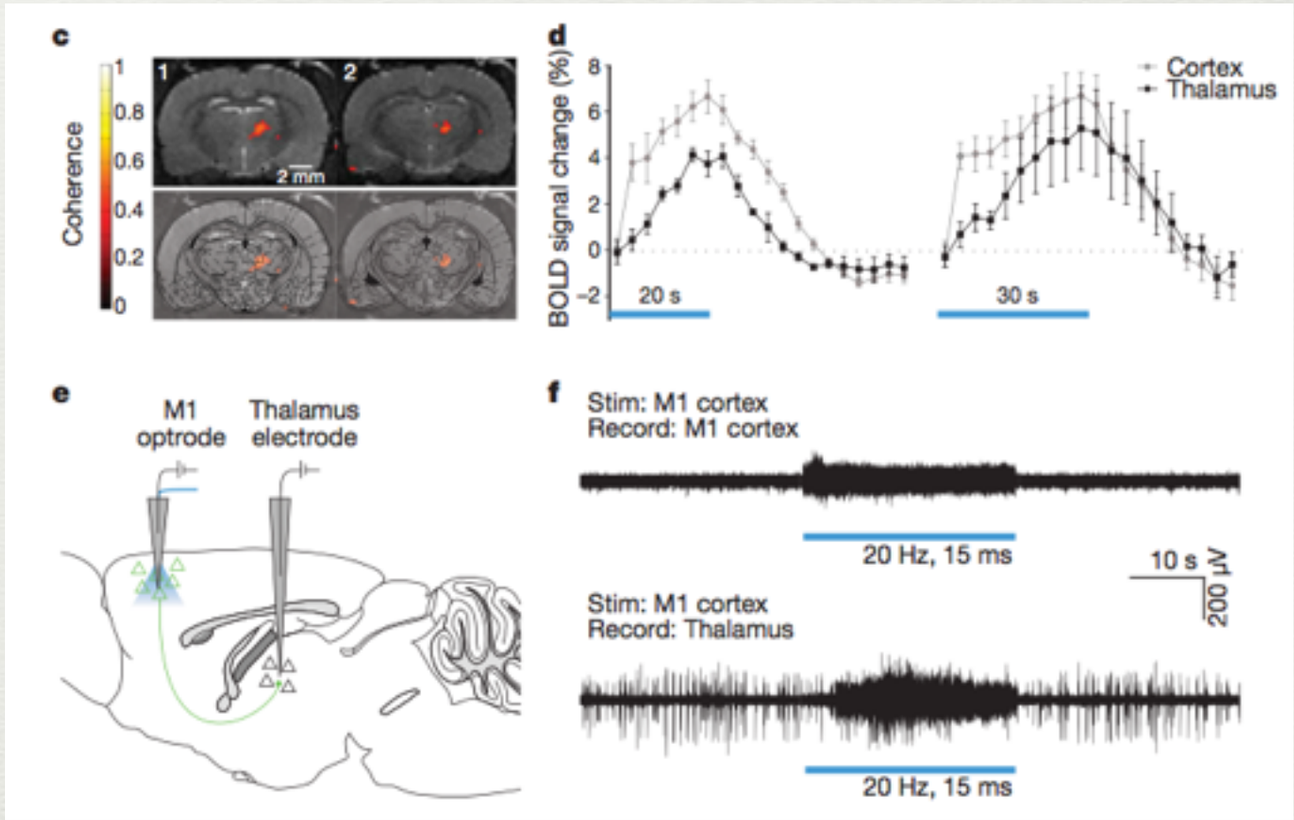


Validity

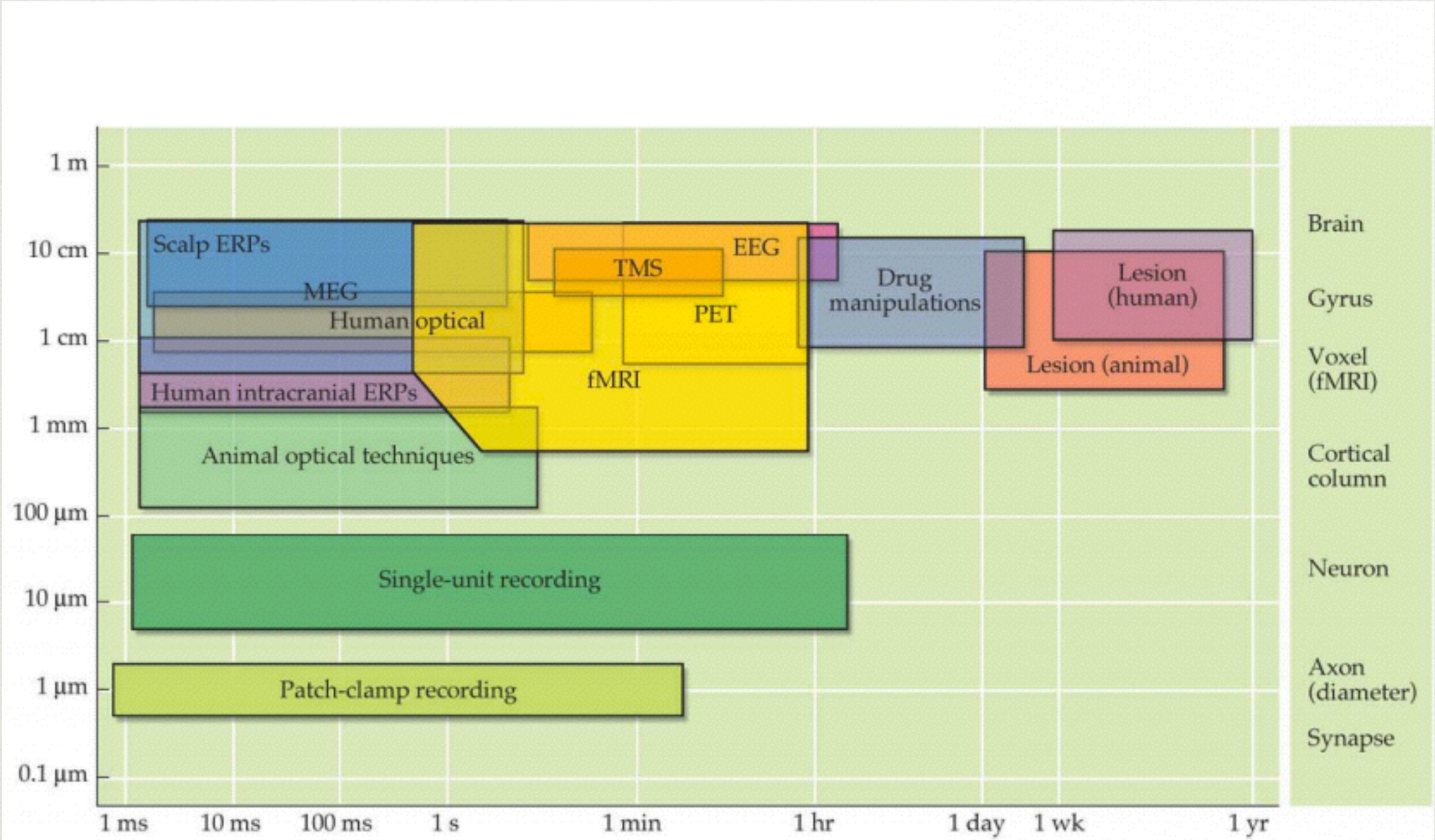
- Logothesis (2002)



- Lee (2010)

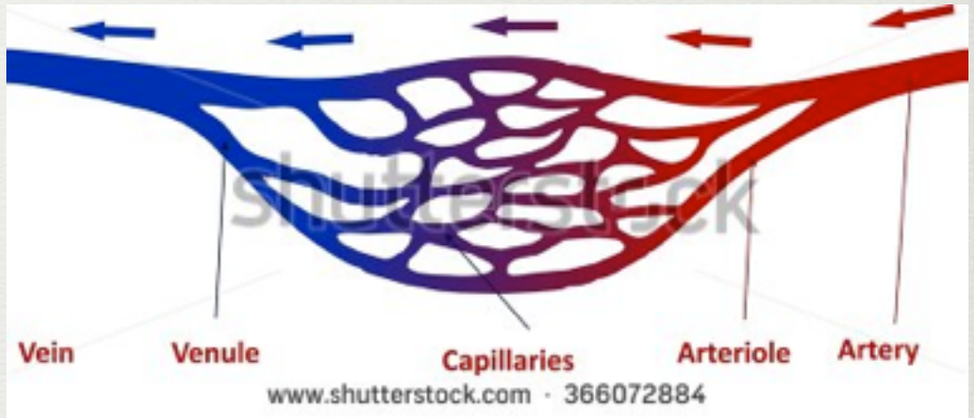


Resolution


1 m
10 cm
1 cm
1 mm
100 µm
10 µm
1 µm
0.1 µm
1 ms
10 ms
100 ms
1 s
1 min
1 hr
1 day
1 wk
1 yr
Scalp ERPs
MEG
Human optical
Human intracranial ERPs
Animal optical techniques
TMS
EEG
PET
fMRI
Drug manipulations
Lesion (human)
Lesion (animal)
Single-unit recording
Patch-clamp recording
Brain
Gyrus
Voxel (fMRI)
Cortical column
Neuron
Axon (diameter)
Synapse

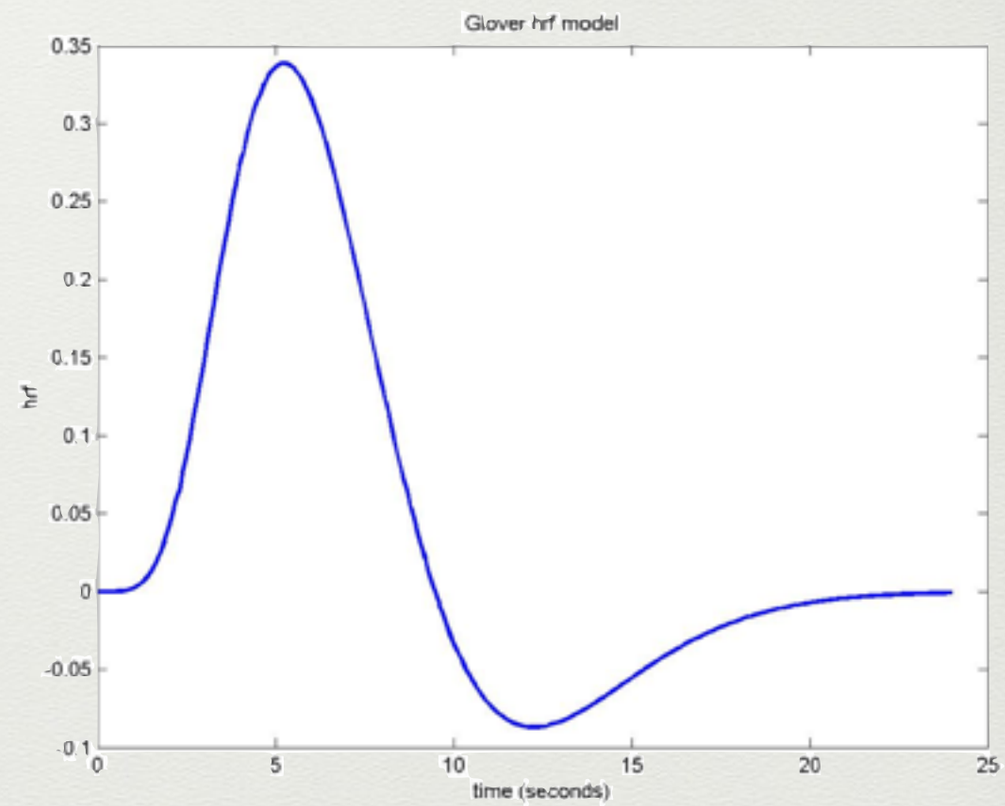
Spatial Resolution

- Typical voxel sizes: 3x3x3 mm (3T) or 1x1x1mm (7T)
- Mean distance between closest cappilary and individual neuron in order of 50 µm (0.05 mm)
- Neural activity influences vessel dilation at scale of milimeters

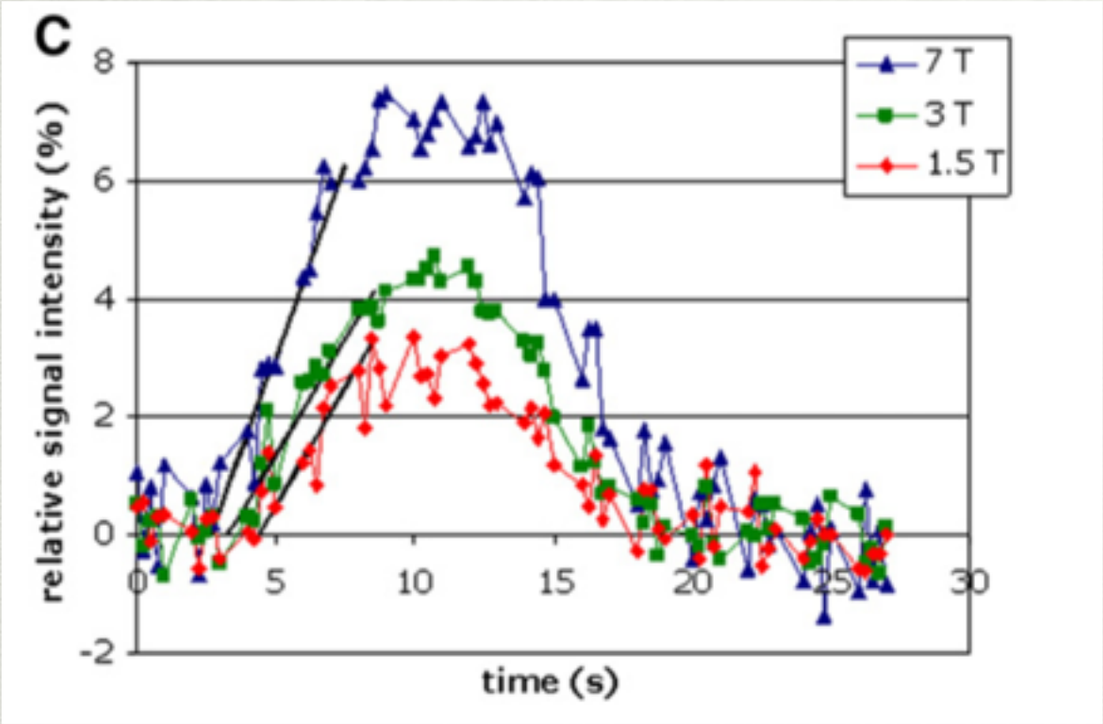


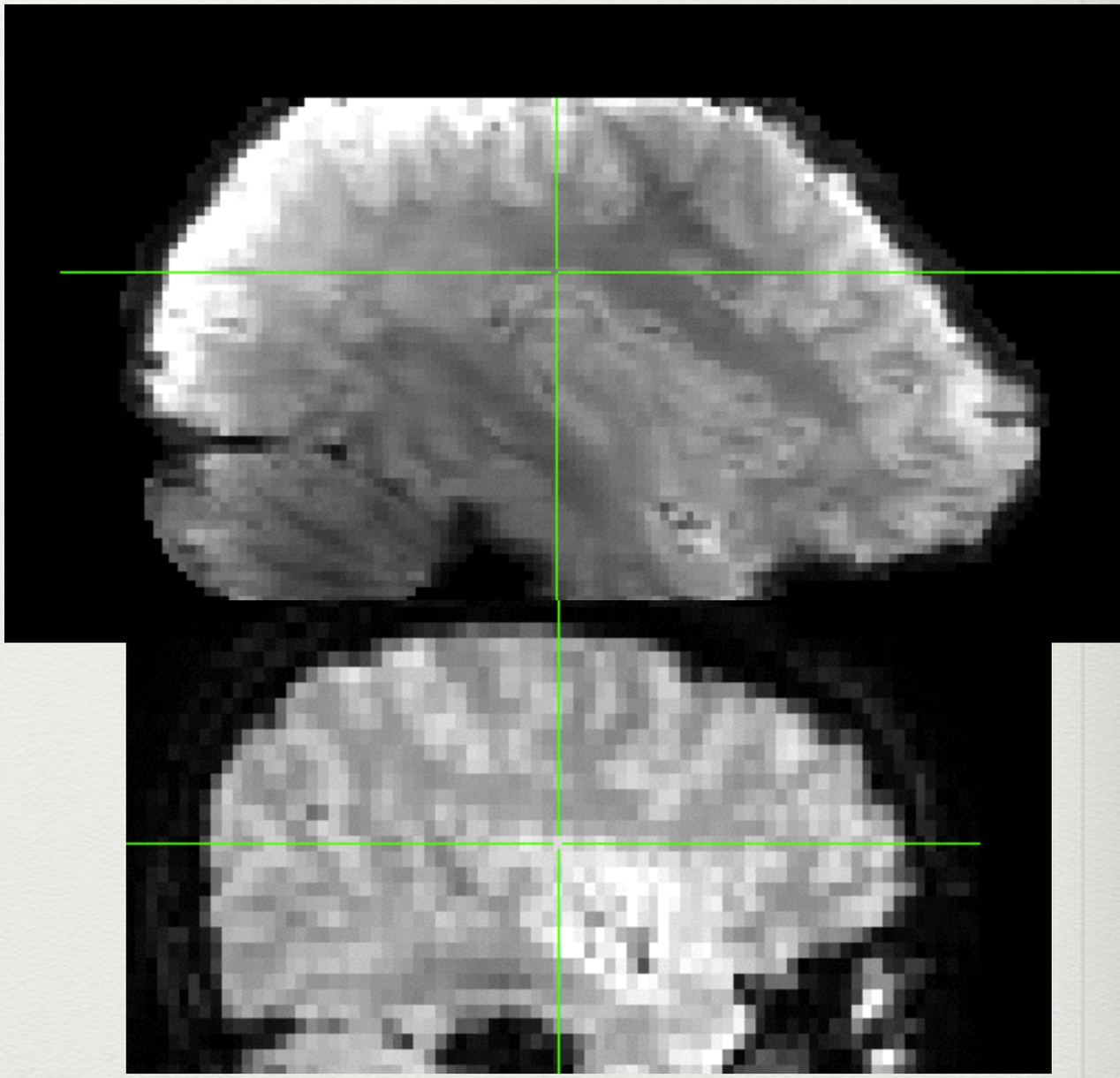
Temporal resolution of fMRI

- BOLD-response is rather sluggish
- Order of seconds
- TR of most scanning sequences is around 2 seconds



Field strengths



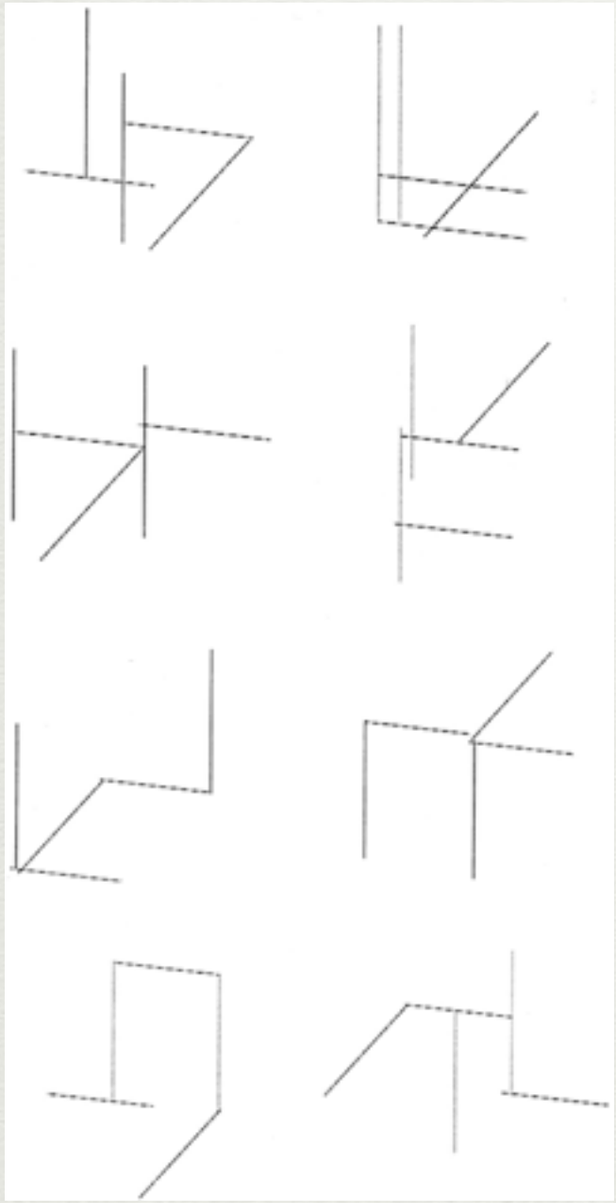


van der Zwaag, 2009

What fMRI can do

- “Sluggish indirect measure of neural activity, with a temporal resolution of 1-3 s and a spatial resolution of 25-30mm3.”
- “Because of its limited temporal and spatial resolution, fMRI is most appropriate for answering questions about gross neural architecture, rather than about neural process.”

What fMRI can do



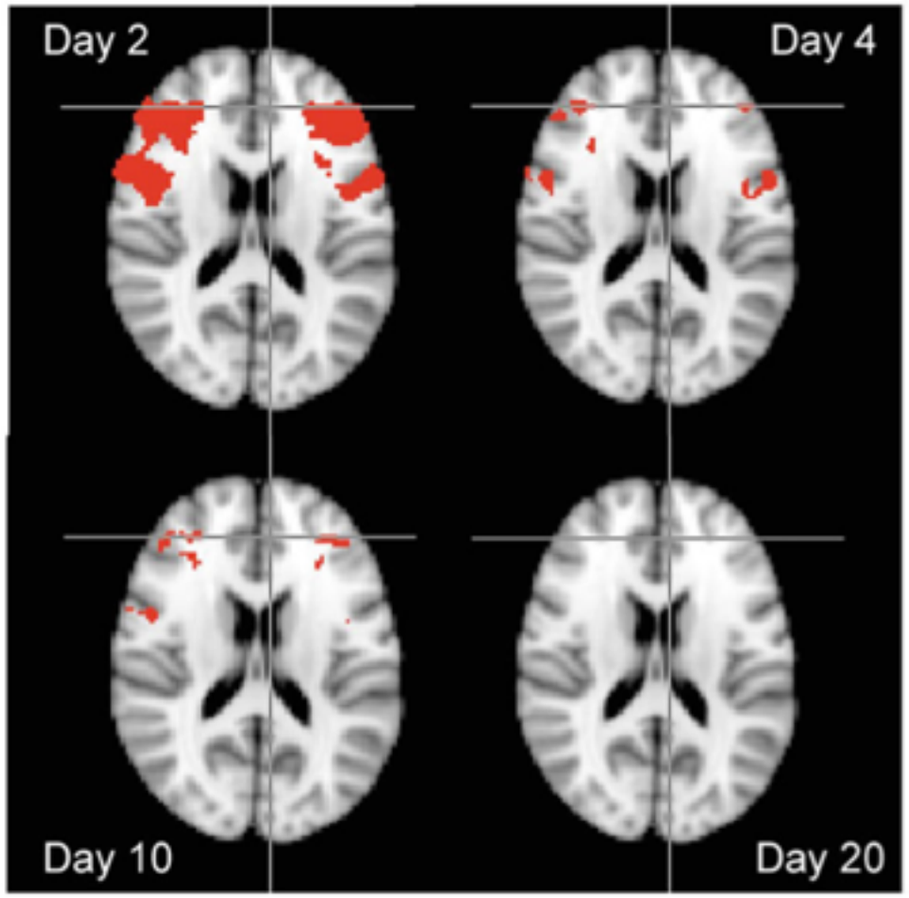
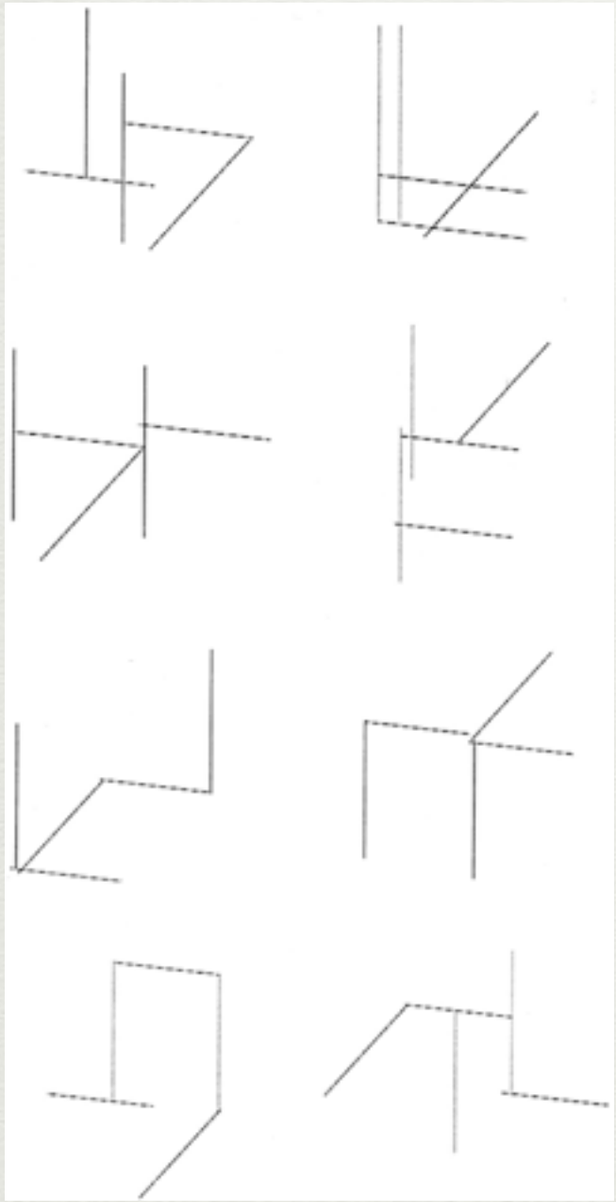
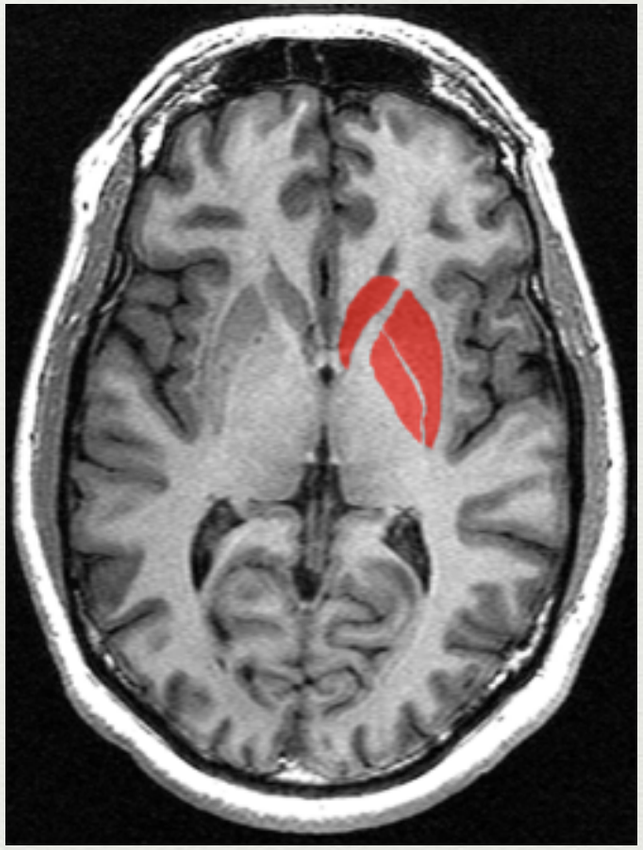


Fig. 5.1 Significant BOLD activation in prefrontal cortex on four different days of training on the same categorization task [8]

Waldschmidt & Ashby (2011)

What fMRI can do





Waldschmidt & Ashby (2011)

What fMRI can do

- “suppose some cognitive theory predicts that the same perceptual and cognitive processes mediate performance in two different tasks. Then this theory should predict similar patterns of activation in an fMRI study of the two tasks”

Lopez-Paniagua & Seger (2011)

What fMRI can do

- In sum
  - Good measure to understand neural and cognitive architecture
  - Less powerful to understand exact underlying neural processes

Possibilities

- Difference in brain activation across conditions
  - Different classes of stimuli
  - Different cognitive states
- Parametric designs
- Effective connectivity patterns
  - PPI
  - DCM
  - Ancestral-graph theory
- Resting state activity
- Model-based analysis

Possibilities

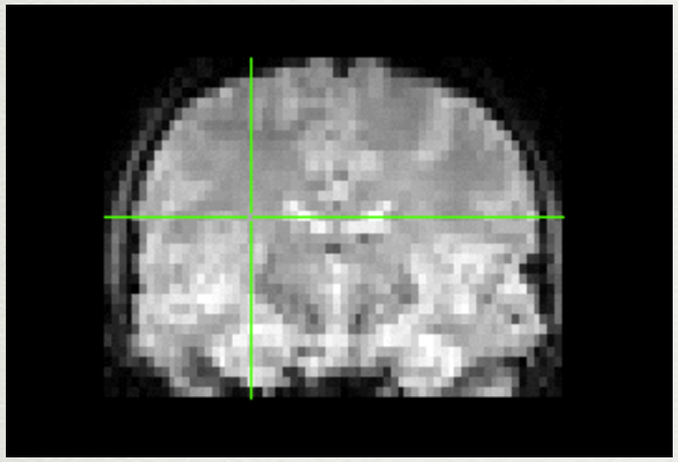
- Difference in brain activation across conditions
  - Different classes of stimuli
  - Different cognitive states
- Parametric designs
- Effective connectivity patterns
  - PPI
  - DCM
  - Ancestral-graph theory
- Resting state activity
- **Model-based analysis**

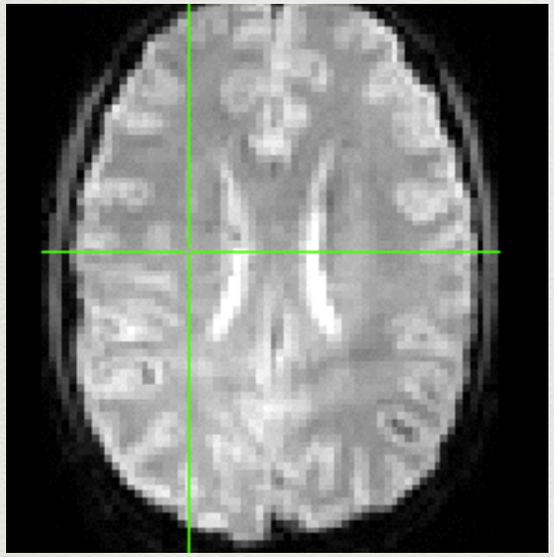
Typical scanning session

- Localizer (2min)
- Structural scan (8-10 min)
- Functional scan (30-60 min)
- Fieldmap (5 min)

Functional

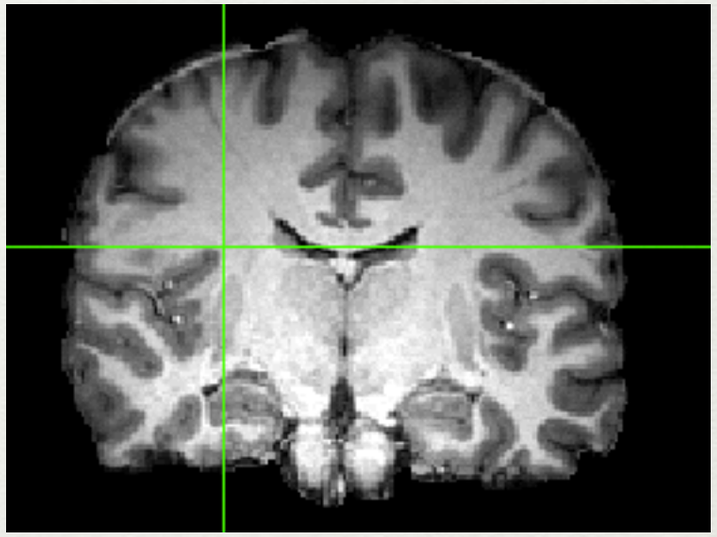
(approx 3mm isotropic)

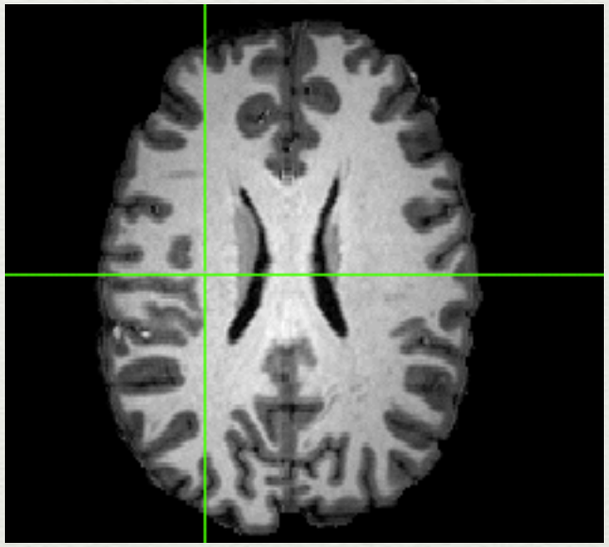




Structural

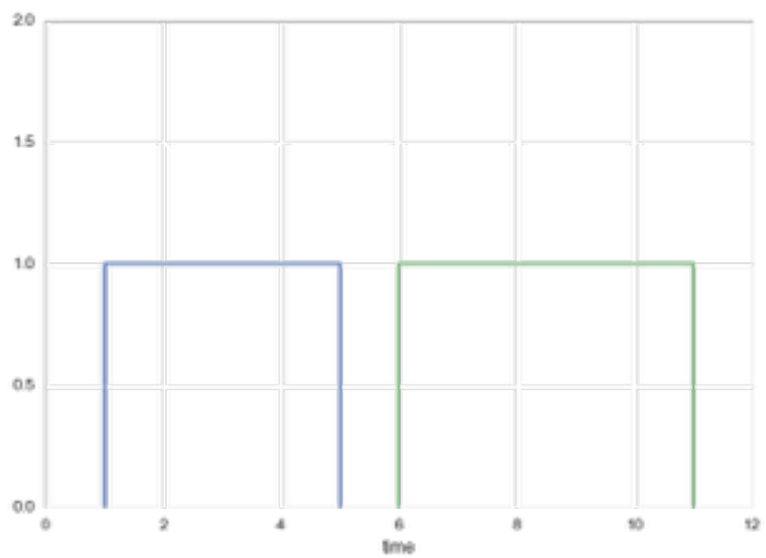
(approx 1mm isotropic)



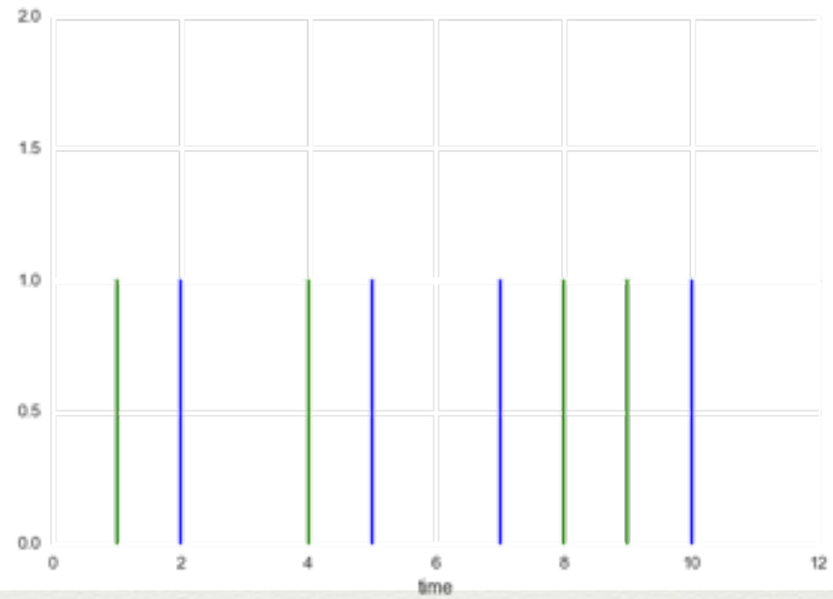


Experimental design:

Block



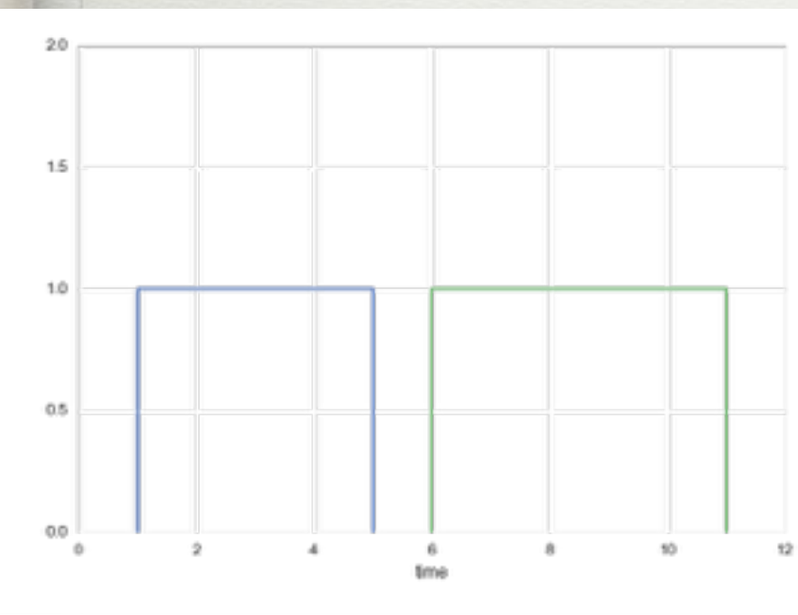
Event-related



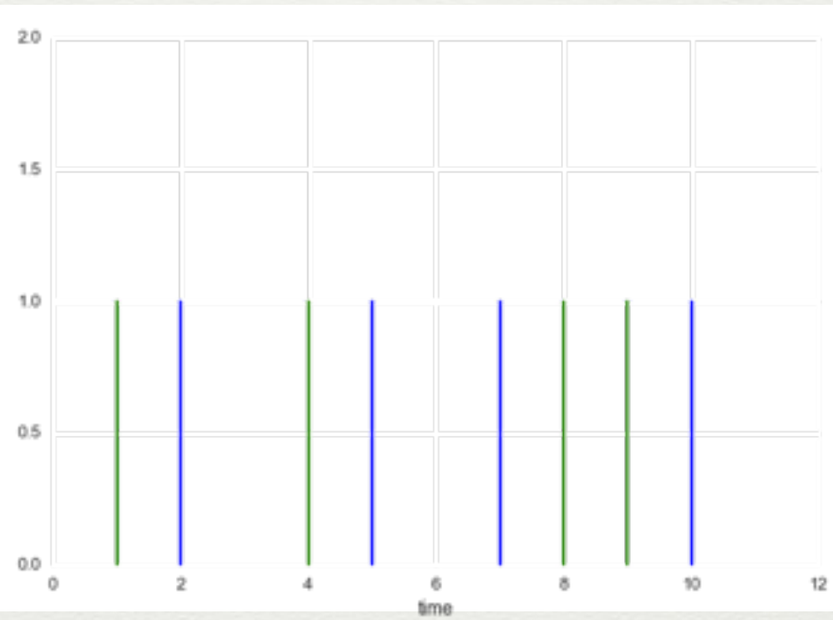
Free-running

Experimental design:

Block



Event-related



Free-running

inter-trial-interval

Goals

- Find differences in neural activation between conditions (contrast)
- Find correlations between fMRI and behavioural measures (parametric design)

Data Analysis

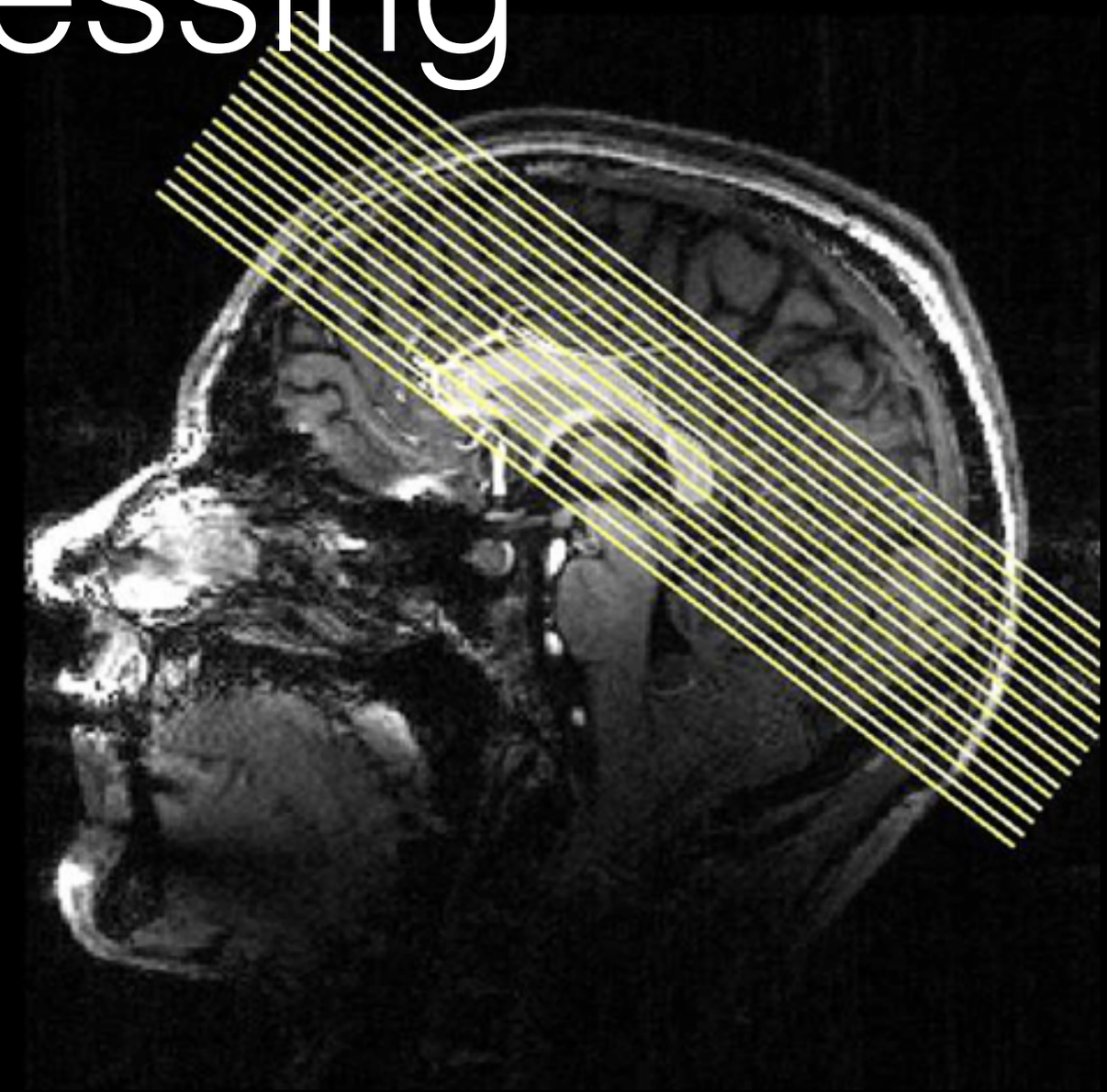
- Very challenging
  - Large amounts of data
  - Spatial and temporal correlation
  - Large amounts of noise

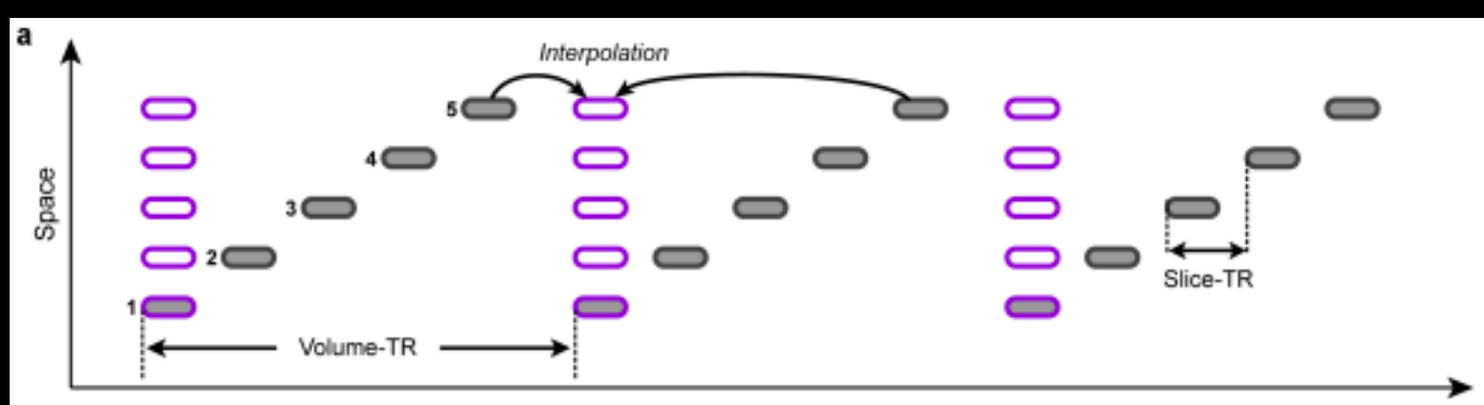
Data Analysis

- Three stages
  - Preprocessing
  - Individual subject analysis
  - Group-wise inference

Preprocessing

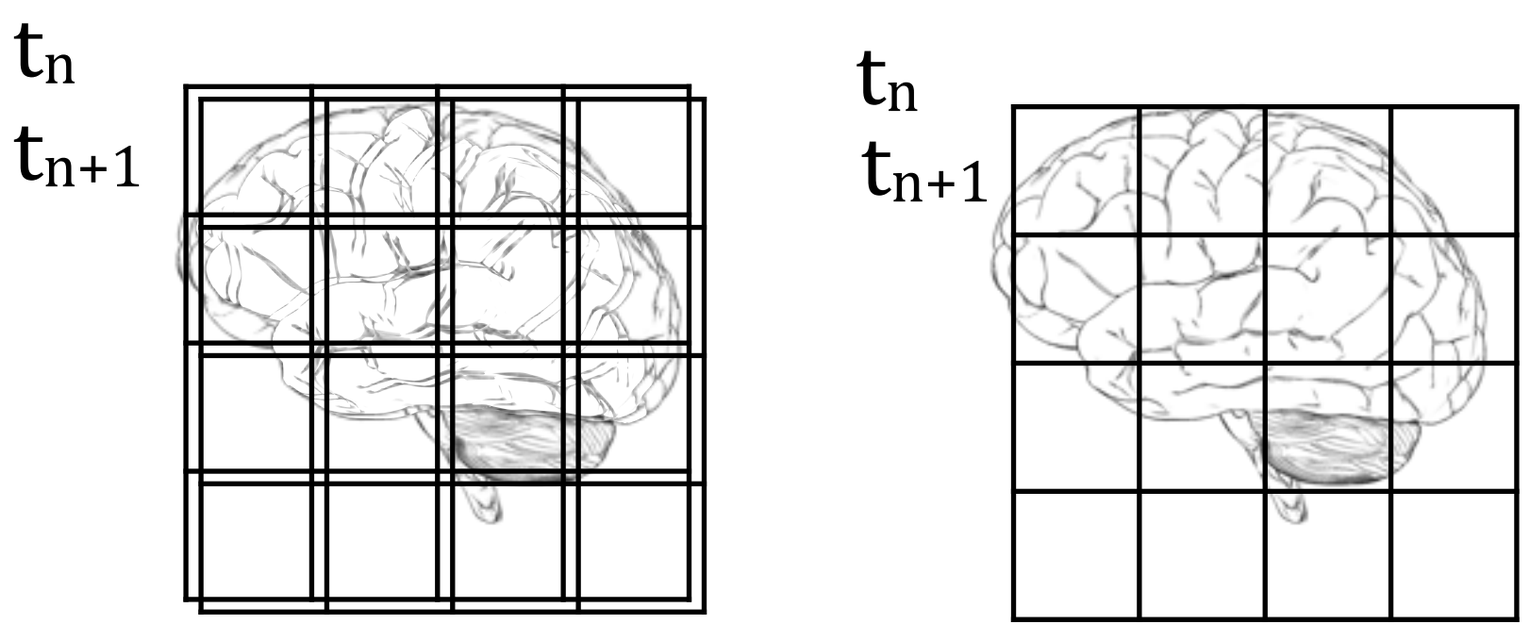
- 1. Slice-time correction





Preprocessing

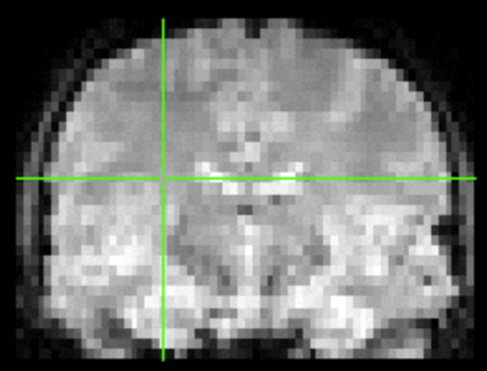
- 2. Motion correction



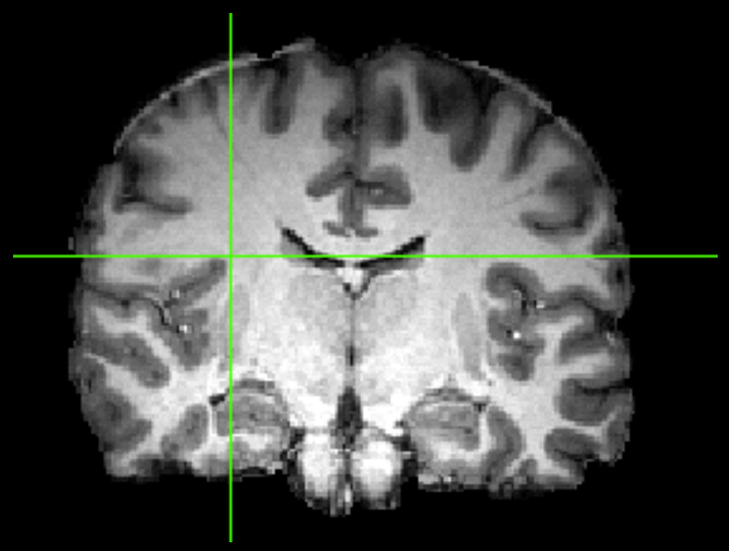
Preprocessing

- 3. Coregistration

Functional
(approx 3mm isotropic)



Structural
(approx 1mm isotropic)

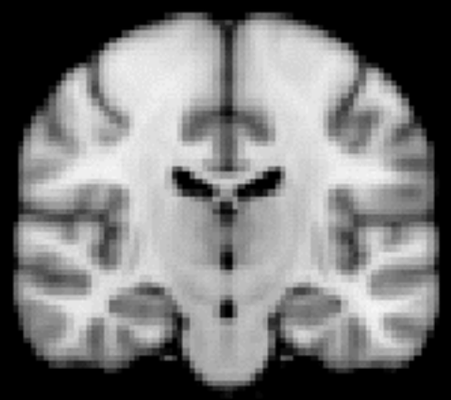


Preprocessing

- 4. Normalization

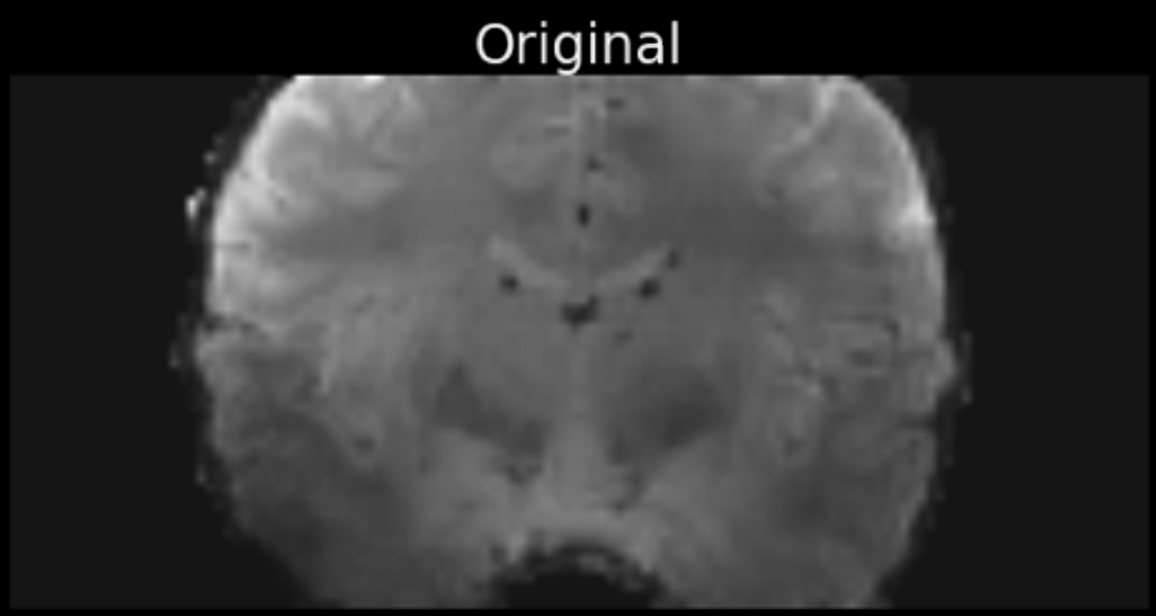
Structural
(approx 1mm isotropic)

MNI152
(0.4/1/2 mm isotropic)



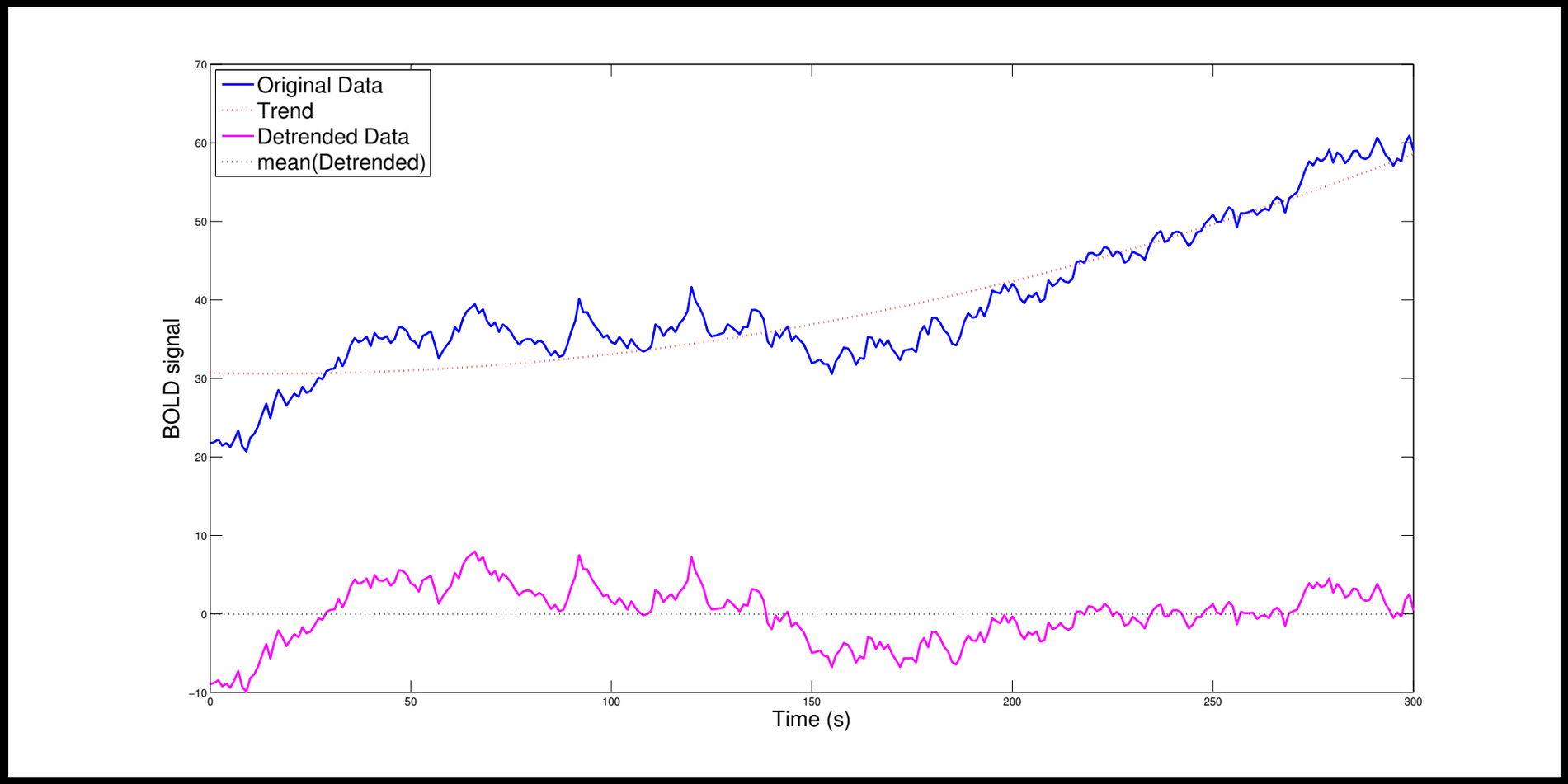
Preprocessing

- 5. Smoothing



Preprocessing

- 6. Temporal filtering

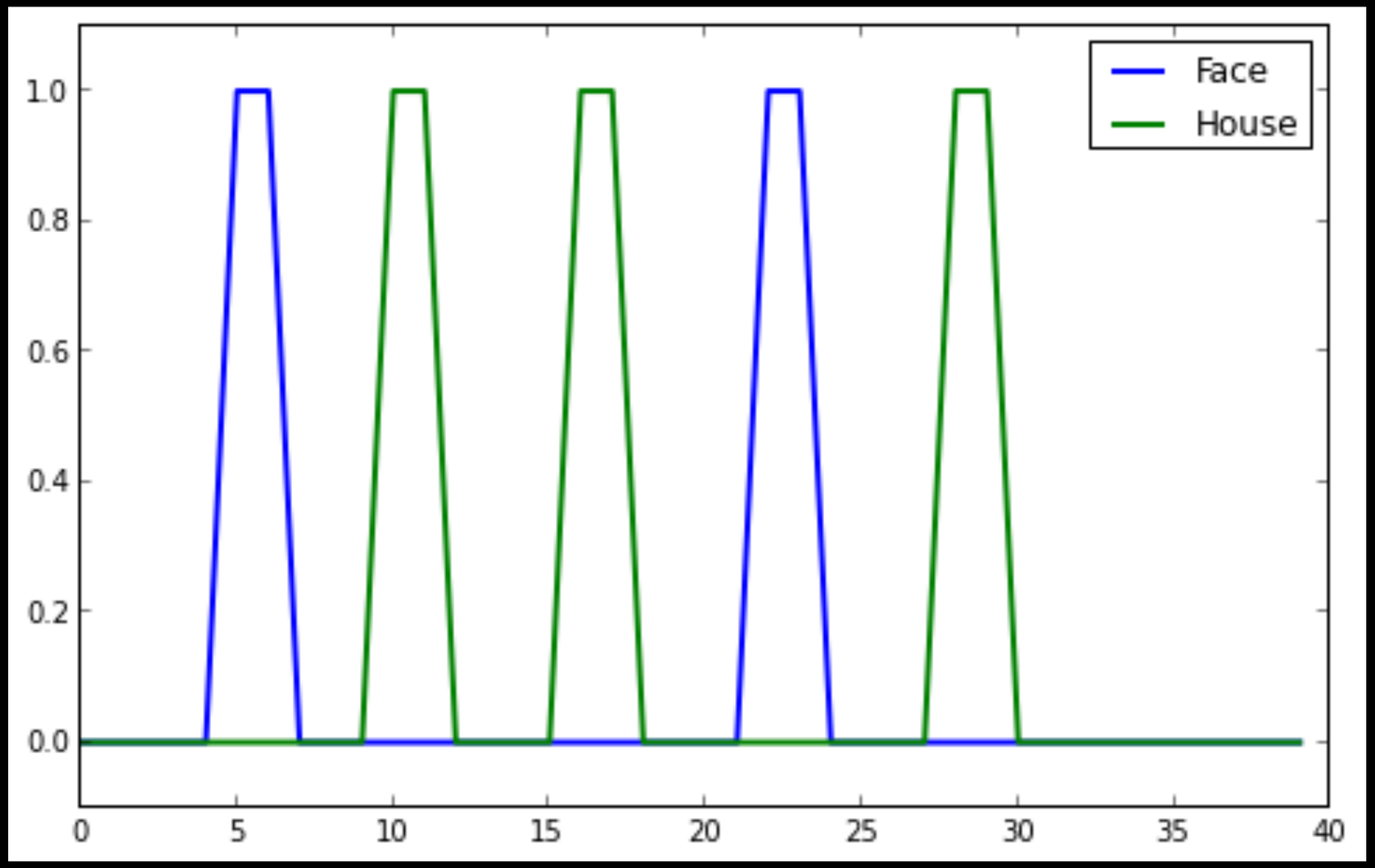


Task-related data analysis

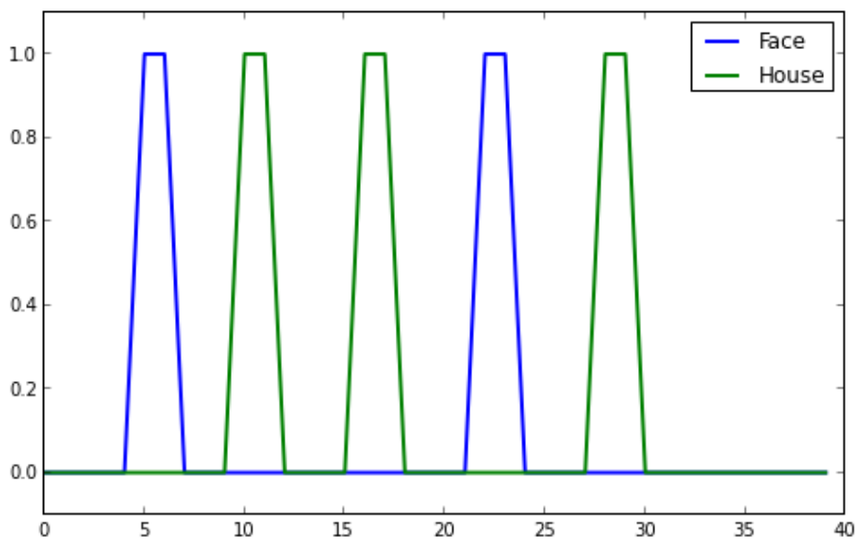
Task-related data analysis

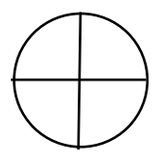
- General idea:
  - “What would a task-related BOLD signal look like?”
  - Correlate this hypothetical signal to all voxels in the brain

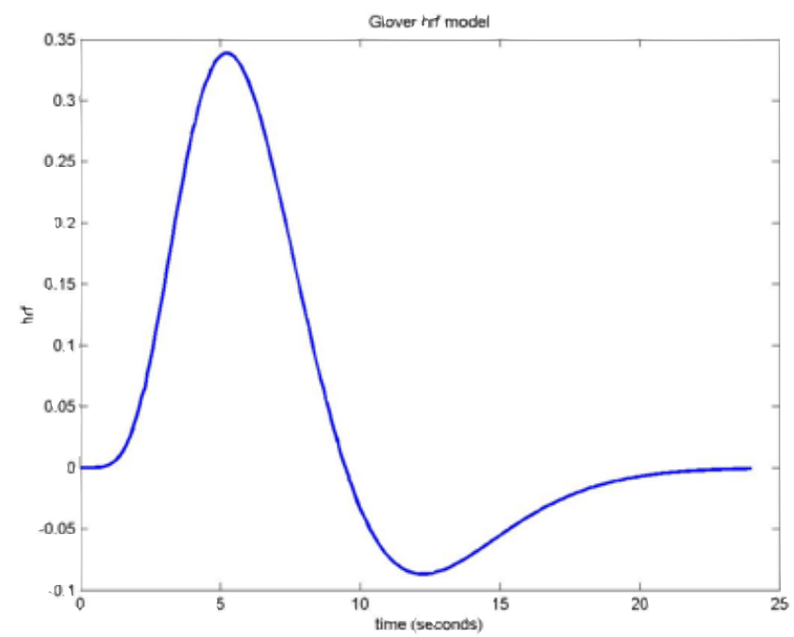
Task-related data analysis

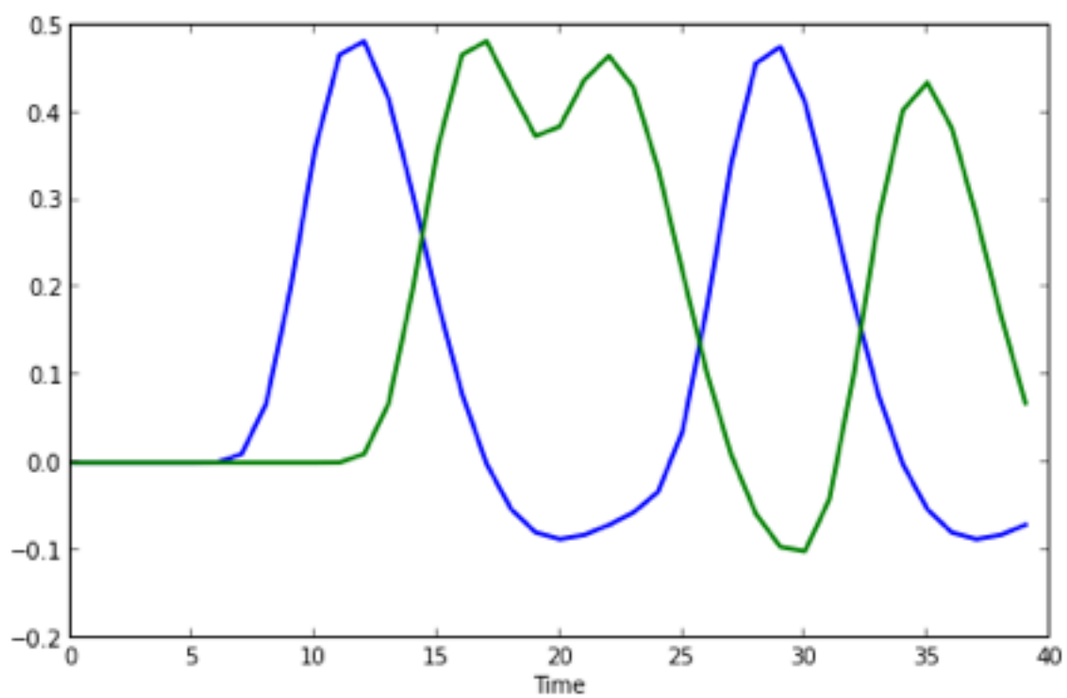


Task-related data analysis





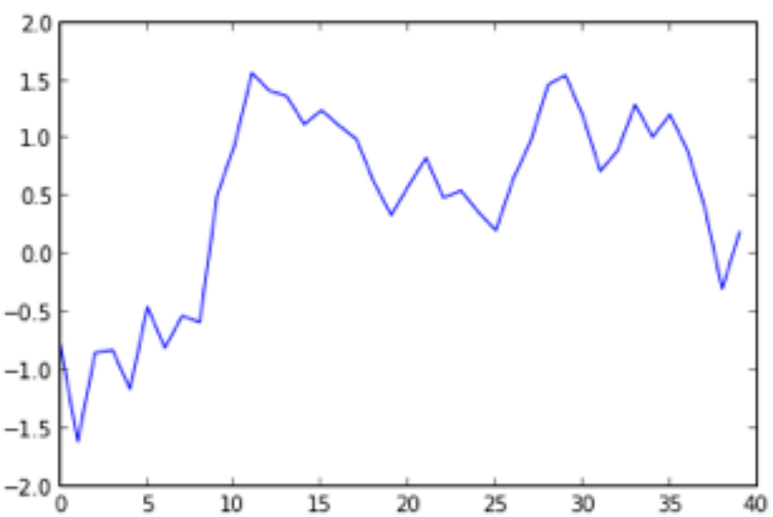




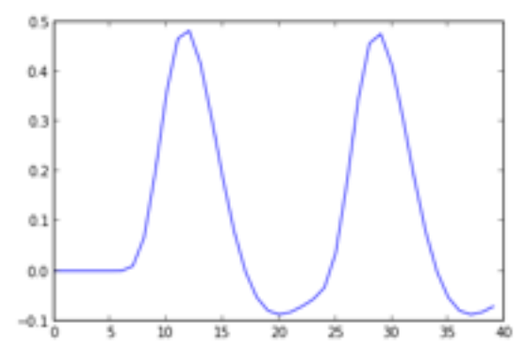
GLM

$$Y_i = \beta X + \varepsilon$$

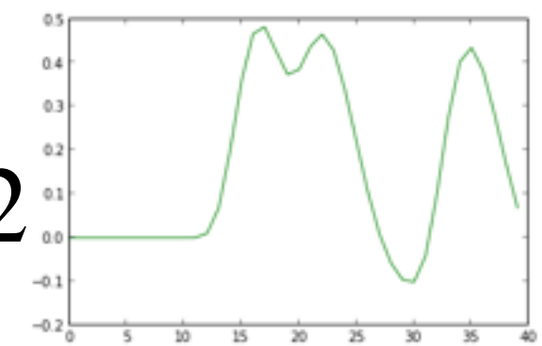
$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots \beta_n X_n + \varepsilon_i$$



$= \beta_0 + \beta_1$

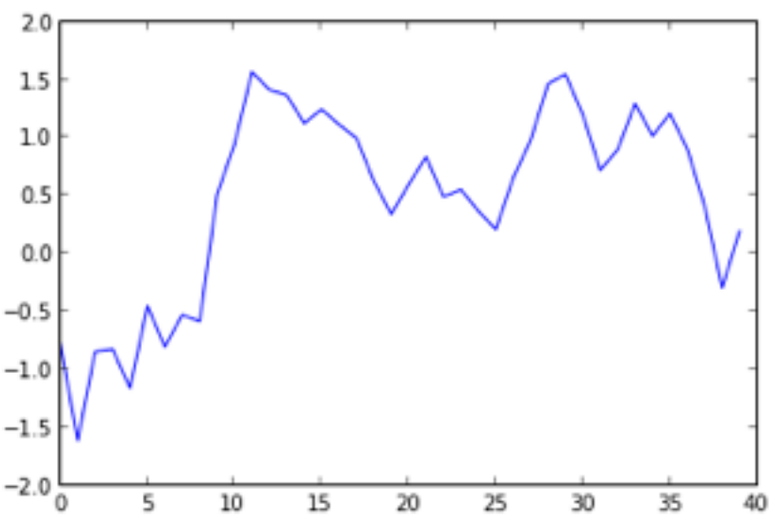


$+ \beta_2$

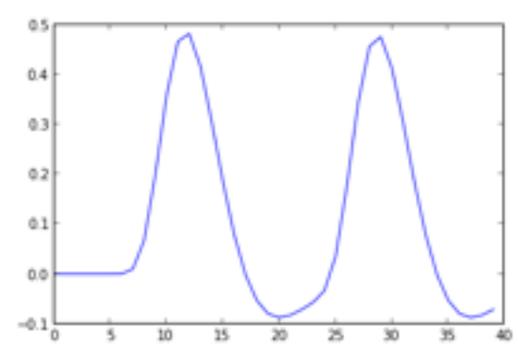


GLM

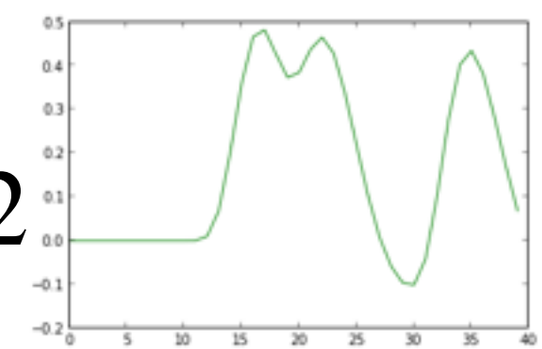
$$\beta_0 = 0$$
$$\beta_1 = 4$$
$$\beta_2 = 3$$



$$= \beta_0 + \beta_1$$



$$+ \beta_2$$



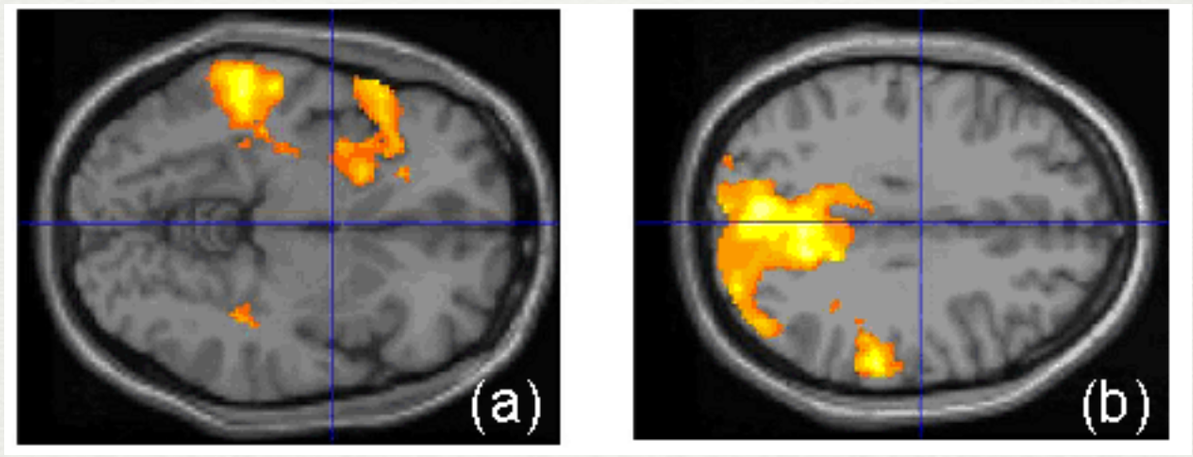
Result

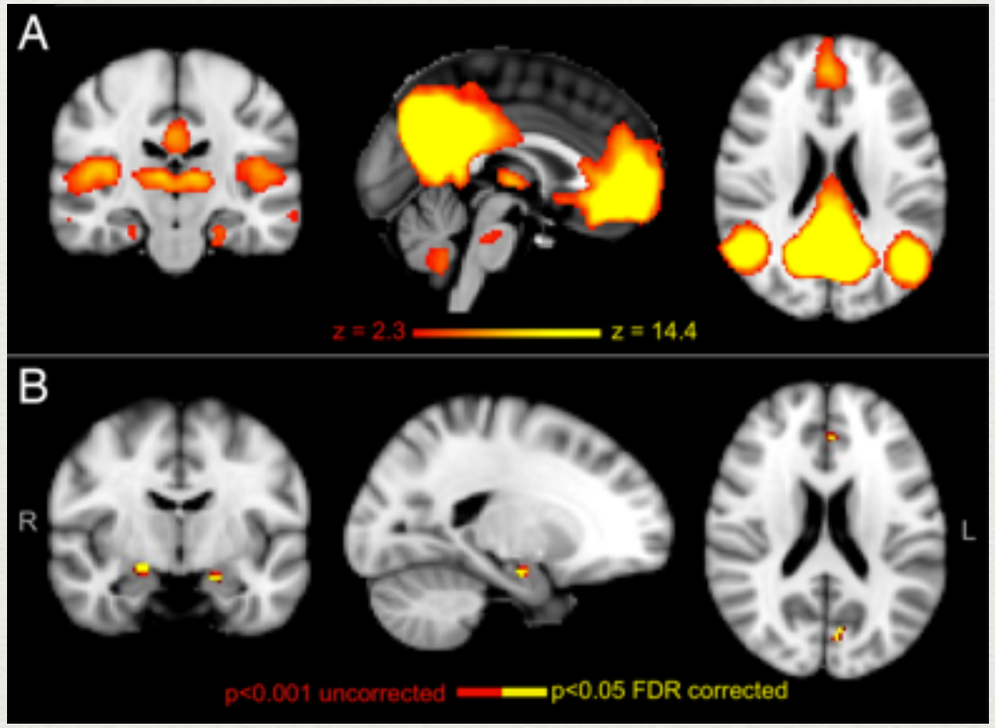
- For every subject a brain map with
  - For every voxel
    - Beta-value for every regressor
    - Variance for every regressor
  - (contrasts)

Group inference

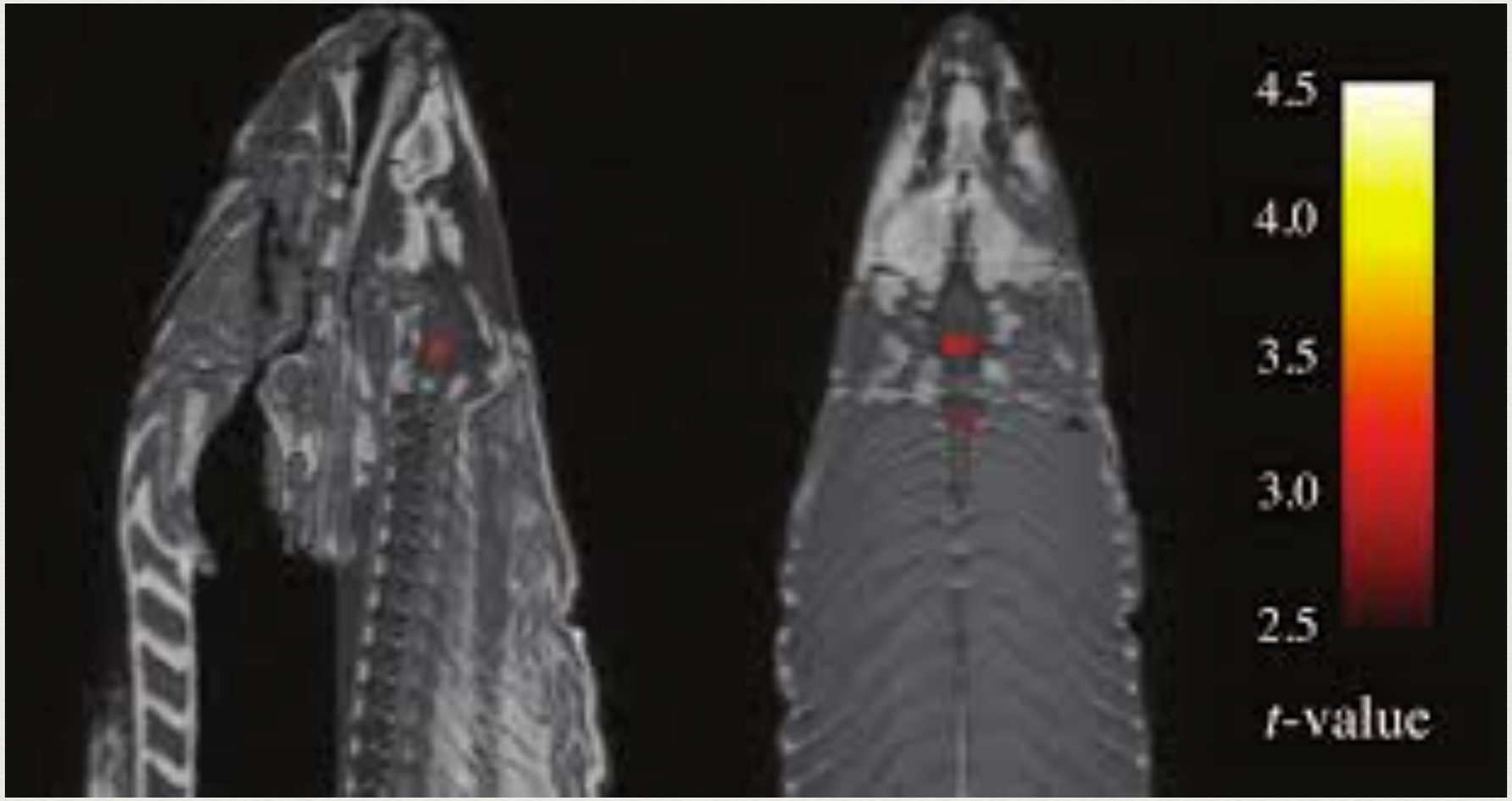
- Combine beta-regressors/contrasts all subjects
- Bring them to a standard space (MNI152)
- Apply a random effects model
  - Individual subject variance is taken into account

Statistical Parametric Mapping





Multiple comparisons



Bennett et al., 2009

Multiple comparisons

- False Discovery Rate (FDR)
- Gaussian Random Field Theory (GRF)

Summary

- Most fMRI experiments are event-related design
  - Principle of *pure insertion*
  - Univariate analysis
    - Which voxels correspond to the inserted cognitive construct?
    - Multiple comparison problem

Overview

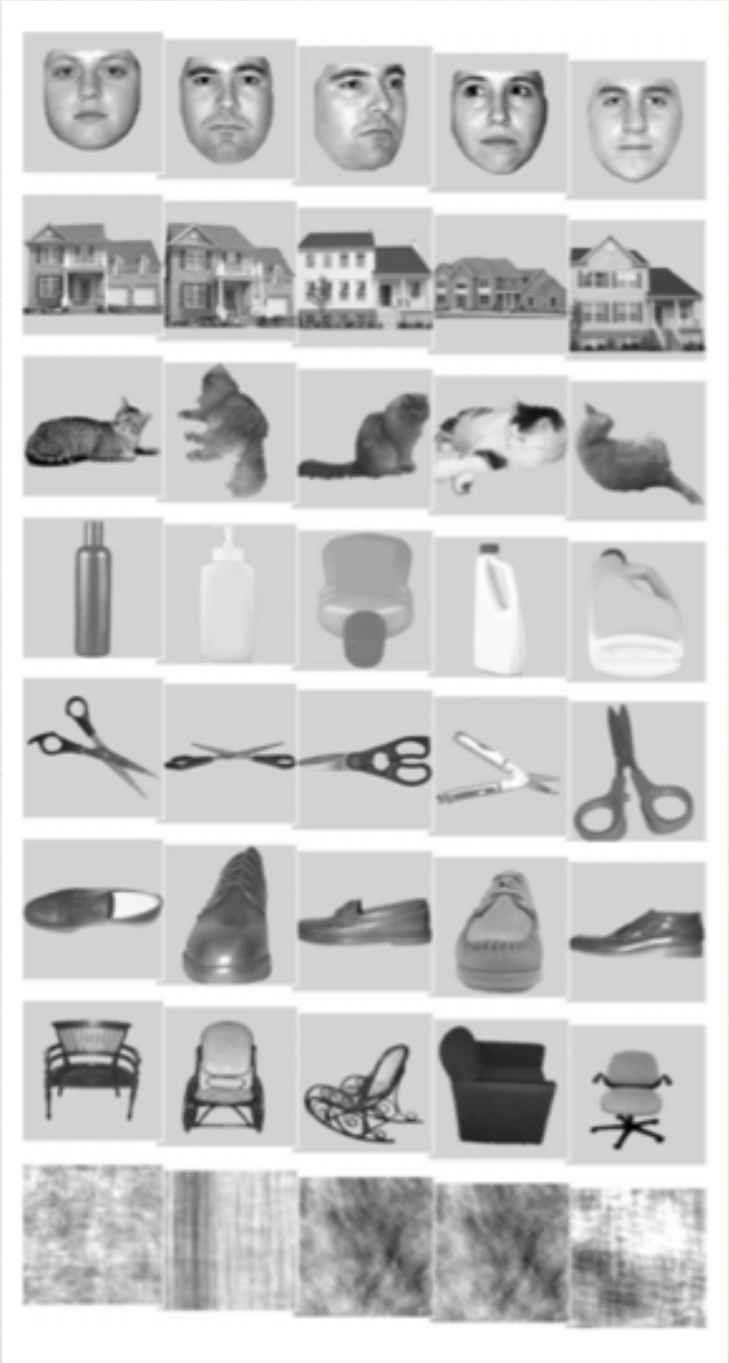
- Functional magnetic resonance imaging (fMRI)
- **Multivariate techniques**
- “How to understand the brain”: linking computational cognitive models to neural data
- Practical: Python for model-based cognitive neuroscience

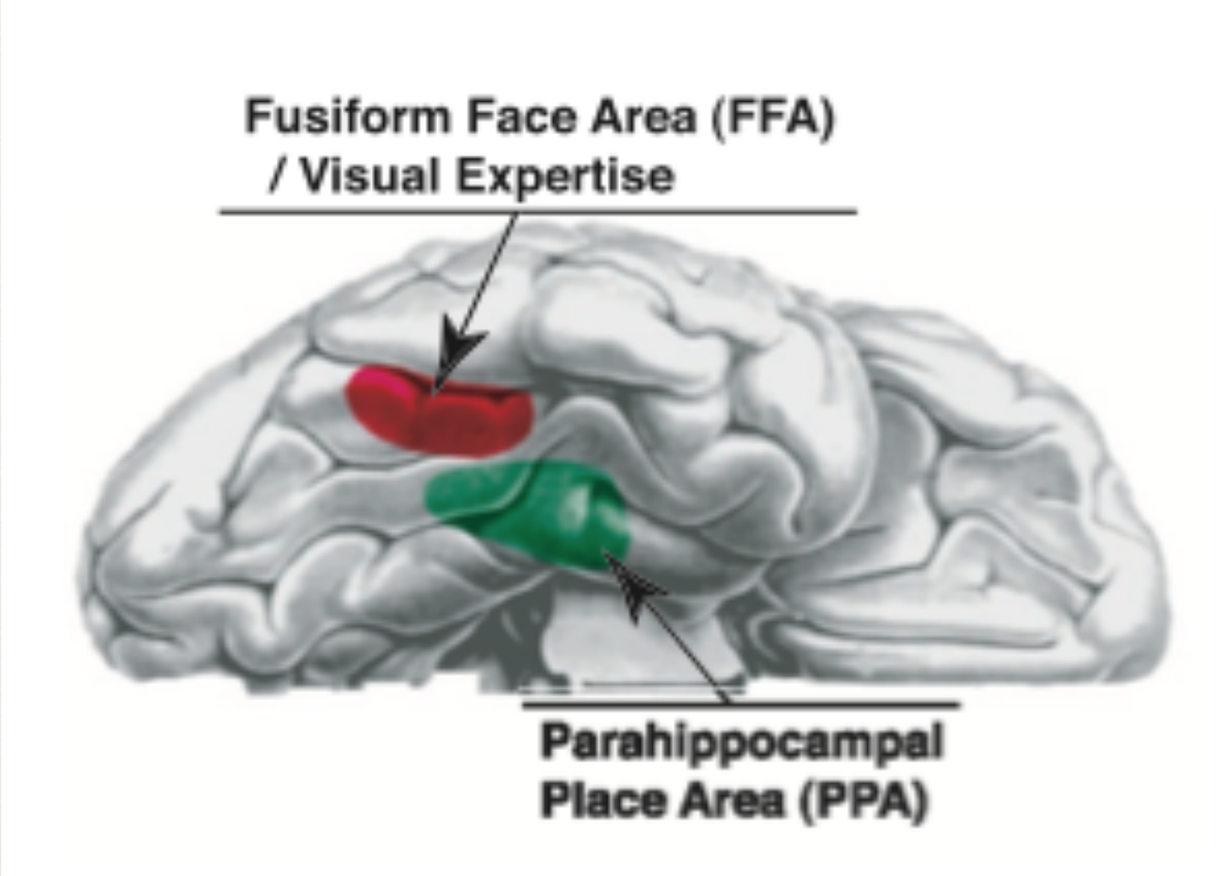
Science
AAAS

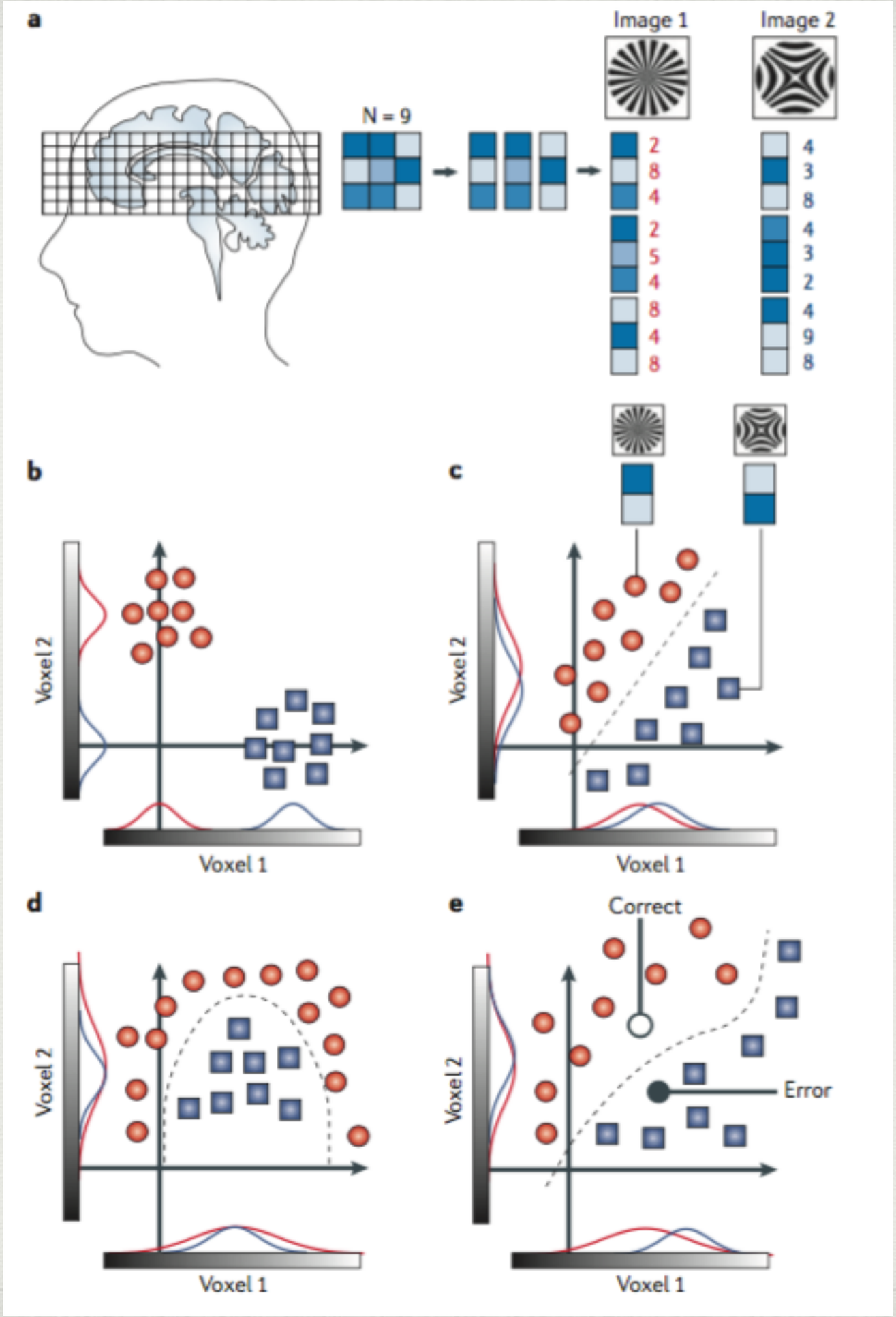
Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex

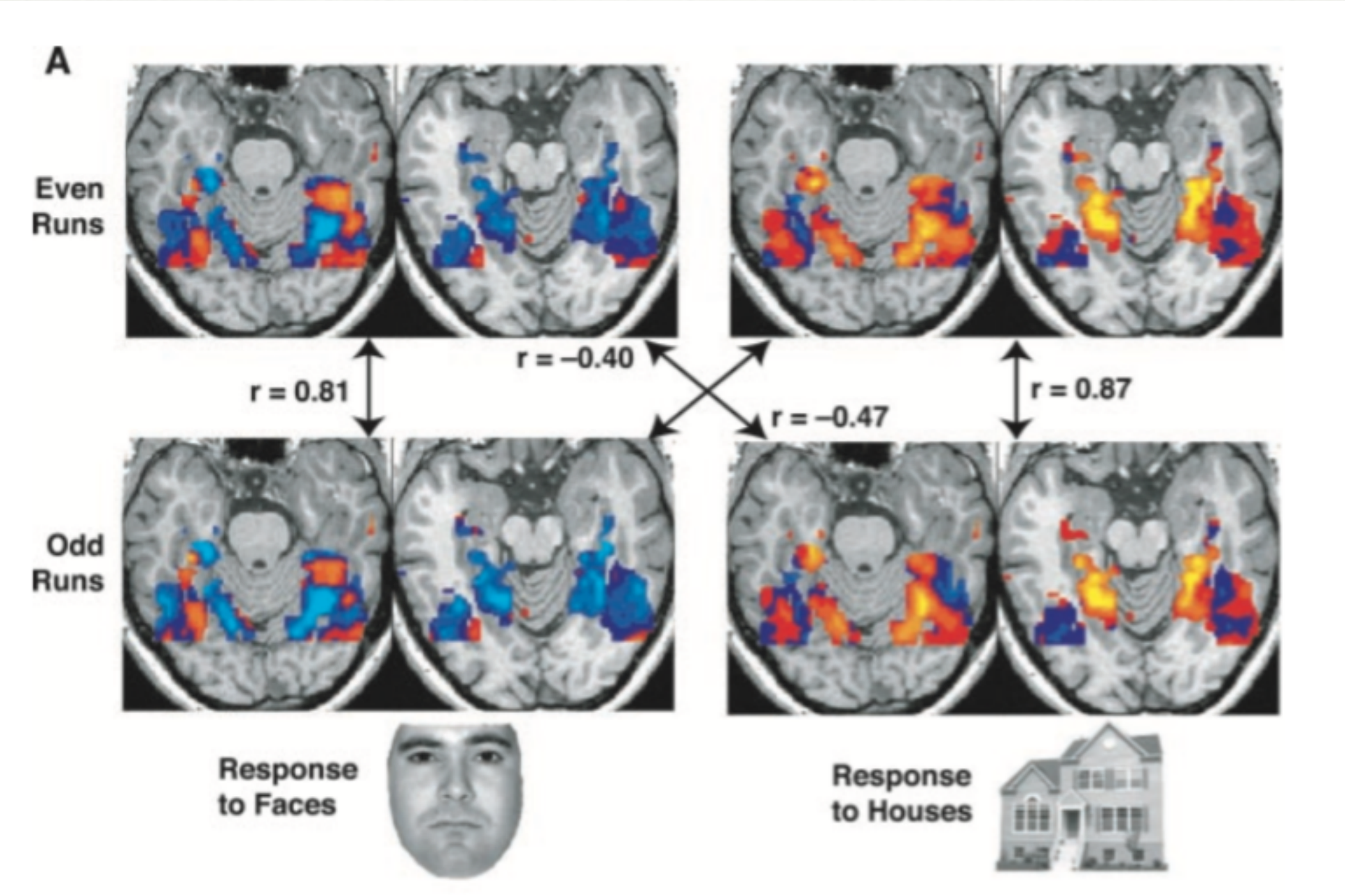
James V. Haxby *et al.*

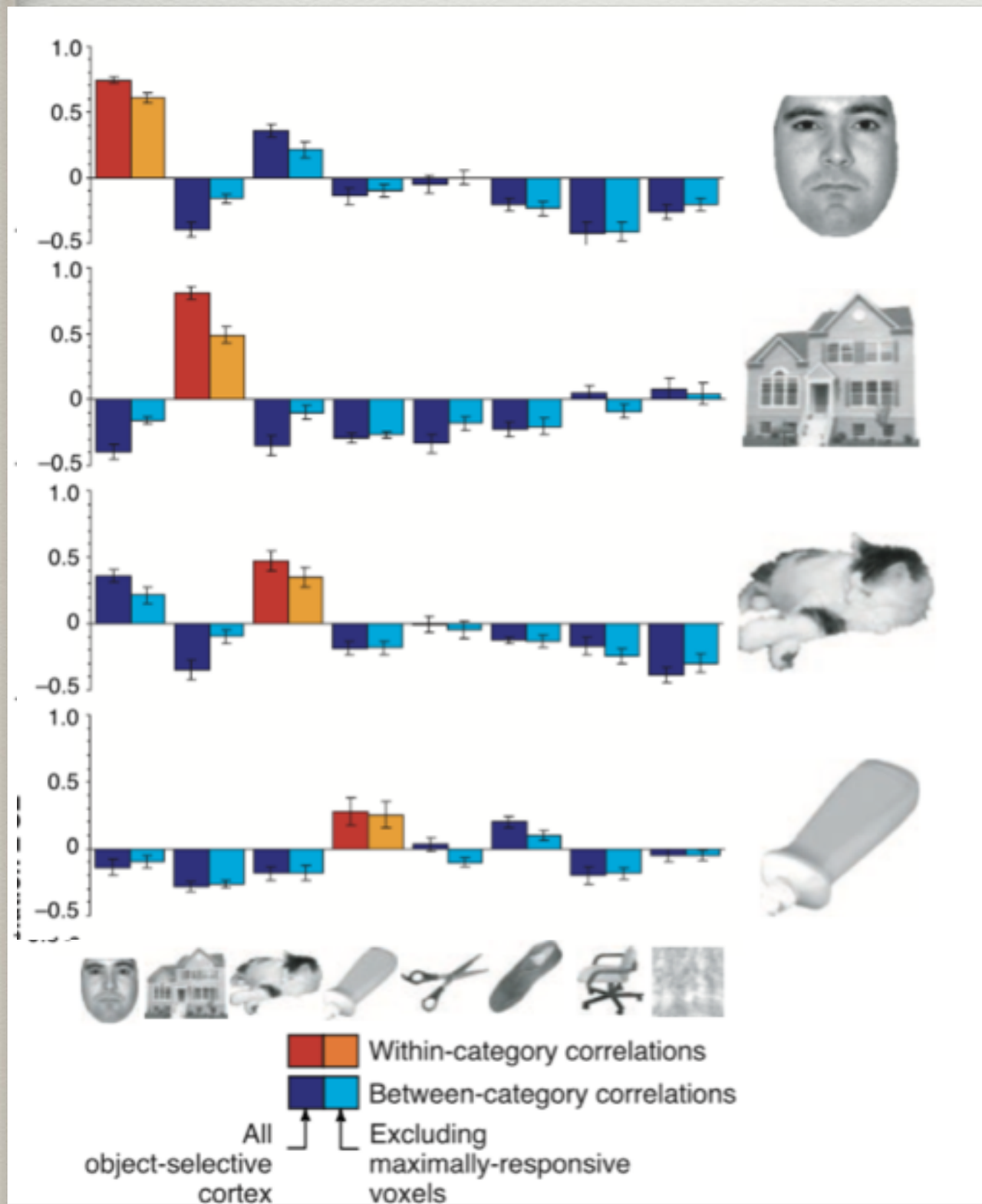
Science **293**, 2425 (2001);
DOI: 10.1126/science.1063736

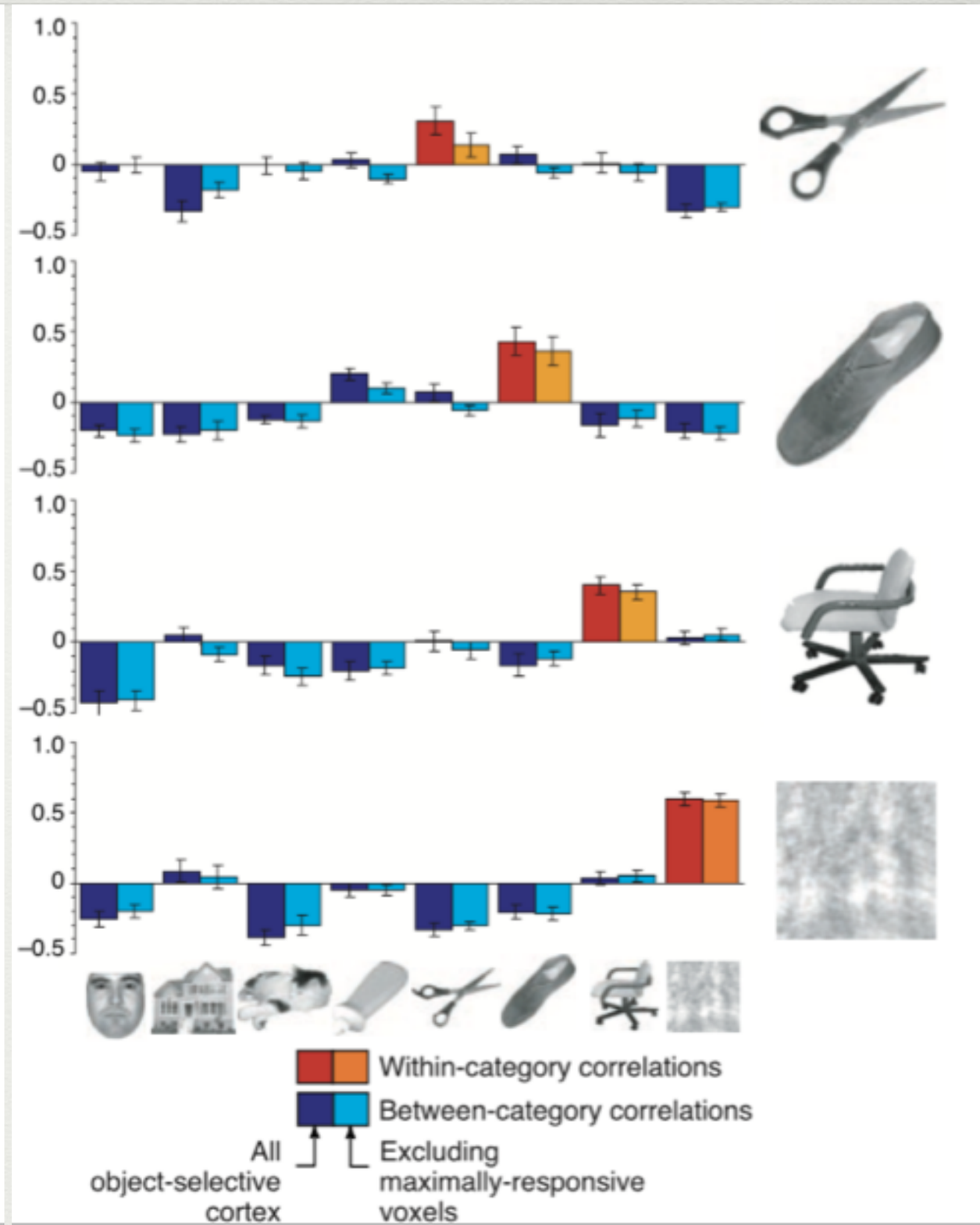





a
Image 1
Image 2
N = 9
2
8
4
2
5
4
8
4
8
4
3
8
4
3
2
4
9
8
b
Voxel 2
Voxel 1
c
Voxel 2
Voxel 1
d
Voxel 2
Voxel 1
e
Correct
Error
Voxel 2
Voxel 1


A
Even Runs
Odd Runs
r = 0.81
r = −0.40
r = −0.47
r = 0.87
Response to Faces
Response to Houses


1.0
0.5
0
−0.5
Within-category correlations
Between-category correlations
All object-selective cortex
Excluding maximally-responsive voxels


1.0
0.5
0
−0.5
Within-category correlations
Between-category correlations
All object-selective cortex
Excluding maximally-responsive voxels

NEUROIMAGING

Decoding mental states from brain activity in humans

John-Dylan Haynes[*‡§] *and Geraint Rees*[‡§]

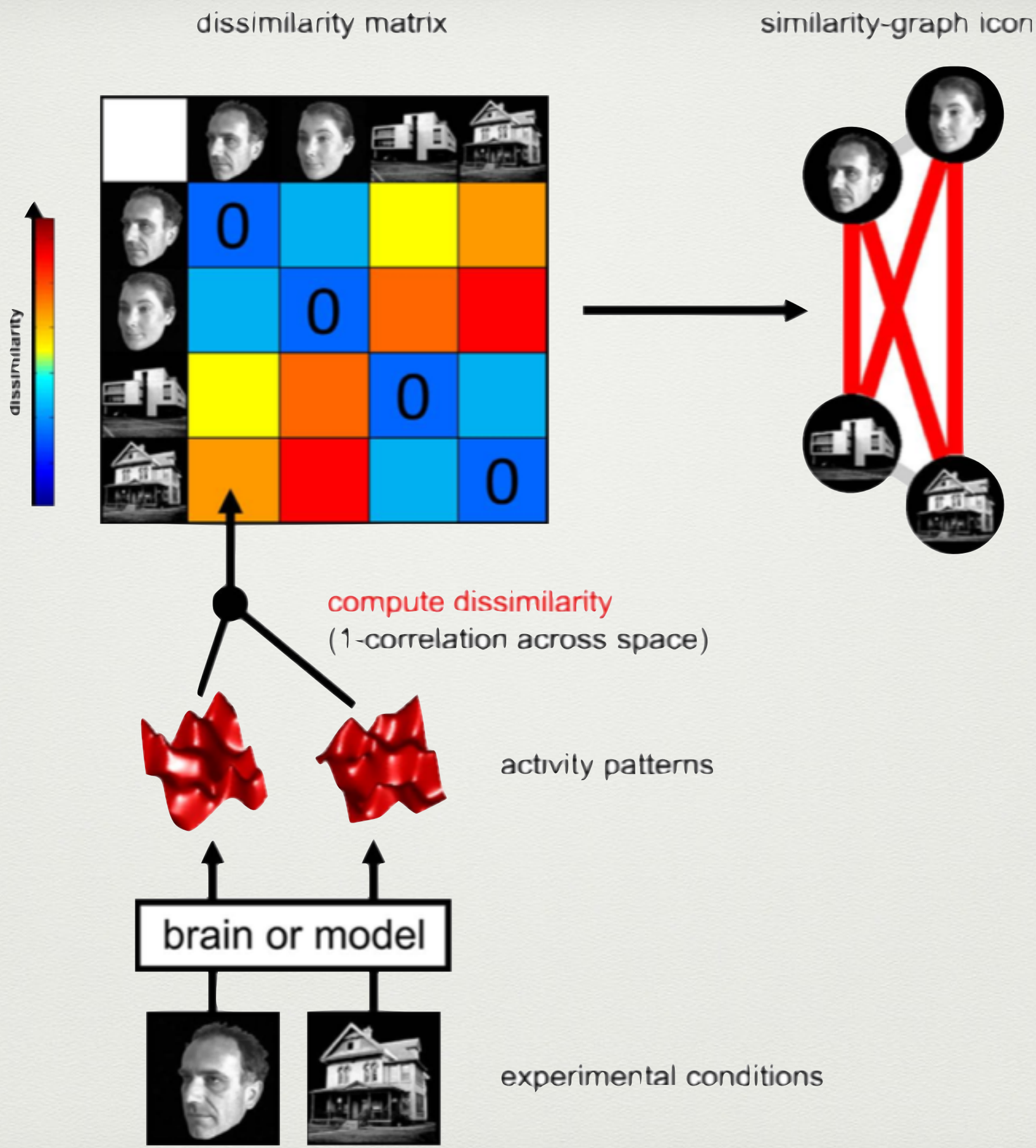
Representational similarity analysis – connecting the branches of systems neuroscience

Nikolaus Kriegeskorte[1,*]***, Marieke Mur***[1,2] ***and Peter Bandettini***[1]

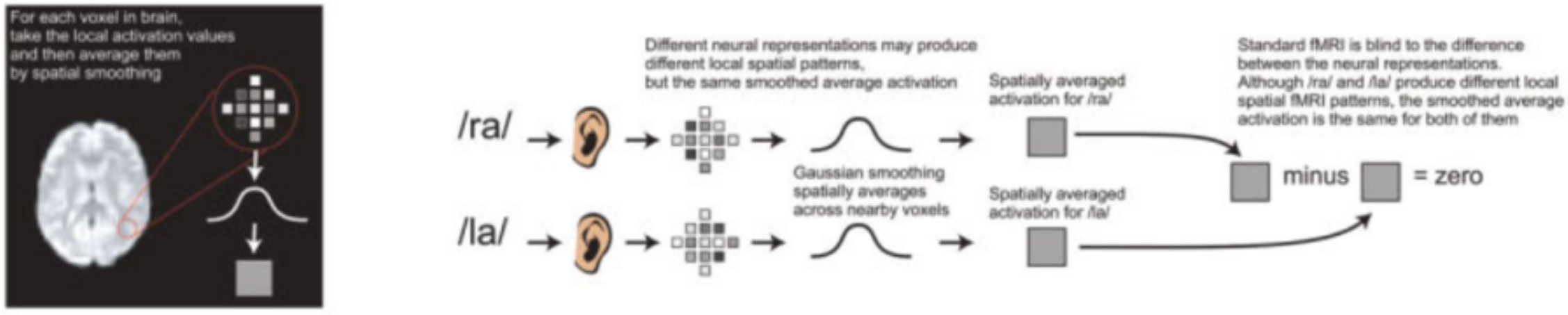
Beyond mind-reading: multi-voxel pattern analysis of fMRI data

Kenneth A. Norman[1]**, Sean M. Polyn**[2]**, Greg J. Detre**[1] **and James V. Haxby**[1]

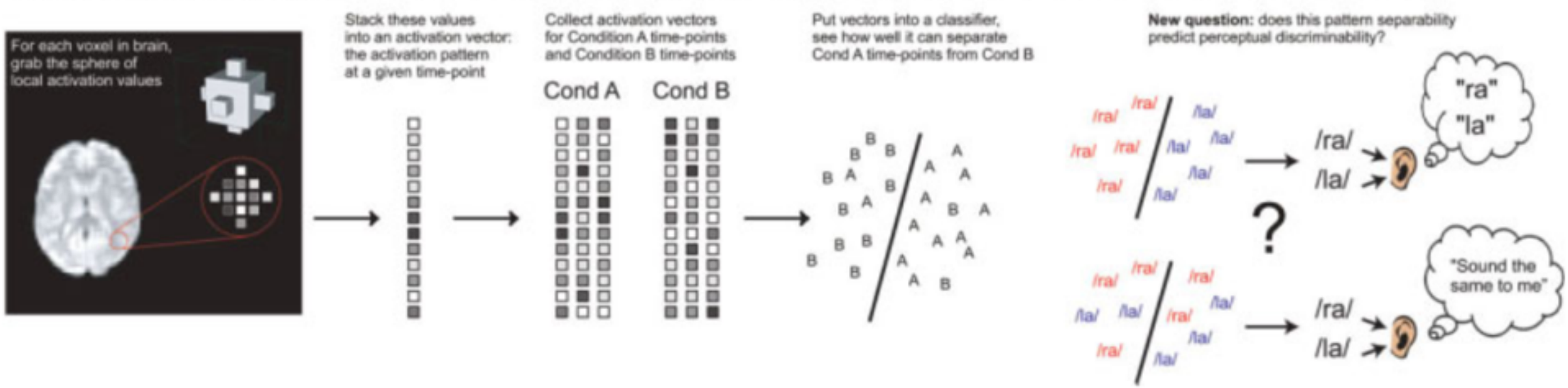
[1] Department of Psychology, Princeton University, Green Hall, Washington Road, Princeton, NJ 08540, USA
[2] Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Philadelphia, PA 19104, USA

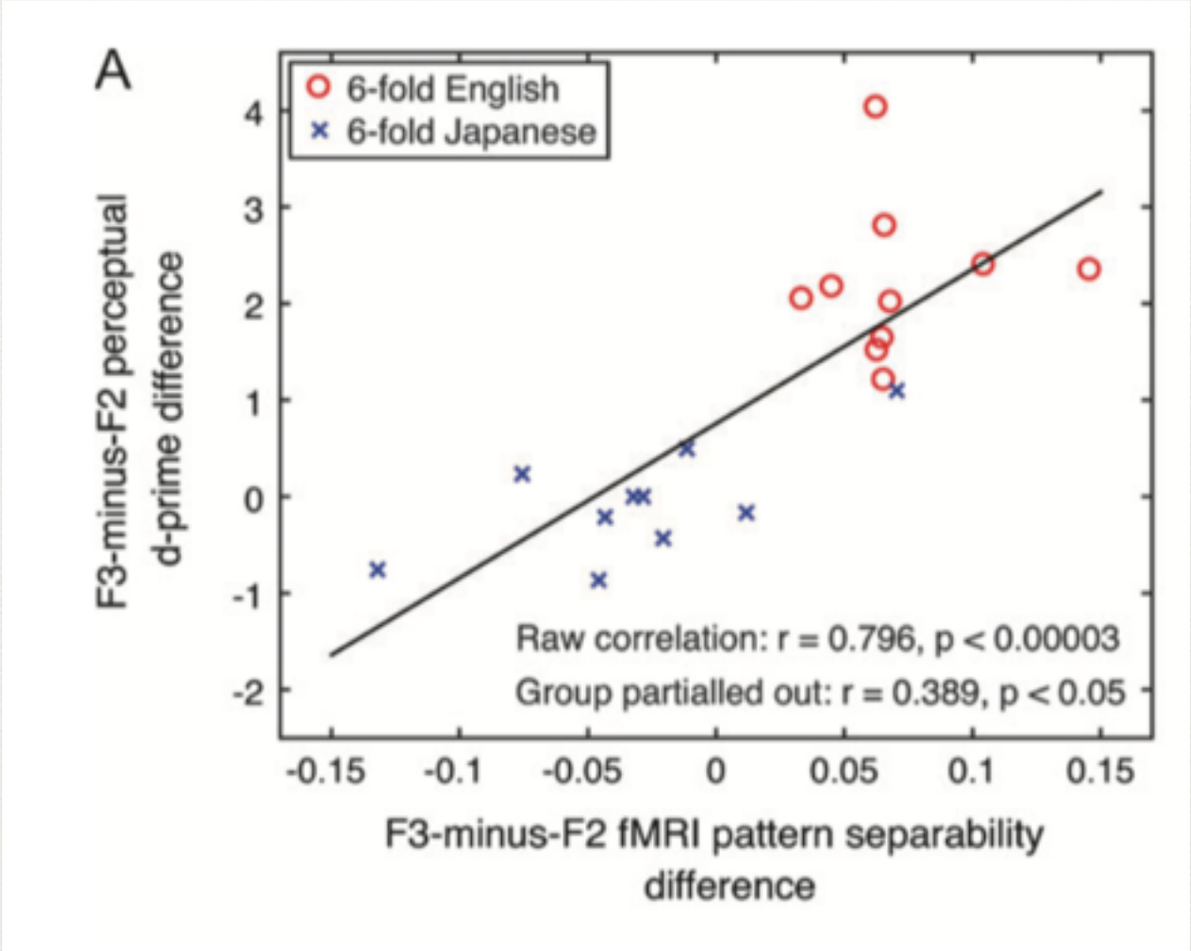

dissimilarity matrix
similarity-graph icon
dissimilarity
0
0
0
0
compute dissimilarity
(1-correlation across space)
activity patterns
brain or model
experimental conditions

A Standard fMRI: representations lost



B Pattern-information fMRI: representations regained. But do they relate to people's behaviour?




A
6-fold English
6-fold Japanese
F3-minus-F2 perceptual d-prime difference
4
3
2
1
0
-1
-2
Raw correlation: r = 0.796, p < 0.00003
Group partialled out: r = 0.389, p < 0.05
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
F3-minus-F2 fMRI pattern separability difference

Overview

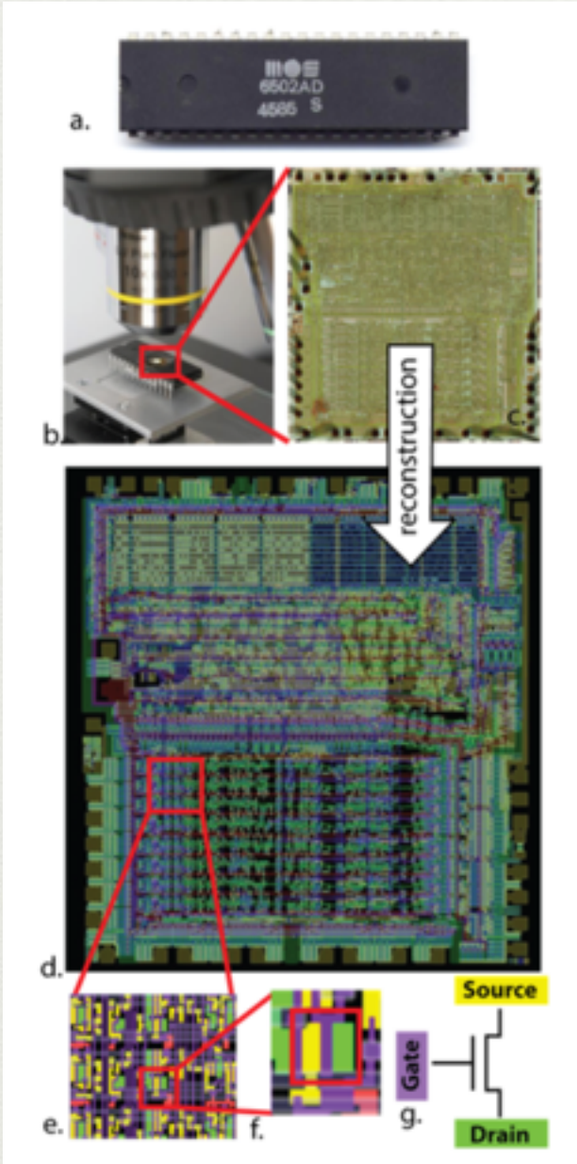
- Functional magnetic resonance imaging (fMRI)
- Multivariate techniques
- **“How to understand the brain”: linking computational cognitive models to neural data**
- Practical: Python for model-based cognitive neuroscience

Could a neuroscientist understand a microprocessor?

Eric Jonas
University of California, Berkeley
jonas@eecs.berkeley.edu

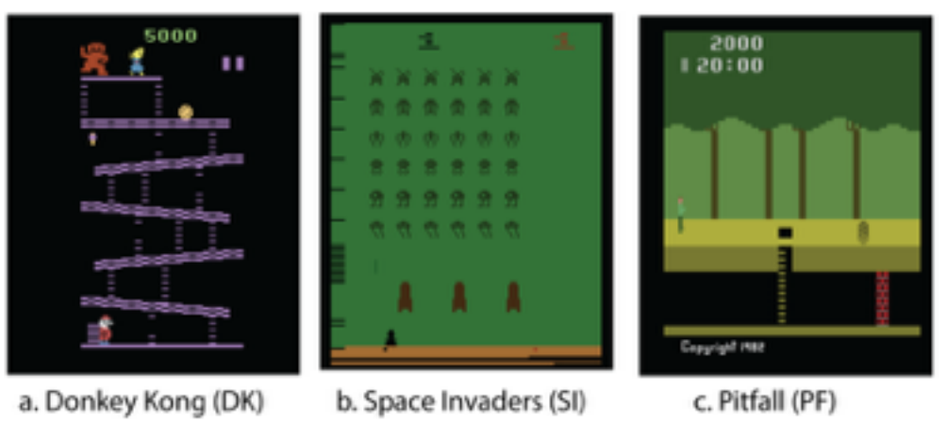
Konrad Kording
Northwestern University
koerding@gmail.com

May 26, 2016



?

>



a. Donkey Kong (DK) b. Space Invaders (SI) c. Pitfall (PF)

Figure 1: ***Example behaviors.*** *We use three classical video games as example behaviors for our model organism – (A) Donkey Kong (1981), (B) Space Invaders (1978), and (C) Pitfall (1981).*

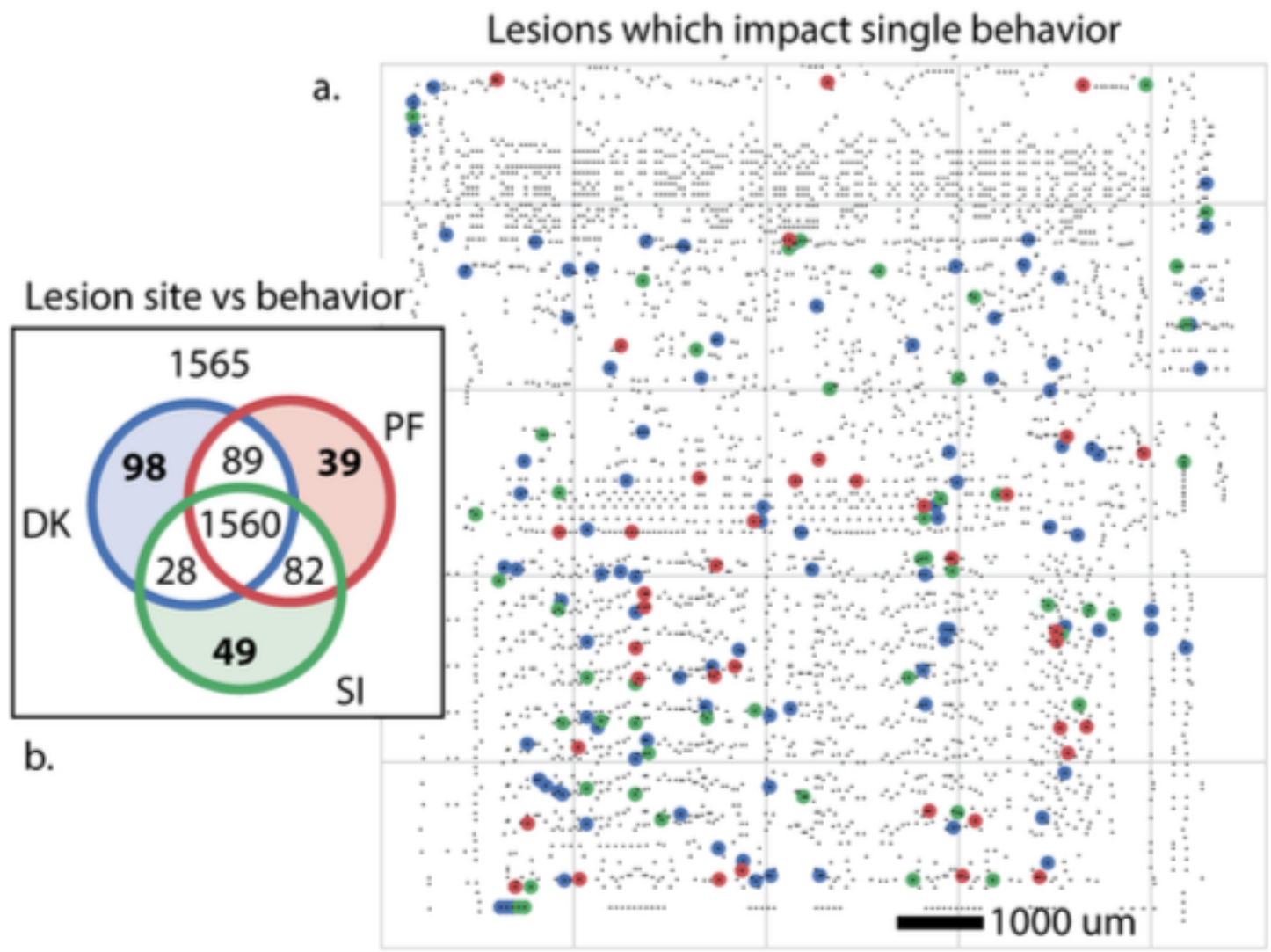


Figure 4: ***Lesioning every single transistor to identify function.*** *We identify transistors whose elimination disrupts behavior analogous to lethal alleles or lesioned brain areas. These are transistors whose elimination results in the processor failing to render the game.* ***(A)*** *Transistors which impact only one behavior, colored by behavior.* ***(B)*** *Breakdown of the impact of transistor lesion by behavioral state. The elimination of 1565 transistors have no impact, and 1560 inhibit all behaviors.*

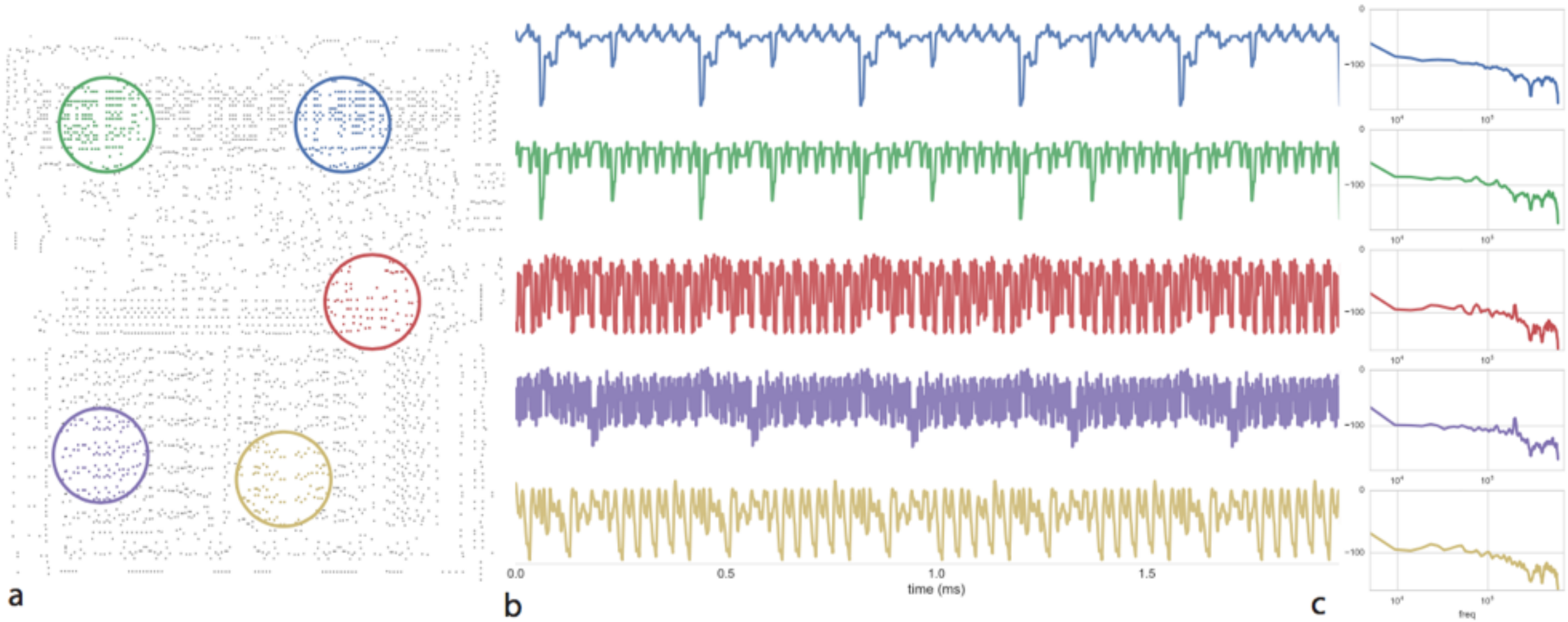


Figure 8: ***Plotting local field potentials to understand network properties.*** *We recorded from the processor during behavior DK. (**A**) Transistor switching is integrated and low-pass filtered over the indicated region. (**B**) local-field potential measurements from the indicated areas. (**C**) Spectral analysis of the indicated LFP regions identifies varying region-specific oscillations or "rhythms"*

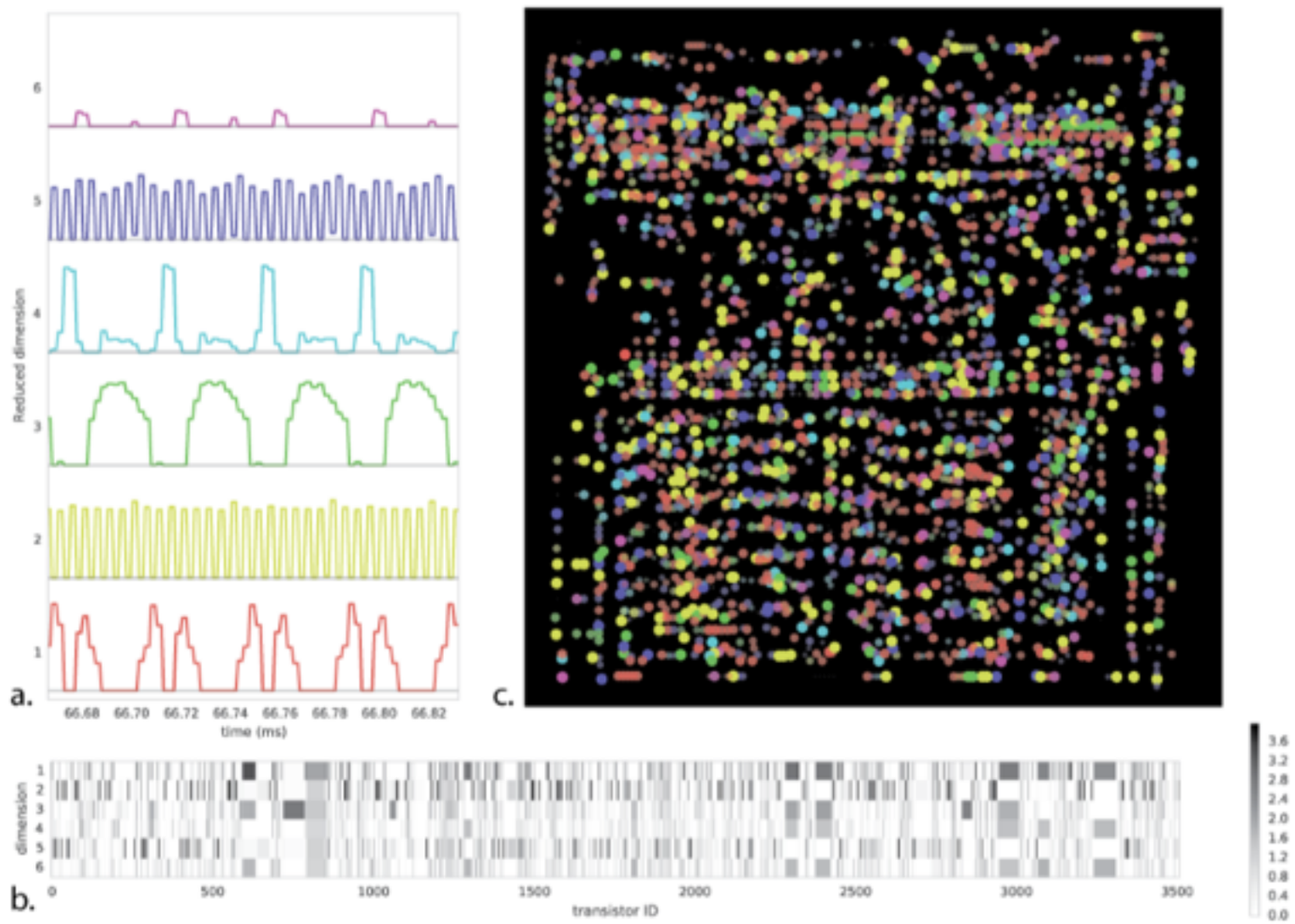


Figure 11: ***Dimensionality Reduction to understand the roles of transistors.*** *We apply non-negative matrix factorization (NMF) to the space invaders (SI) task. (**A**) shows the six*

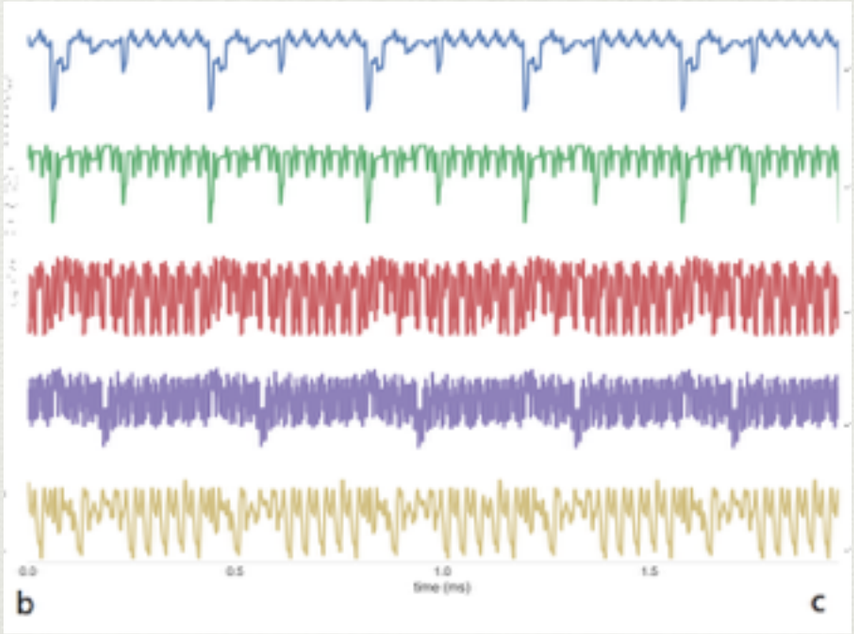
Jonas & Kording

- *"We show that [popular data analysis methods from neuroscience] reveal interesting structure in the data but do not meaningfully describe the hierarchy of information processing in the processor. This suggests that current approaches in neuroscience may fall short of producing meaningful models of the brain."*

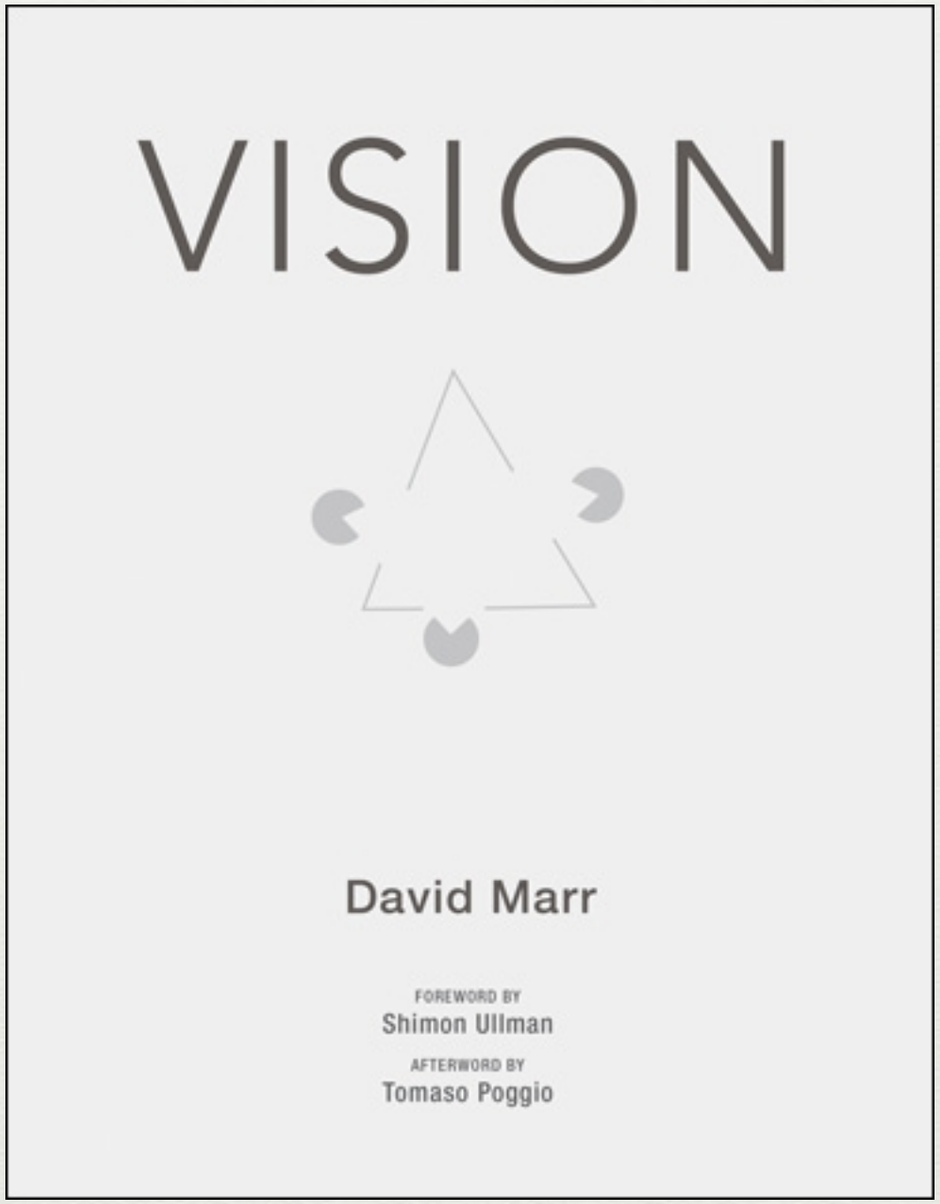
Can we stop doing neuroscience?

This is not all of neuroscience

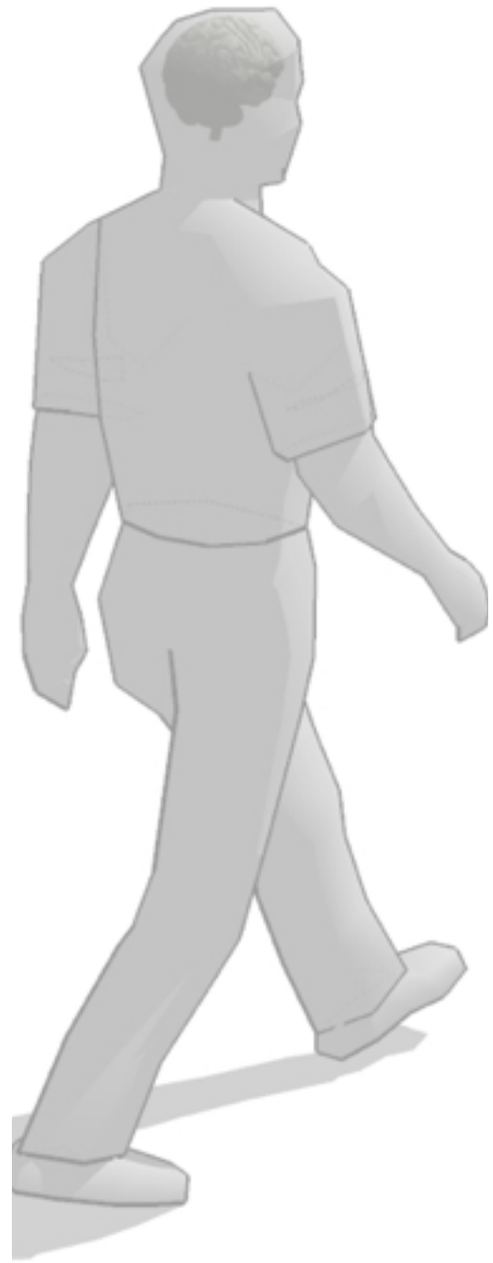
- Neuroscience entails more than just recording neurons
- "Behavior" is more than *booting* or *not booting up* a game
- We have **experimental psychology** to help us set up clever experiments to disentangle different functions.
- We have **mathematical psychology** to formally quantify psychological processes
- **Model-based** neuroscience



David Marr


VISION
David Marr
FOREWORD BY
Shimon Ullman
AFTERWORD BY
Tomaso Poggio

Three levels of description *(David Marr, 1982)*



Three levels of description *(David Marr, 1982)*

Computational

Why do things work the way they do?
What is the goal of the computation?
What are the unifying principles?

maximize:

$$R_t = r_{t+1} + r_{t+2} + \cdots + r_T$$

Three levels of description *(David Marr, 1982)*

Computational

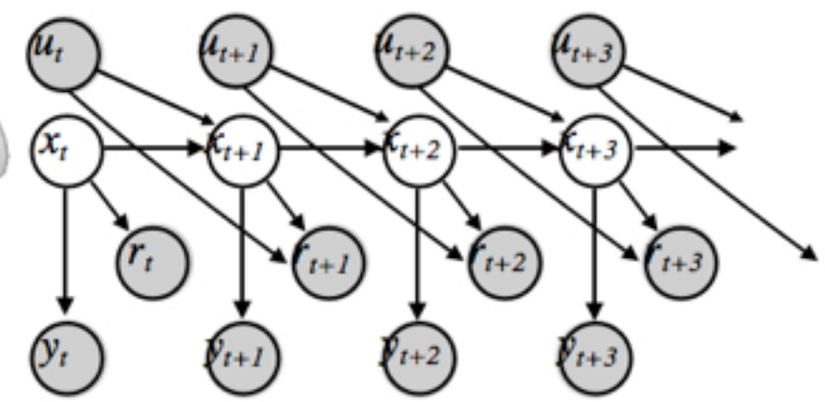
Why do things work the way they do?
What is the goal of the computation?
What are the unifying principles?

maximize:

$$R_t = r_{t+1} + r_{t+2} + \cdots + r_T$$

Algorthmic

What representations can implement such computations?
How does the choice of representations determine the algorithm?



Three levels of description *(David Marr, 1982)*

Computational

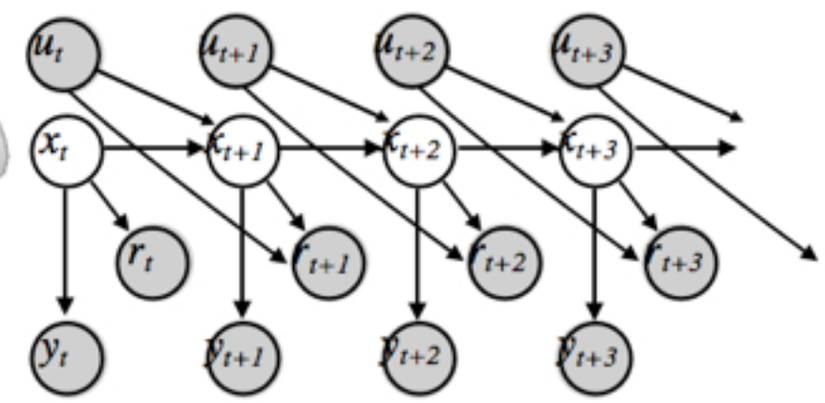
Why do things work the way they do?
What is the goal of the computation?
What are the unifying principles?

maximize:

$$R_t = r_{t+1} + r_{t+2} + \cdots + r_T$$

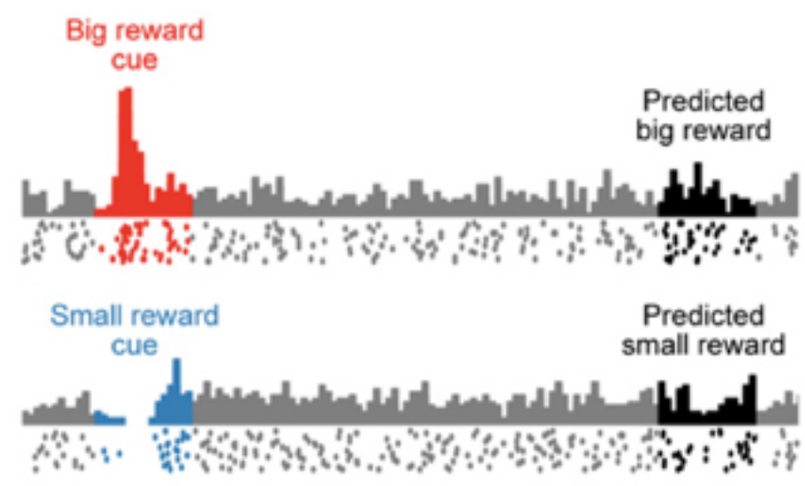
Algorthmic

What representations can implement such computations?
How does the choice of representations determine the algorithm?



Implementational

How can such a system be built in hardware?
How can neurons carry out the computations?



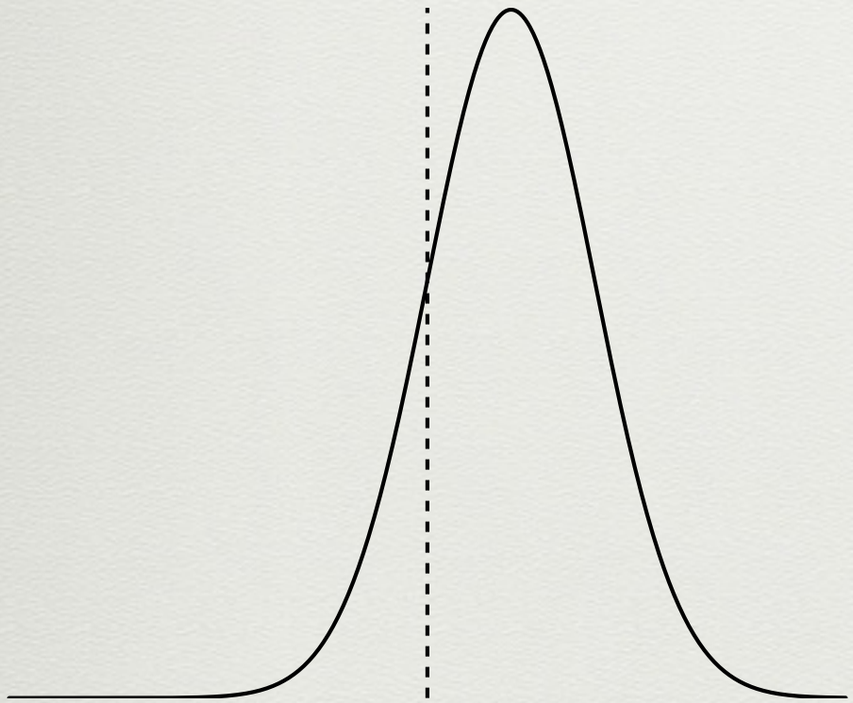
Robinson (1992)

- *"Trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feather: it just cannot be done. In order to understand bird flight, we have to understand aerodynamics; only then do the structure of feathers and the different shapes of birds' wings make sense"*

Making a decision

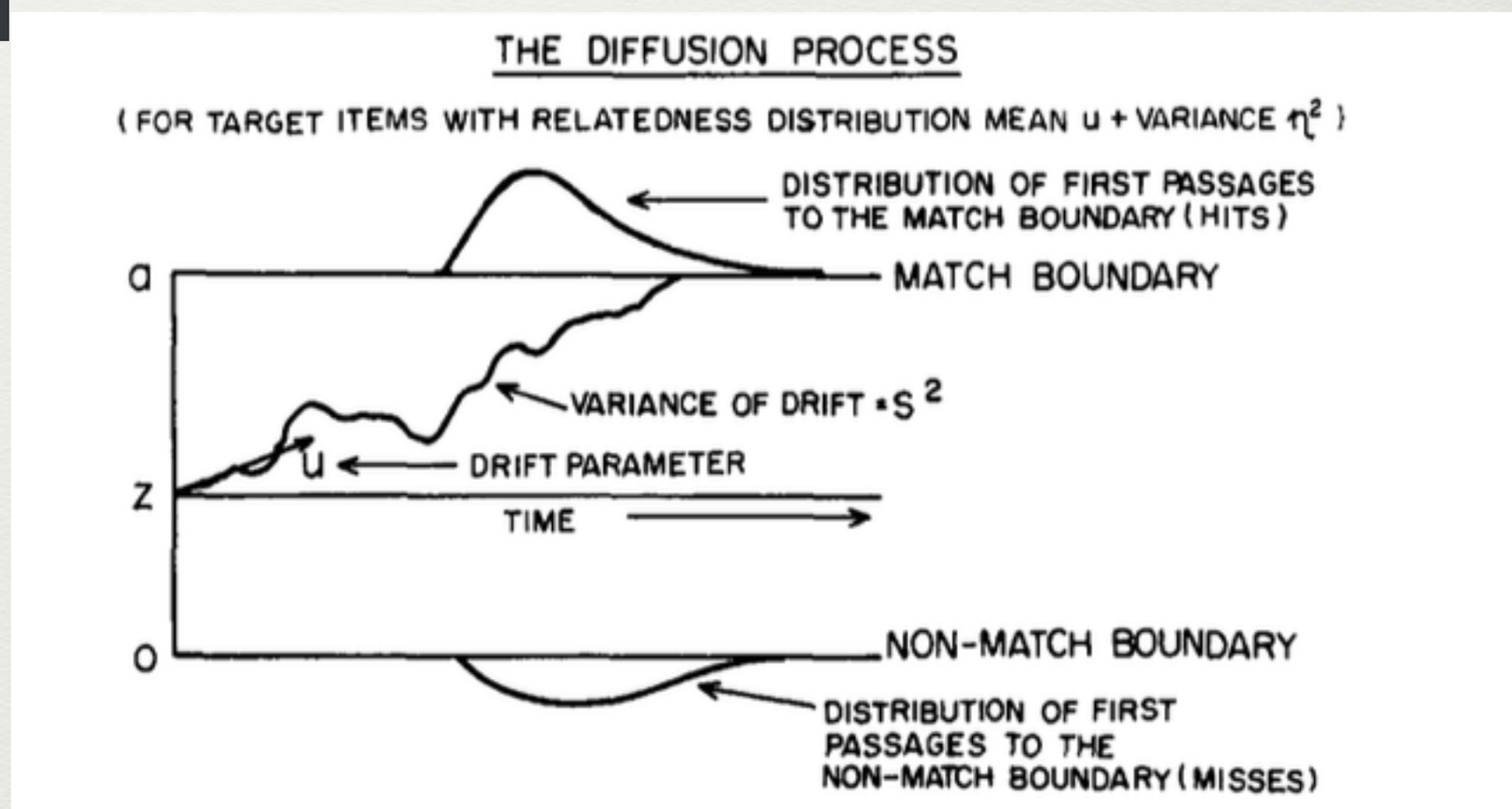
Making a decision

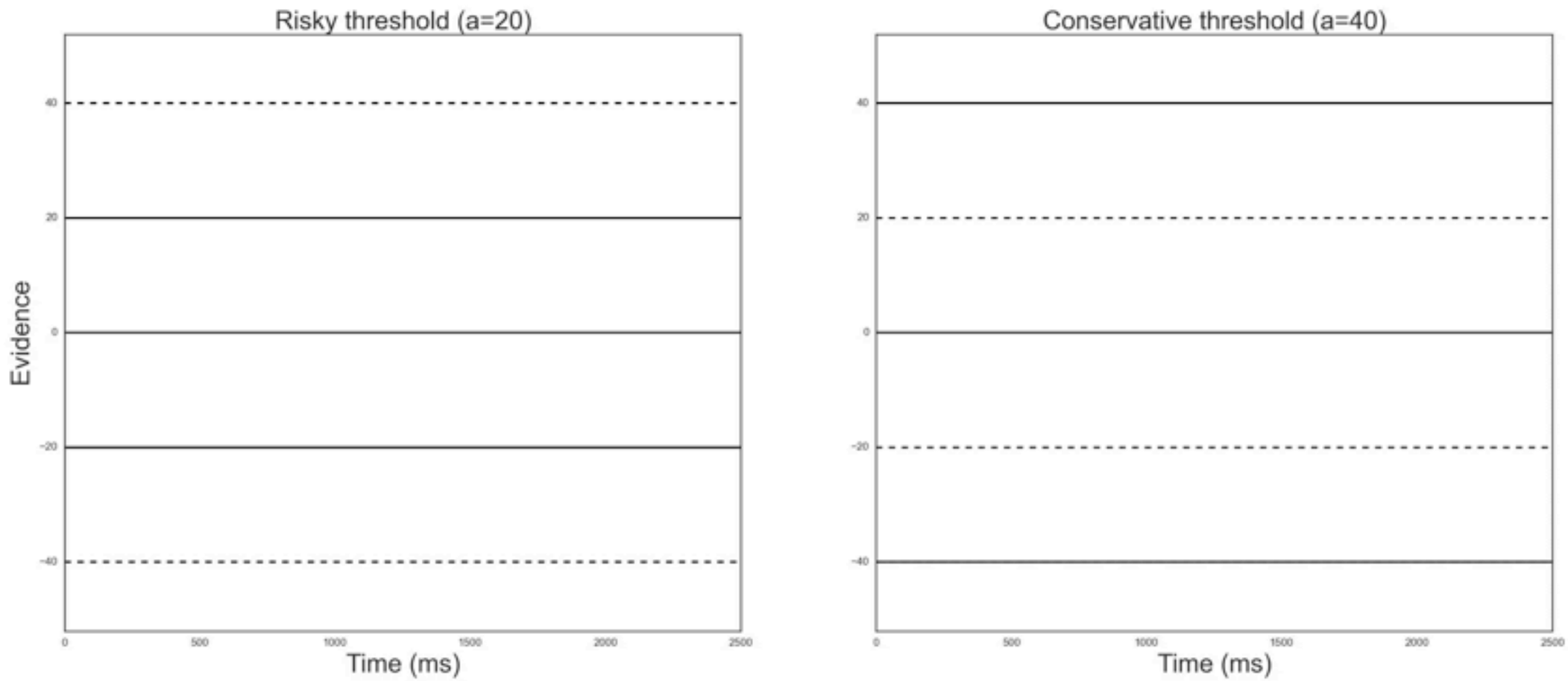
- *We **sample evidence** from our environment and make a **decision** when we have **enough evidence** in favour of one particular **option**.*



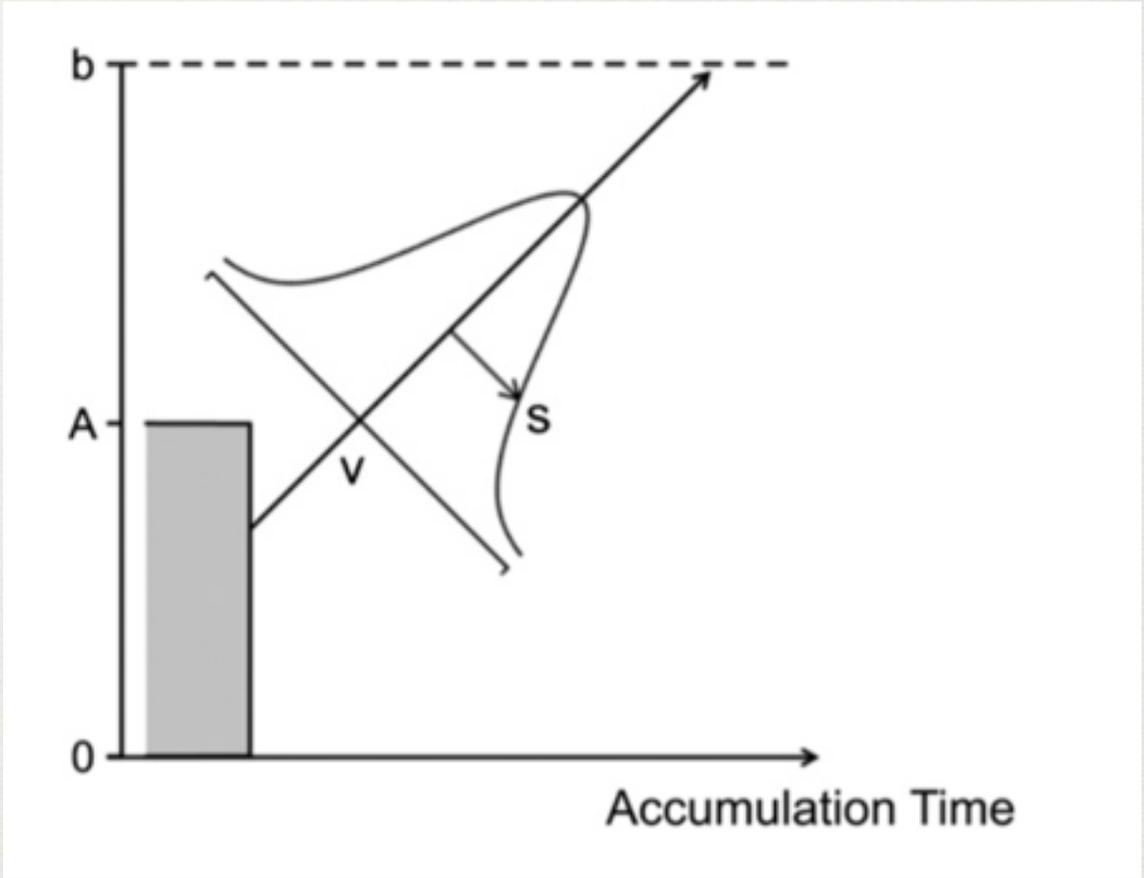
$\int$

Ratcliff (1978)
Diffusion Decision Model

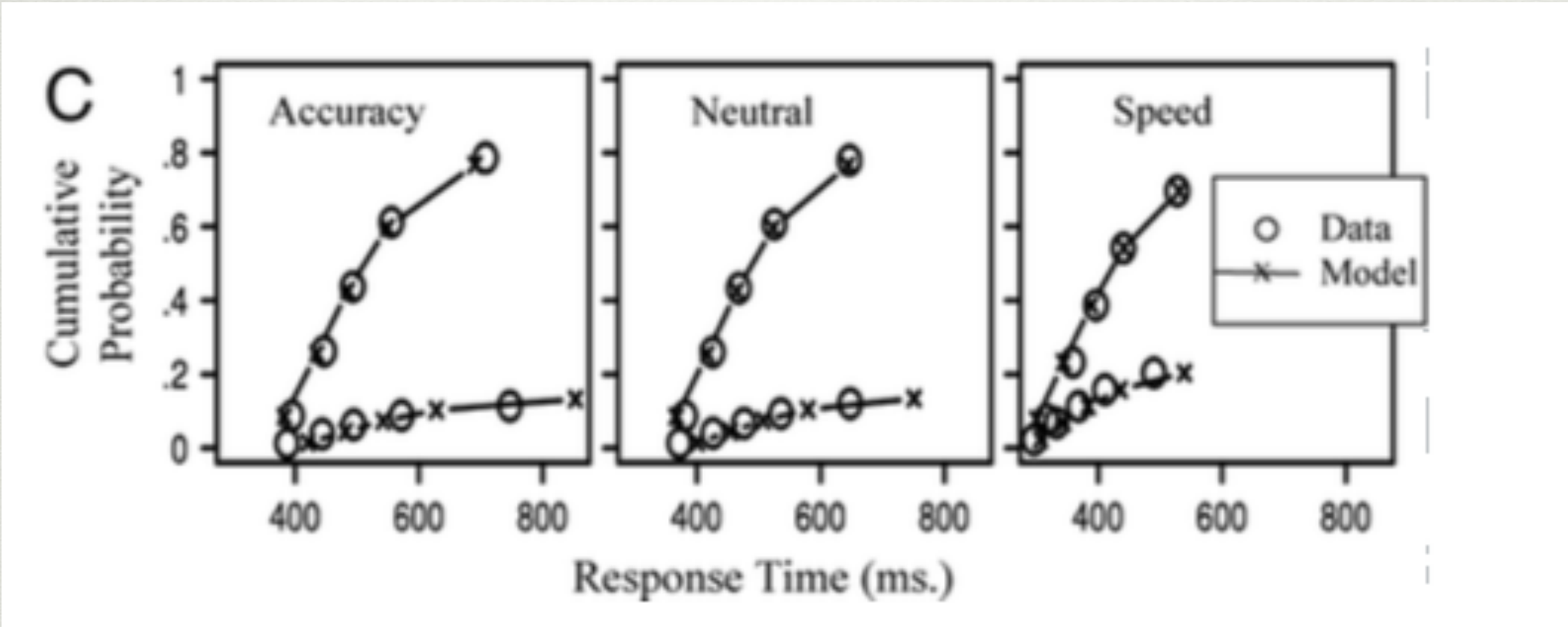


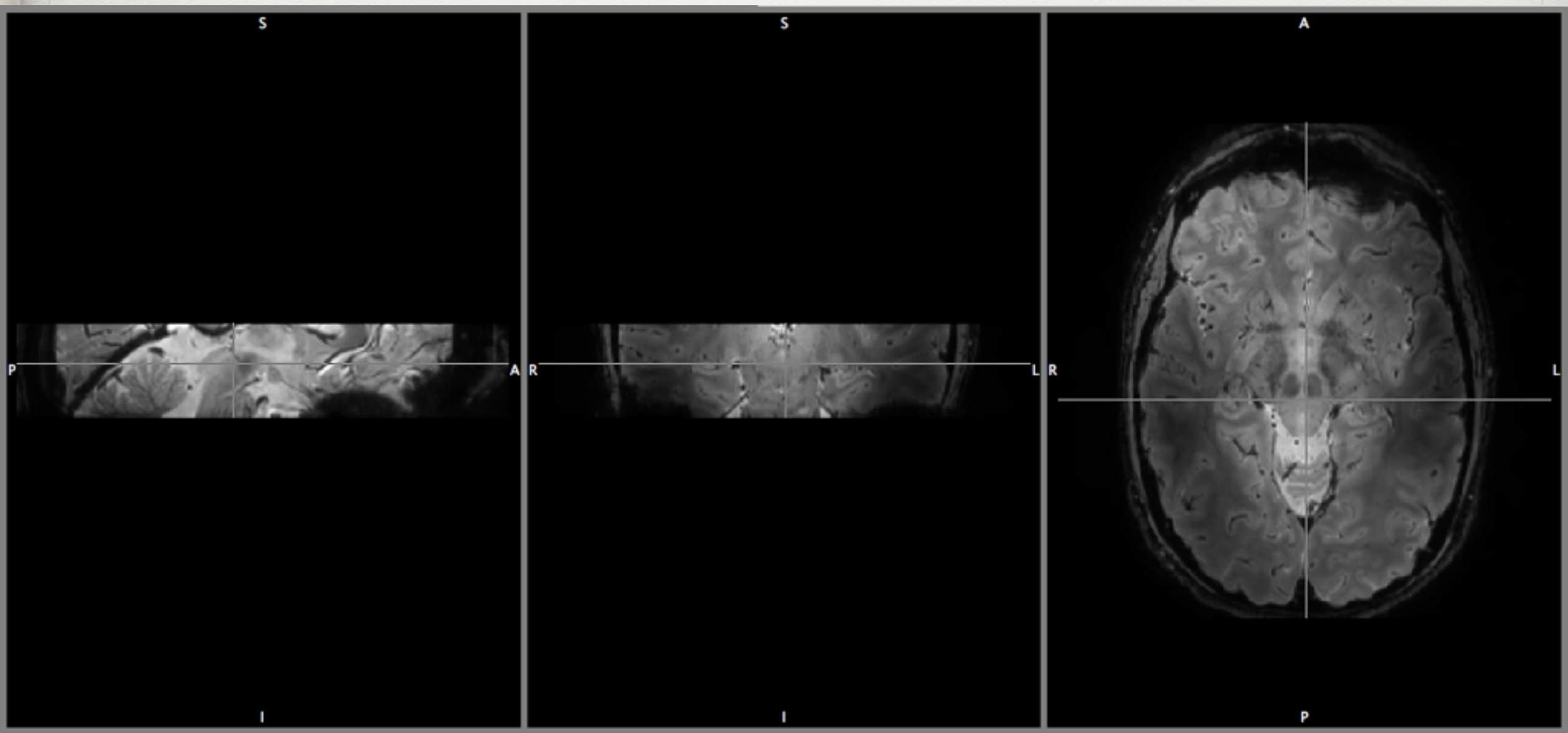

Risky threshold (a=20)
Conservative threshold (a=40)
Evidence
Time (ms)
Time (ms)
40
20
0
-20
-40
0
500
1000
1500
2000
2500

LBA model



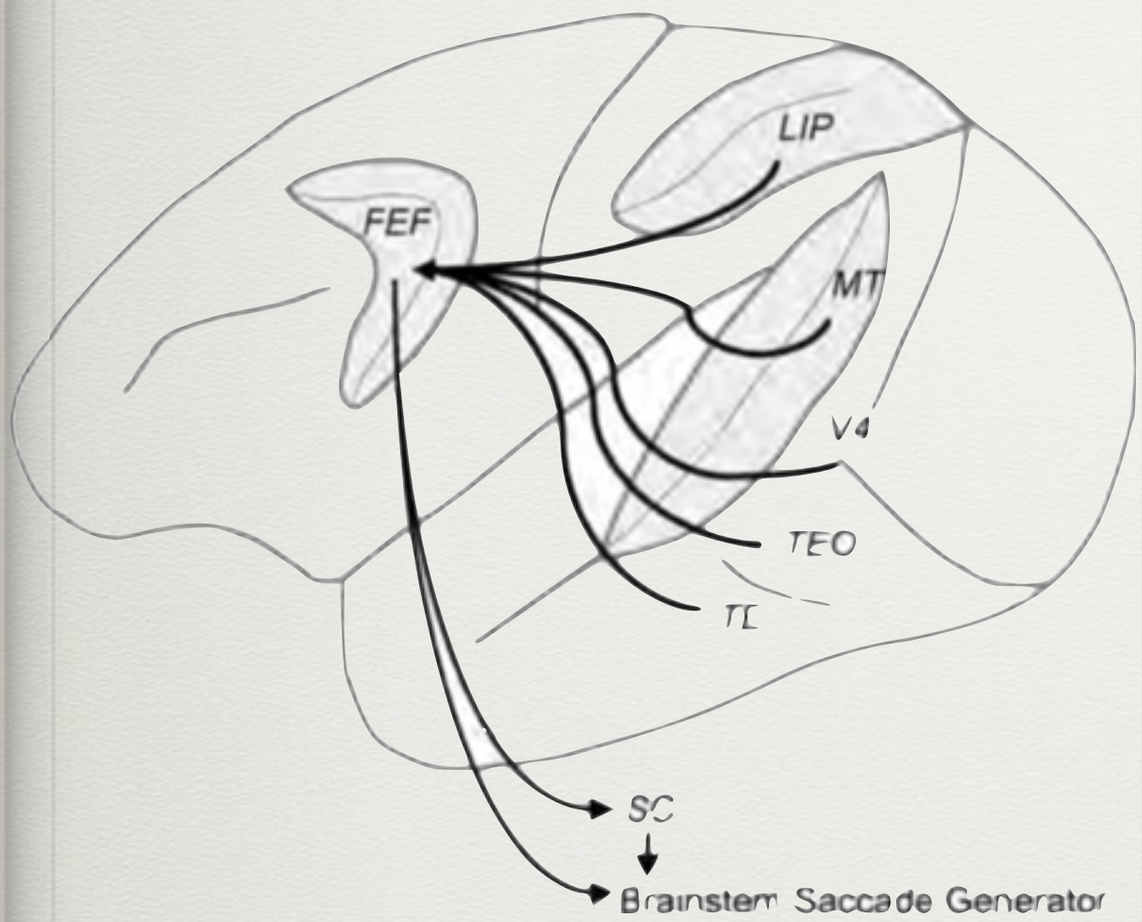
Brown & Heathcote, 2008

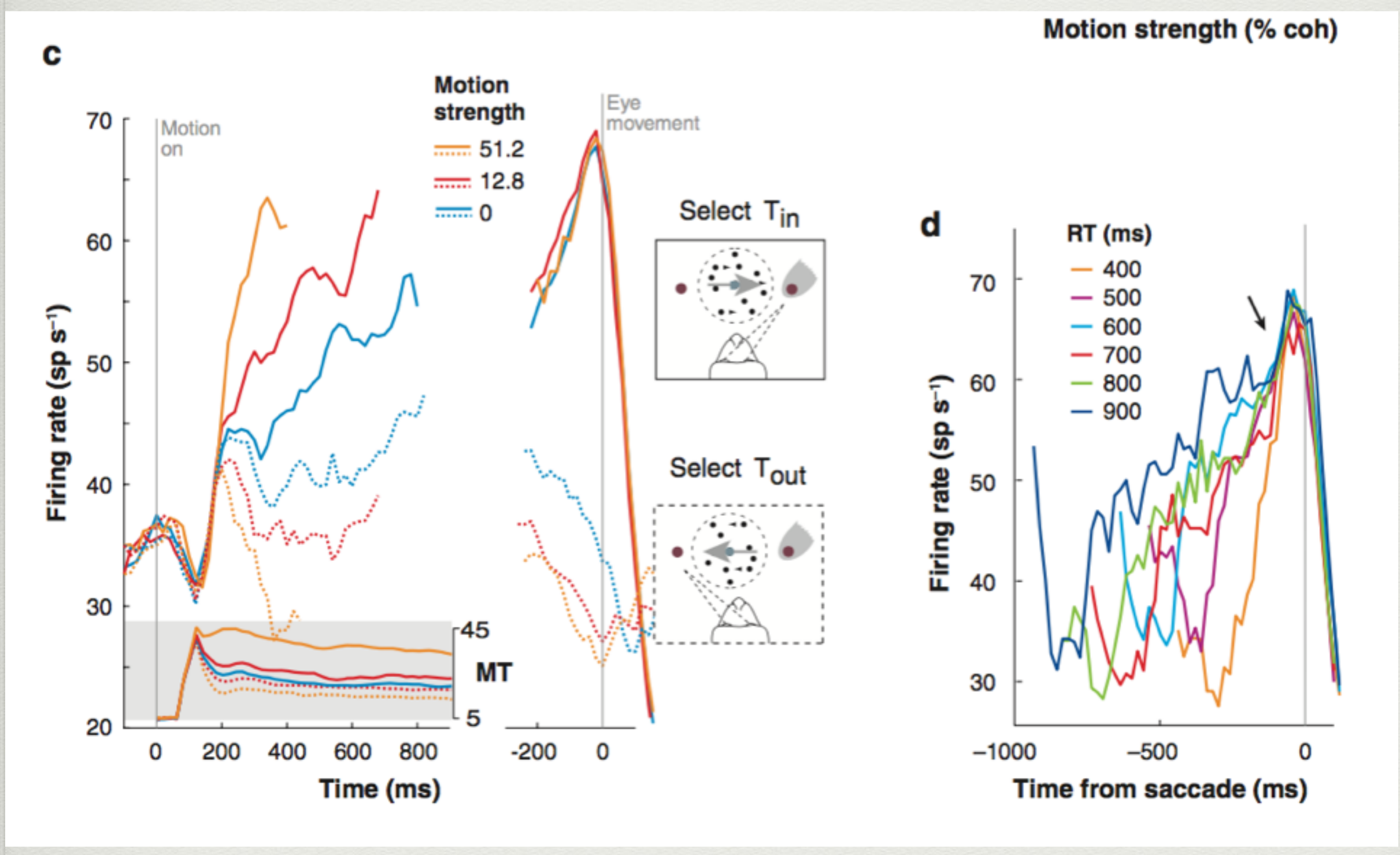

C
Cumulative Probability
1
.8
.6
.4
.2
0
Accuracy
Neutral
Speed
Data
Model
400
600
800
Response Time (ms.)


S
P
A
I
S
R
L
I
A
R
L
P

How to link?

- **Qualitatively**
- Quantitatively
- Joint model

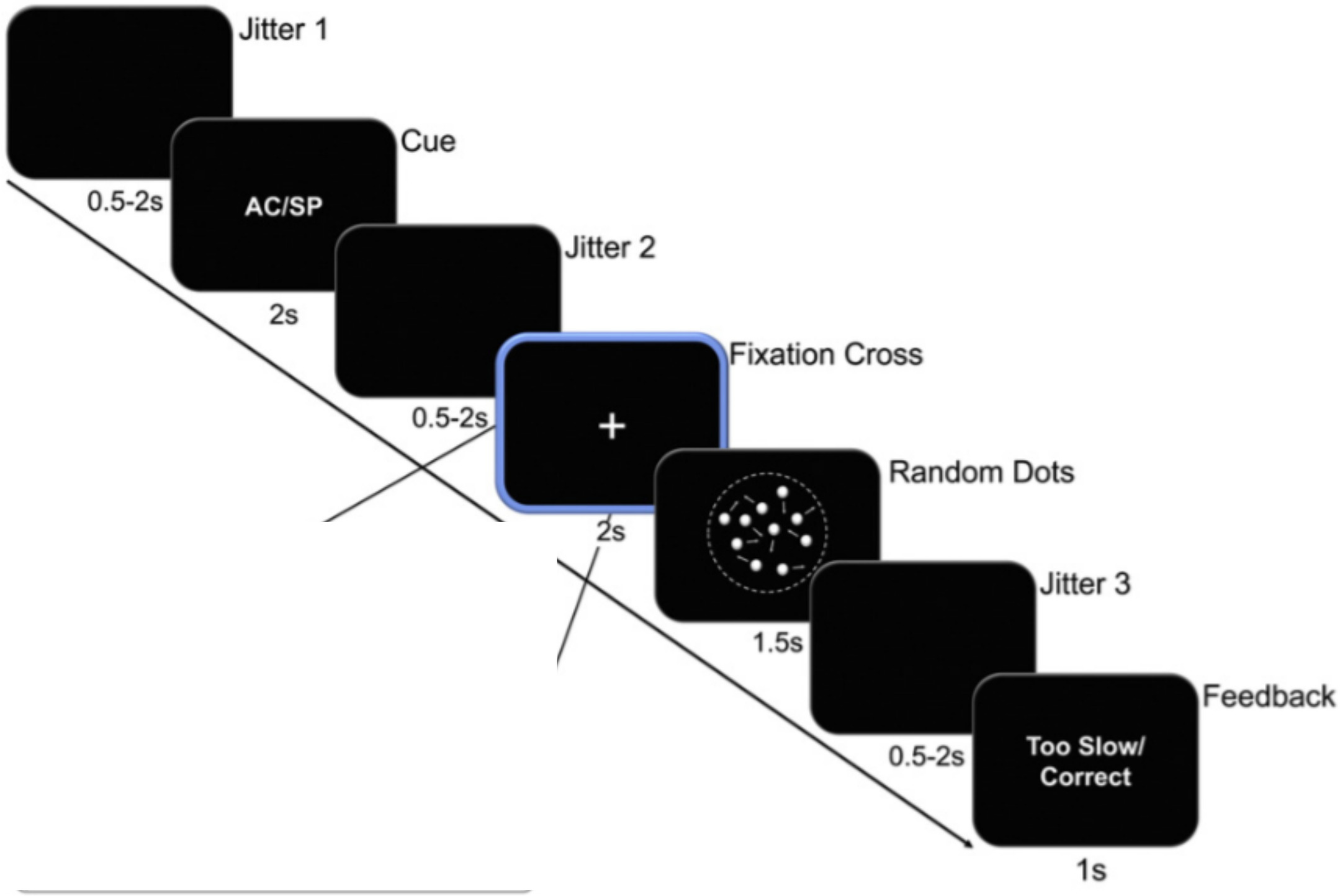

LIP
FEF
MT
V4
TEO
TE
SC
Brainstem Saccade Generator



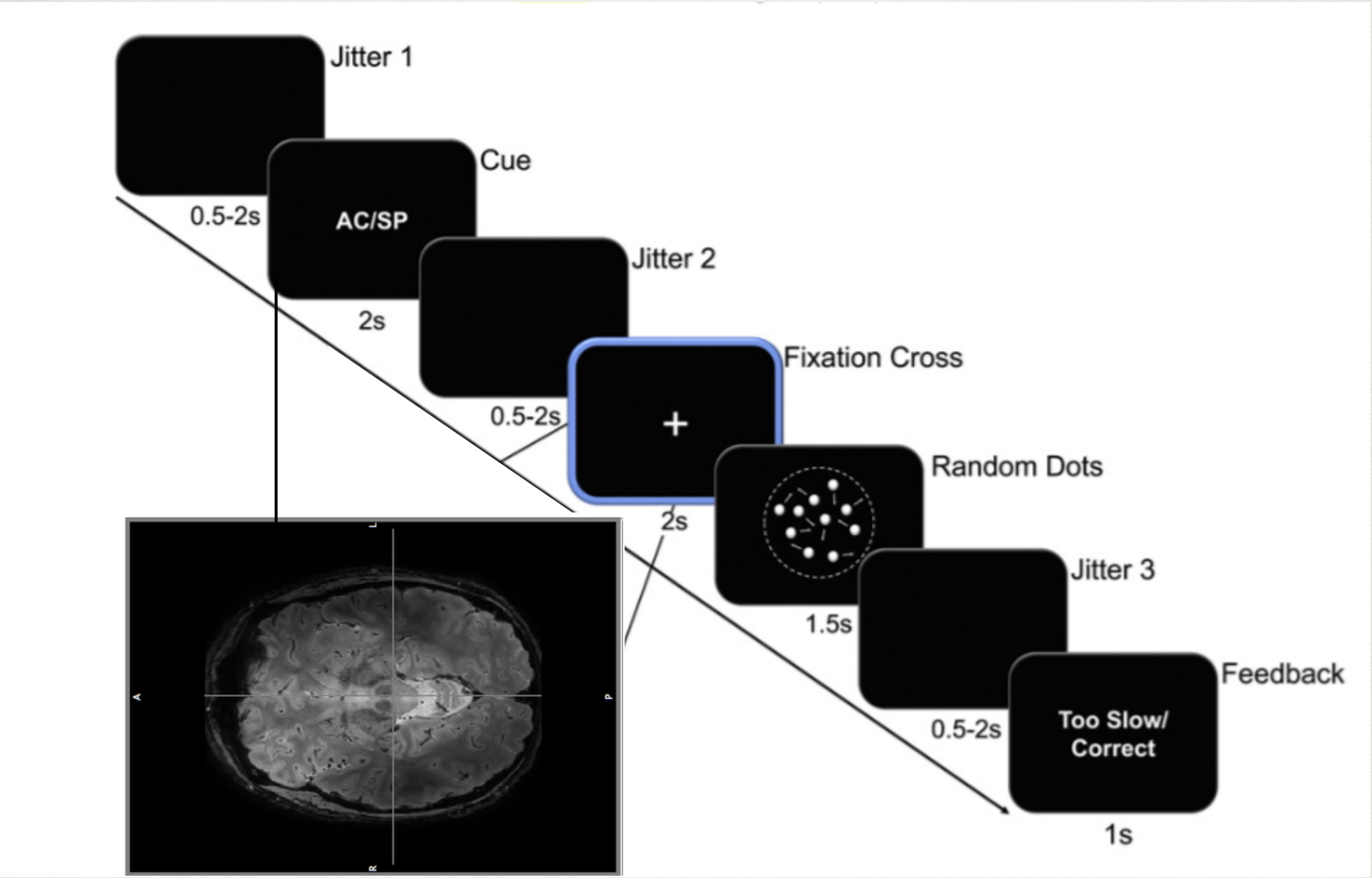
Gold & Shadlen, 2007

How to link?

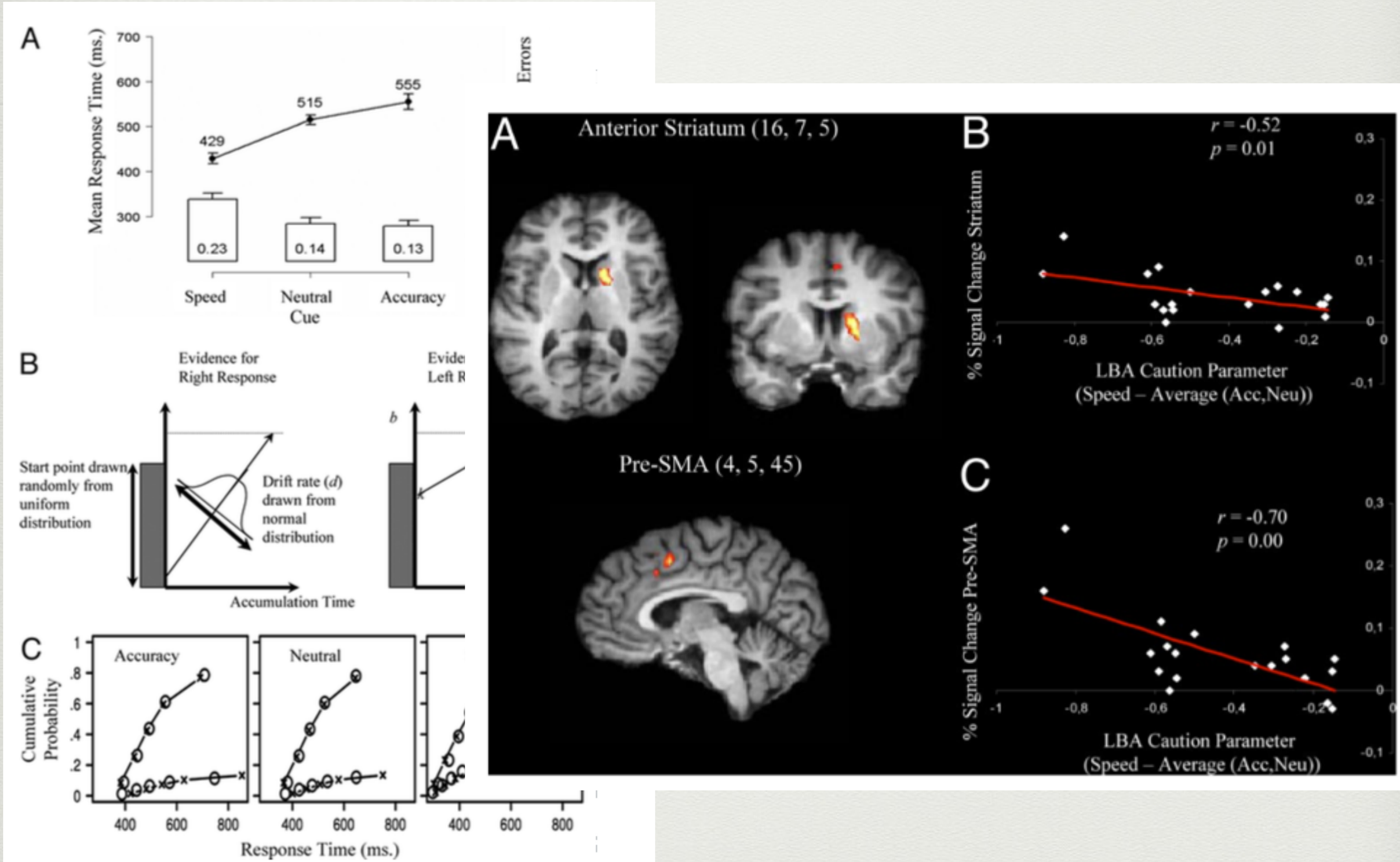
- Qualitatively
- **Quantitatively**
  - **Individual differences**
  - Within-subject variability
  - Stimulus variability
- Joint model



Forstmann et al., 2008



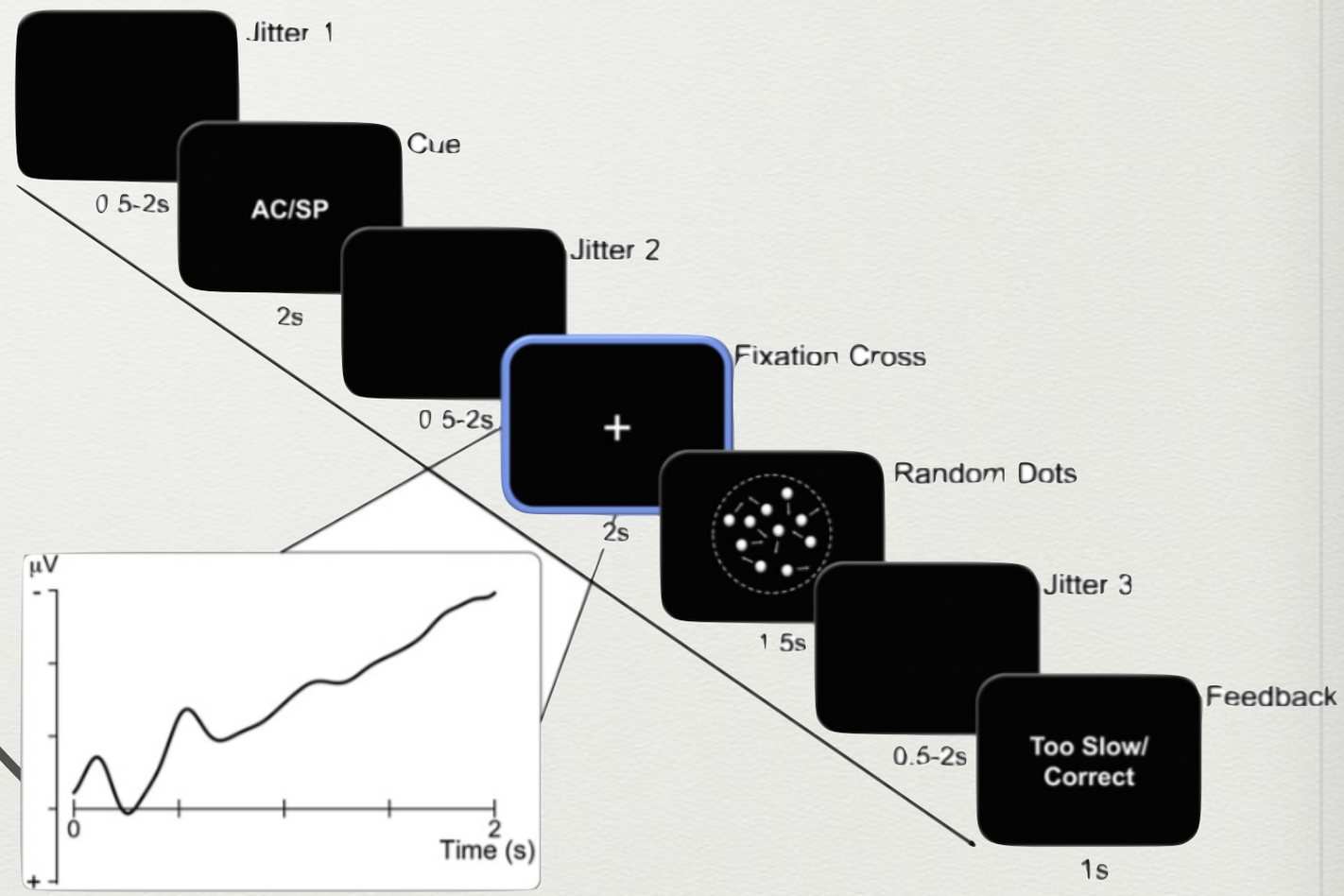
Forstmann et al., 2008



Forstmann et al., 2008

How to link?

- Qualitatively
- **Quantitatively**
  - Individual differences
  - **Within-subject variability**
  - Stimulus variability
- Joint model



FCZ
(pre-SMA)

Boehm et al., 2014

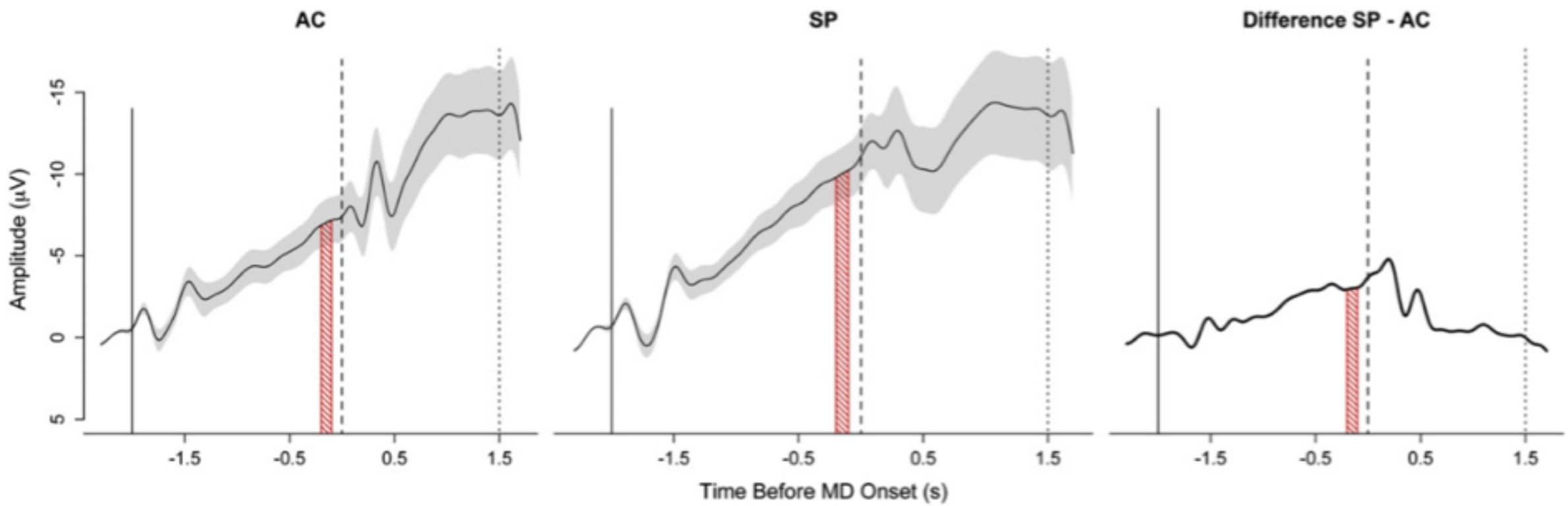


Fig. 4. EEG data aligned to the onset of the moving dots, averaged across participants for SP, AC and the difference waveform SP–AC, low-pass filtered at 5 Hz. The red striped area indicates the measurement window for the mean amplitude. The gray shaded area indicates 1 SE. The lines indicate fixation onset (solid), moving dots onset (dashed), and moving dots offset (dotted).

Boehm et al., 2014

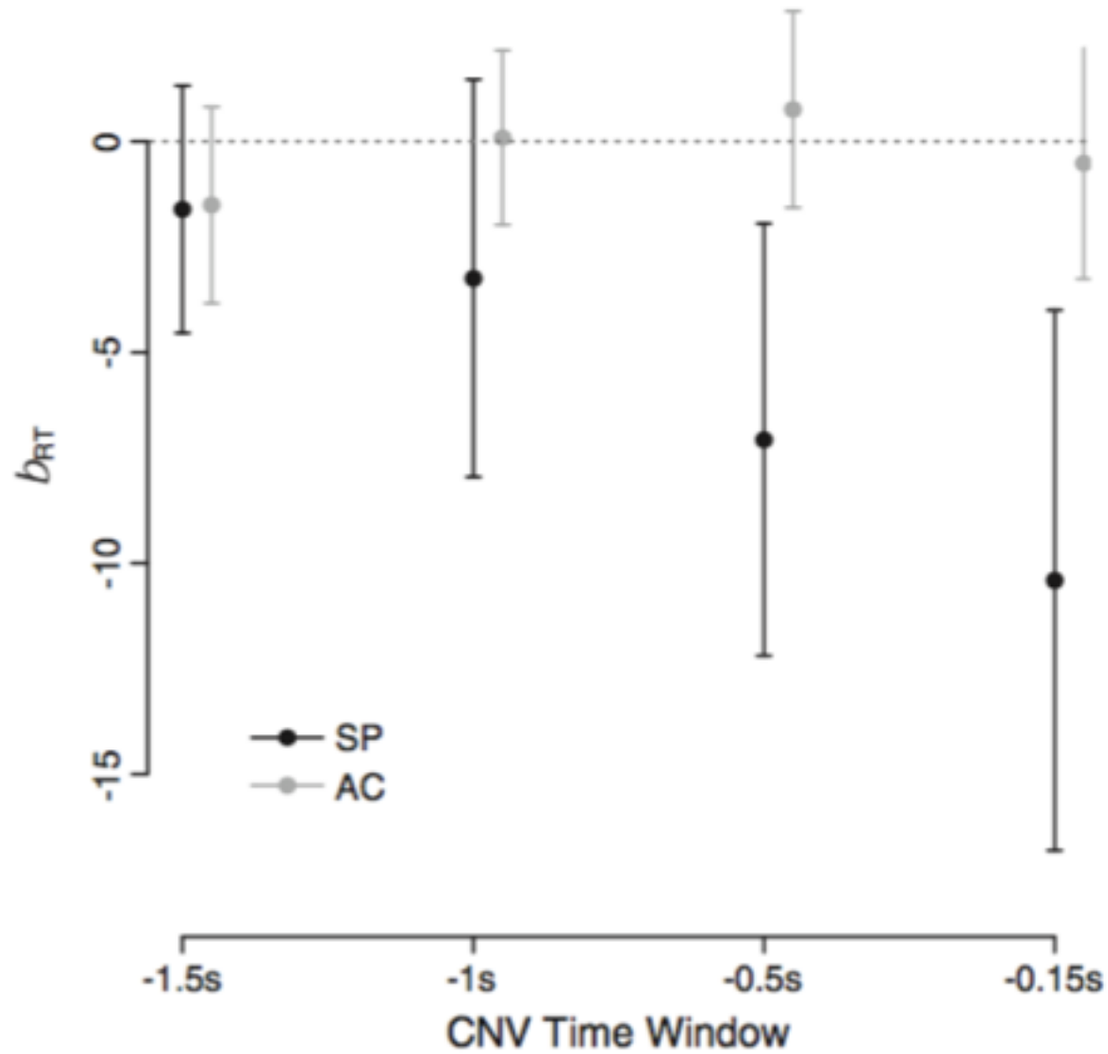


Fig. 6. Development of the beta regression weights for RT in SP and AC as the CNV built up. Beta weights are based on the RT × Cond model. Regression weights for AC were obtained by recoding AC to 0 and SP to 1 in the LME model. Error bars show approximate 95% confidence intervals based on the t-statistics computed by the lme4 R-package (Bates et al., 2013).

Boehm et al., 2014

Within subject-variability

- Within subject-variability is the main workhorse for model-based cognitive neuroscience, especially value-based decision making ("neuroeconomics")
- Relating within-subject/trial-to-trial variability to sequential sampling models is harder, because there are no trial-to-trial estimates of the parameters.
- Possible solutions:
  - Median split on neural data (Ratcliff et al., 2009)
  - Bayesian regression trial-to-trial measures (HDDM; Wiecki et al., 2013)
  - Estimate covariance structure hyperparameters (Turner et al., 2013)

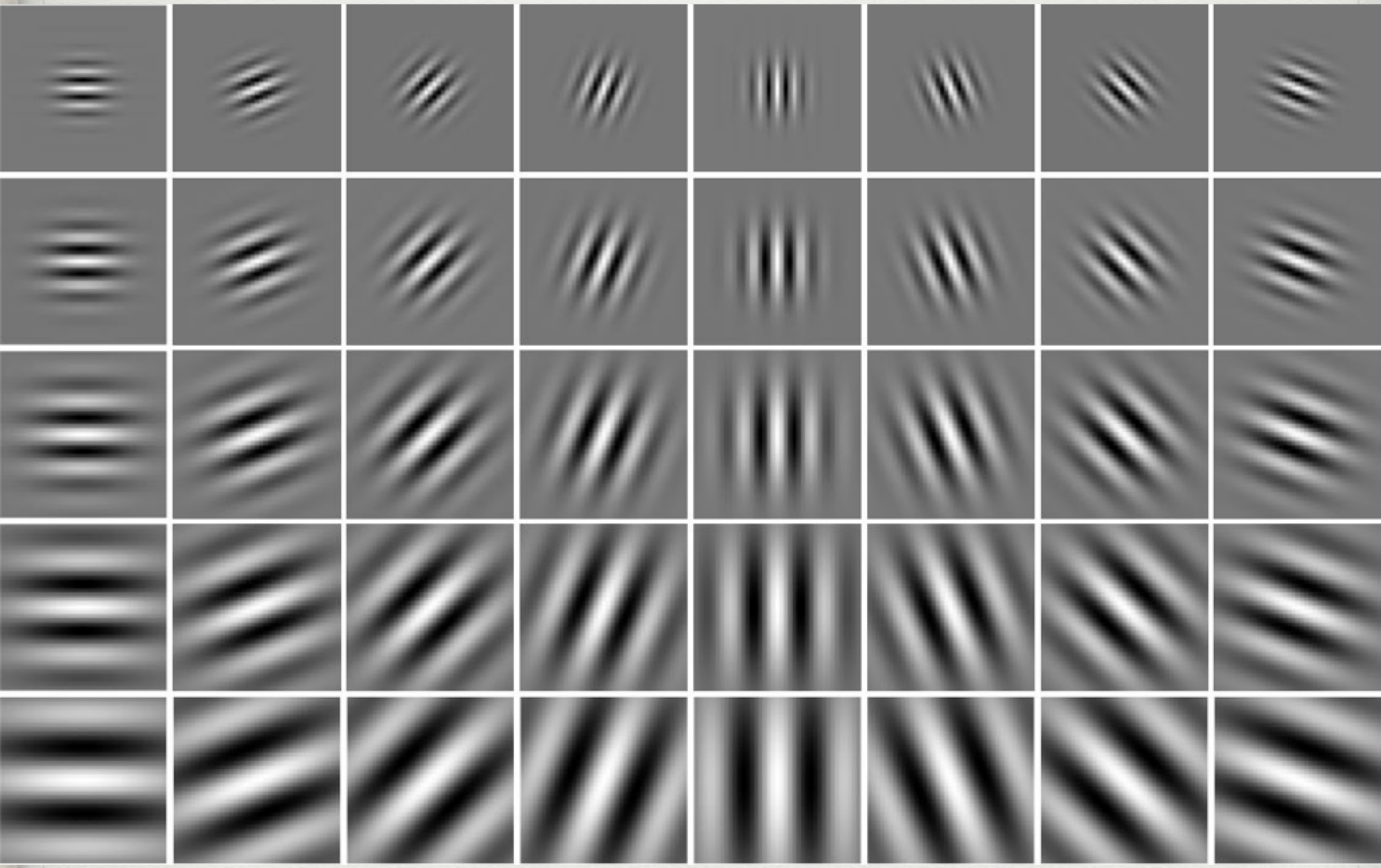
How to link?

- Qualitatively
- **Quantitatively**
  - Individual differences
  - Within-subject variability
  - **Stimulus variability**
- Joint model

Stimulus variability

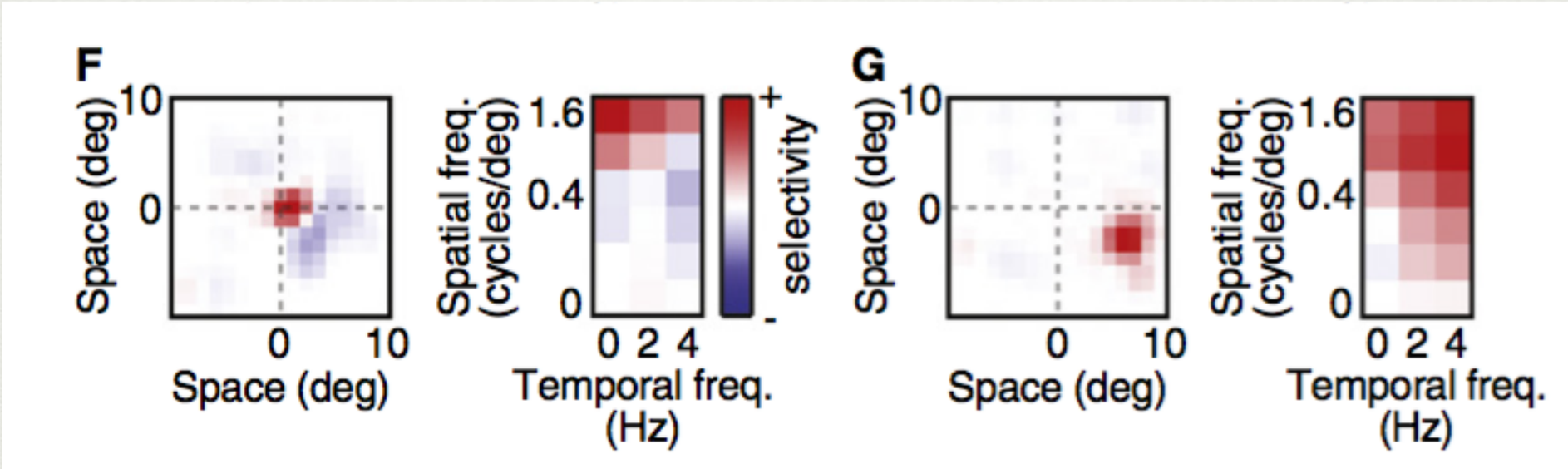
- Visual science (David Marr)
- Encoding models: stimulus -> brain activity
  - P(brain_activity | stimulus)
- Bayes' rule can invert these models
  - P(stimulus | brain_activity) ~= P(brain_activity | stimulus) P(stimulus)

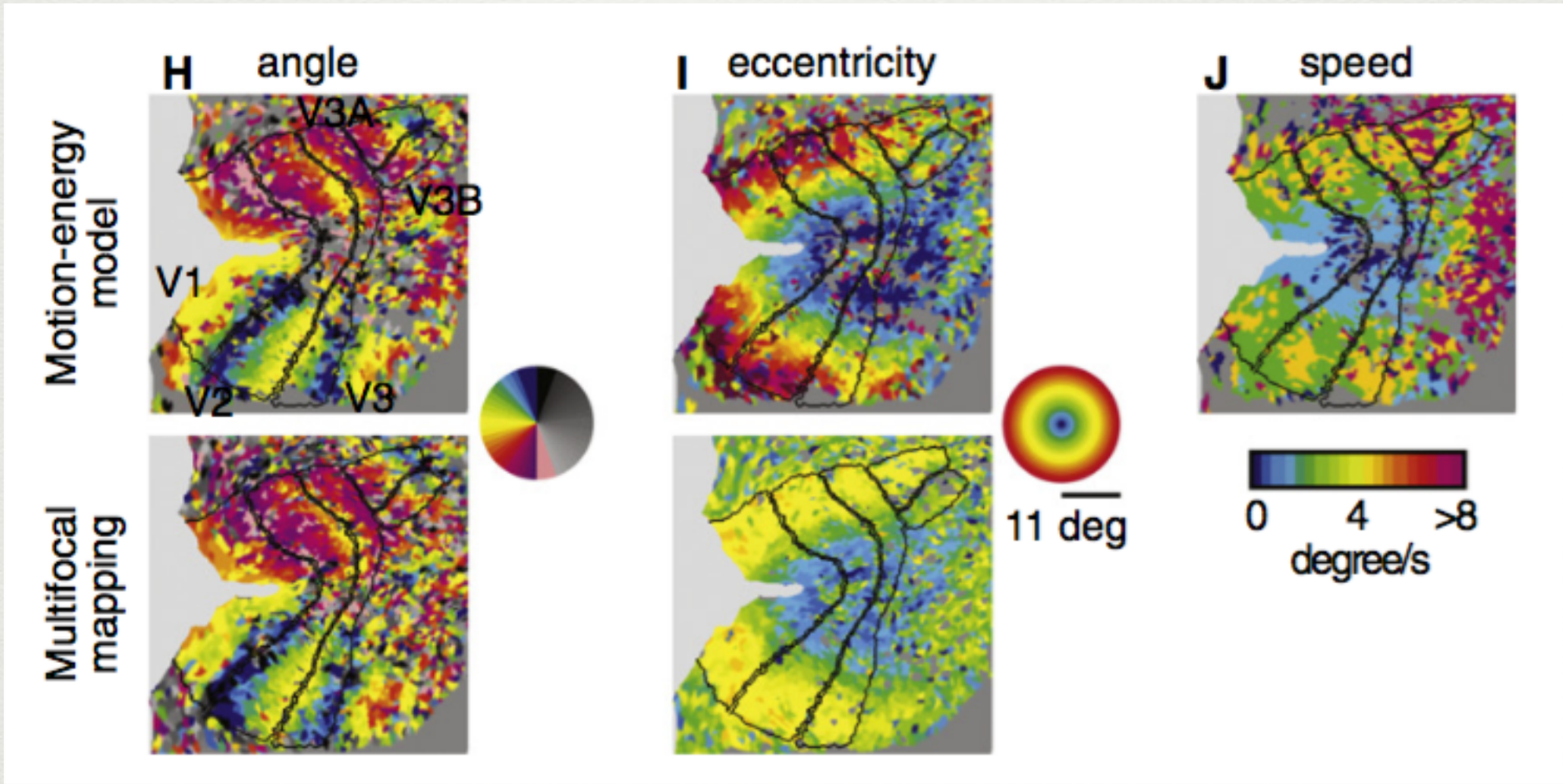
Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies



voxel 1

voxel 2




H
angle
I
eccentricity
J
speed
Motion-energy model
Multifocal mapping
V3A
V3B
V1
V2
V3
11 deg
0
4
>8
degree/s

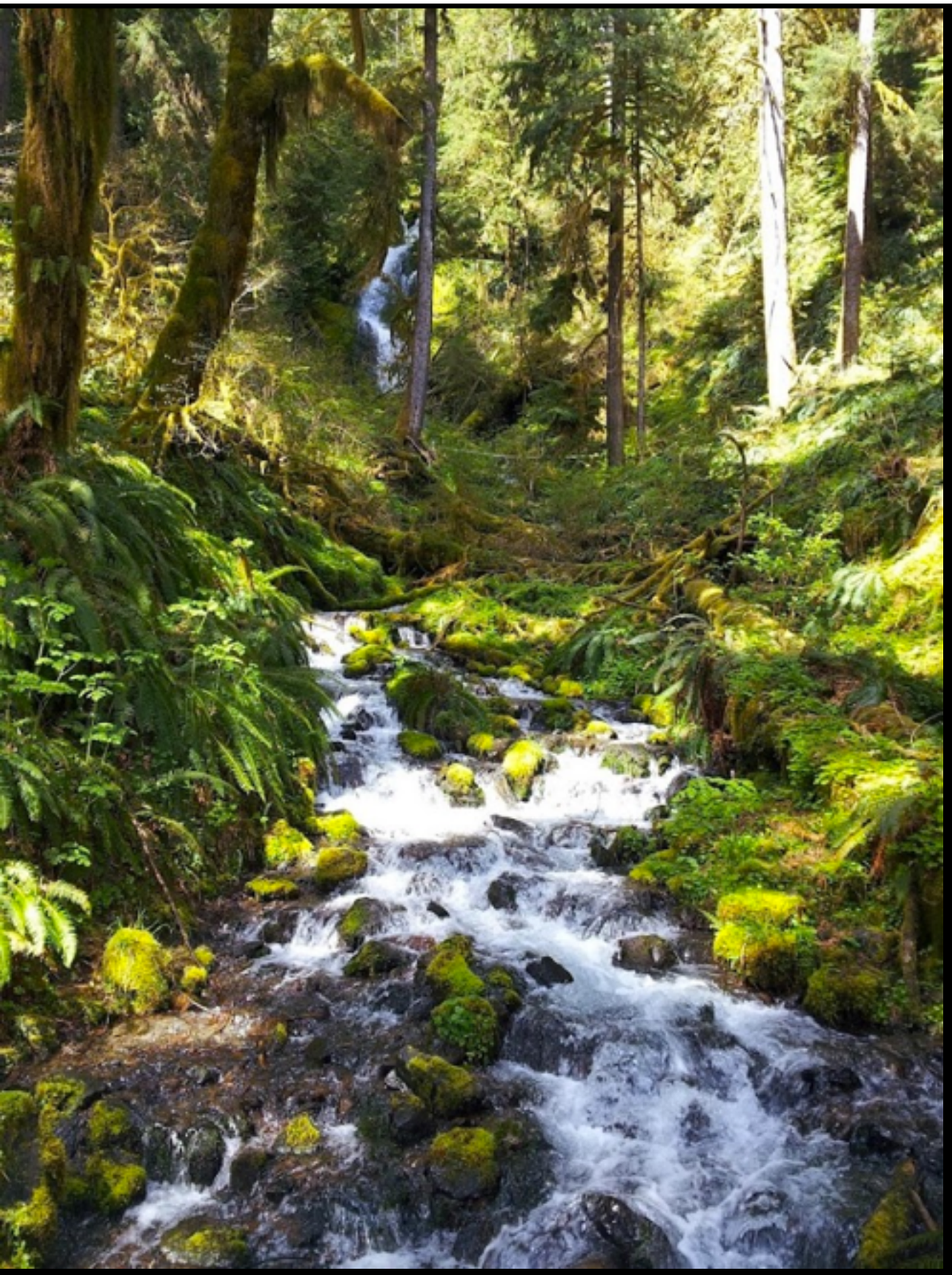
Reconstructing visual experiences from brain activity evoked by natural movies

Shinji Nishimoto, An T. Vu, Thomas Naselaris, Yuval Benjamini, Bin Yu, Jack L. Gallant

Supplemental movie S1

Inceptionism: Going deeper into Neural Networks

12 dec. 2008–18 jun. 2015



LETTER

doi:10.1038/nature14236

Human-level control through deep reinforcement learning

Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

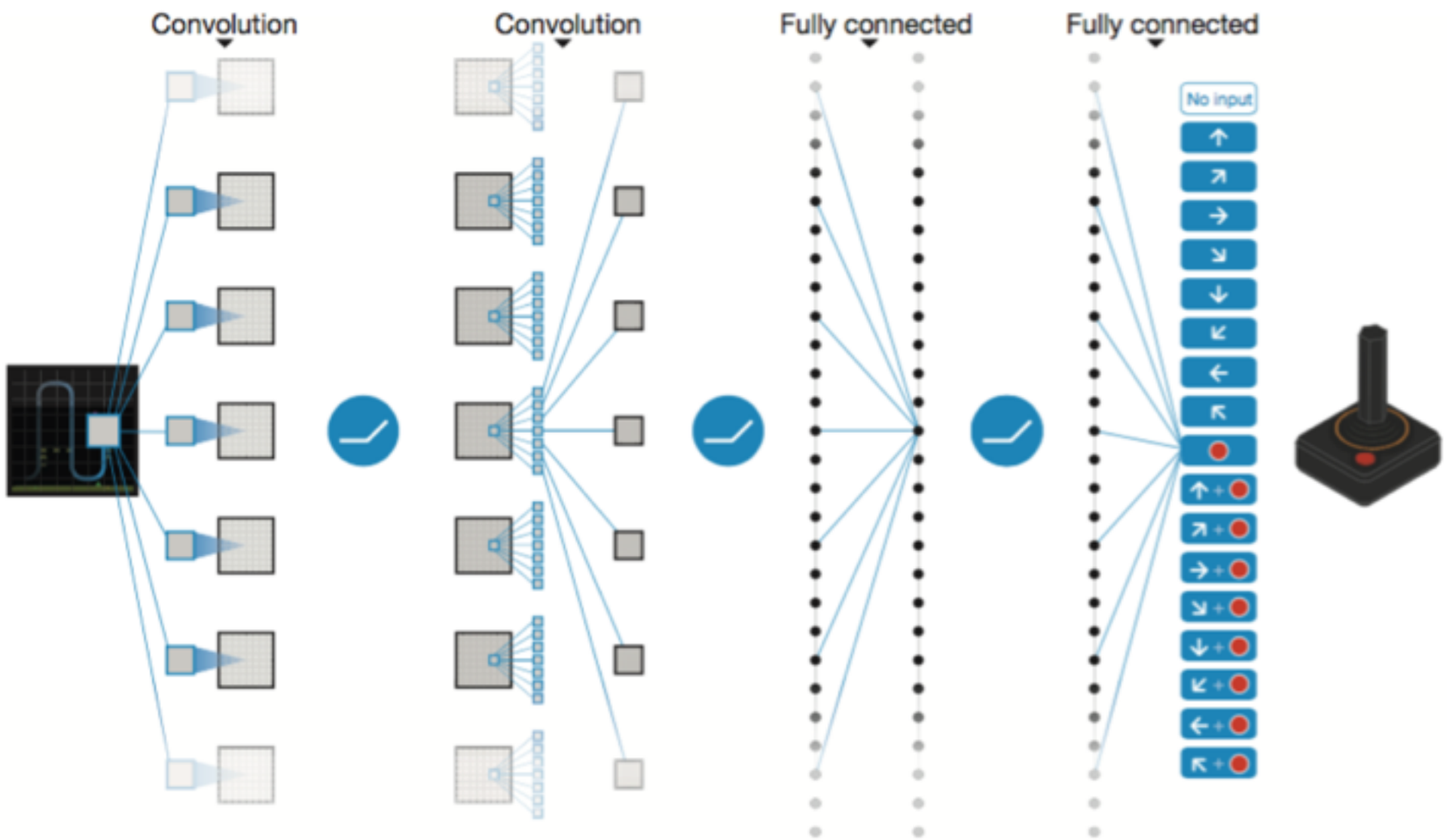
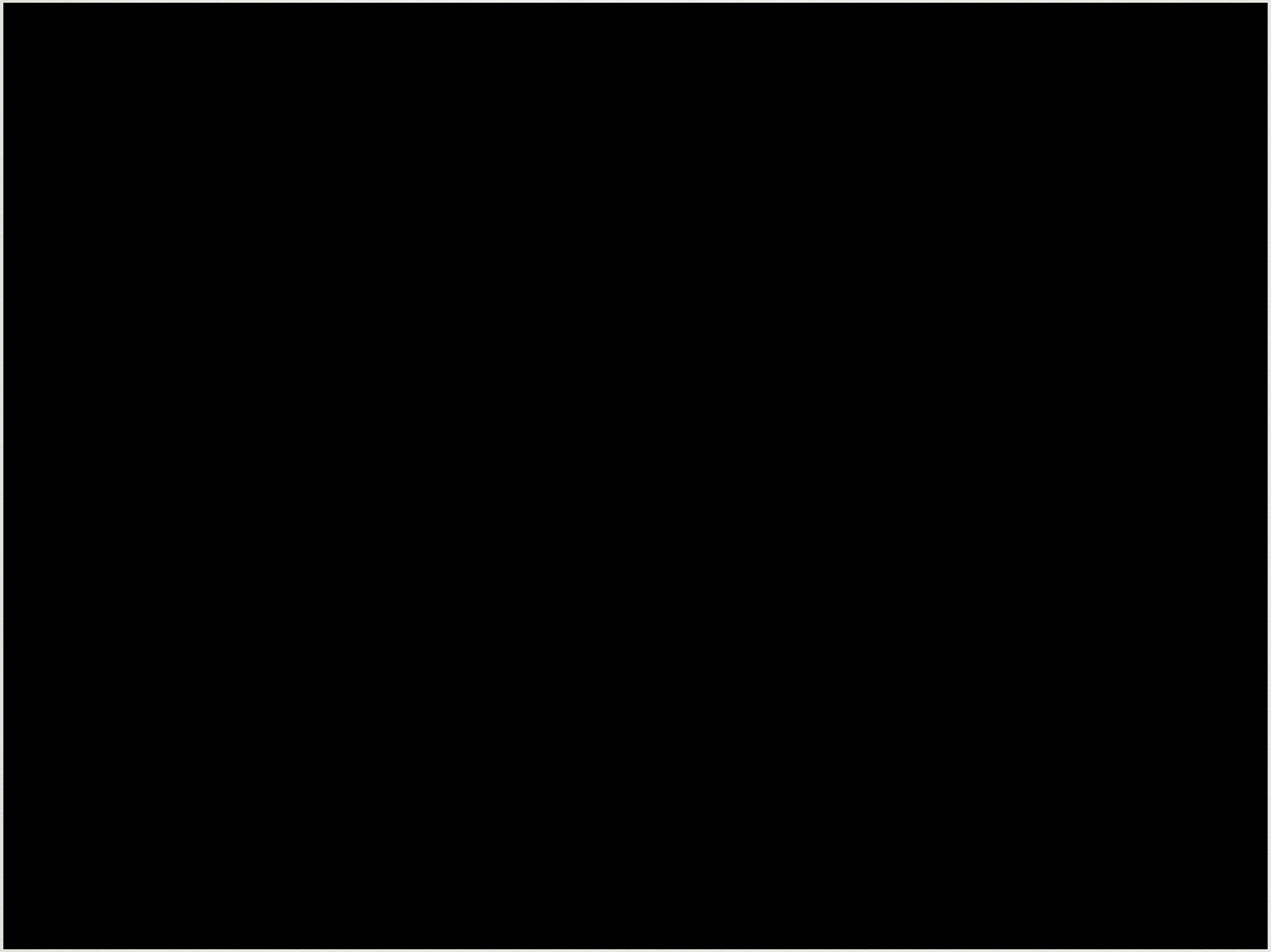


Figure 1 | Schematic illustration of the convolutional neural network. The details of the architecture are explained in the Methods. The input to the neural network consists of an $84 \times 84 \times 4$ image produced by the preprocessing map ϕ, followed by three convolutional layers (note: snaking blue line symbolizes sliding of each filter across input image) and two fully connected layers with a single output for each valid action. Each hidden layer is followed by a rectifier nonlinearity (that is, $\max(0,x)$).



Deep Neural Networks Reveal a Gradient in the Complexity of Neural Representations across the Ventral Stream

Umut Güçlü and Marcel A. J. van Gerven
Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands

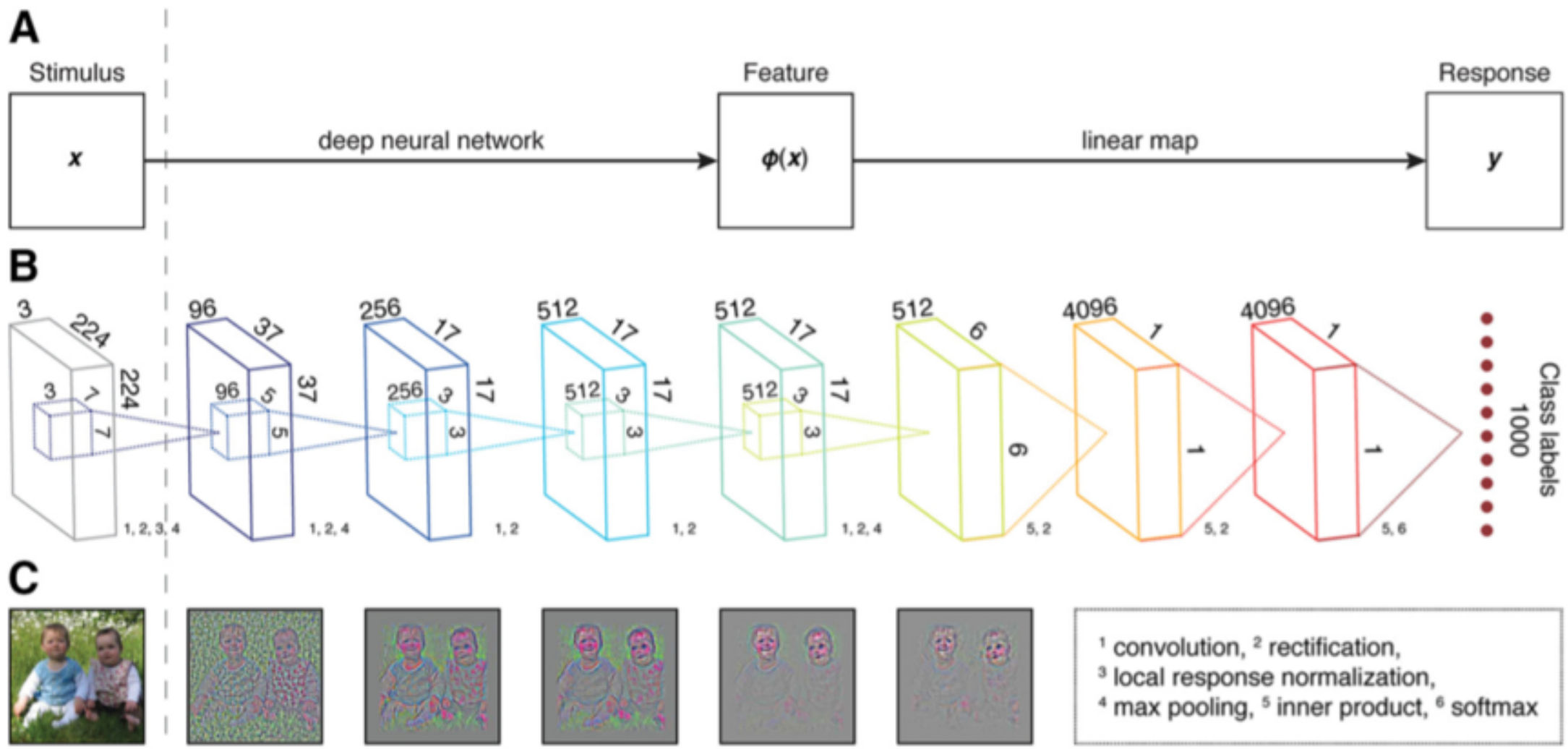
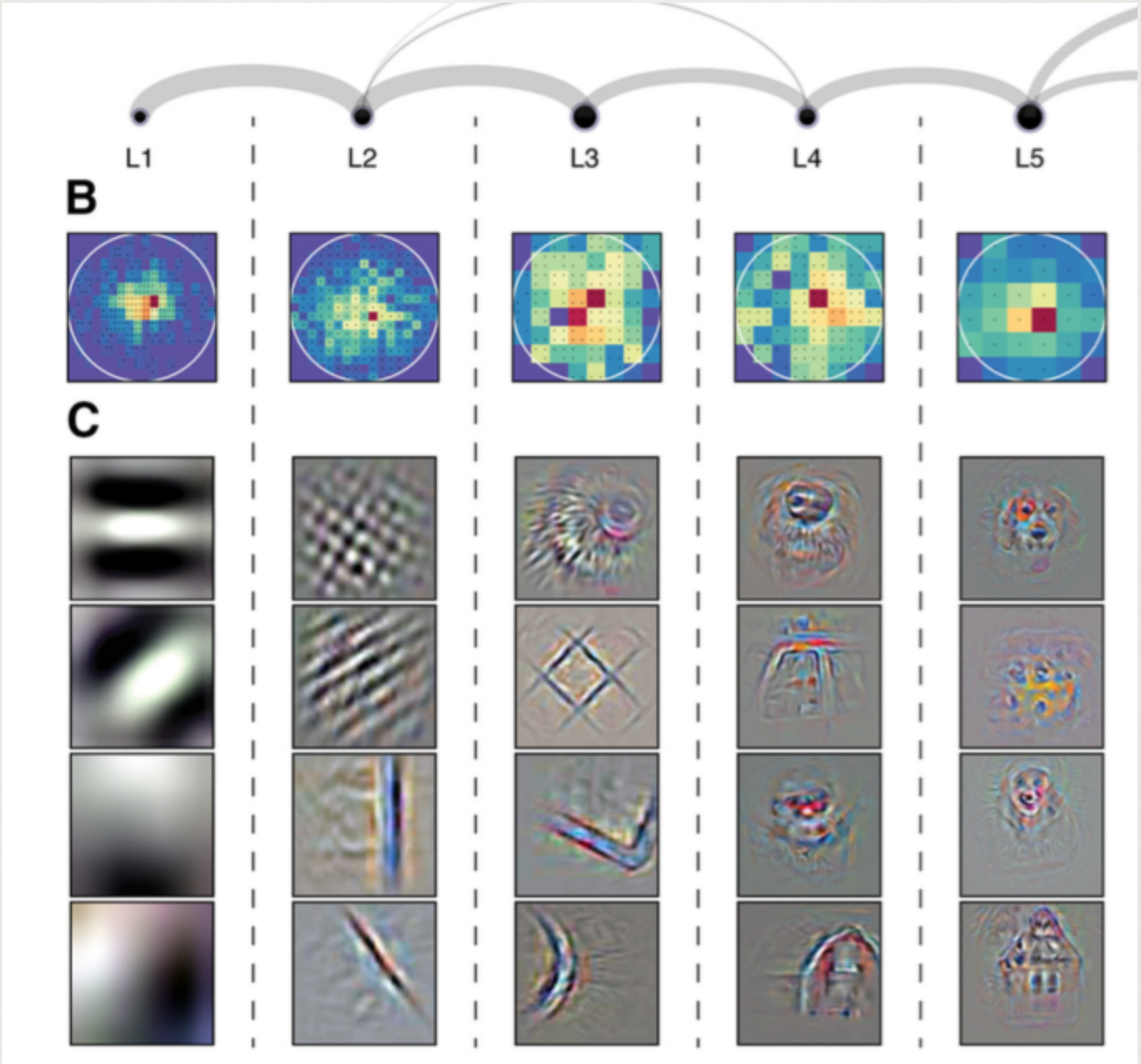
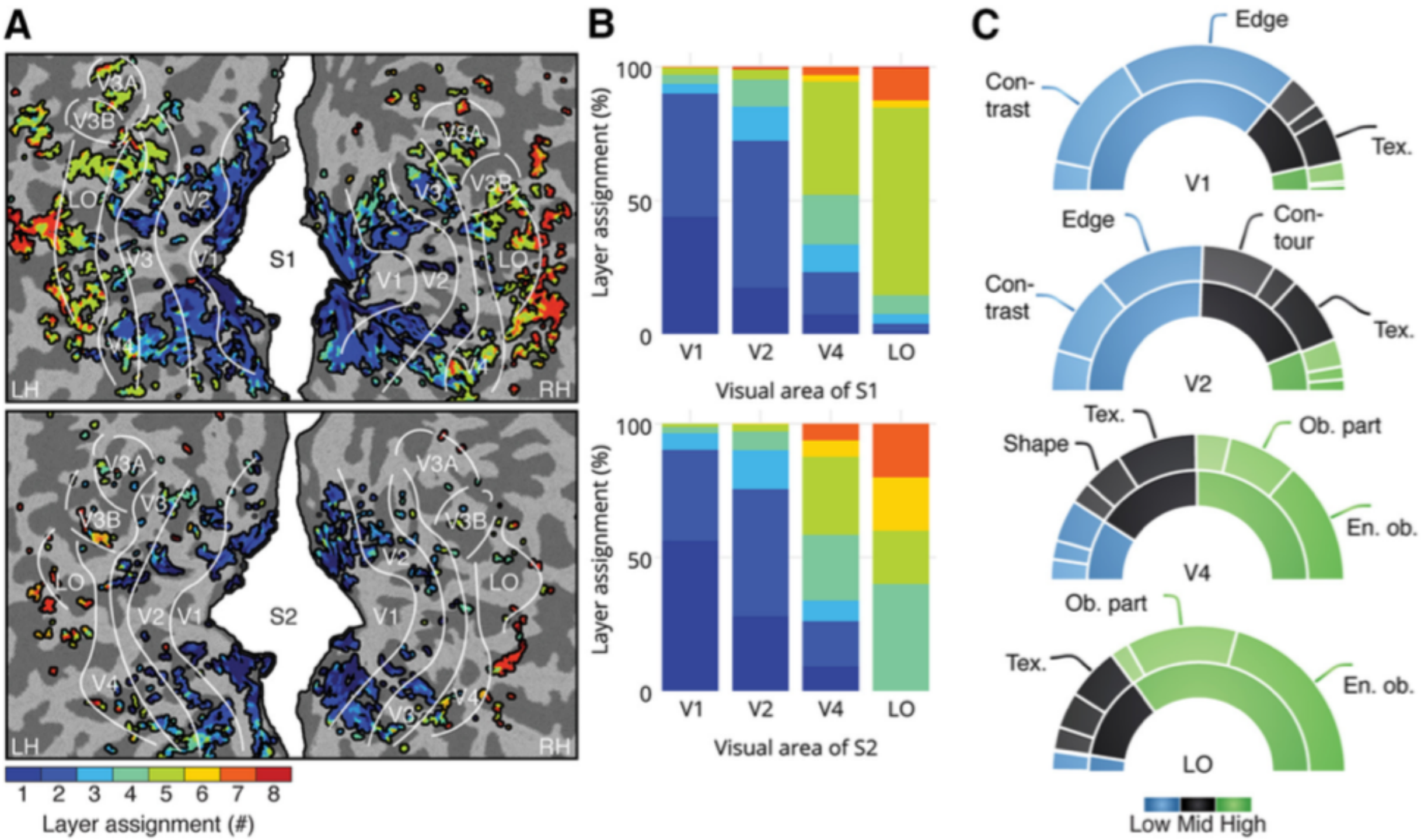


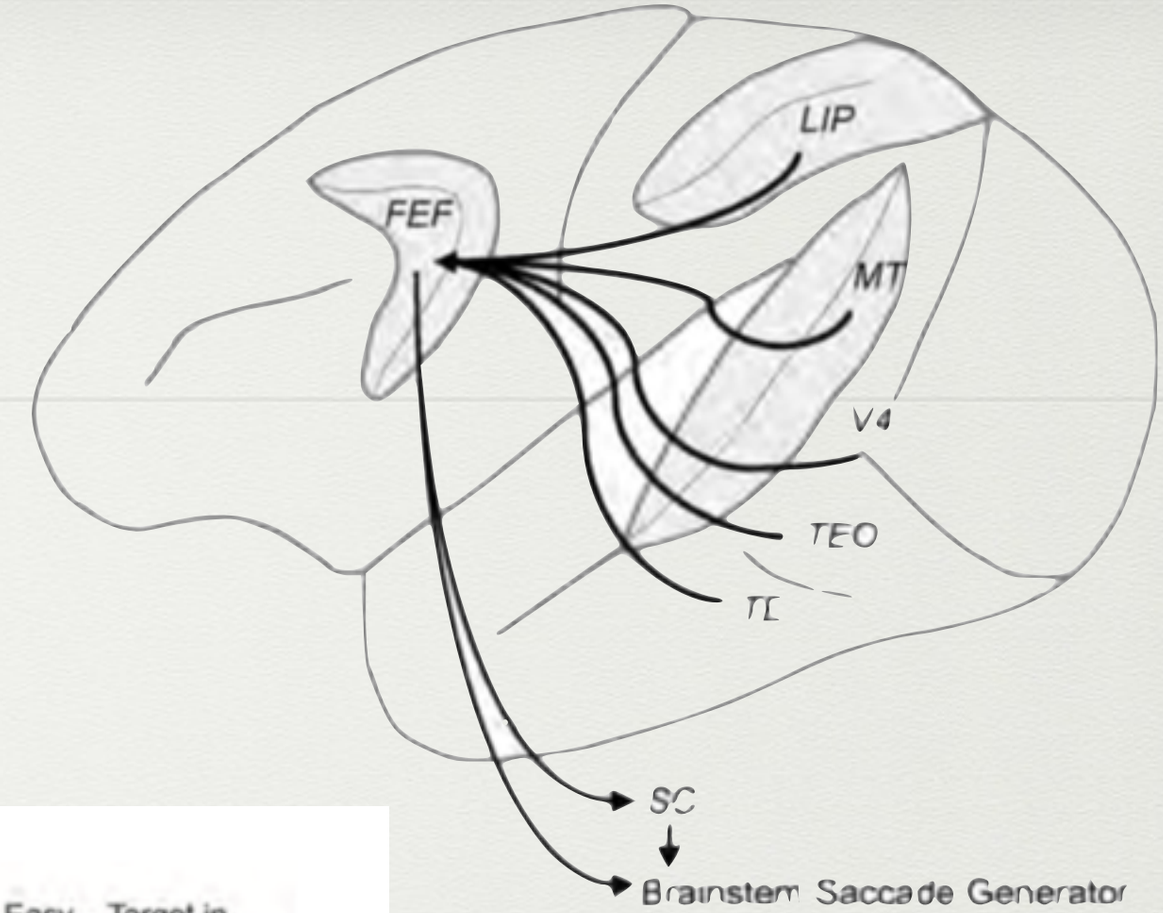
Figure 1. DNN-based encoding framework. ***A***, Schematic of the encoding model that transforms a visual stimulus to a voxel response in two stages. First, a deep (convolutional) neural network transforms the visual stimulus (***x***) to multiple layers of feature representations. Then, a linear mapping transforms a layer of feature representations to a voxel response (***y***). ***B***, Schematic of the deep neural network where each layer of artificial neurons uses one or more of the following (non)linear transformations: convolution, rectification, local response normalization, max pooling, inner product, and softmax. ***C***, Reconstruction of an example image from the activities in the first five layers.

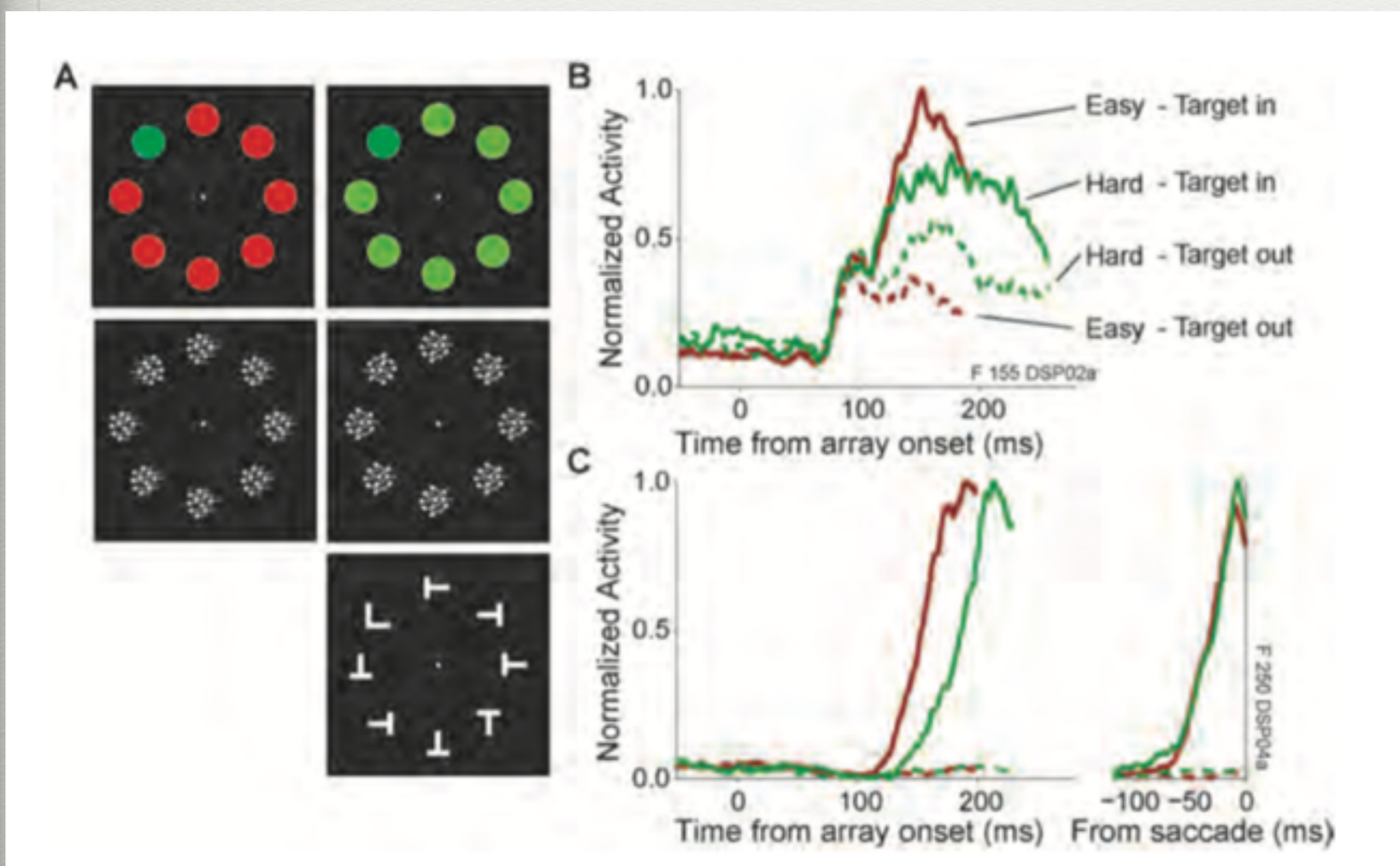

L1
L2
L3
L4
L5
B
C


A
V3A
V3B
LO
V2
V3
V1
S1
V4
LH
RH
S2
1 2 3 4 5 6 7 8
Layer assignment (#)
B
Layer assignment (%)
100
50
0
V1 V2 V4 LO
Visual area of S1
Visual area of S2
C
Edge
Con-
trast
Tex.
V1
Con-
tour
V2
Shape
Ob. part
En. ob.
V4
LO
Low Mid High

How to link?

- Qualitatively
- Quantitatively
  - Individual differences
  - Within-subject variability
  - Stimulus variability
- **Joint model**

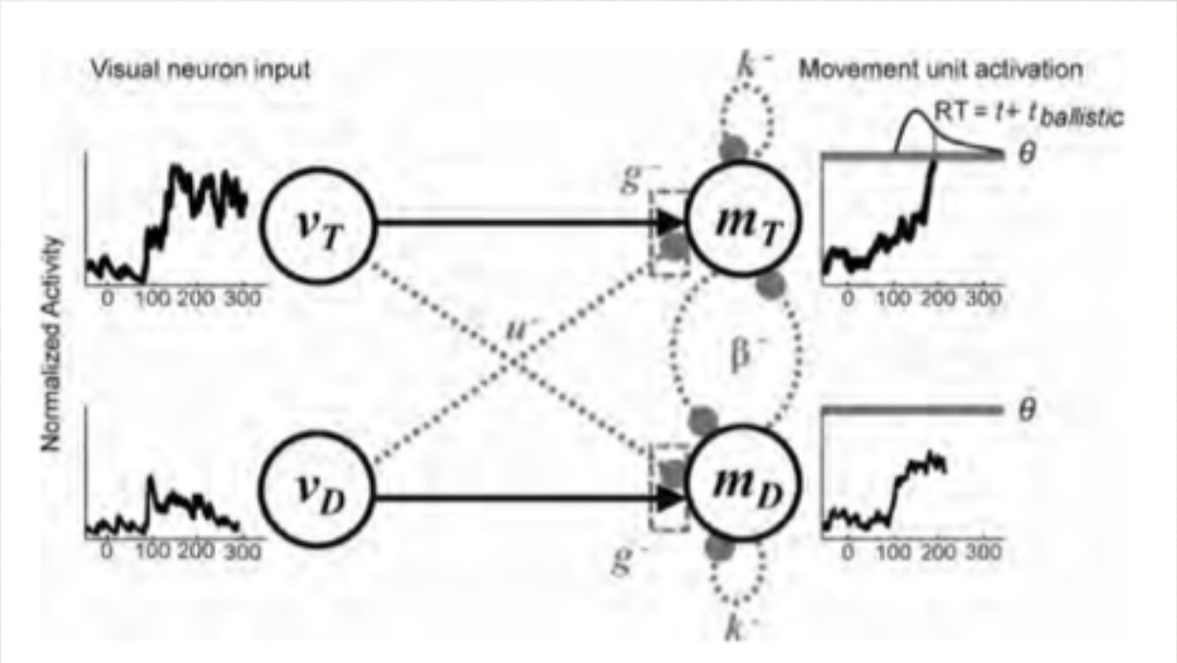




Purcell et al., 2010

Joint model Purcell

- Uses activity in *visual neuron* to predict both *behaviour* and *neural activity in FEF*
- Can dissociate between different integrator models that make identical predictions about behaviour



Summary

- Formal, computational mode of cognition are crucial to interpret neural recordings
- There exist multiple strategies to link cognitive models and neural data, ranging from very loose (qualitative), tighter (quantitative), to very tight (joint models).

Python for neuroscience

- Model-based neuroscientists have to do
  - Computational modelling of behavior
  - Statistical modelling of behavior
  - Statistical modelling of neural data

Possibilities

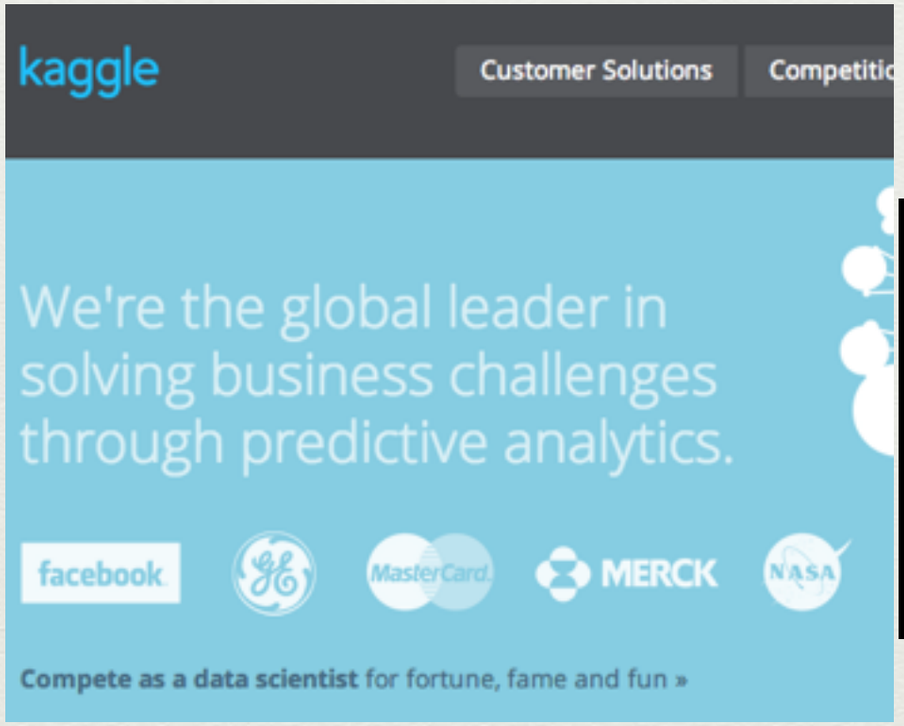
MATLAB

python

python

- Very good language design
  - Fully object-oriented
  - explicitly designed to be concise and easy-to-read
- Most-used programming language after R
- Lot of momentum, especially in neurosciences
- Extensibility
  - Cython
  - Modules
  - RPy
- Open source, rapid development github-model

Python in the Real World



Python in Science

statsmodels

NumPy

Nipype:
Neuroimaging in Python
Pipelines and Interfaces

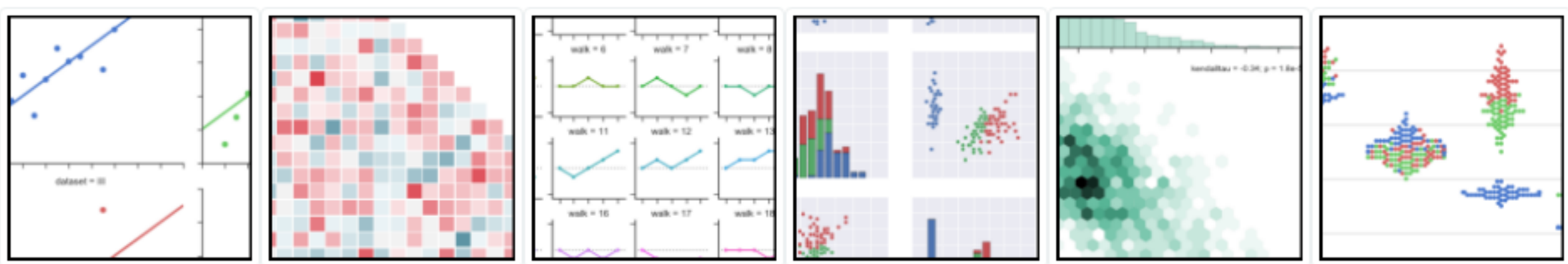
HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python

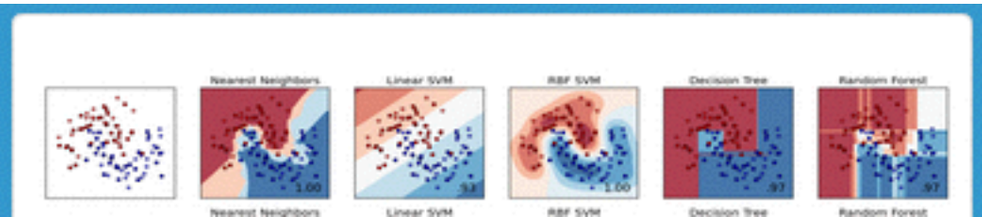
http://www.fron
10.3389/fninf.2

Thomas V. Wiecki*[*†], Imri Sofer[†] and Michael J. Frank

Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI, USA

Seaborn: statistical data visualization





scikit-learn

Machine Learning in Python

Today

- Introduction to IPython Notebook
- Introduction to broad array of Python packages that are useful for model-based neuroscientists
- Set of tutorials for you to work on