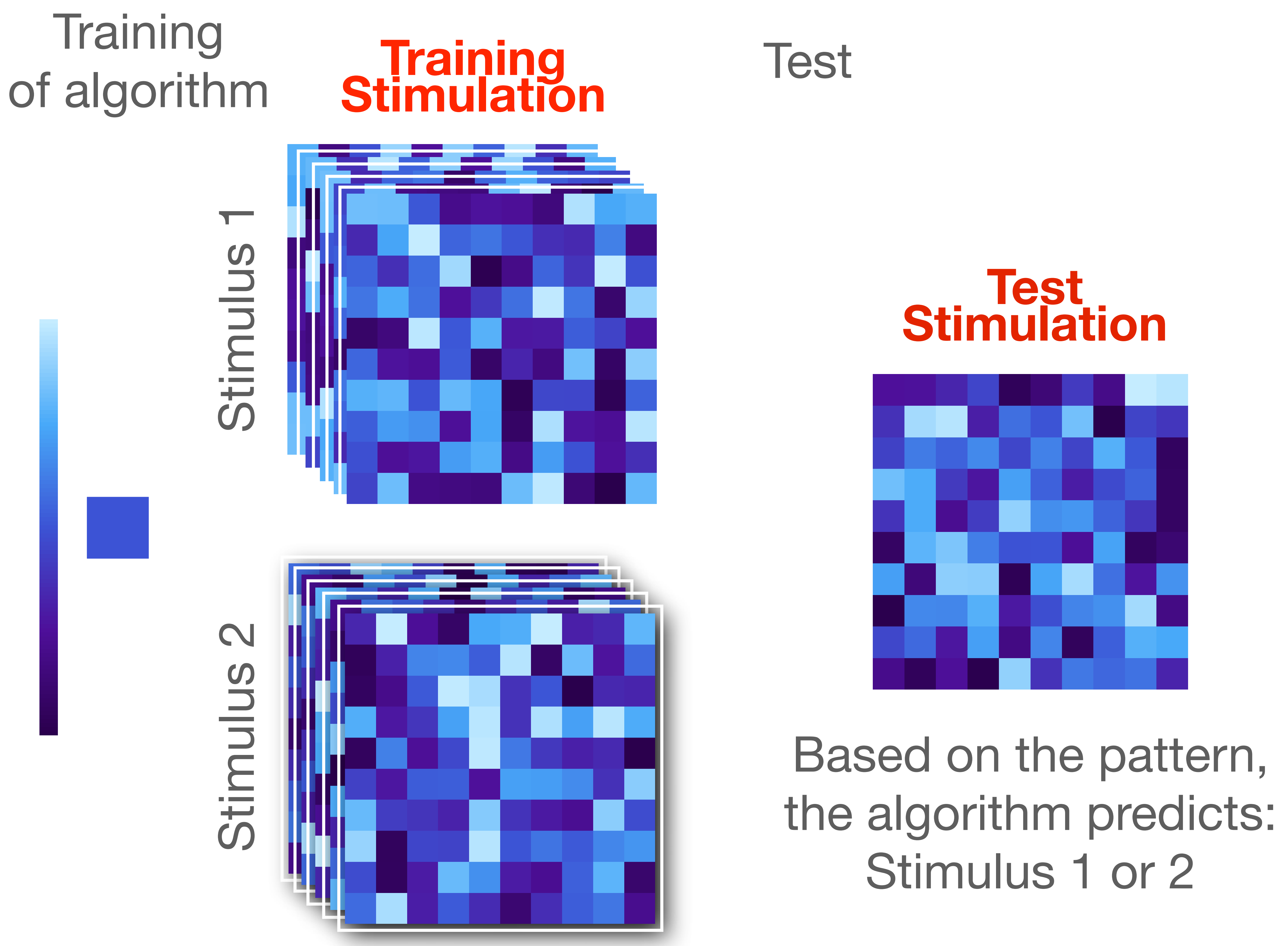
Encoding Models (Receptive Fields in fMRI)

Brain Imaging 2022

Decoding

Not mean signal intensity, but the **pattern** of activity in a certain region of interest is used to classify what the brain state is.



Performance of algorithm expressed as accuracy

gauge on information content in a specific region of interest

Decoding

The classifier draws a relation between the input space (conditions, images, sounds, etc) and the patterns of brain activation found in a given region

If the pattern of activations contains information regarding the input space, the classifier can successfully decode.

Decoding

*Picking up on the feature space **is left to the machine-learning algorithm**, and with sophisticated algorithms it is **difficult to understand** what structure the algorithm is basing its judgments on.*

Strong debates:
How do these patterns relate to neural mechanisms?

Is fMRI sampling sub-voxel information, or merely basing itself on the global structure within a region?

Decoding

We don't care about what each voxel is doing, but only about the distribution of values ('pattern') across a certain region.

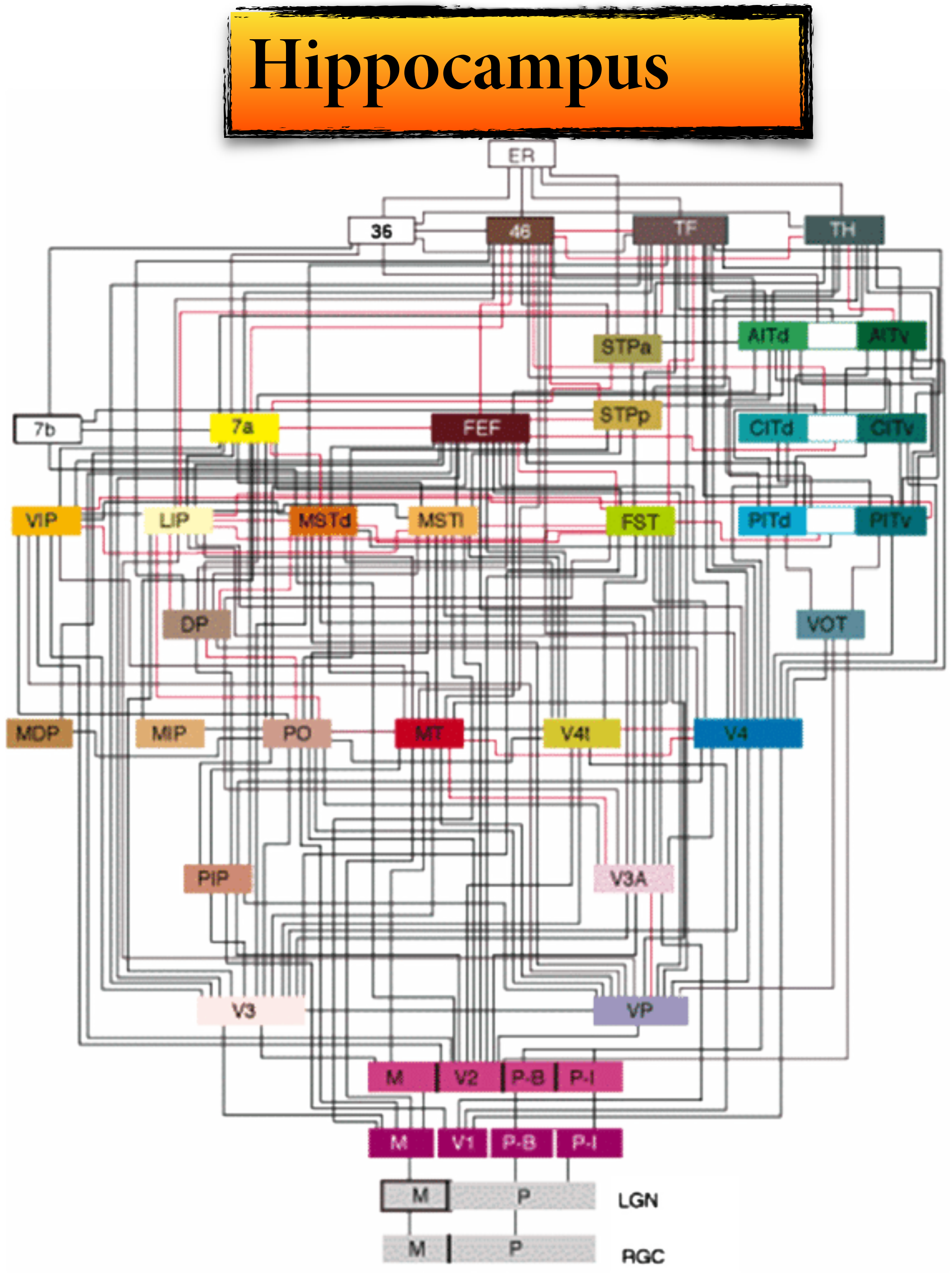
*Perhaps, **the single voxel is important.***

Perhaps, if we value the single voxel, we may
learn more about the brain's underlying mechanisms that lead to patterns...

Encoding

*If we can find what a neuron/sensor/voxel responds to, we will have a model of how this unit encodes information: **an encoding model***

*Once we know this, we can reverse this process, and see for a whole group of neurons/sensors/voxels, what their pattern of activations tells us about conditions, stimuli, etc. - **encoding-model based decoding***



Arrow of:

Increased Abstraction

Increased Invariance

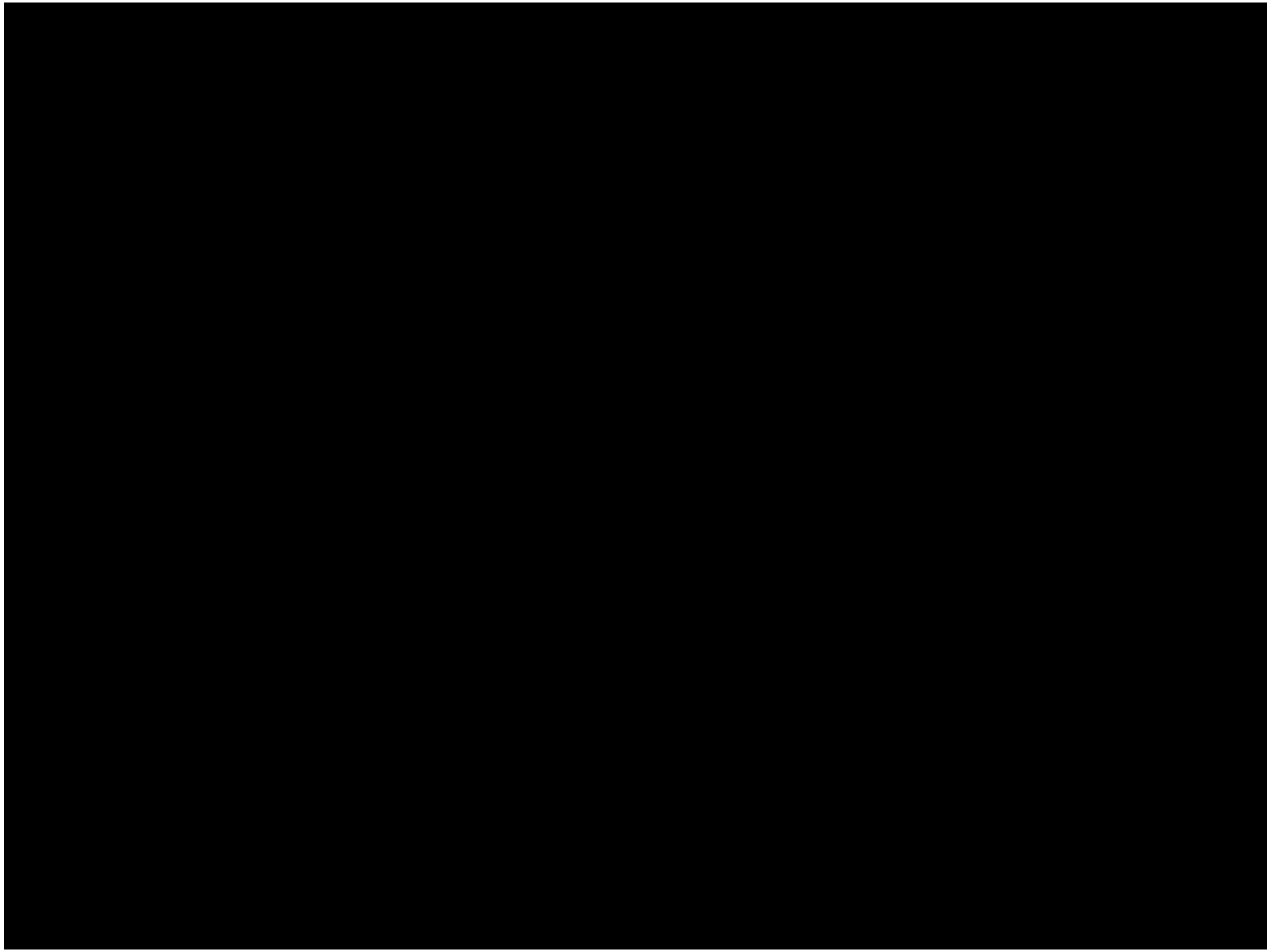
Increased Specialization

Increased Multi-Sensory Integration

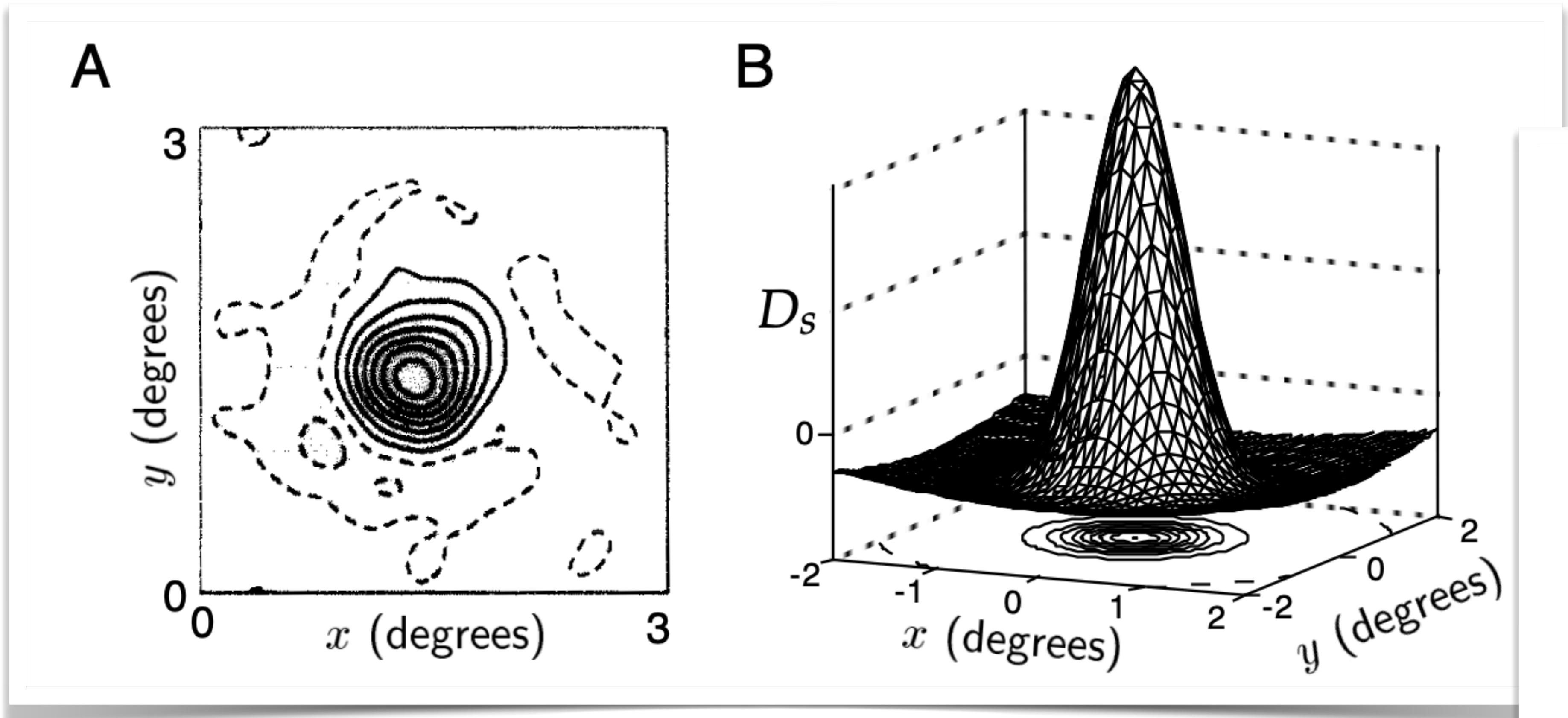
Increased Temporal Integration

Increased Action-Perception Integration

Hubel & Wiesel

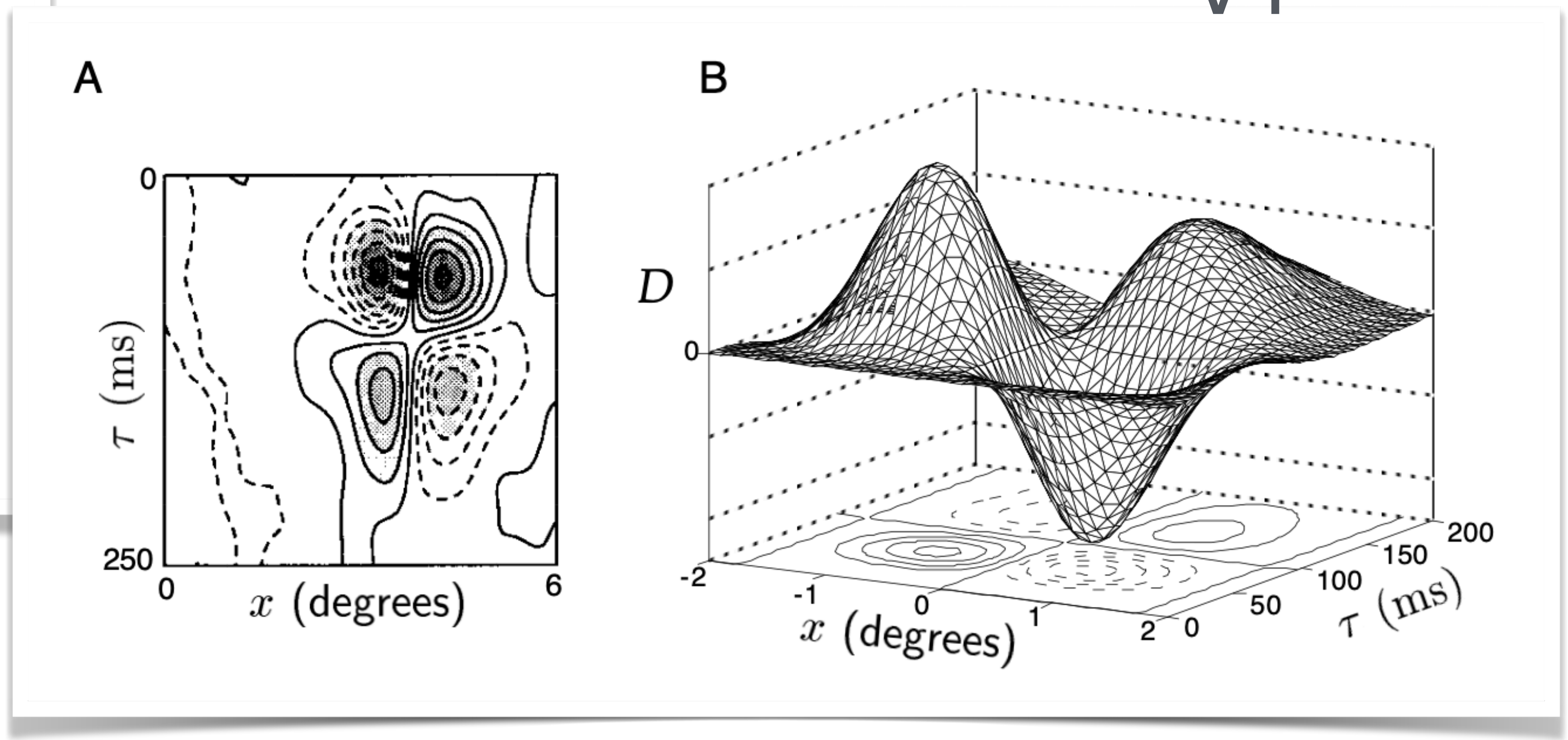


Visual Receptive Fields

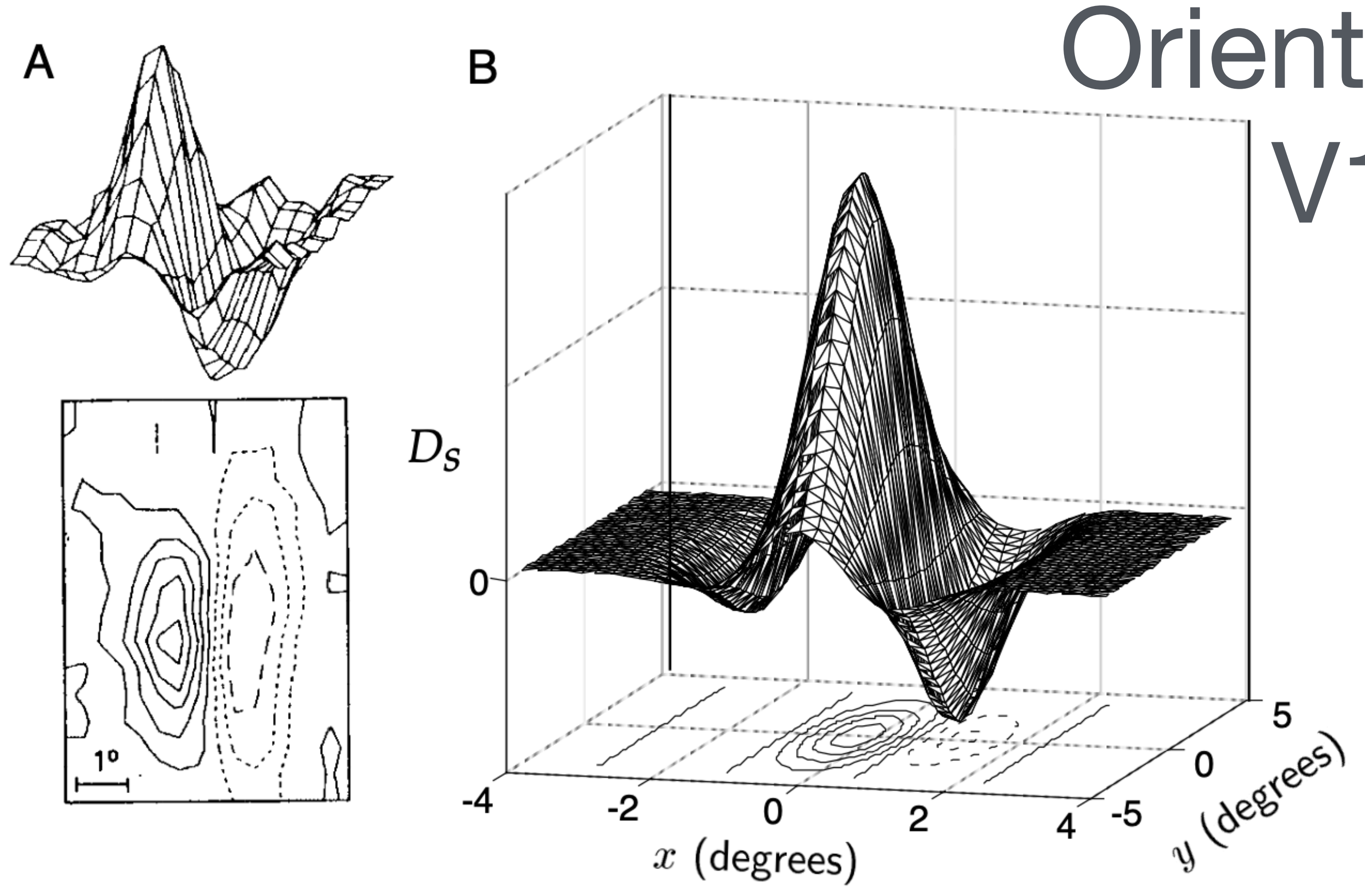


Lateral Geniculate Nucleus
LGN

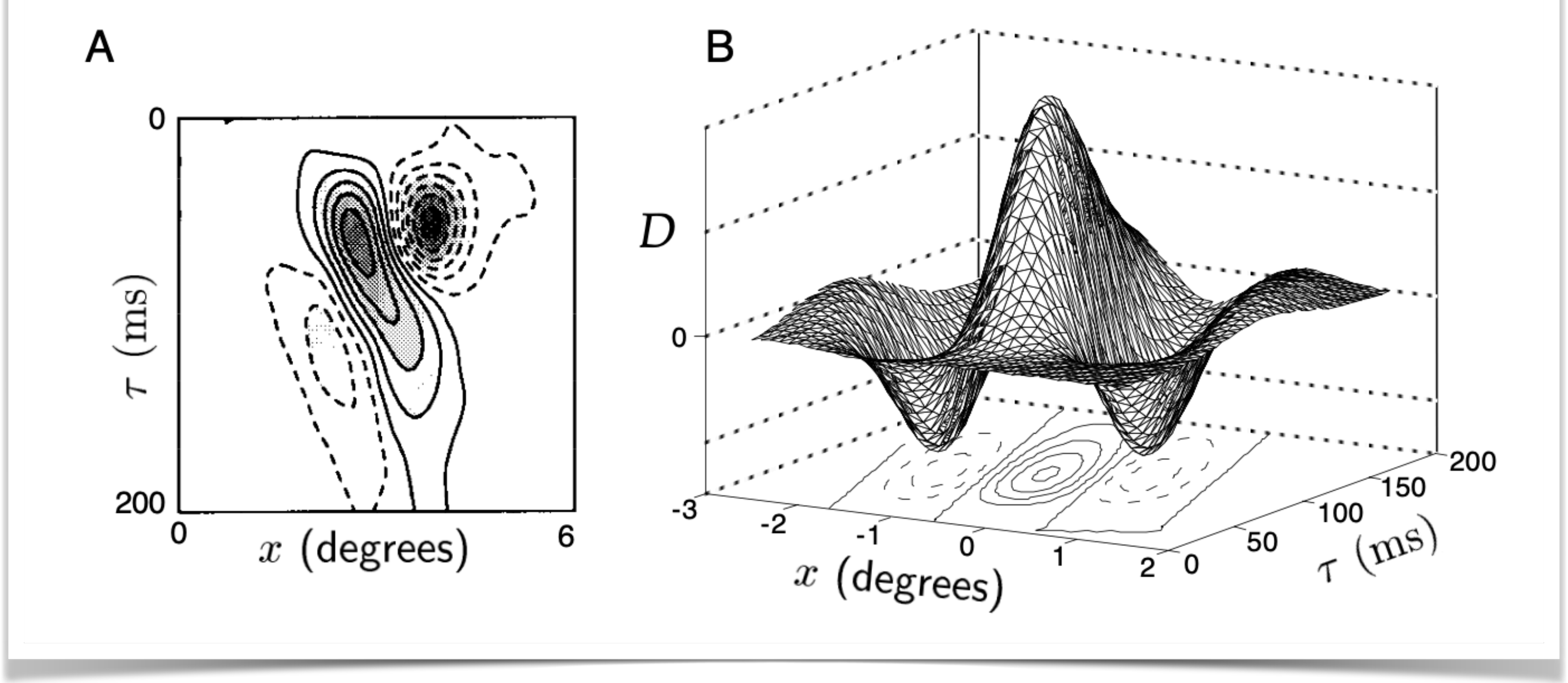
Flicker
V1

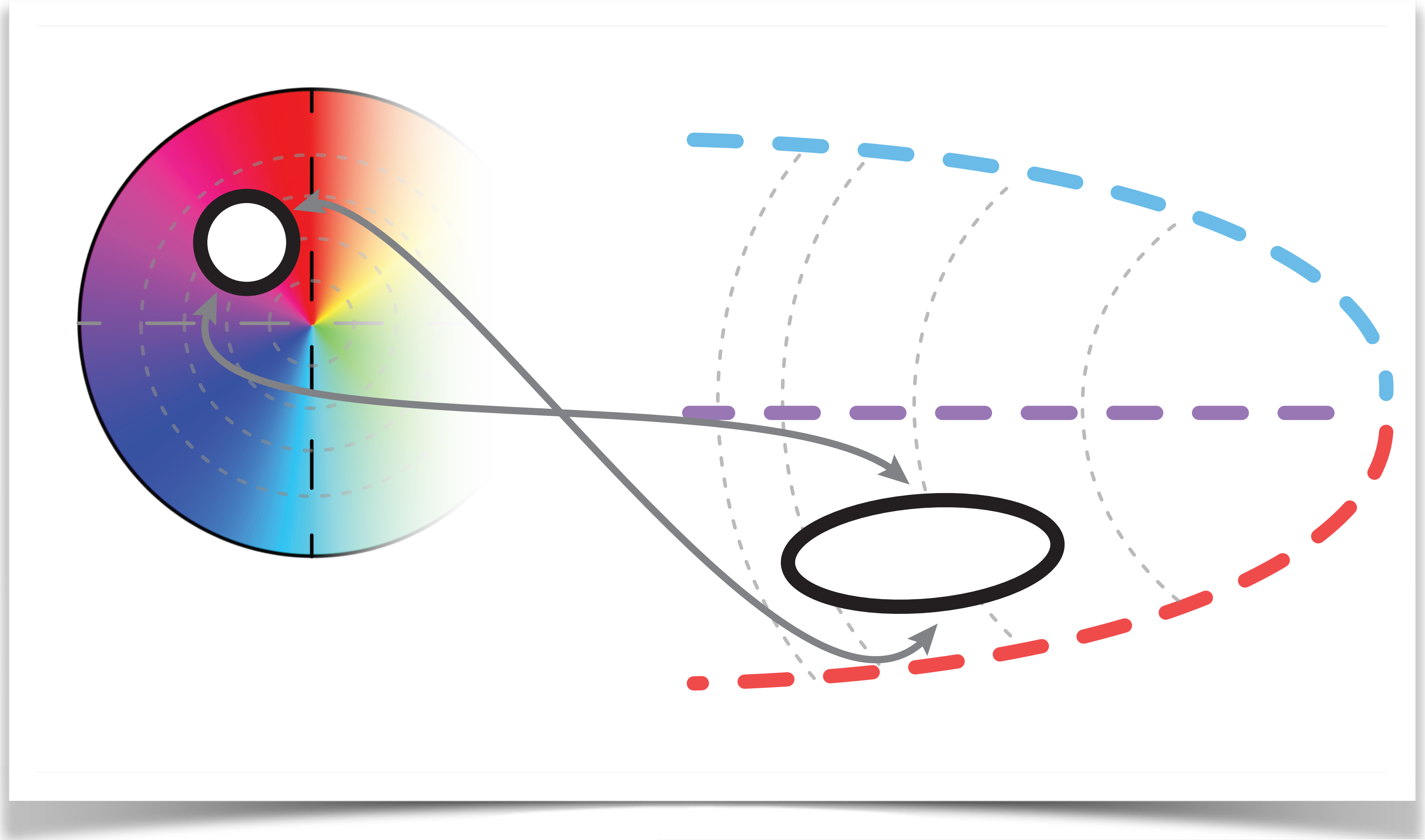


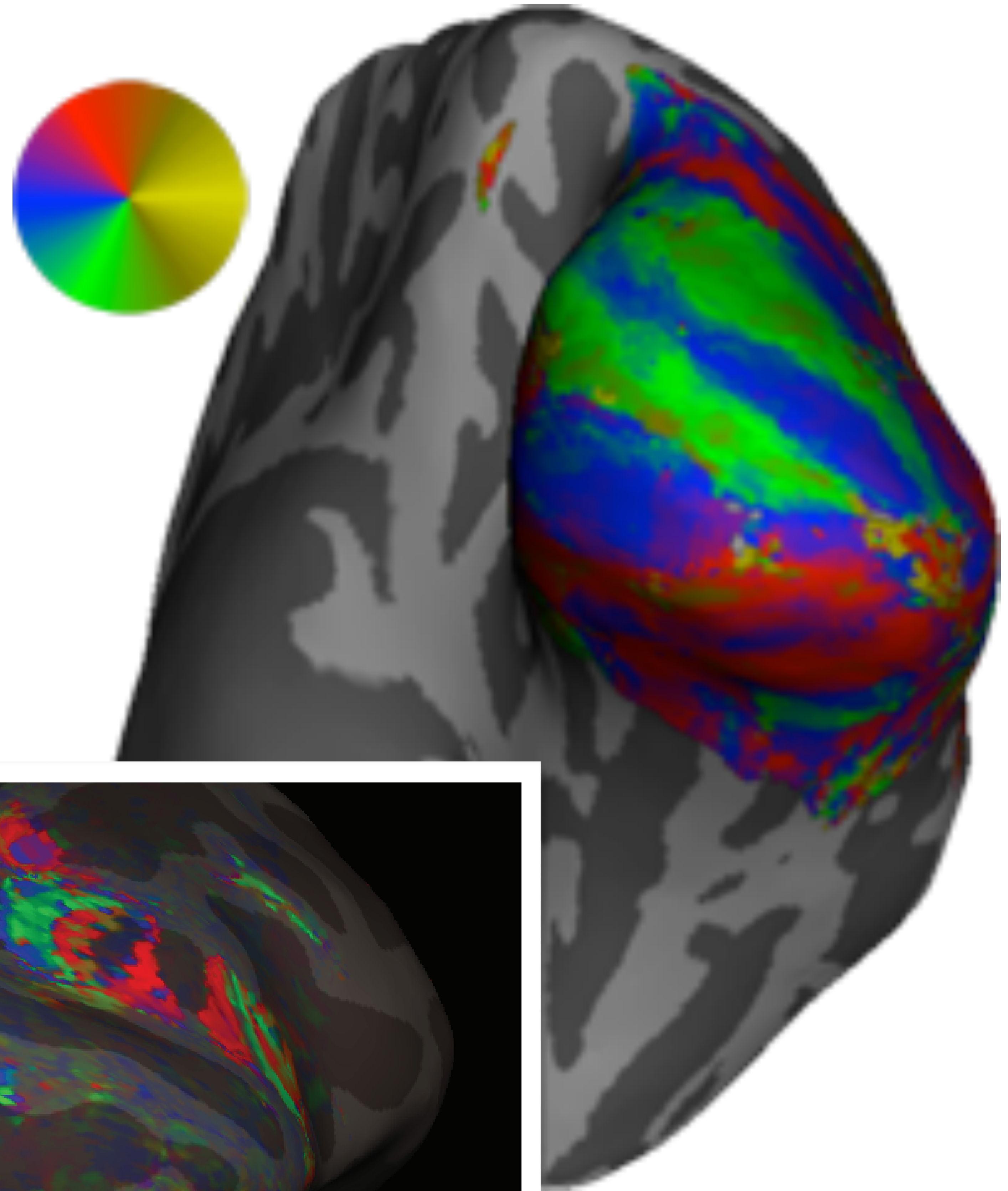
Orientation
V1

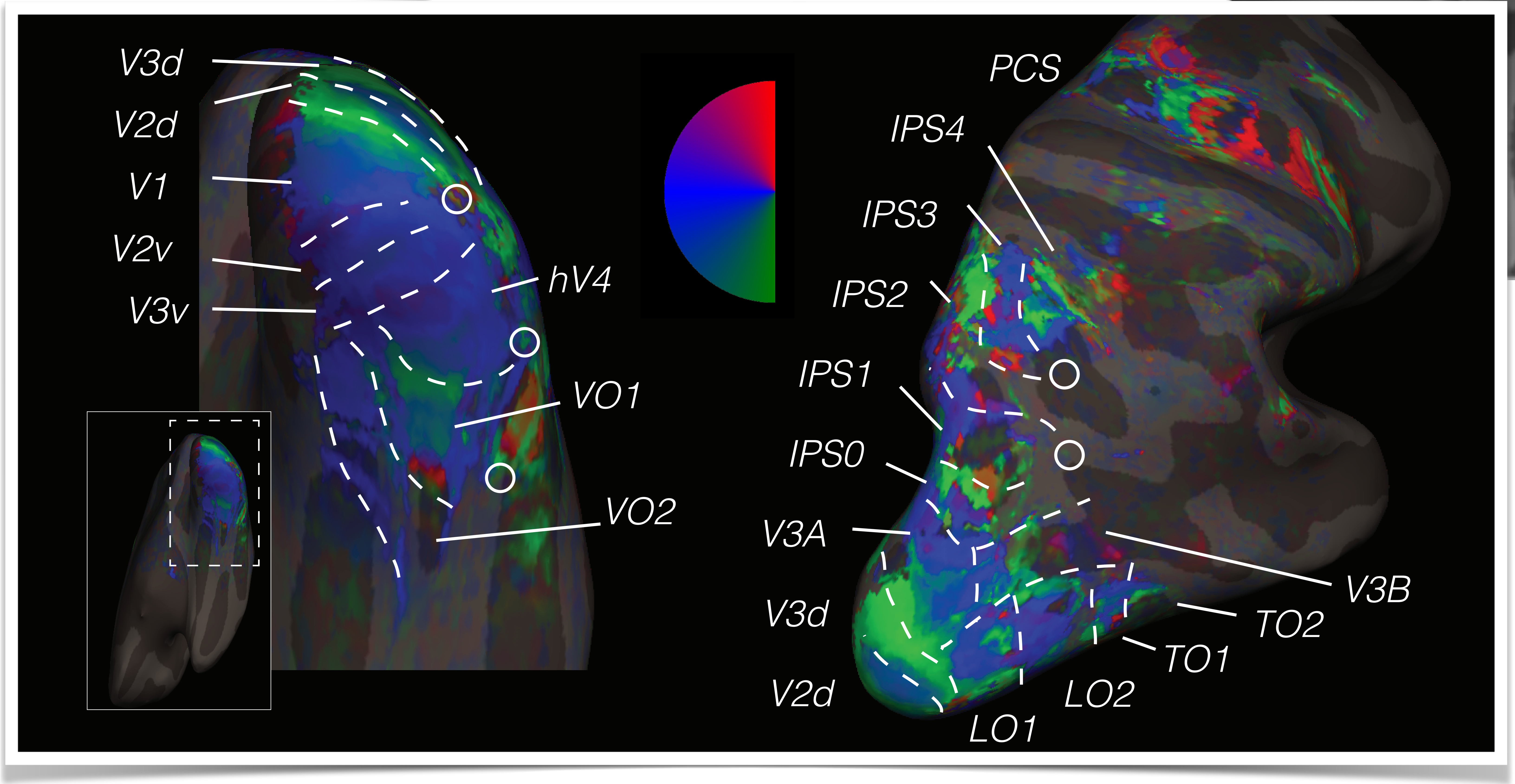


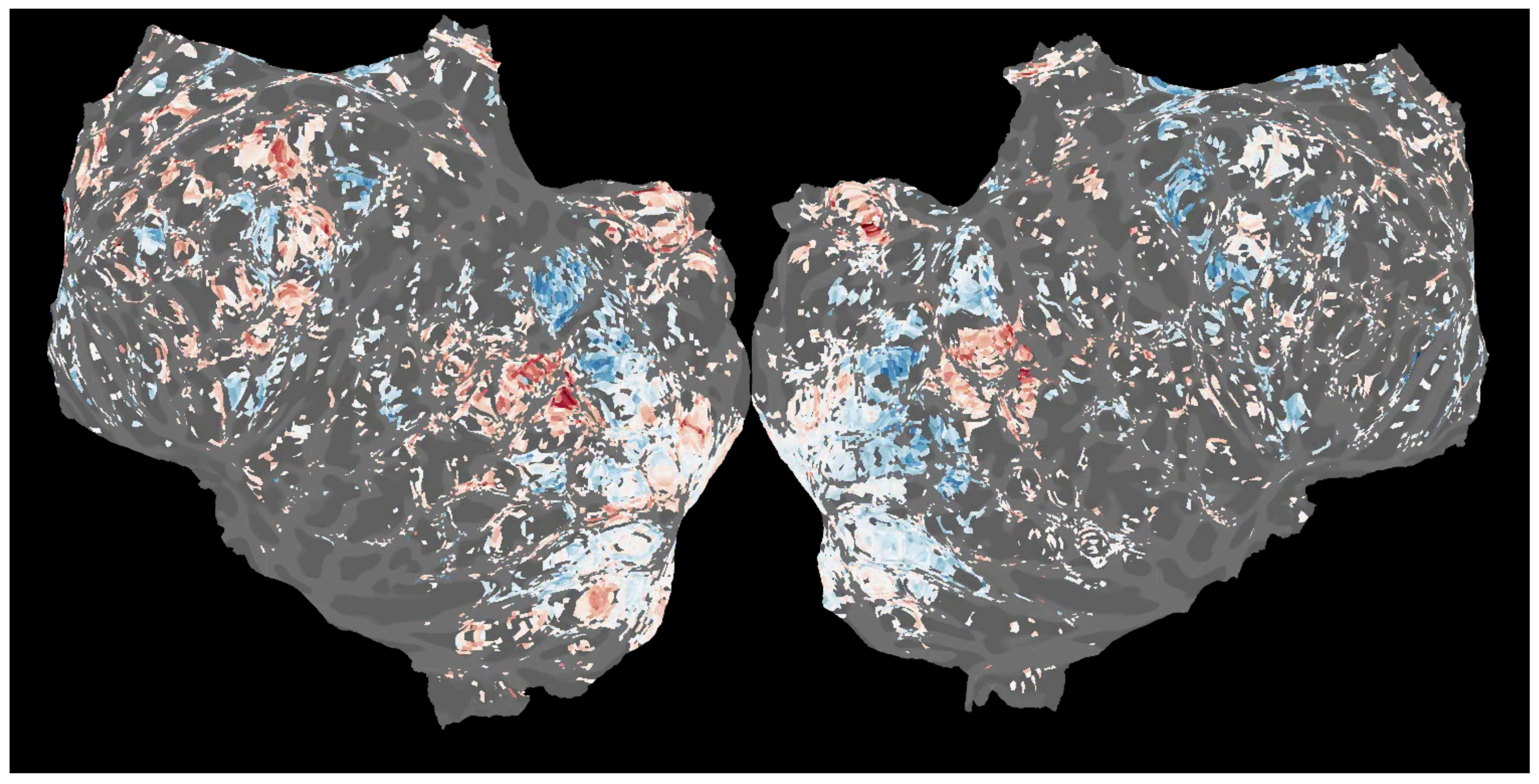
Motion
V1 & MT



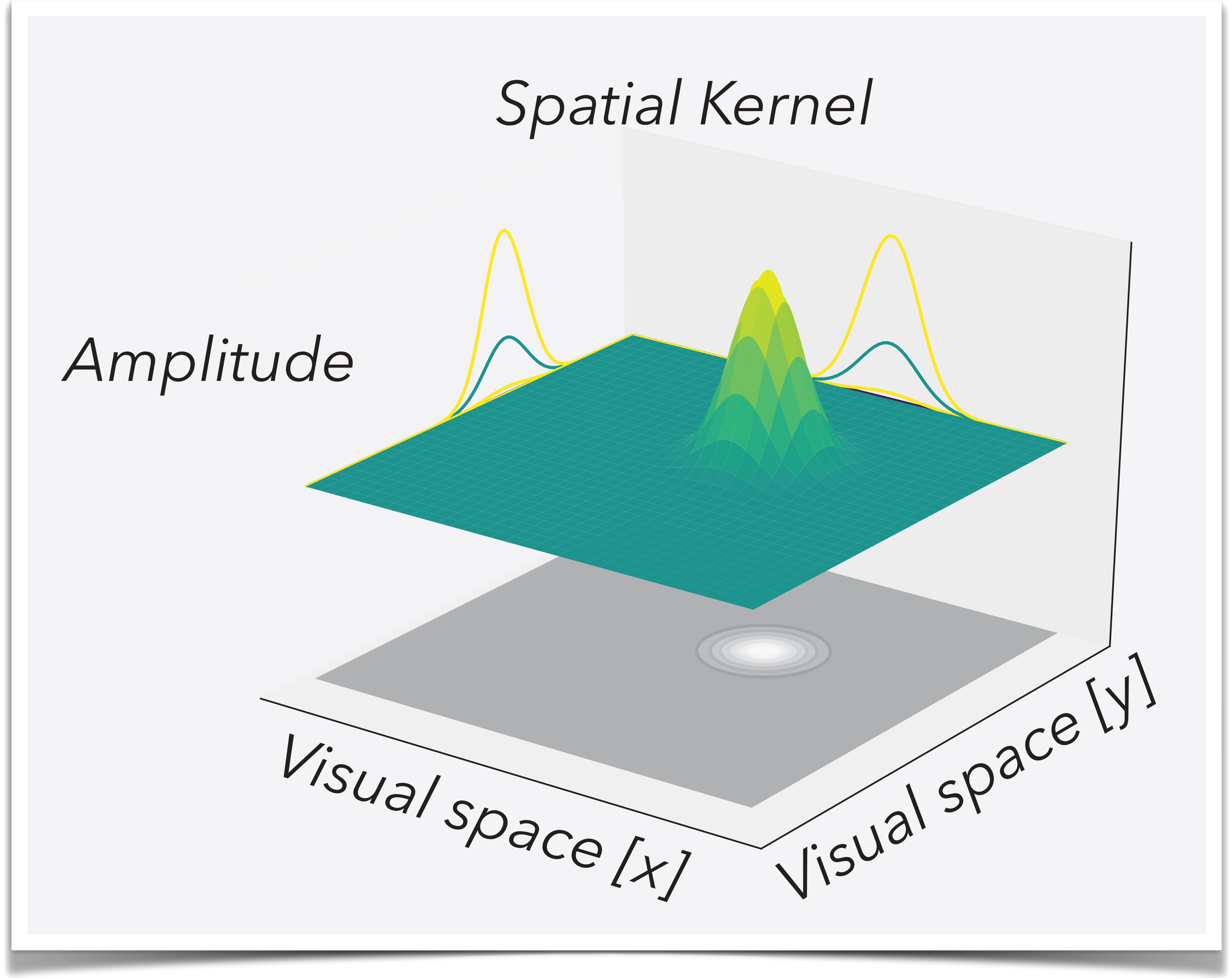





V3d
V2d
V1
V2v
V3v
hV4
VO1
VO2
PCS
IPS4
IPS3
IPS2
IPS1
IPS0
V3A
V3d
V2d
LO1
LO2
TO1
TO2
V3B



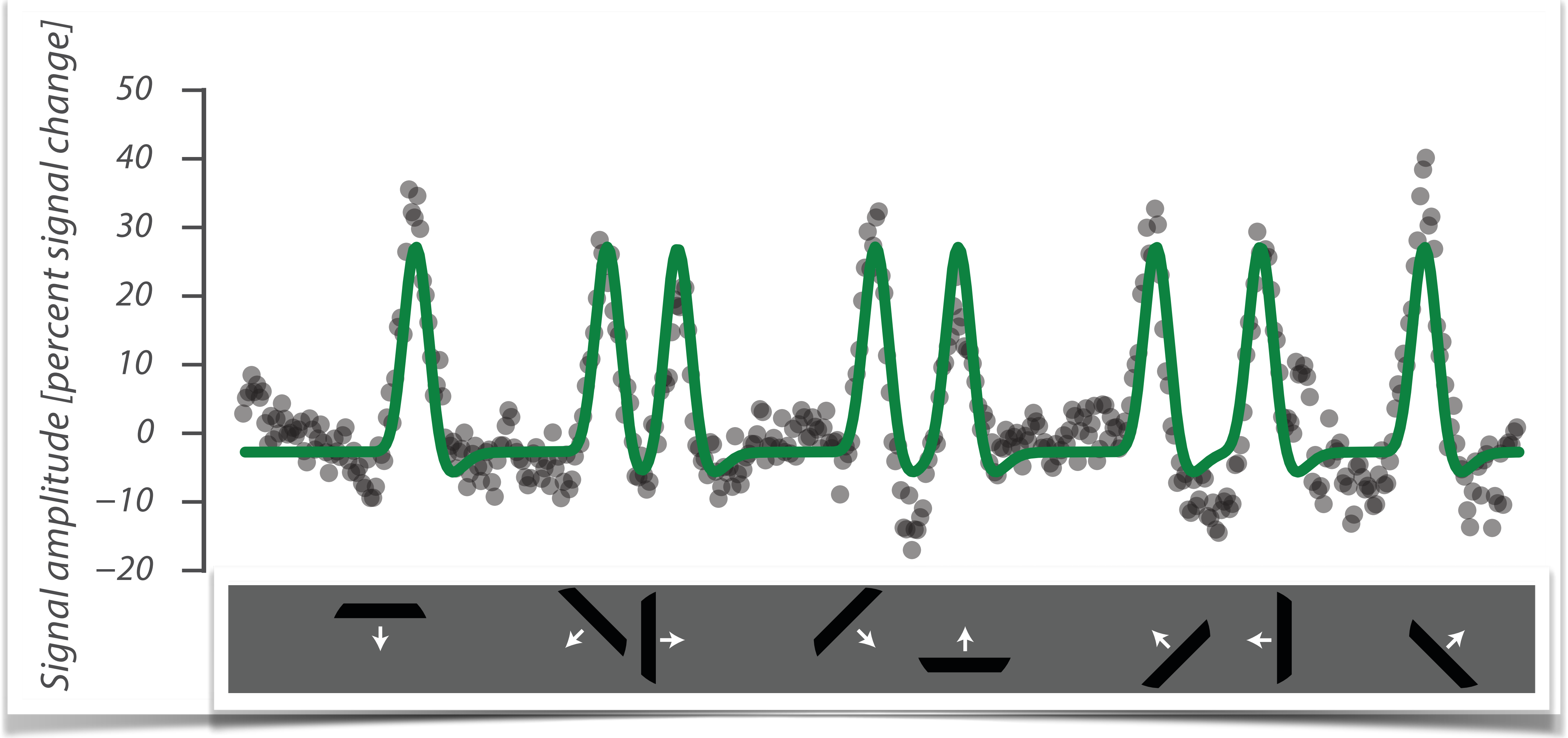
*Structured single-voxel BOLD time course: we can **fit PRF parameters***

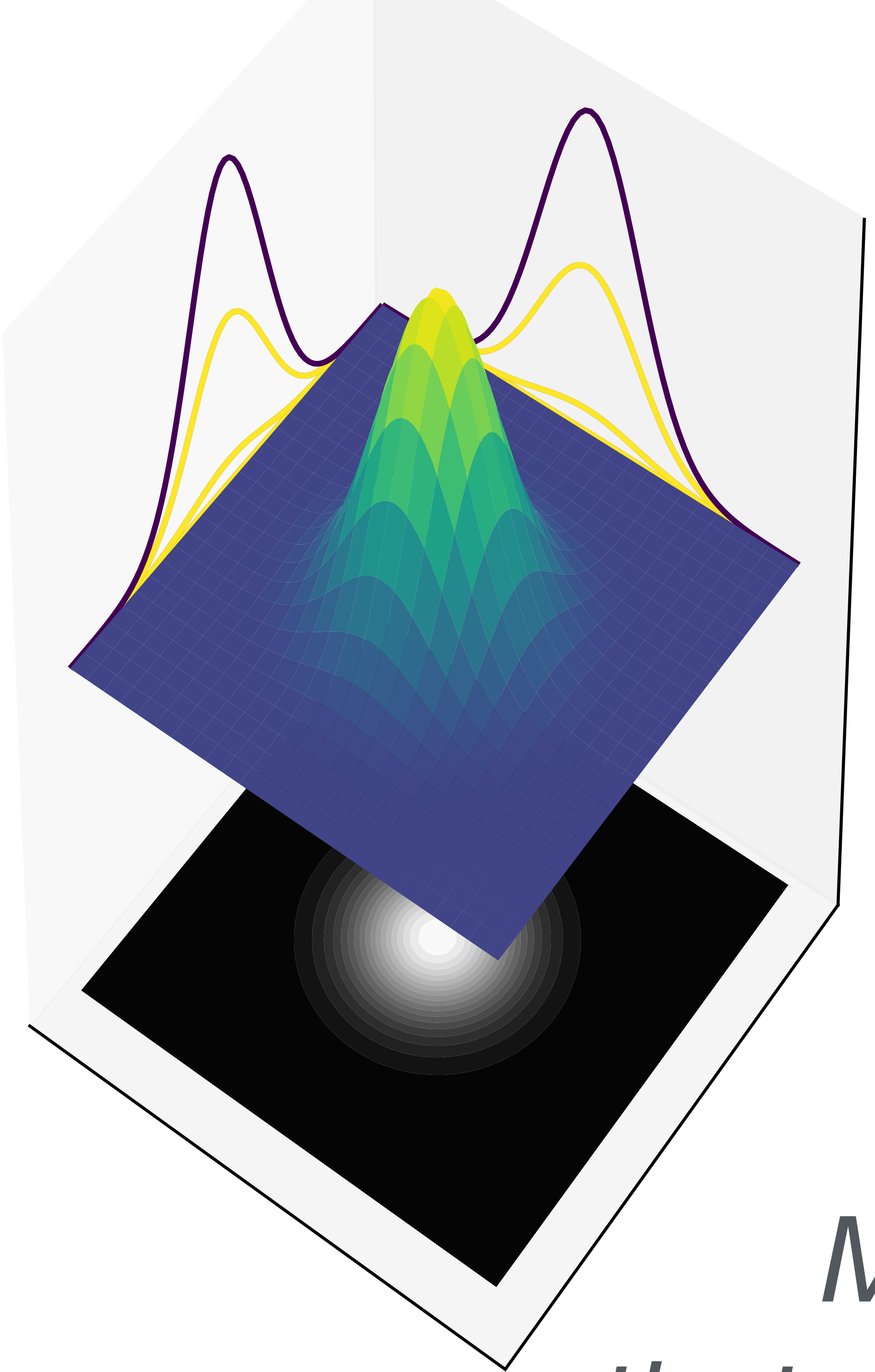


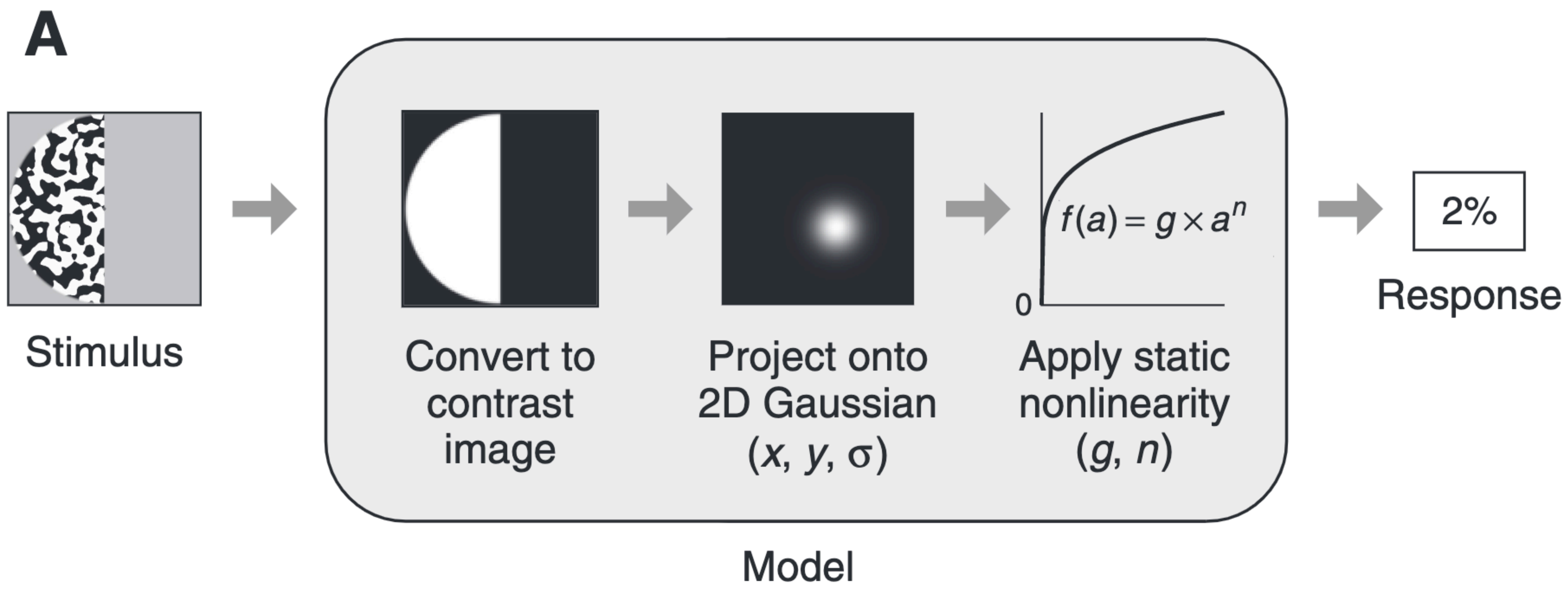
Position

$$g(x_0, y_0, \sigma) = exp(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}),$$

Size

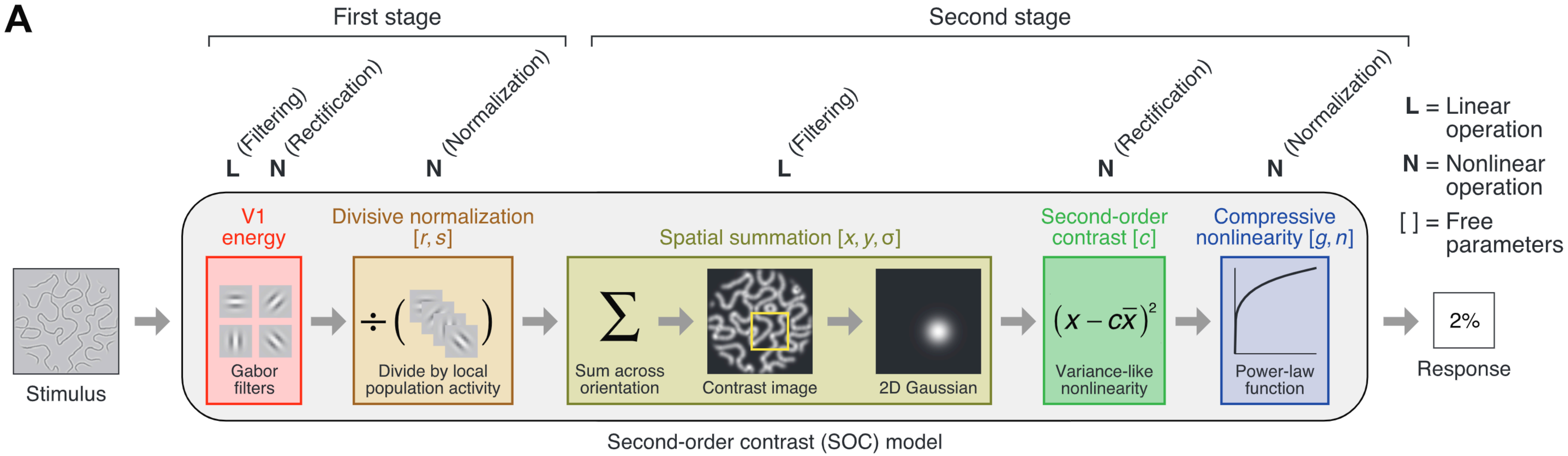






Add "Compressive" Nonlinearity

Make more and more complicated models that explain more and more of the BOLD responses



Decoding from an encoding model

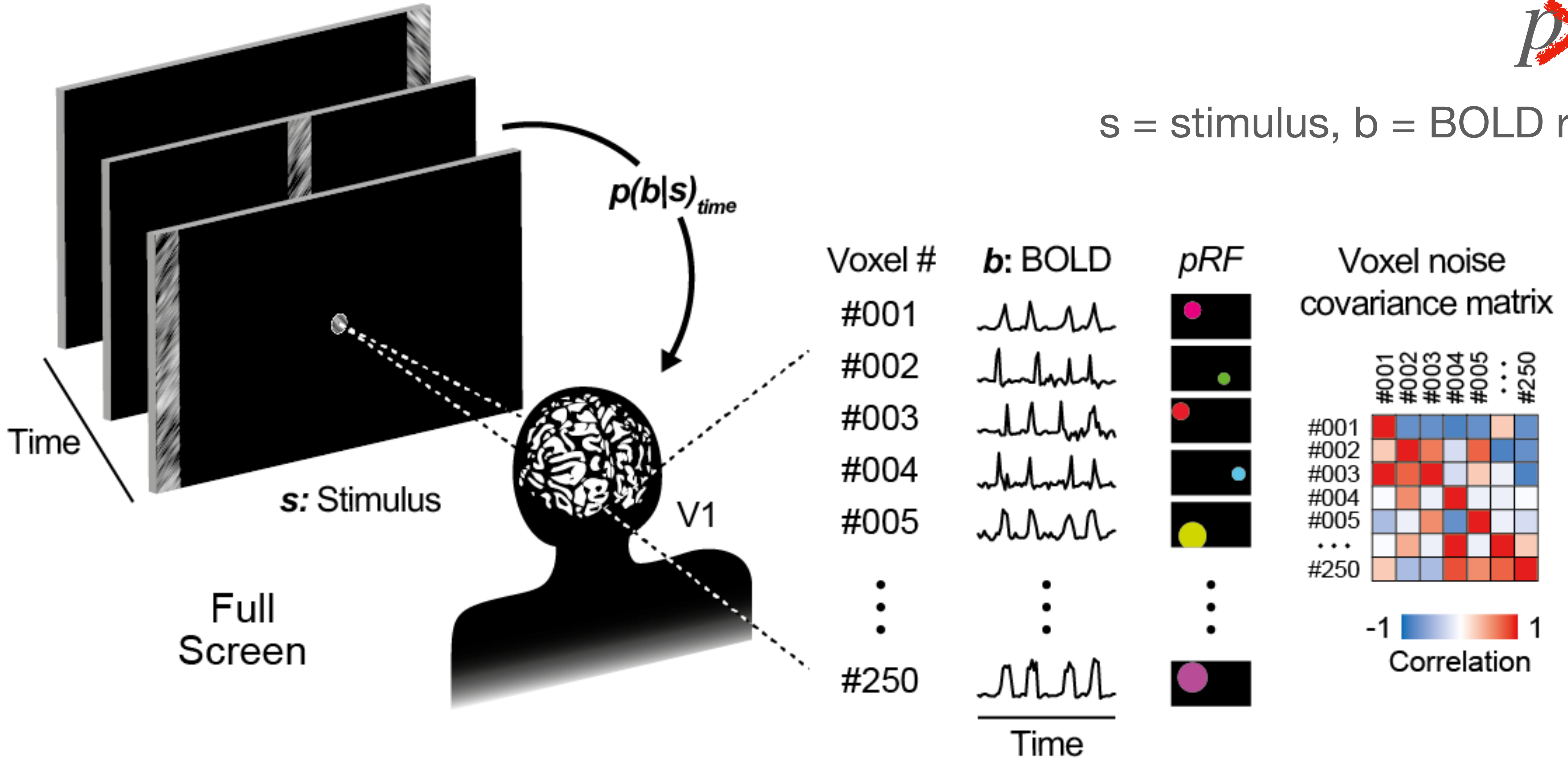
What we want to know, the probability of a stimulus given a BOLD response pattern

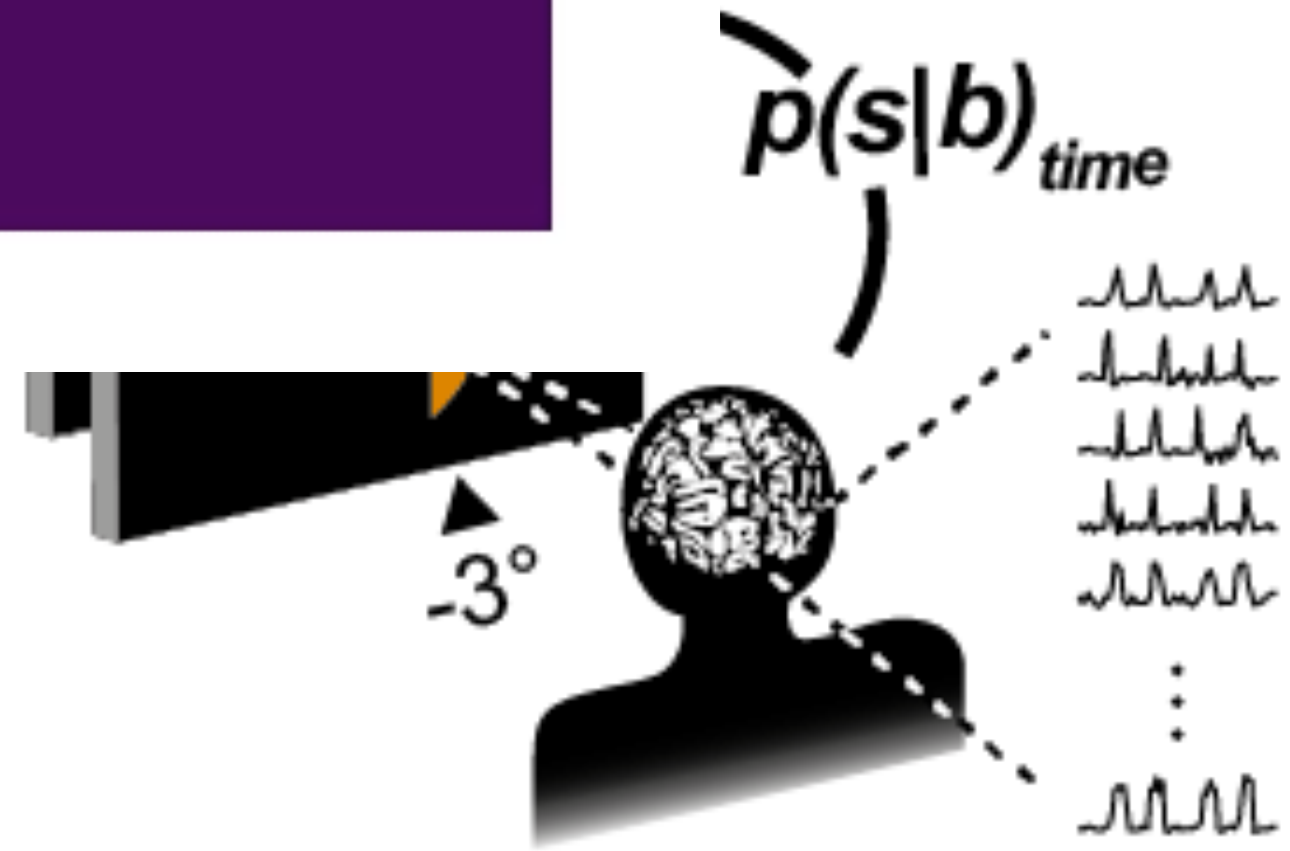
DECODING

We can model/calculate the probability of a BOLD response pattern given a stimulus

$$p(s \mid b) = \frac{p(b \mid s)\,\cancel{p(s)}}{\cancel{p(b)}}$$

s = stimulus, b = BOLD response pattern





Encoding

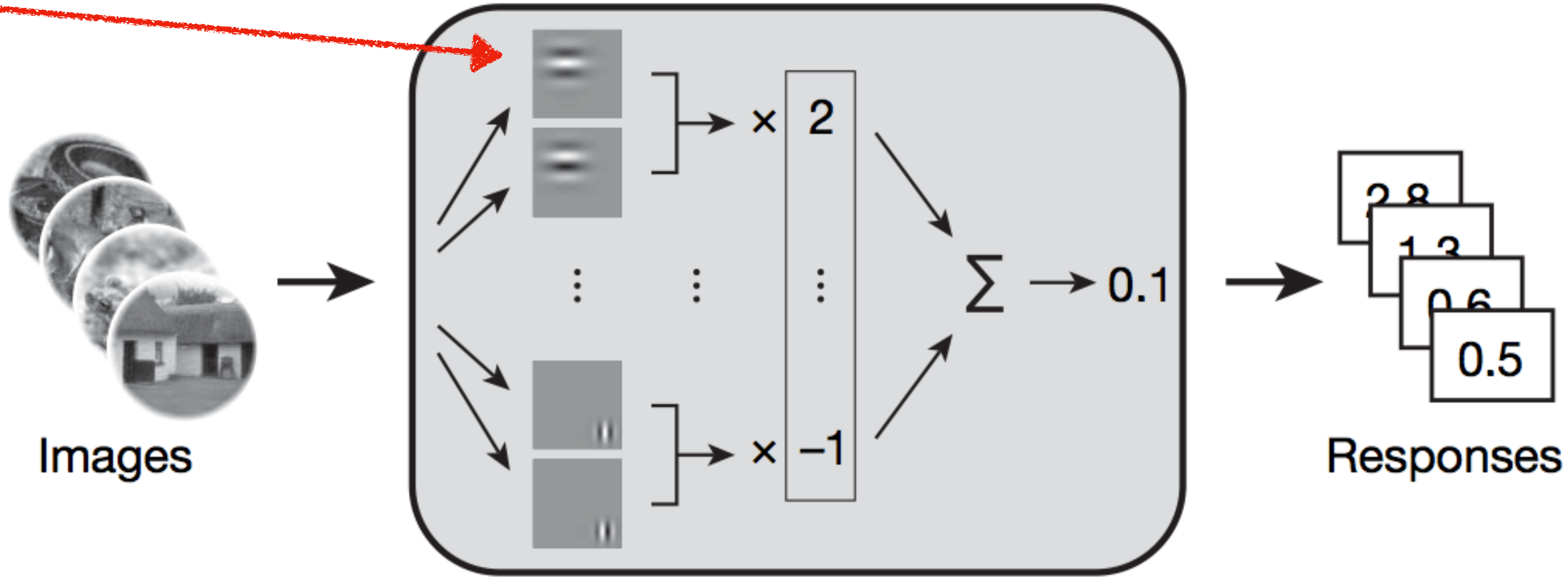
What is the ‘receptive field’ of a voxel?

More complex receptive fields

Stage 1: model estimation

Estimate a receptive-field model for each voxel

Many different models and their predictions

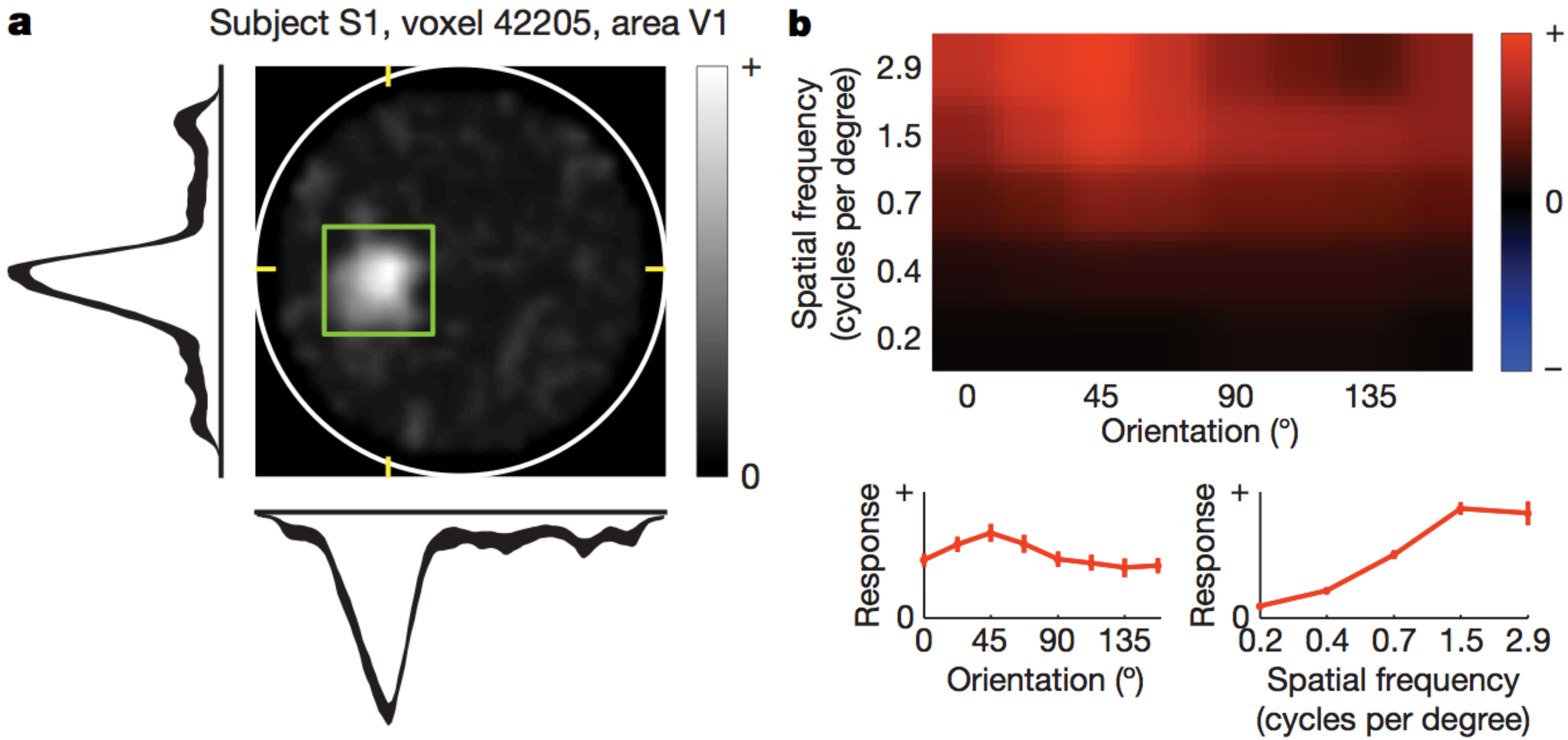


Receptive-field model for one voxel

Estimate a huge amount of parameters, using ***penalized regression***

Encoding

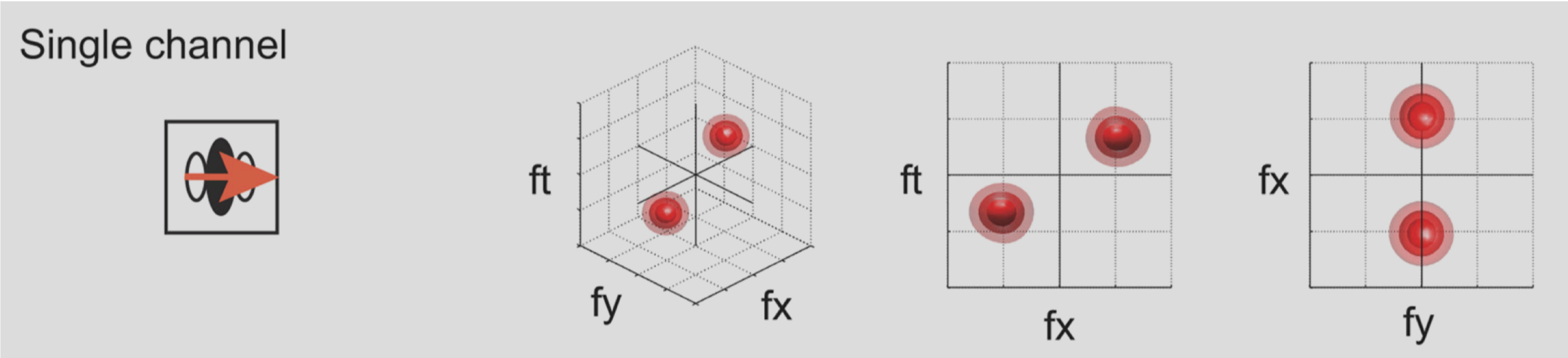
What is the ‘receptive field’ of a voxel?



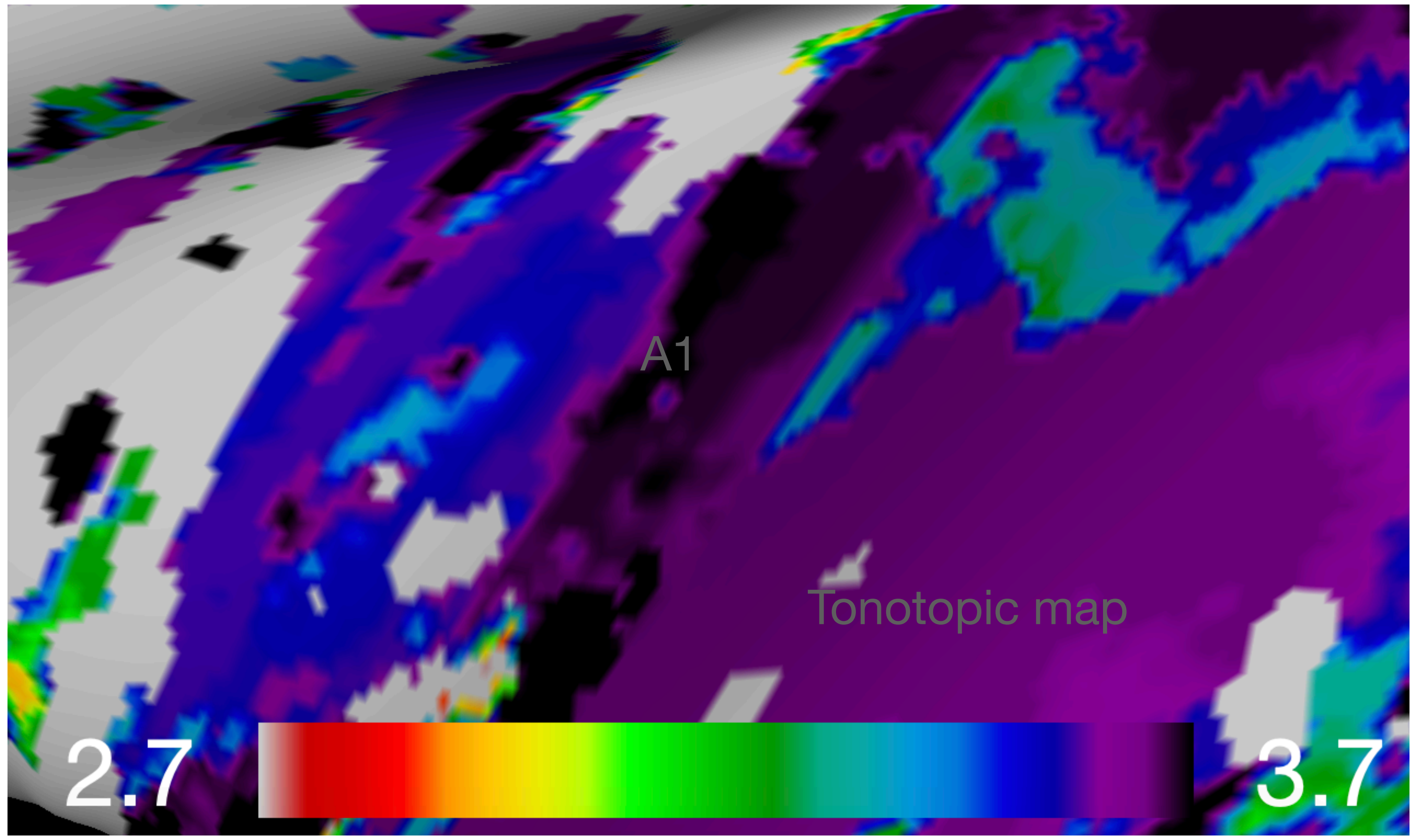
Extend this to movies: Add spatiotemporal receptive field structure

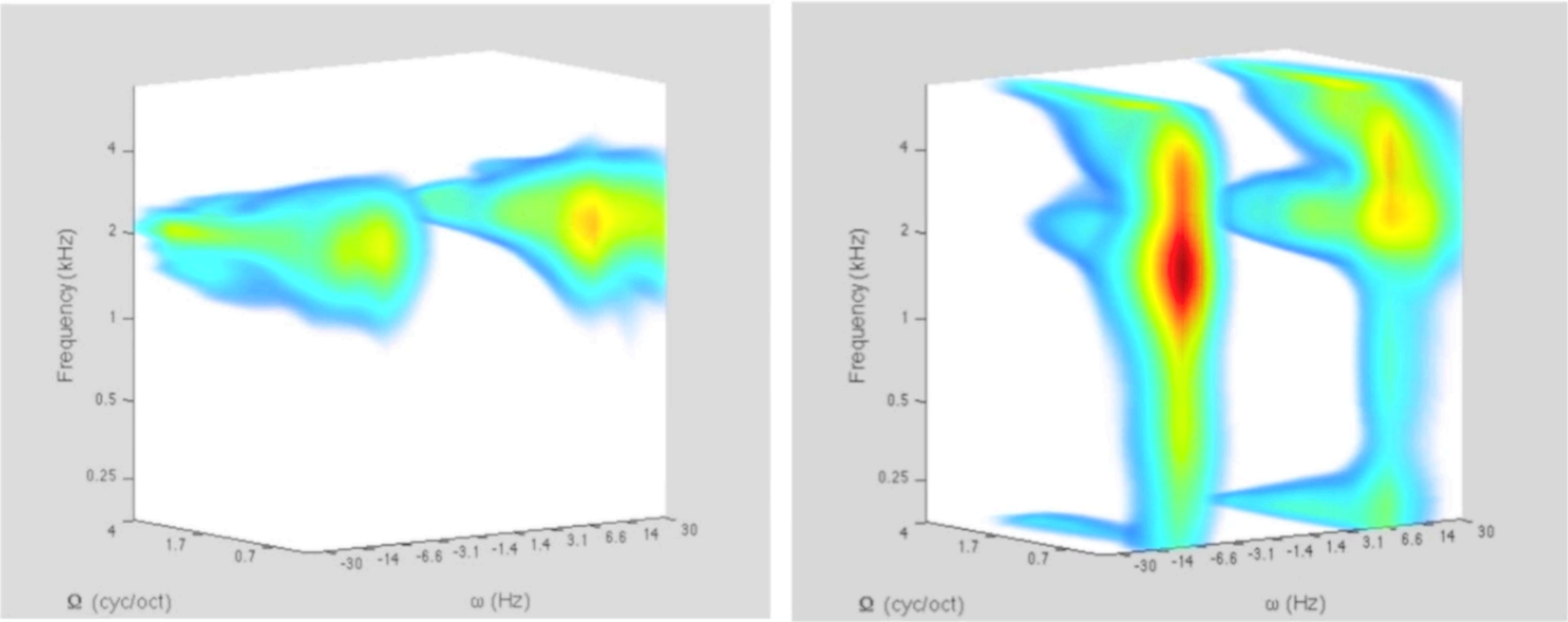
Presented clip

Clip reconstructed from brain activity



And, this works also for sounds

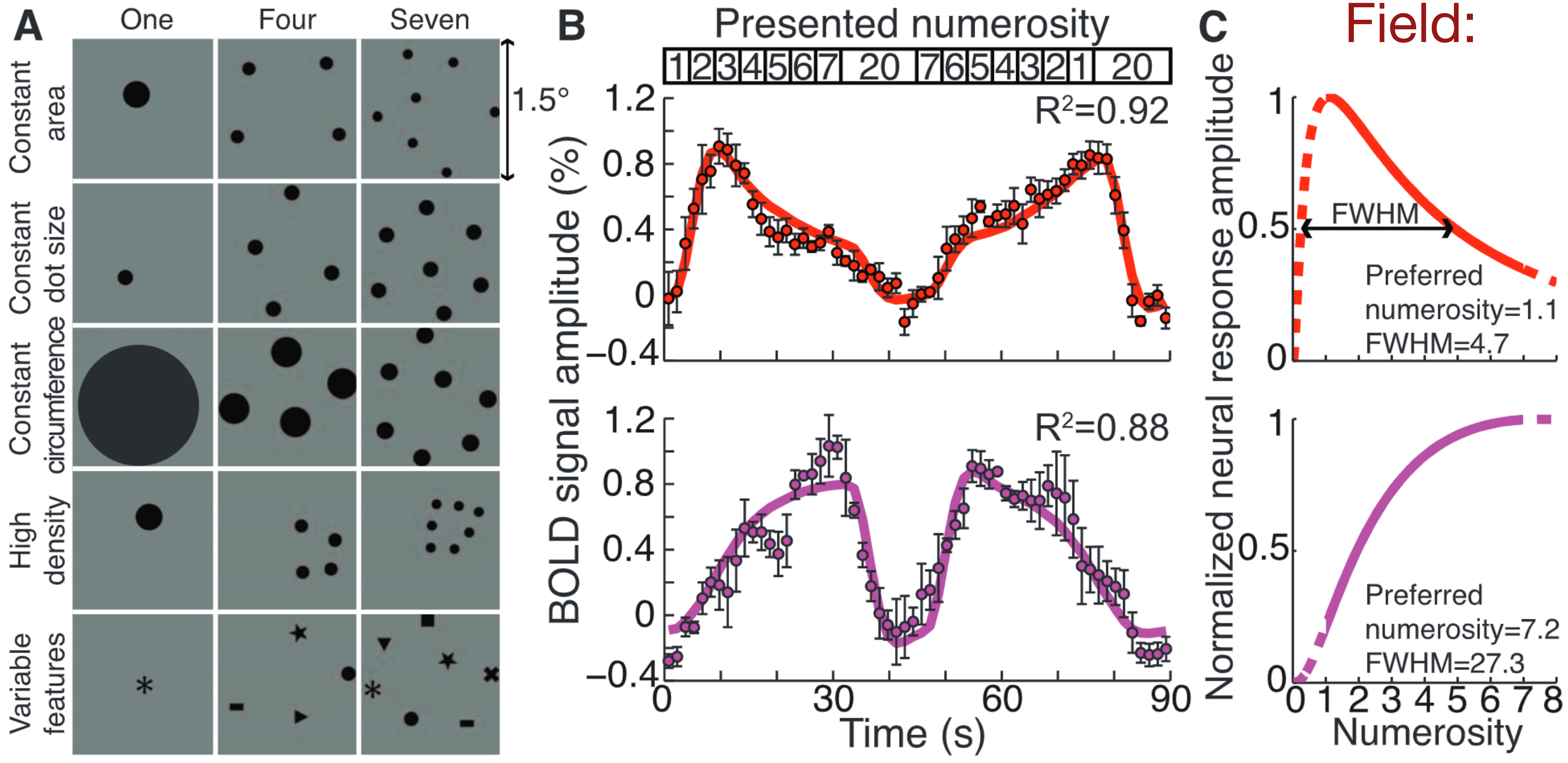




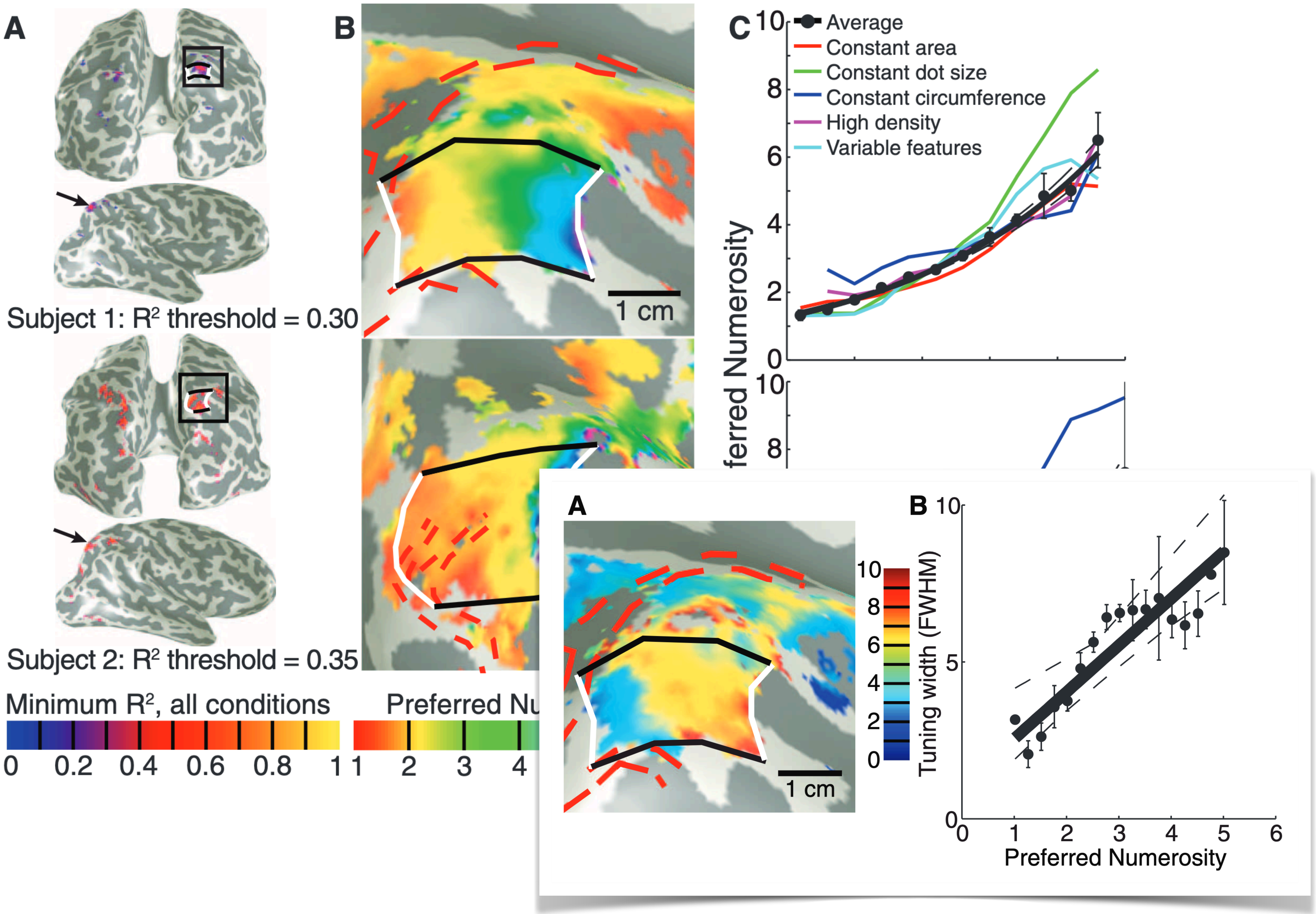
But more…

Also more abstract dimensions are represented using receptive fields, found using fMRI

Receptive Field:

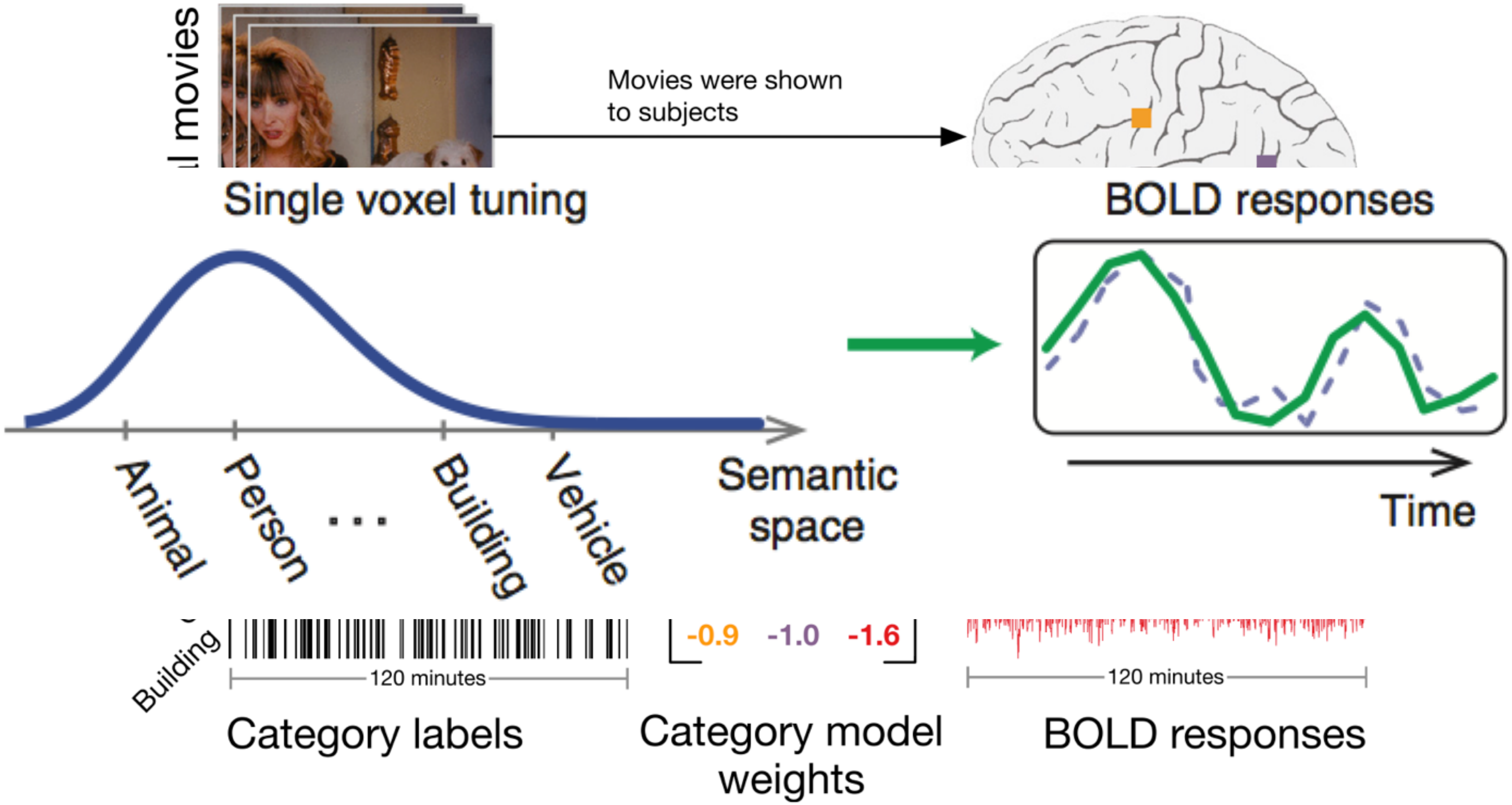


These receptive fields are layed out on maps…



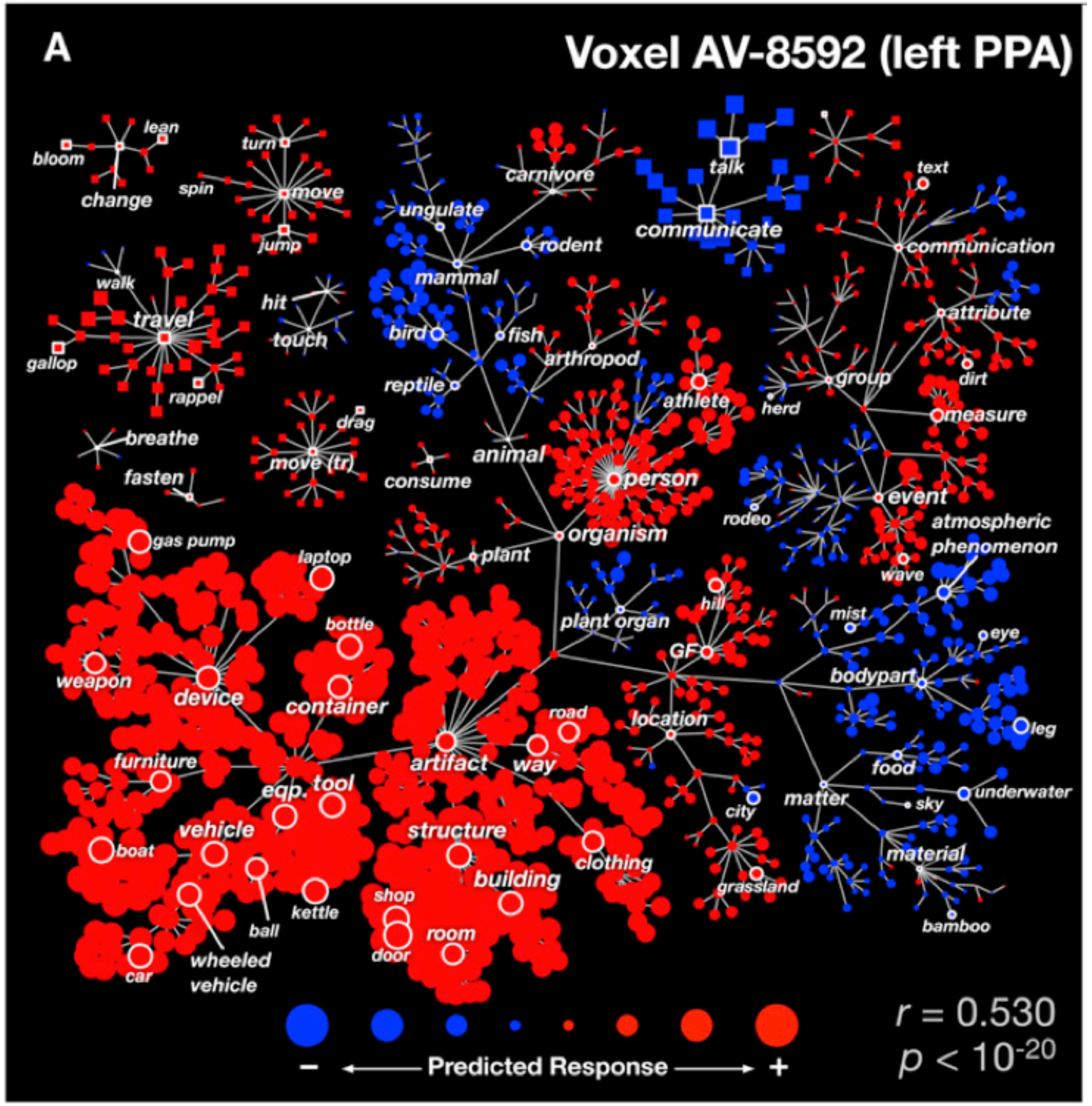
Encoding in higher spaces

A voxel's receptive field doesn't have to be retinotopic, but can live in a more abstract (semantic) space.

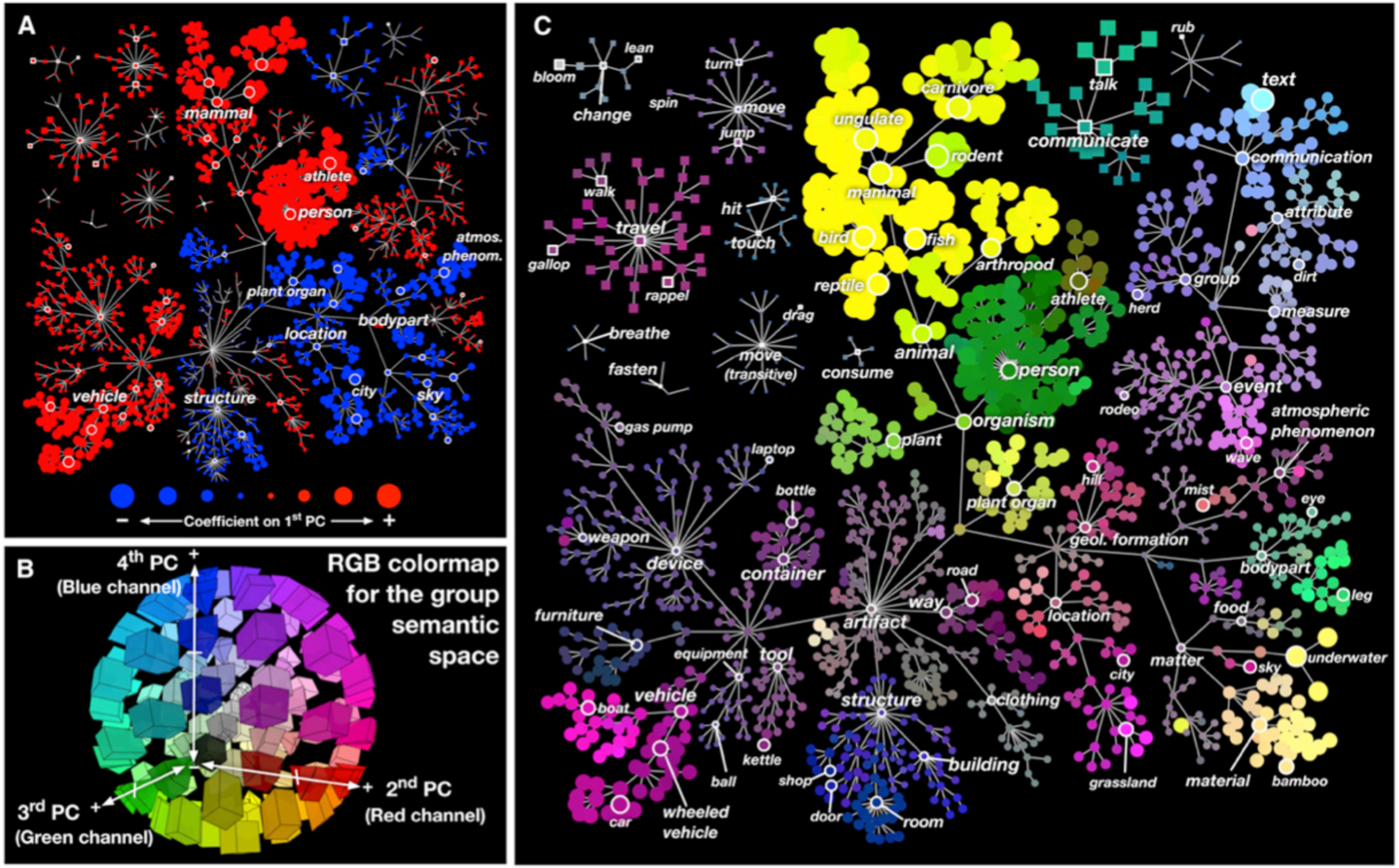


Methods

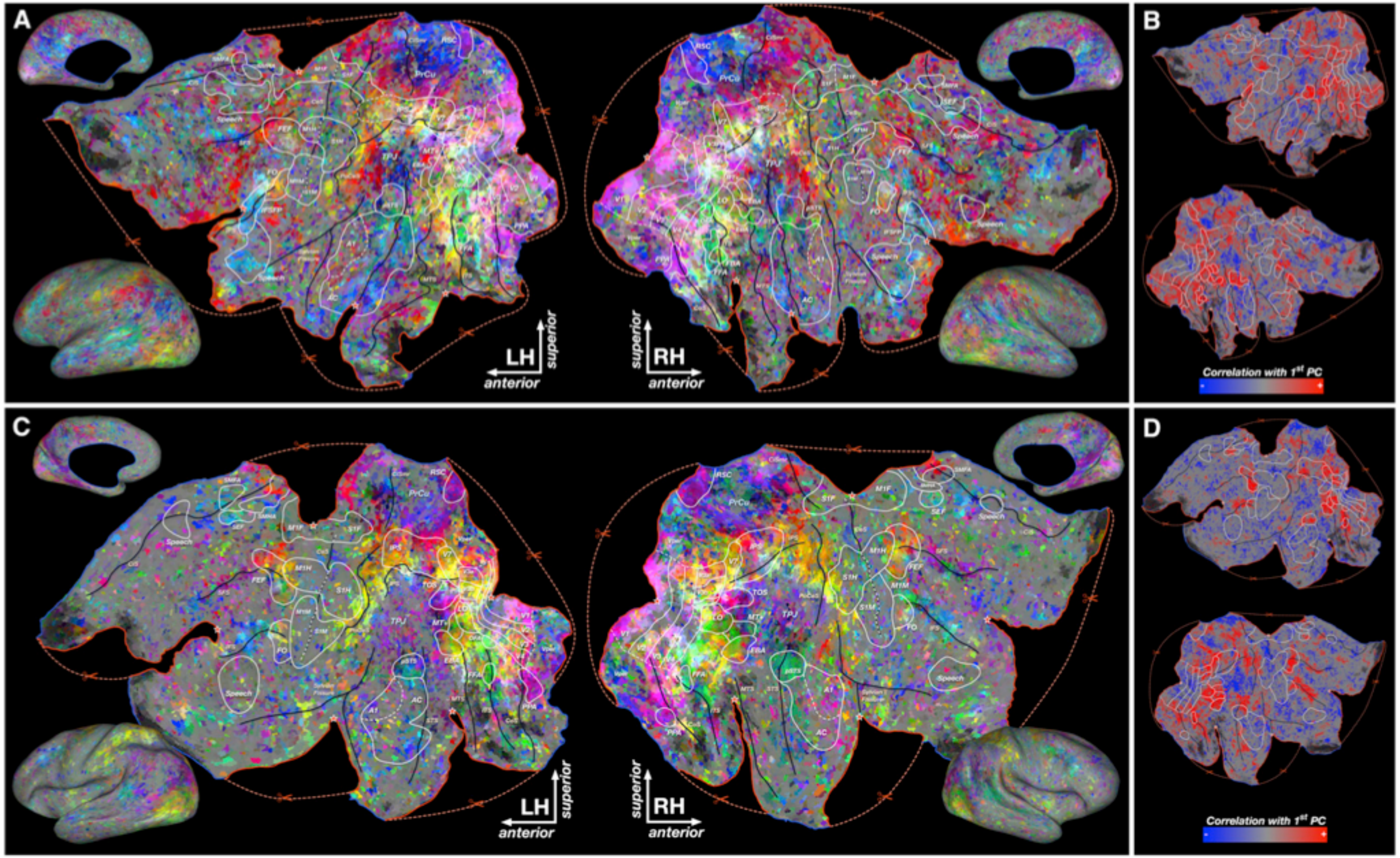
Single-voxel tuning curve in a semantic space



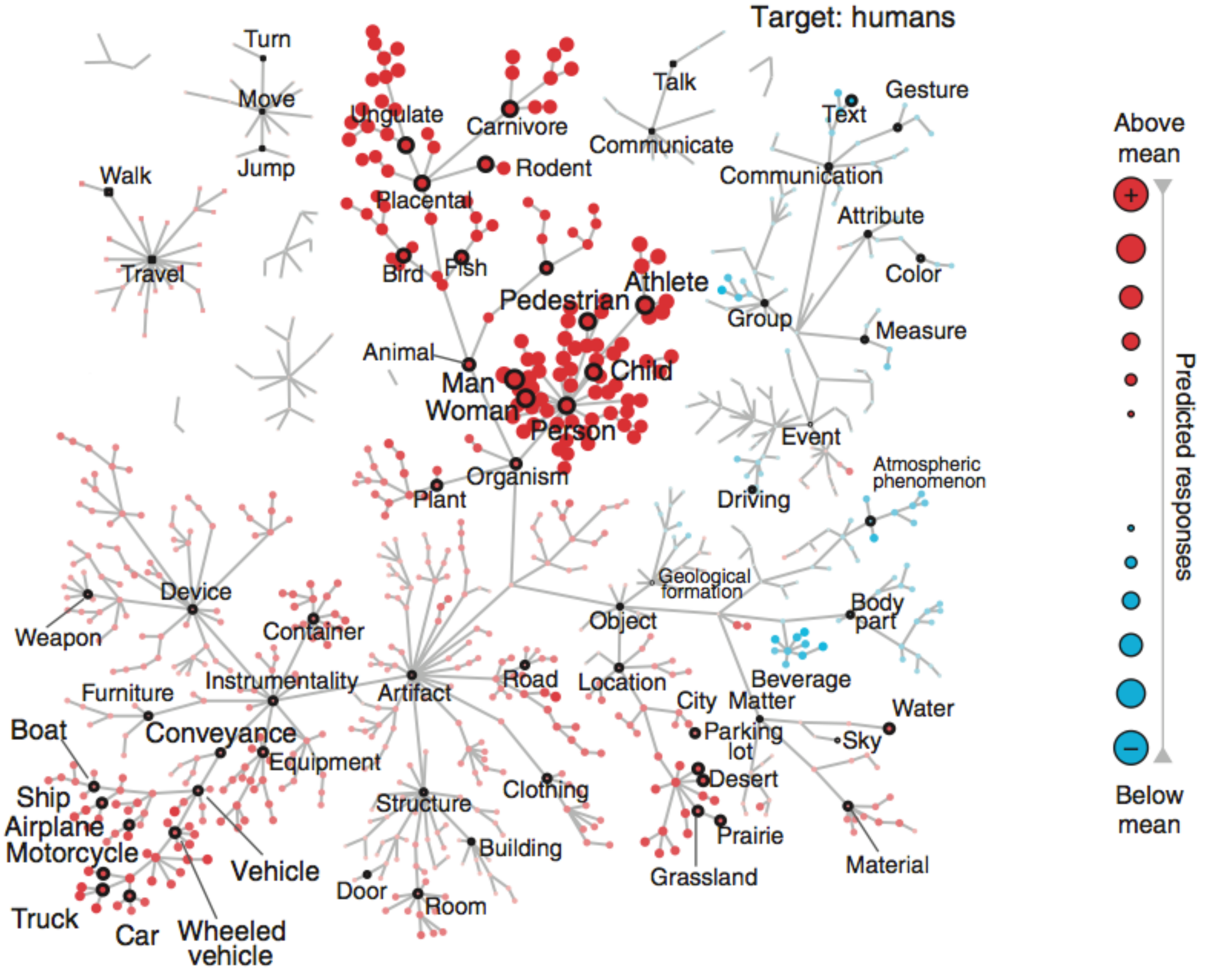
- Doing … f voxels by reg…
- Dimer…
- Projec…



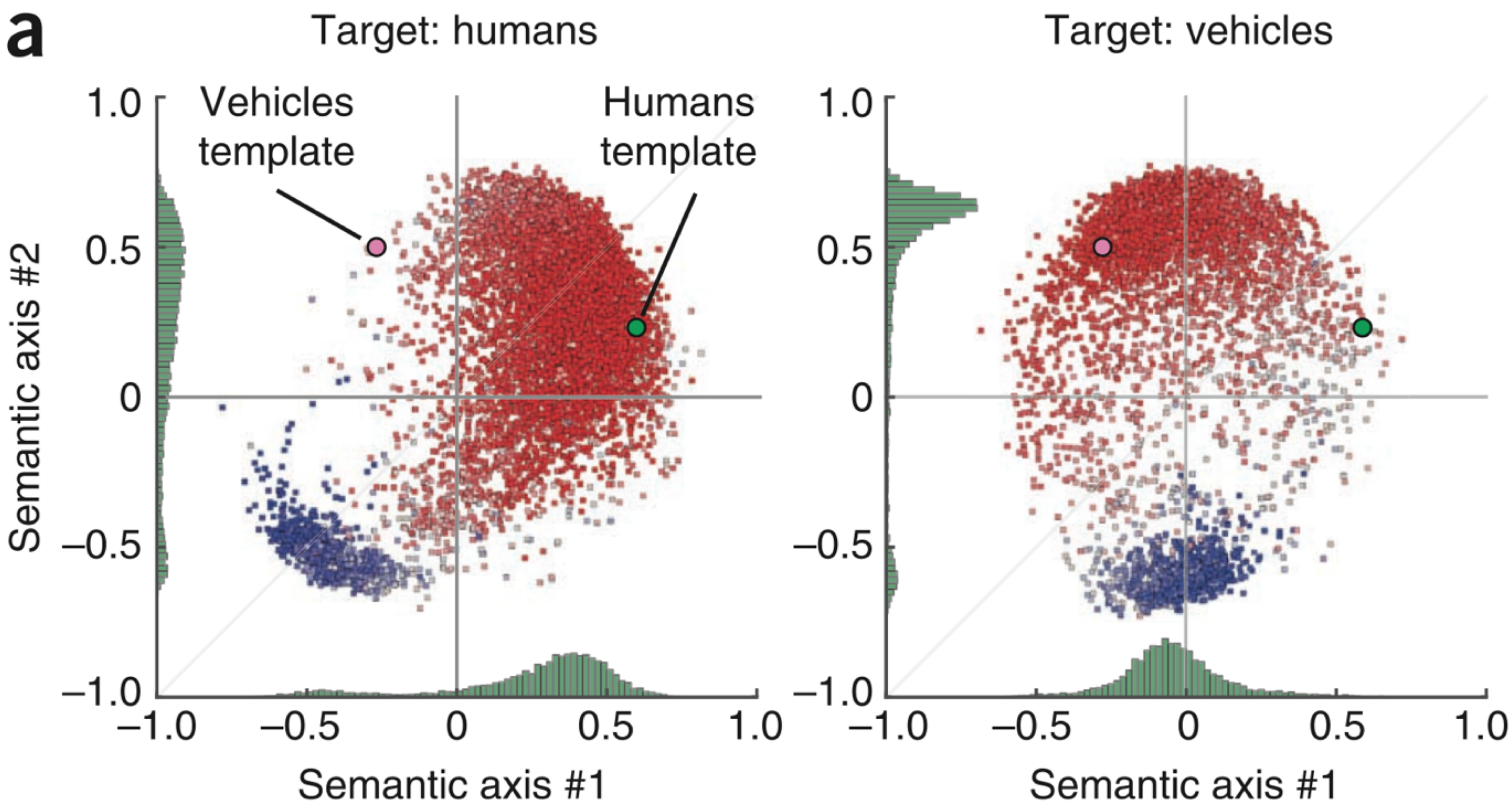
- Project



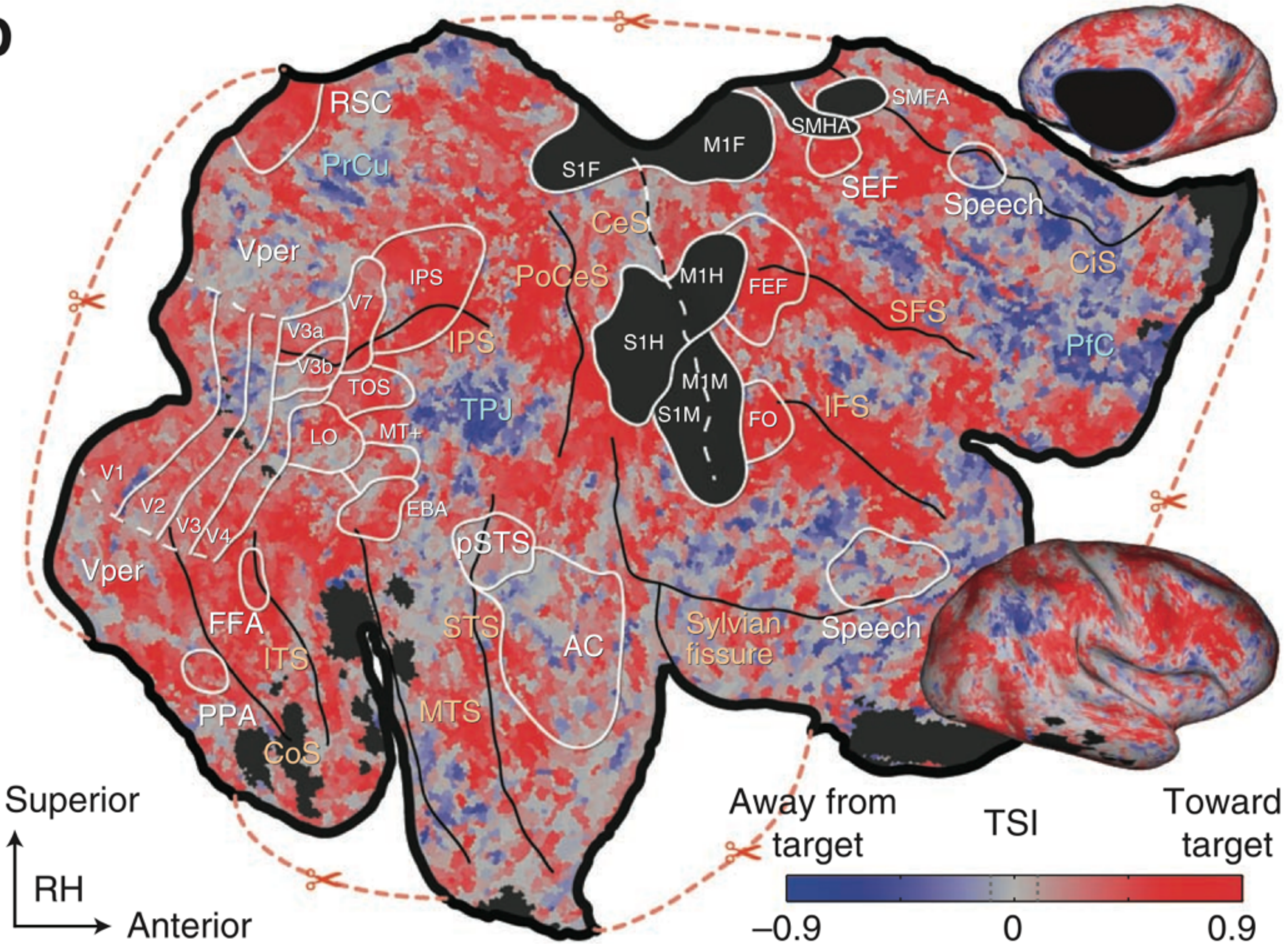
Attention shifts tuning of voxels



Attention shifts tuning of voxels

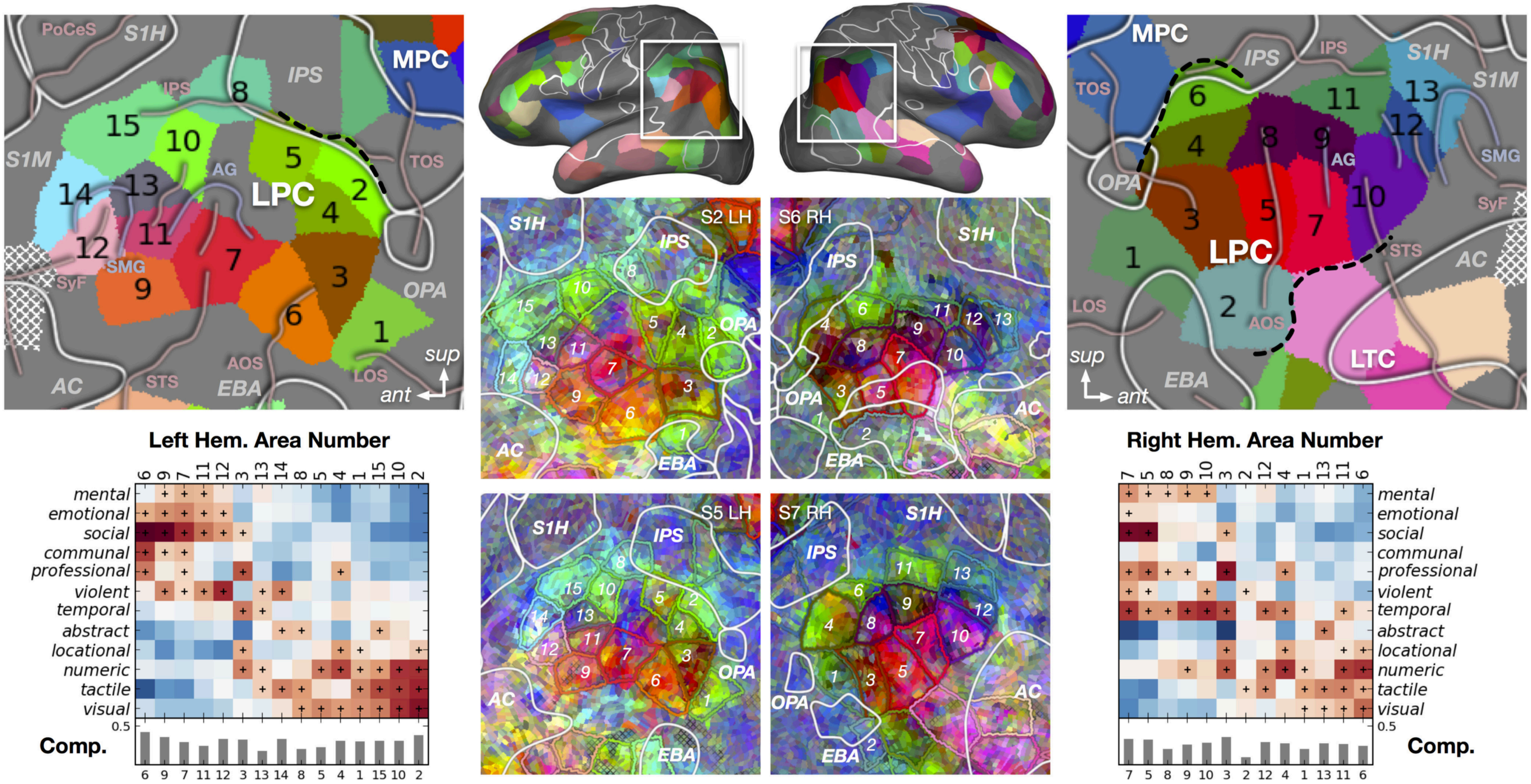


b

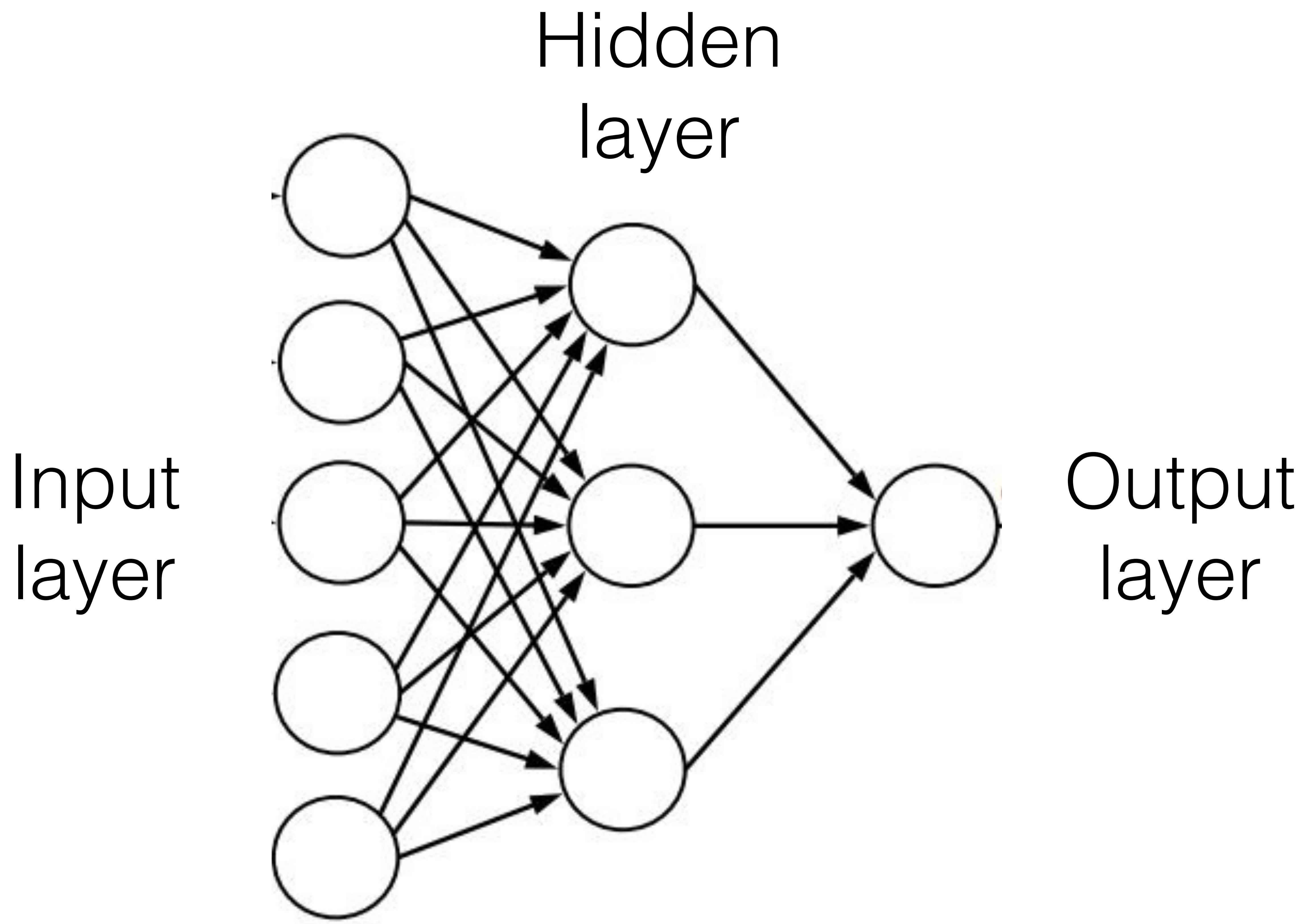


RSC
PrCu
Vper
IPS
V7
V3a
V3b
TOS
IPS
TPJ
LO
MT+
V1
V2
V3
V4
EBA
pSTS
Vper
FFA
ITS
PPA
CoS
STS
MTS
AC
S1F
M1F
SMFA
SMHA
SEF
Speech
CeS
PoCeS
M1H
FEF
S1H
M1M
S1M
FO
SFS
IFS
CiS
PfC
Sylvian fissure
Speech
Superior
RH
Anterior
Away from target
TSI
Toward target
–0.9
0
0.9

Create atlases based on information-processing preferences.



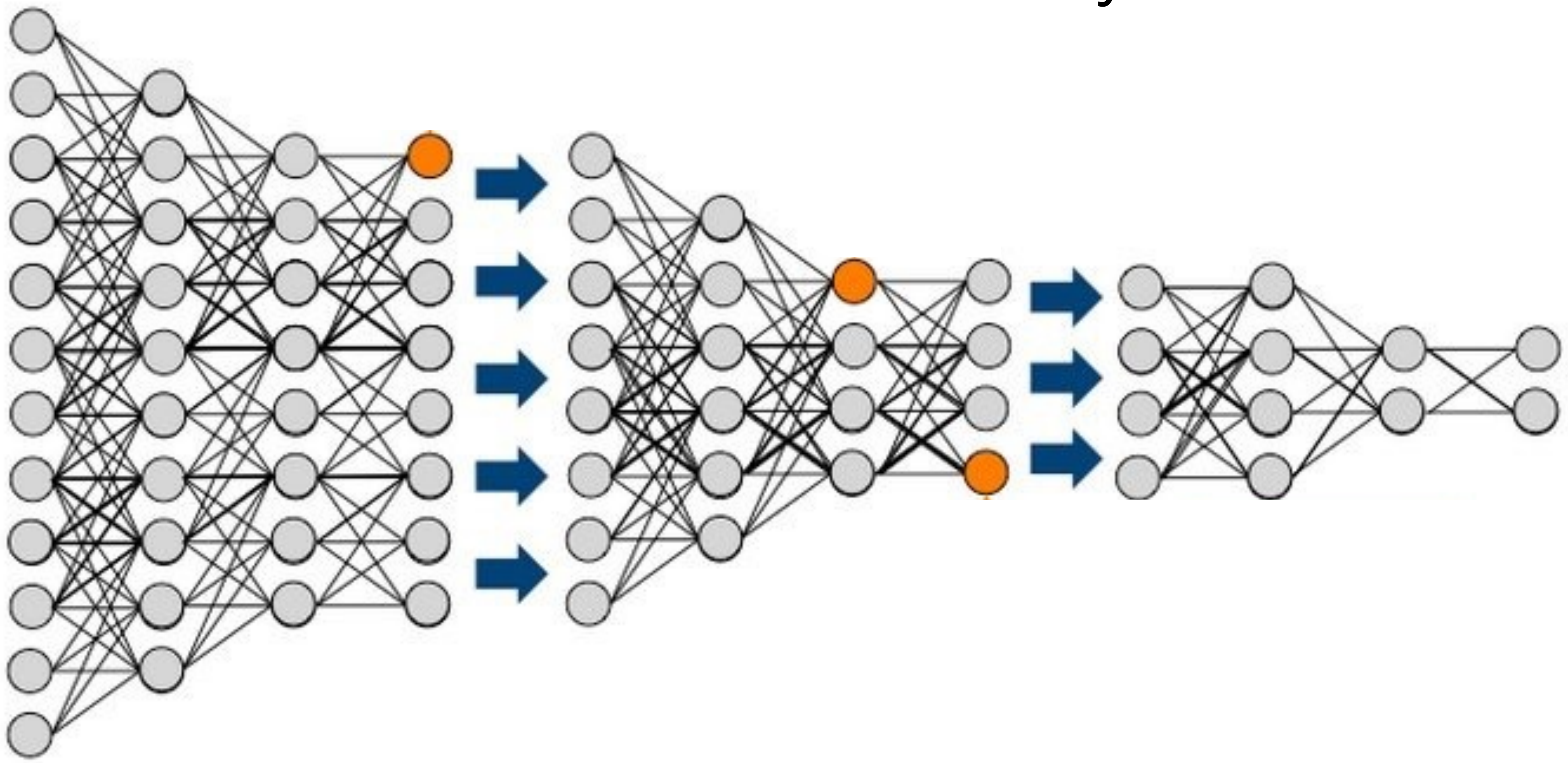
Neural Networks



Convolutional Neural Network (CNN)

Number of hidden layers

Input layer

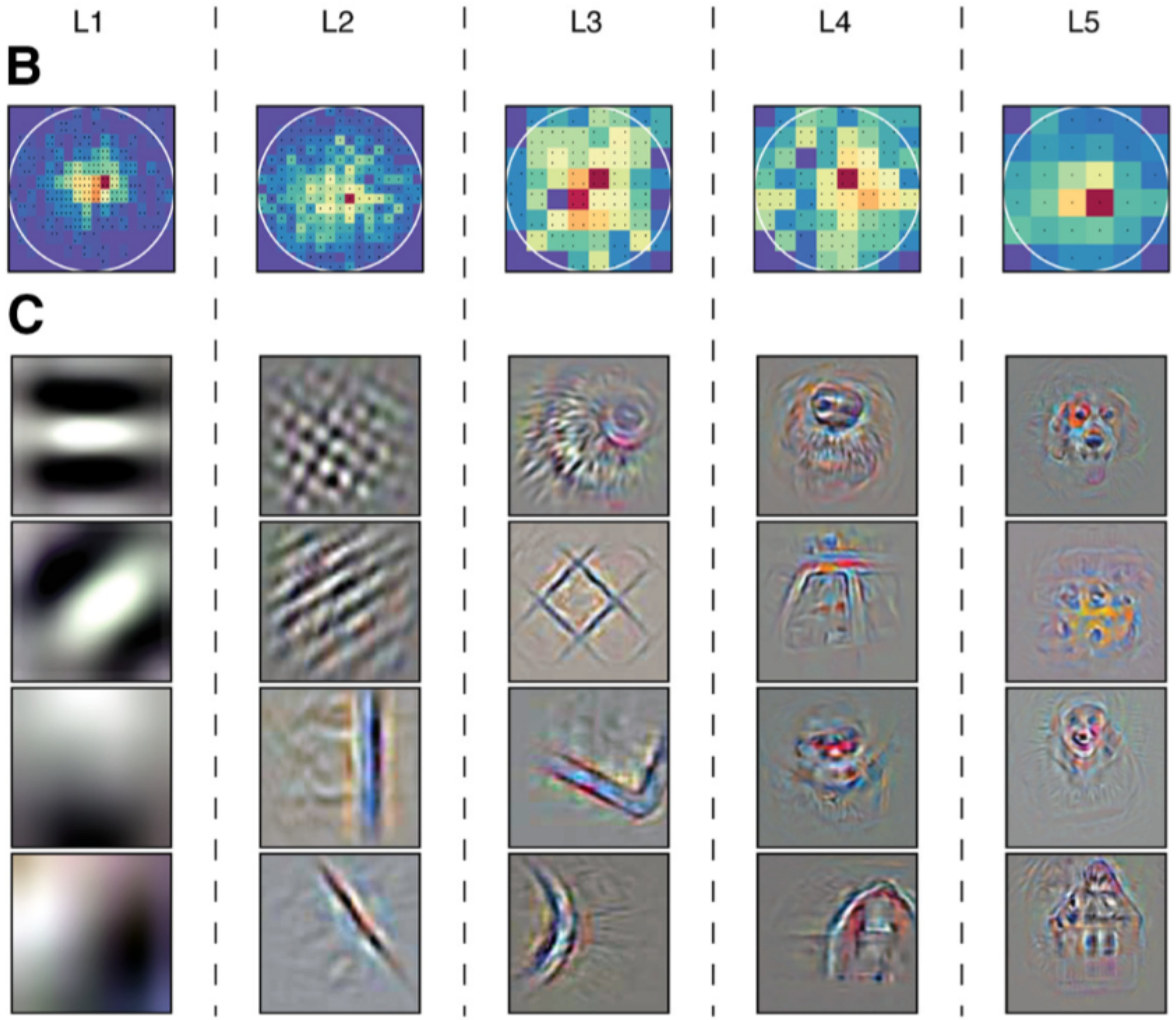


Output layer

Object recognition: CNN

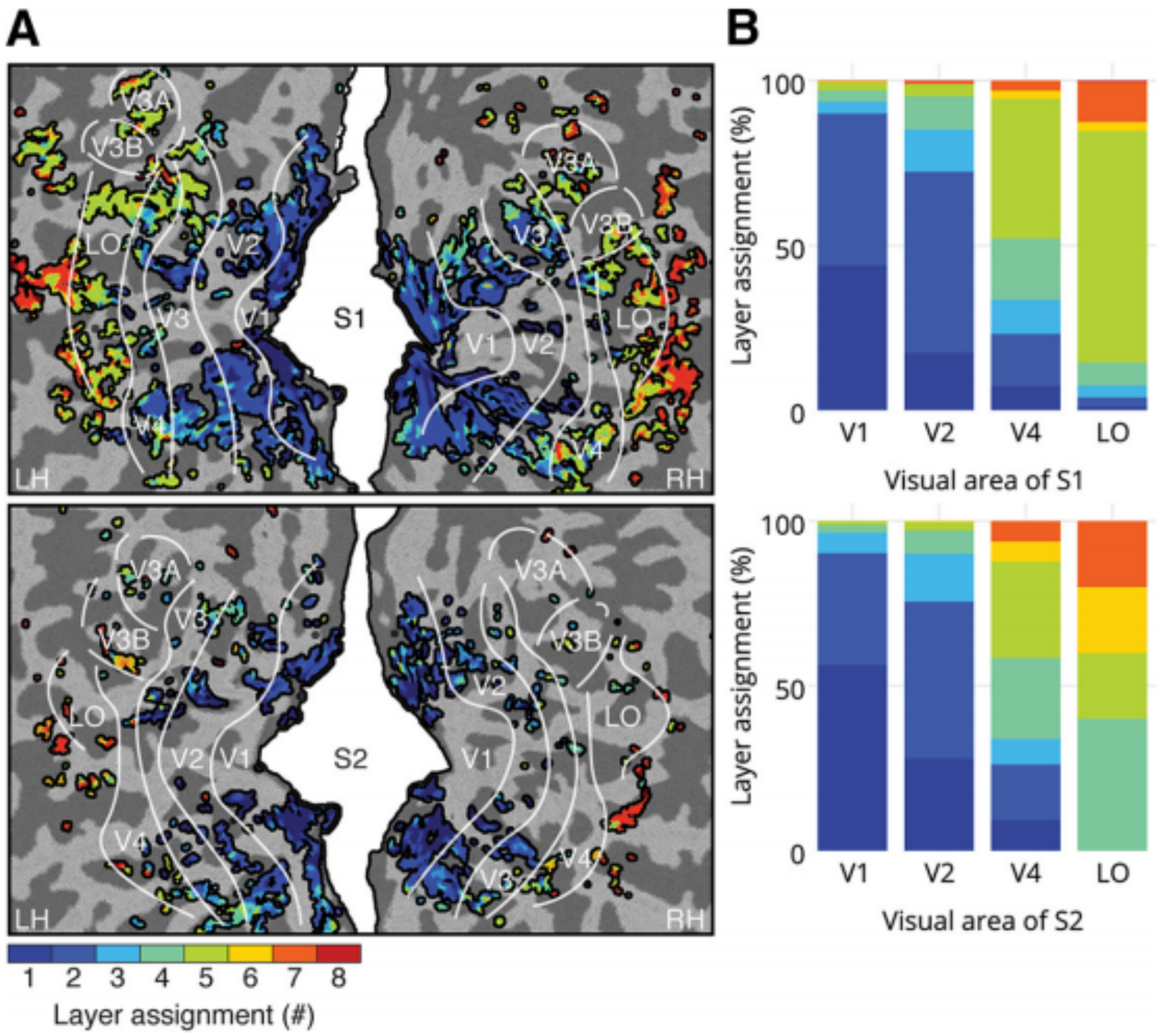
- CNNs are trained on millions of images (ImageNet).

CNN as encoding model



Güçlü U, van Gerven MAJ. 2015

Neural Networks as encoding model



Güçlü U, van Gerven MAJ. 2015

the use of encoding models for neuroscientists

1. compare different encoding models

2. explore model behaviour under different conditions

3. use fMRI voxels as ‘electrodes’,
perform physiology

Encoding models as a basis for decoding

Although principally focused on the single voxel's responses, doing this encoding model estimation for all voxels allows one to investigate the patterns of activation across brain regions.

But now, because we know what single-voxel activations mean, these patterns are informative regarding how the brain as a whole treats information.