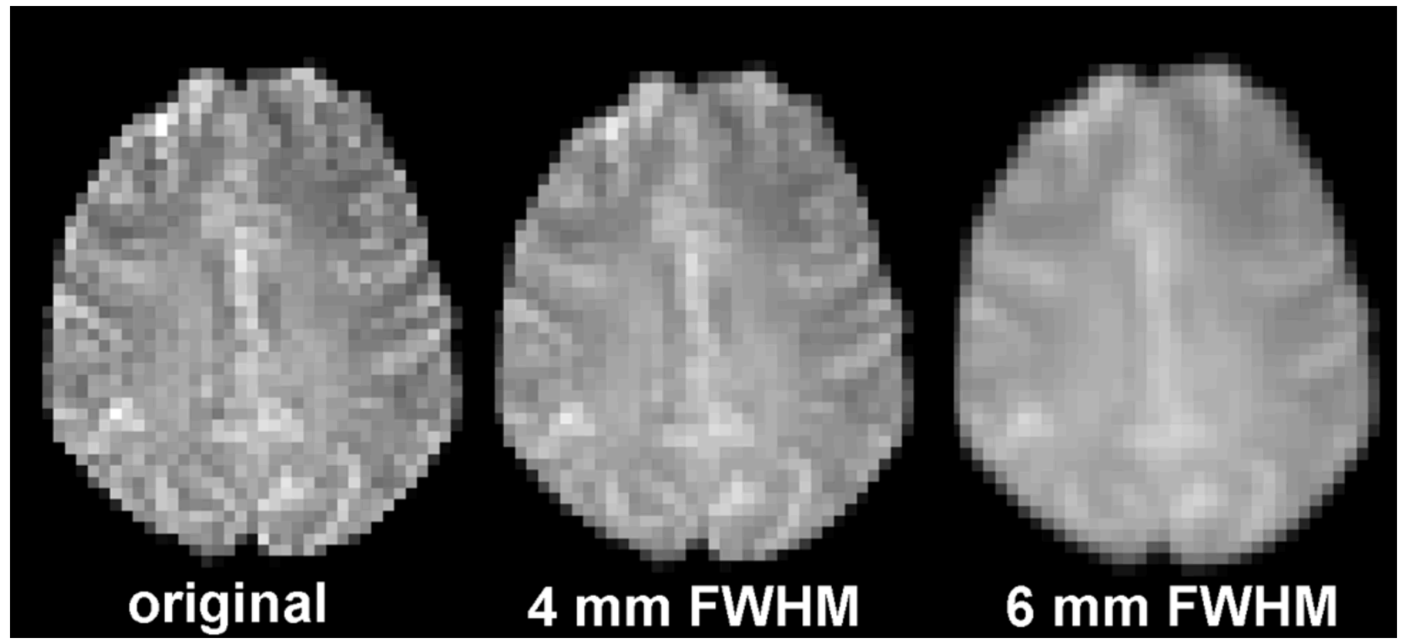
Encoding & Decoding in fMRI

Brain Imaging 2024/5

What does fMRI allow us to say?

What does fMRI allow us to say?

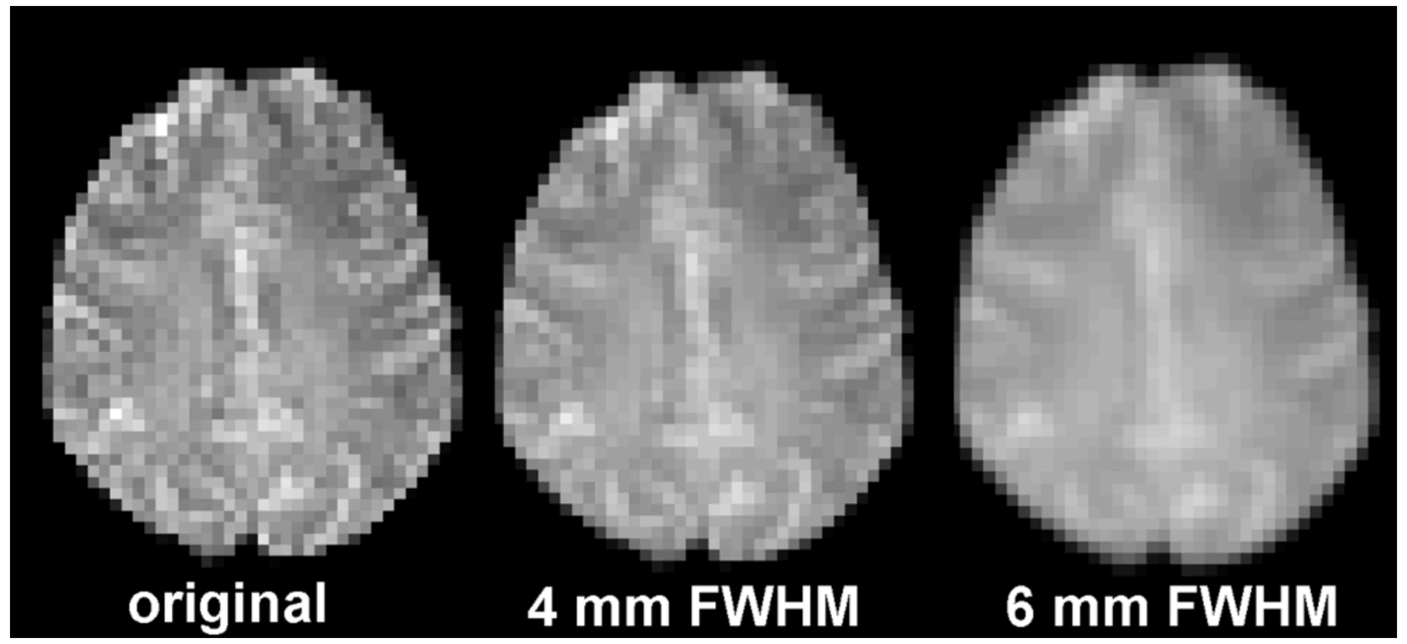


- Historically, we had bad recordings, so smoothing will increased SNR and helped us get results.

What does fMRI allow us to say?

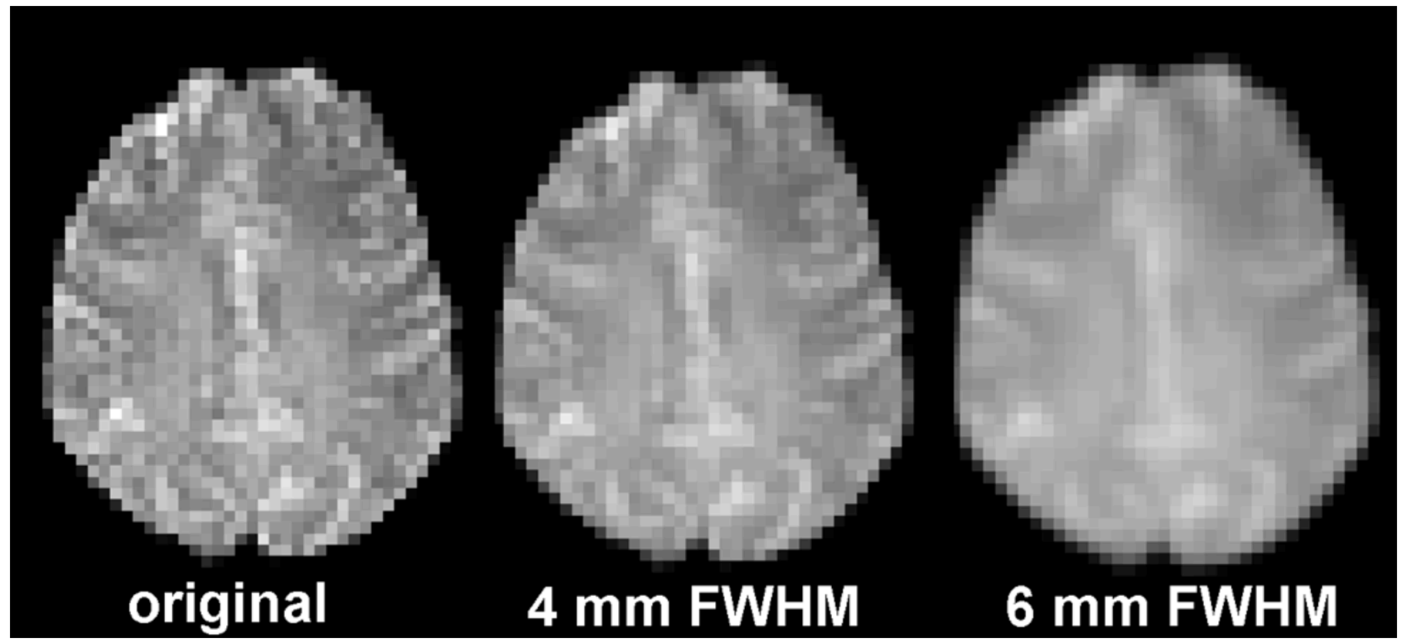
- Historically, we had bad recordings, so smoothing will increased SNR and helped us get results.

- ***BUT: won't everything necessarily look like a BLOB?*** *We're throwing away a lot of information….*



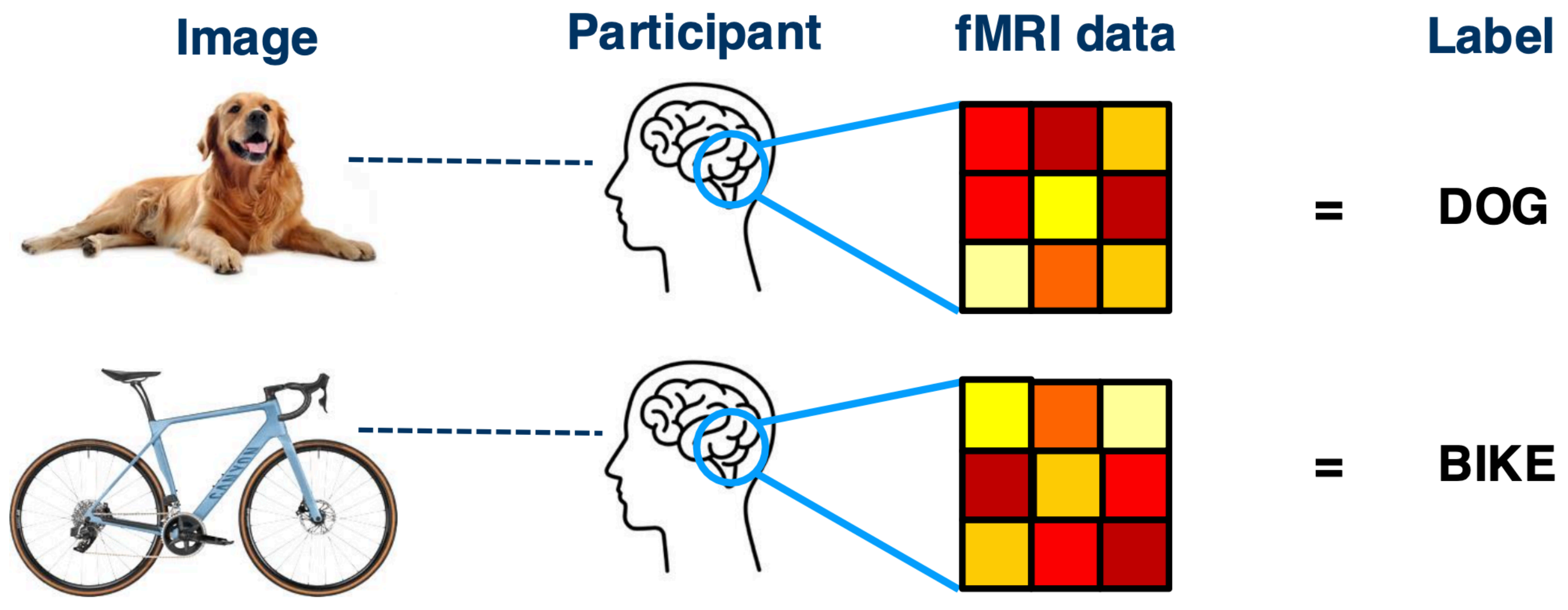
What does fMRI allow us to say?

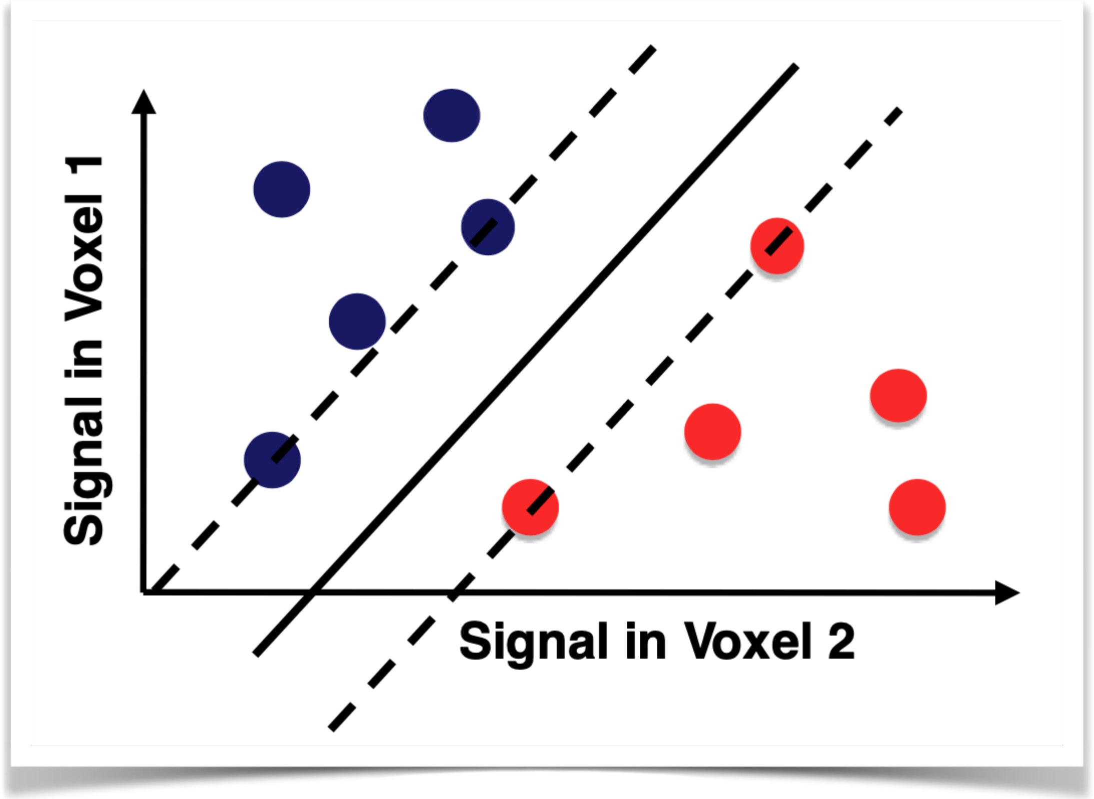
- Historically, we had bad recordings, so smoothing will increased SNR and helped us get results.



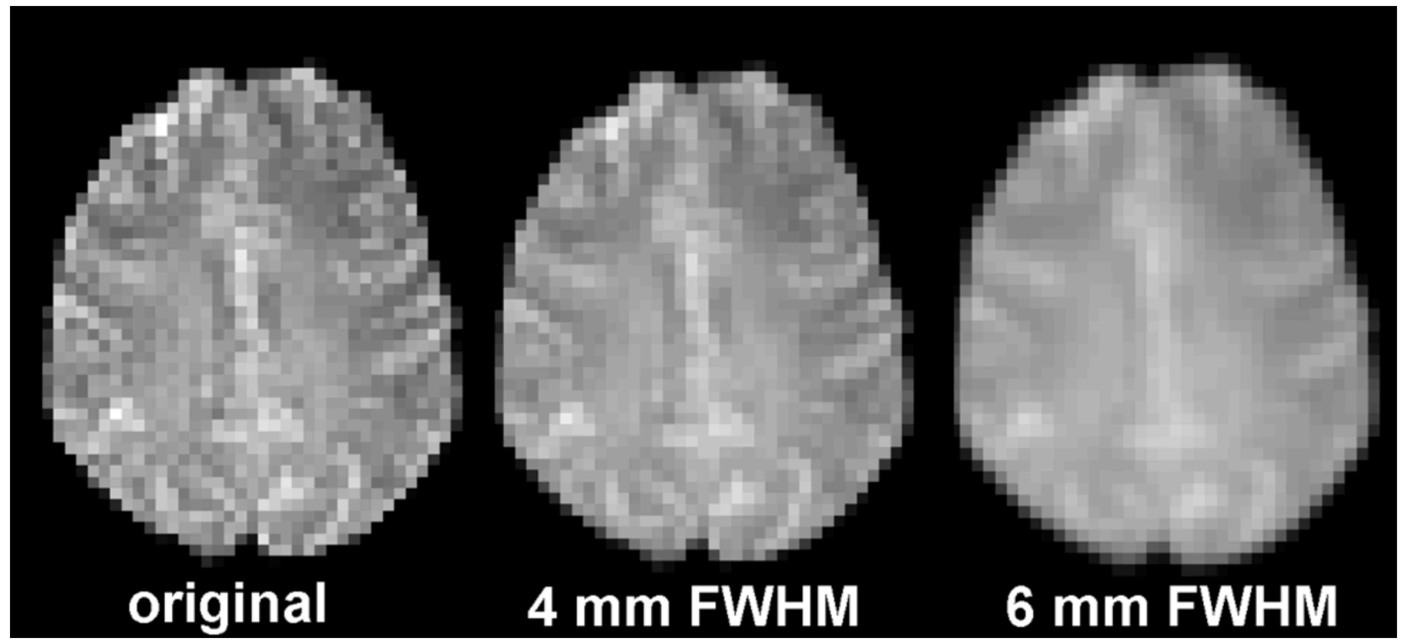
- ***BUT: won't everything necessarily look like a BLOB?***
  We're throwing away a lot of information….

- Analysing patterns across voxels will also increase SNR by pooling across voxels, and help us get results.

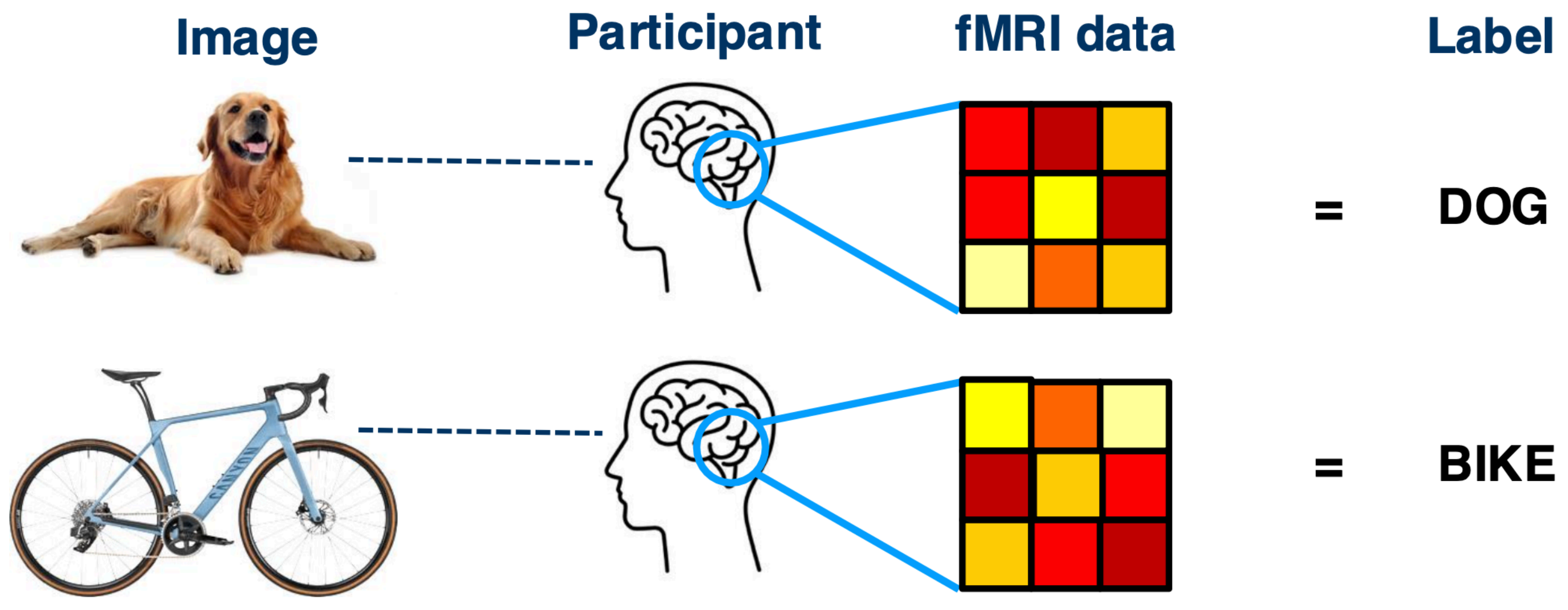


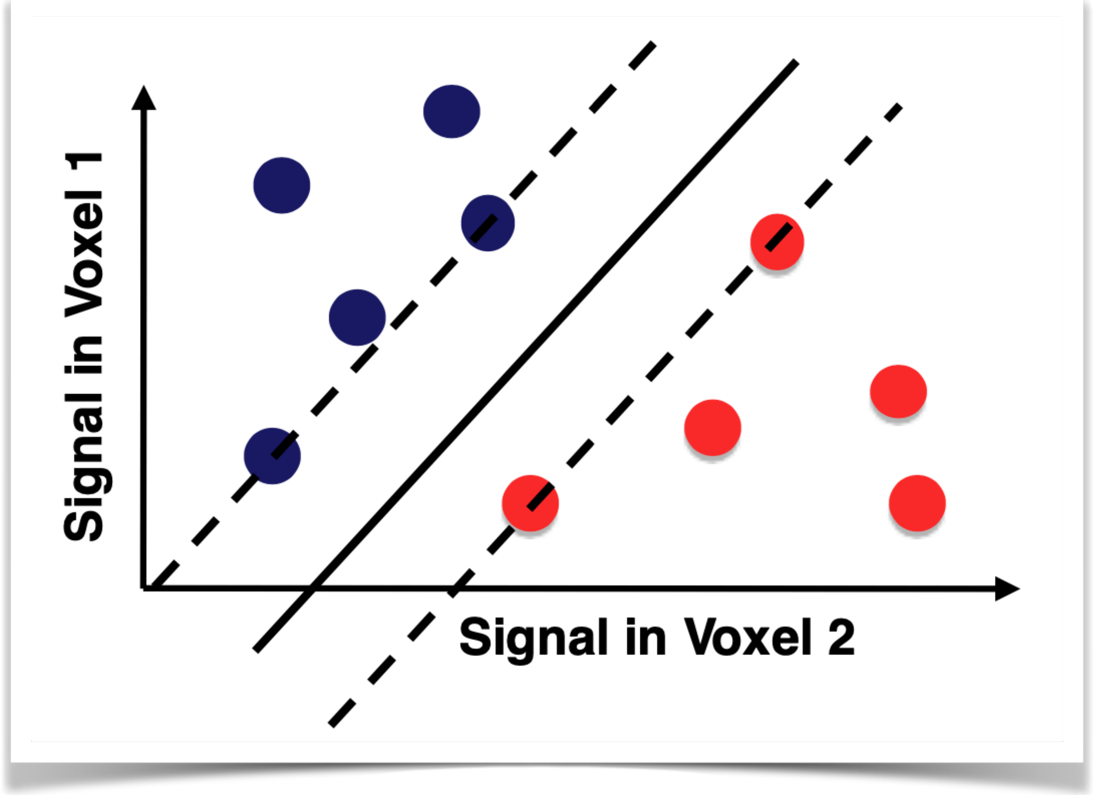


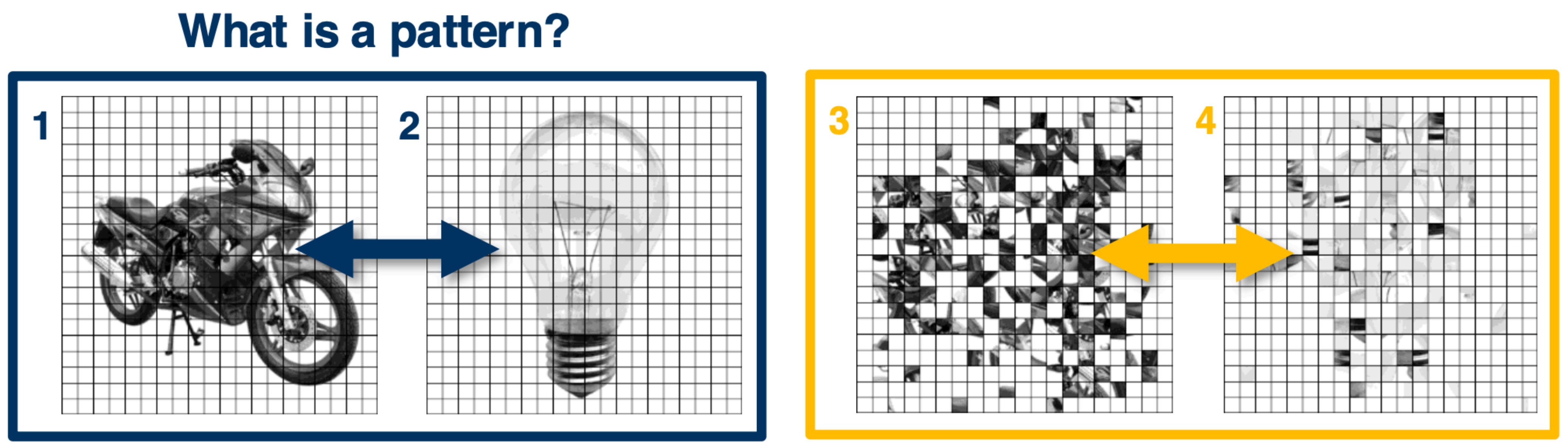
What does fMRI allow us to say?



- Historically, we had bad recordings, so smoothing will increased SNR and helped us get results.
- ***BUT: won't everything necessarily look like a BLOB?***
  We're throwing away a lot of information….
- Analysing patterns across voxels will also increase SNR by pooling across voxels, and help us get results.
- ***BUT: what does a pattern of responses across voxels mean?***
  No explicit model of what process generates this response pattern…

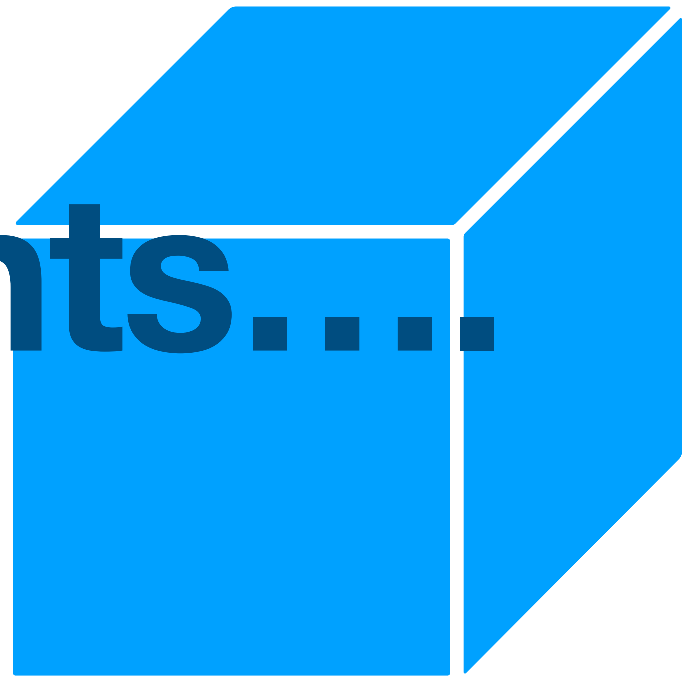






Let’s think deeply about our measurements….

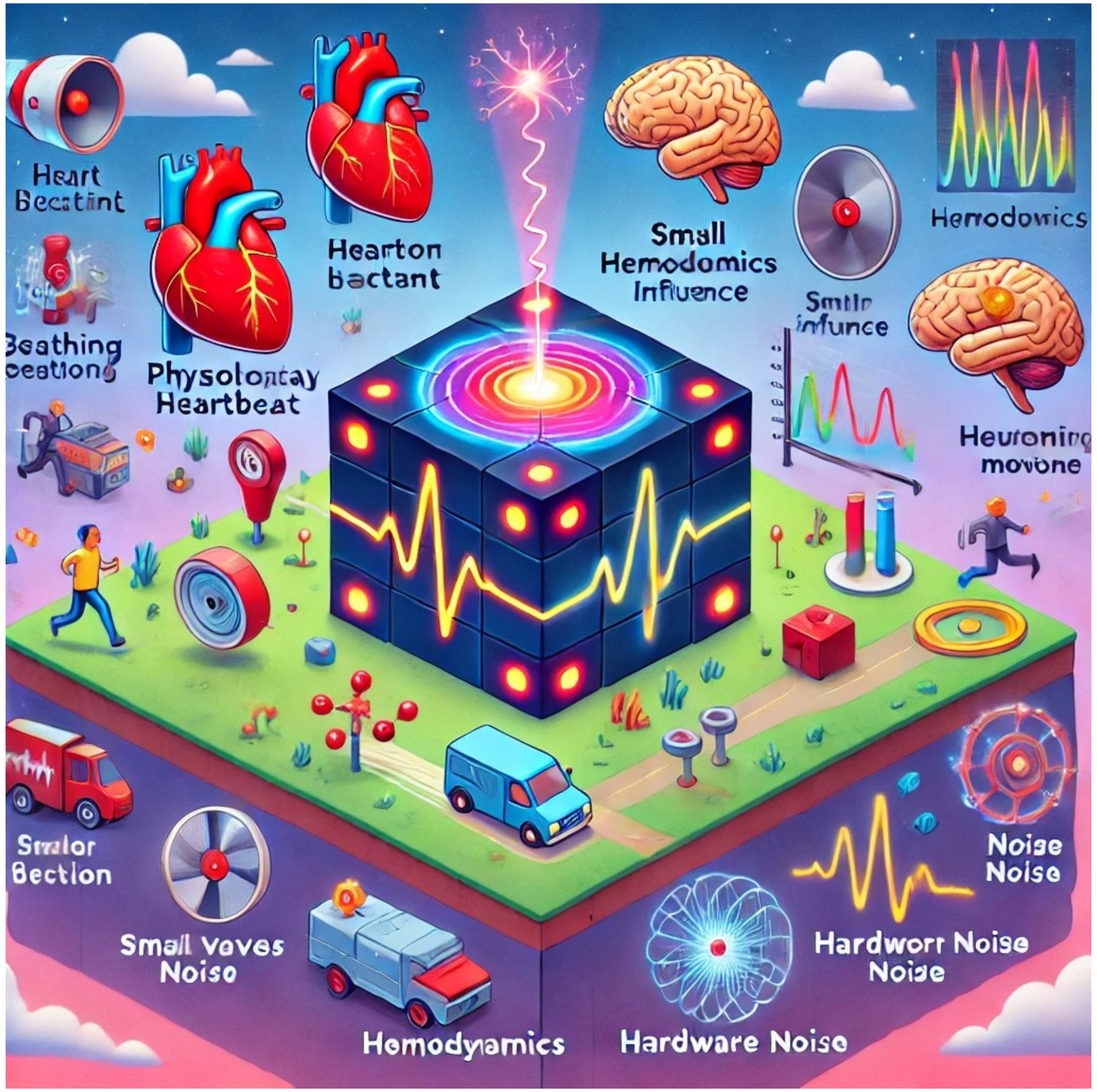
What information does a voxel’s signal contain?



Let’s think deeply about our measurements….

What information does a voxel’s signal contain?

- ***Physiological & nuisance stuff:***
  heartbeat, blood pressure, breathing, movement of the participant, hardware noise

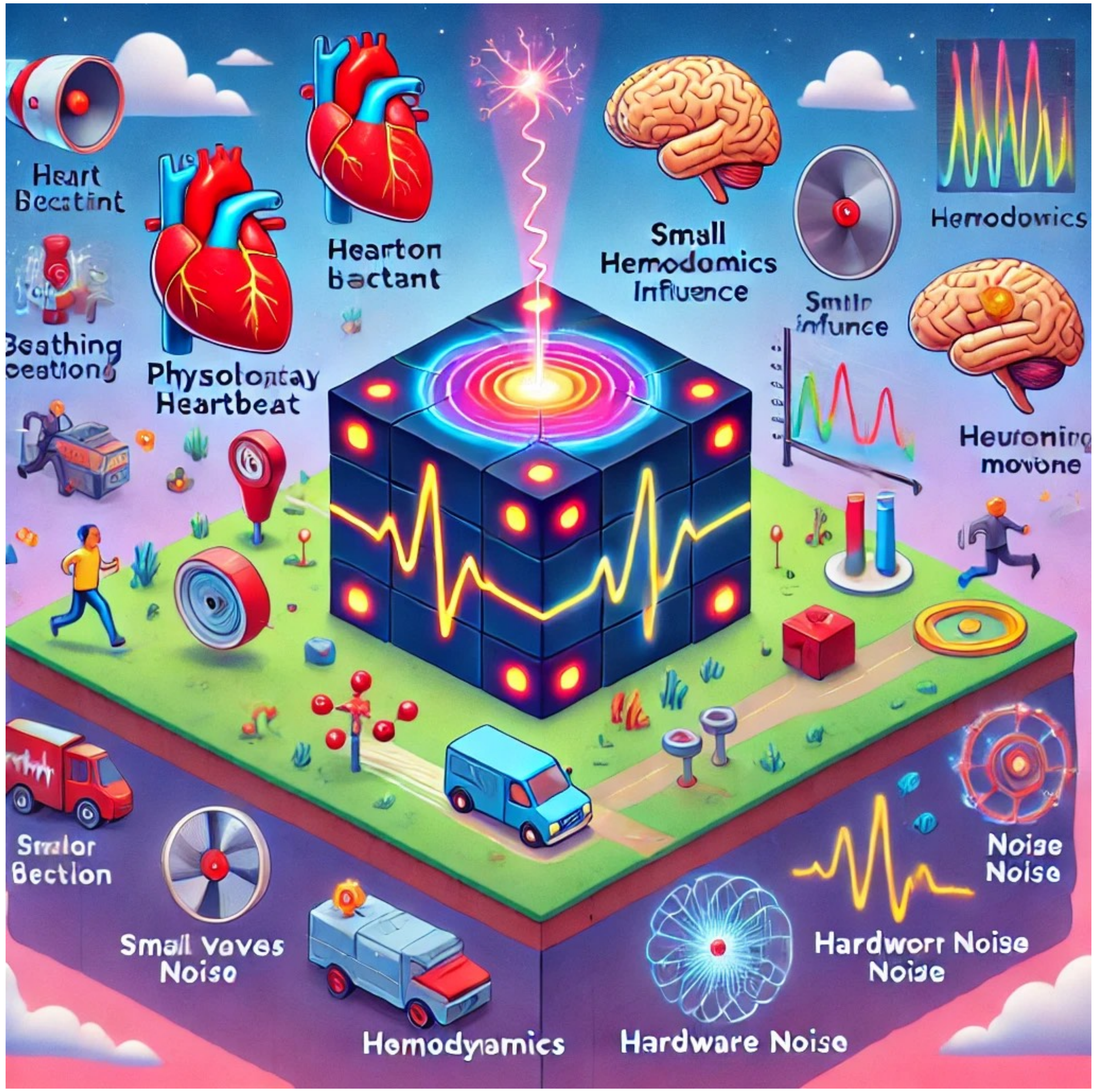


Let’s think deeply about our measurements….

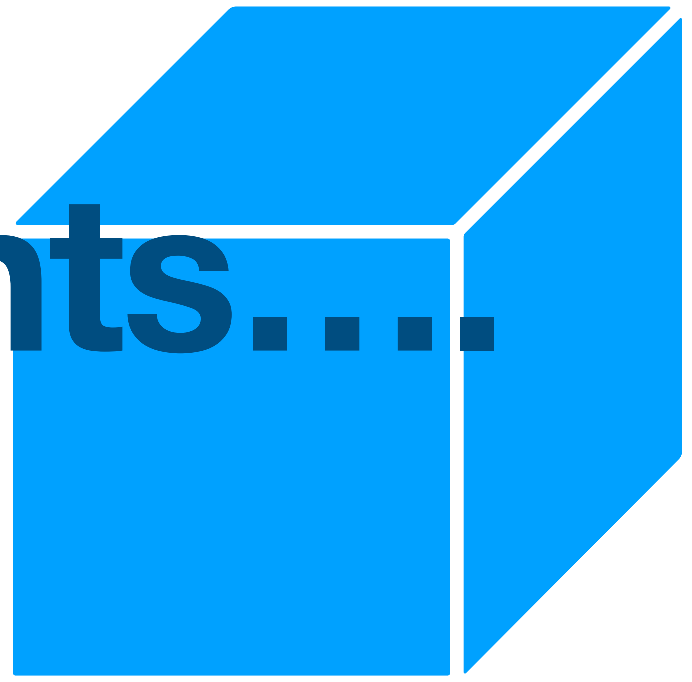
What information does a voxel’s signal contain?

- ***Physiological & nuisance stuff:***
  heartbeat, blood pressure, breathing, movement of the participant, hardware noise

- ***Neurons firing!***
  Interpret our measurements as ***neurons firing***?



Let’s think deeply about our measurements….

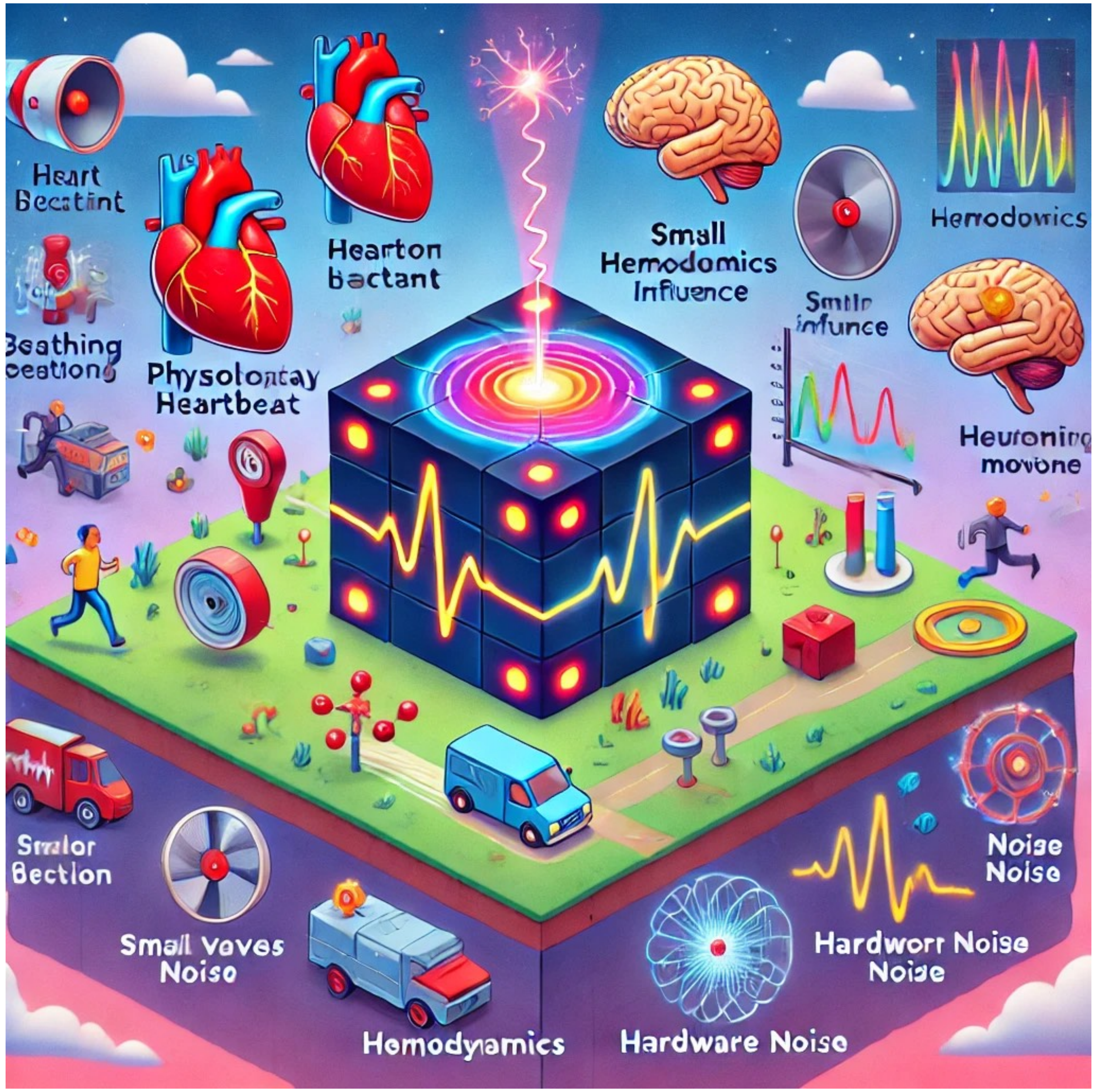


What information does a voxel’s signal contain?

- ***Physiological & nuisance stuff:***
  heartbeat, blood pressure, breathing, movement of the participant, hardware noise

- ***Neurons firing!***
  Interpret our measurements as ***neurons firing***?

- With recent improvements in fMRI signal fidelity, we want to:



Let’s think deeply about our measurements….

What information does a voxel’s signal contain?

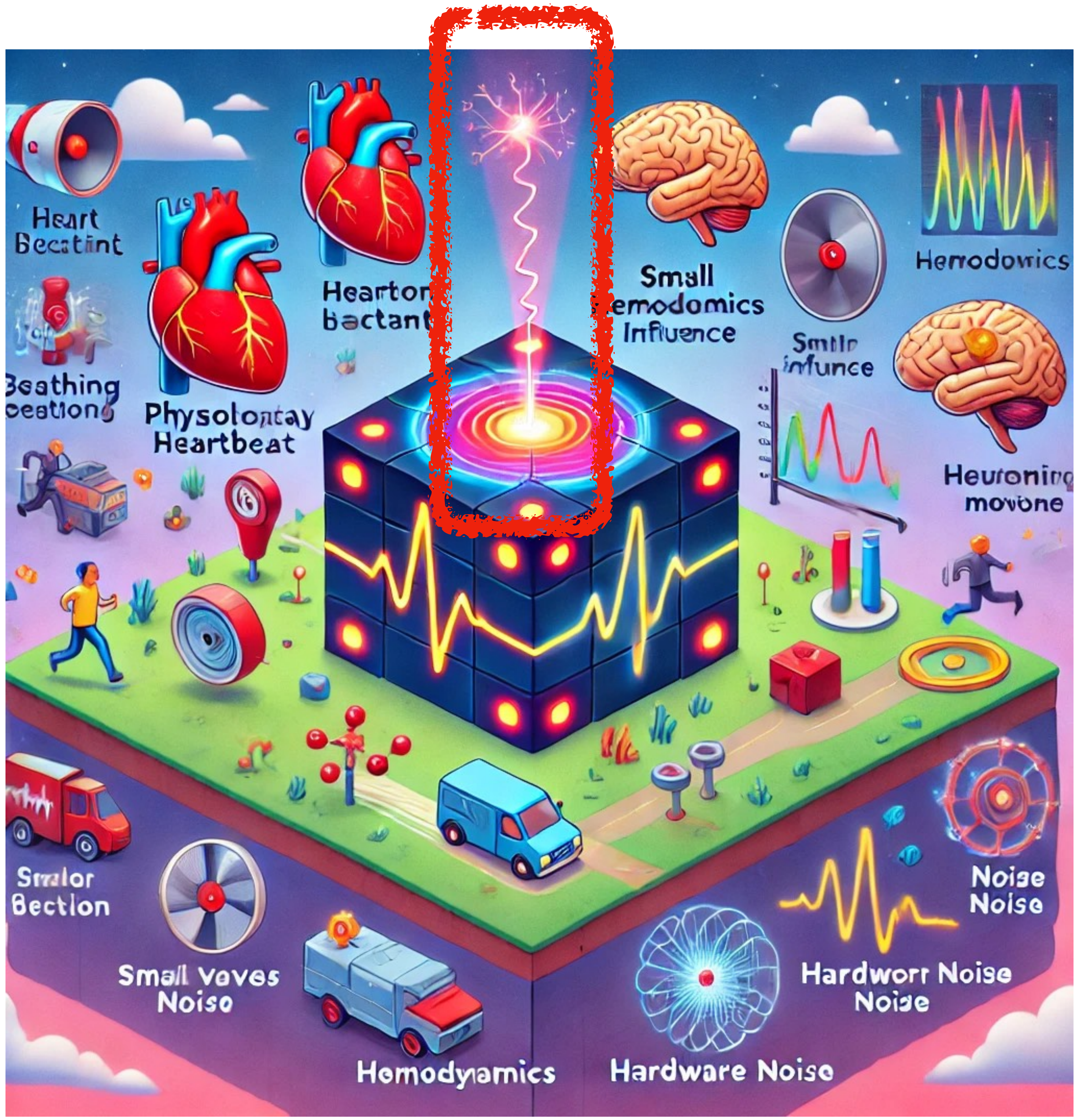
- ***Physiological & nuisance stuff:***
  heartbeat, blood pressure, breathing, movement of the participant, hardware noise

- ***Neurons firing!***
  Interpret our measurements as ***neurons firing***?

- With recent improvements in fMRI signal fidelity, we want to:

1. *Make explicit models of what neurons in **single voxels** are doing*

2. *how this processing generates **patterns across voxels**.*



Receptive fields?

Hubel & Wiesel

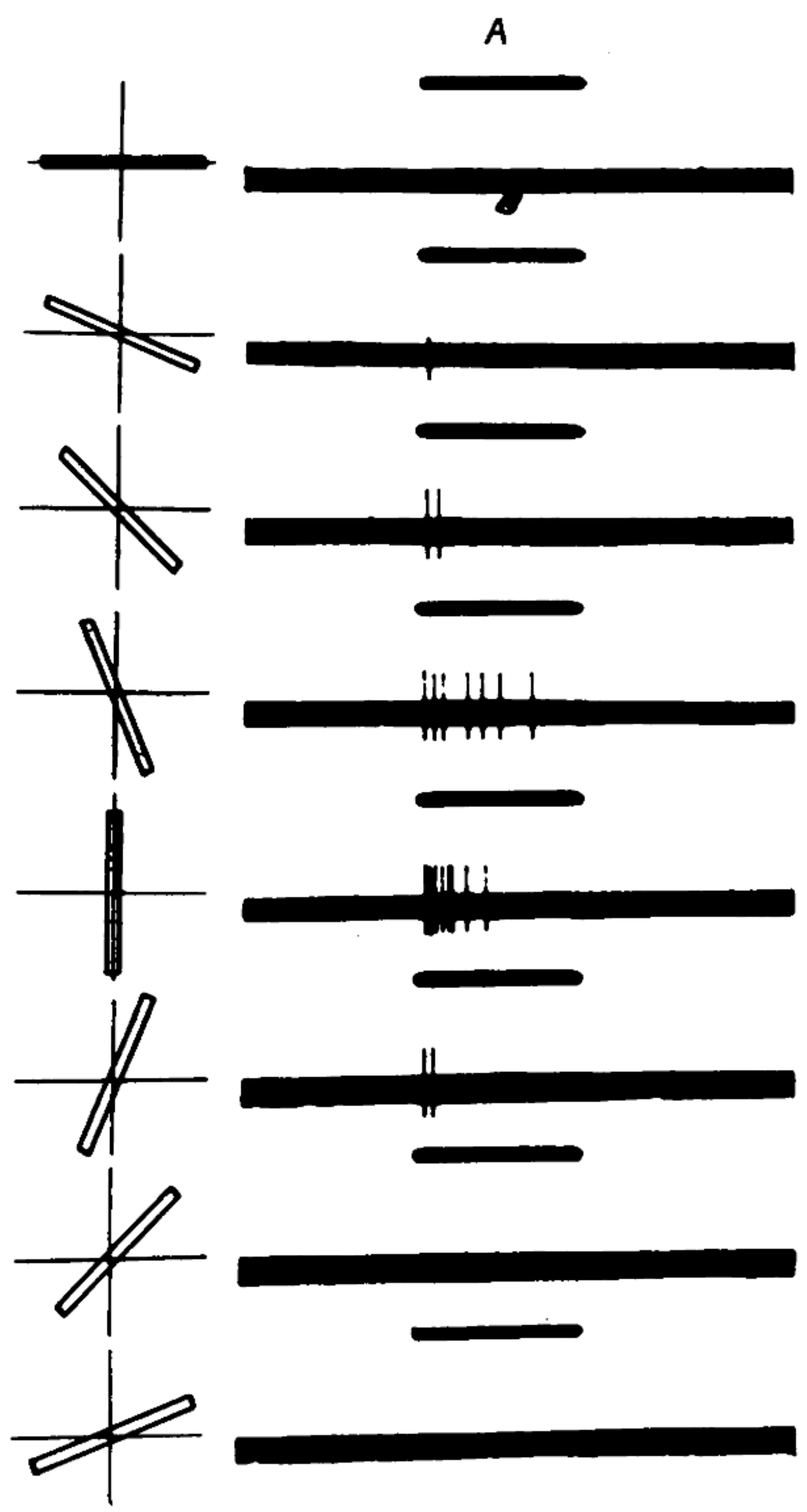
How does a single neuron ‘encode’ visual information?

Hubel & Wiesel

How does a single neuron ‘encode’ visual information?

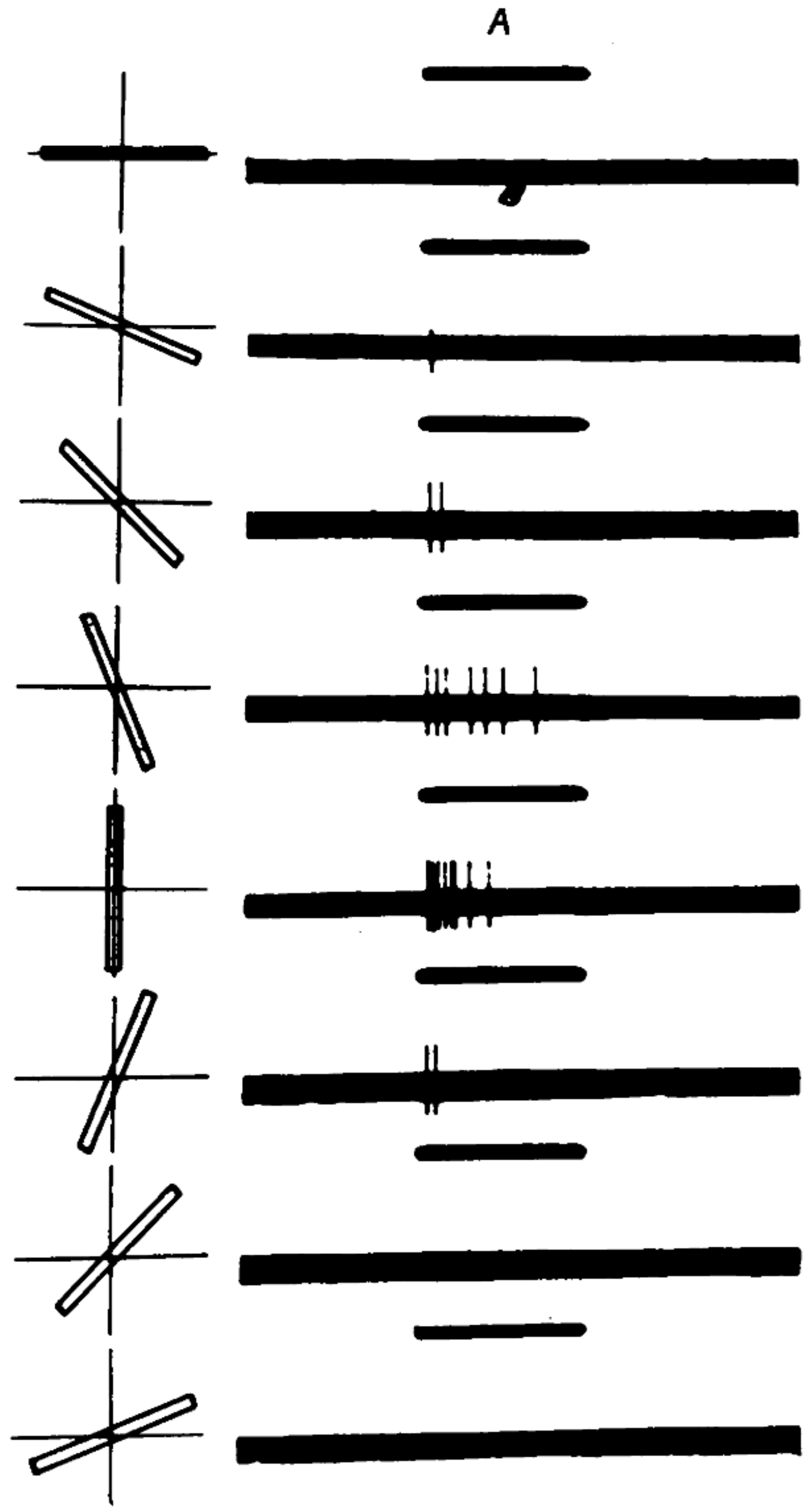
Hubel & Wiesel

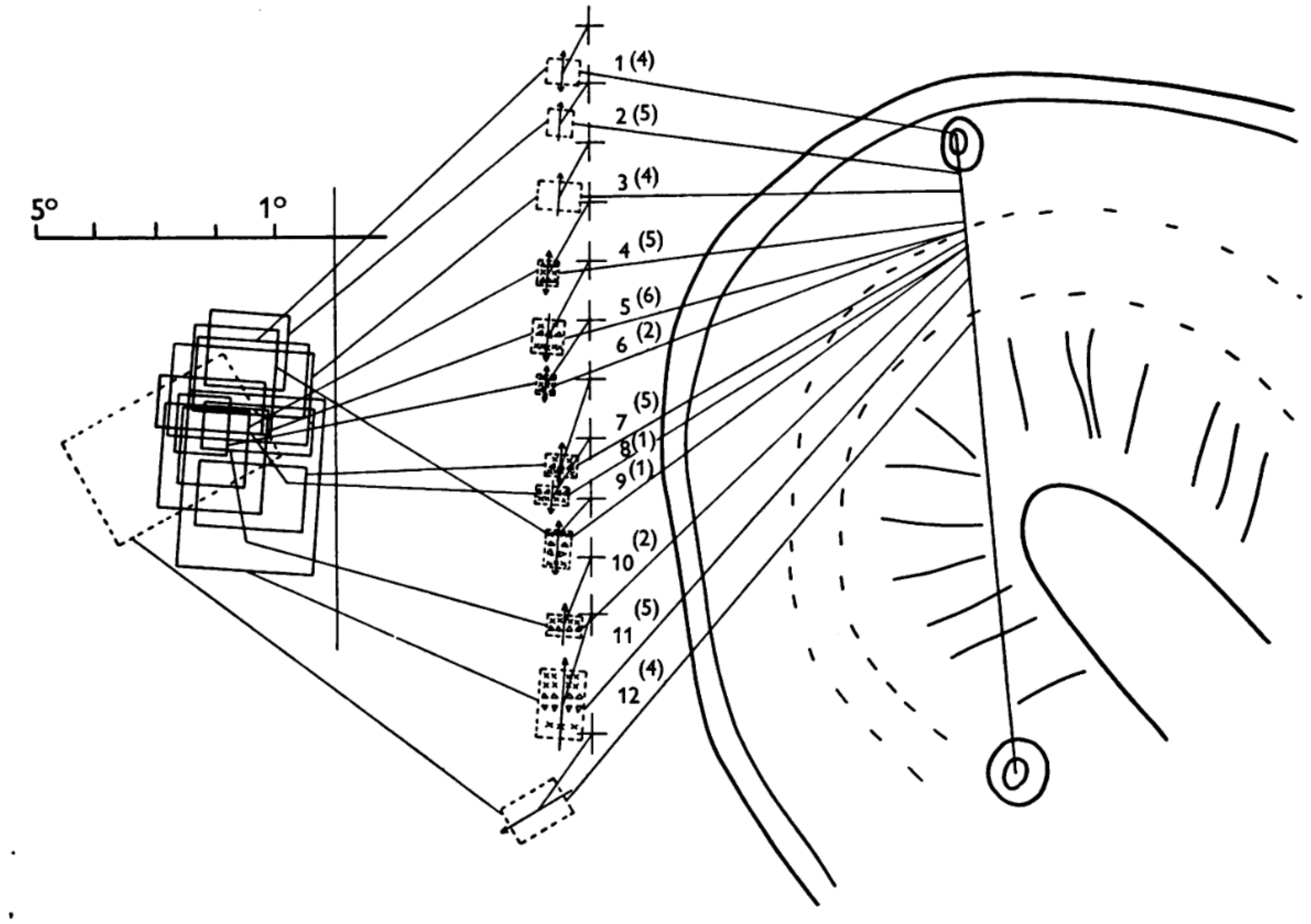
How does a single neuron ‘encode’ visual information?



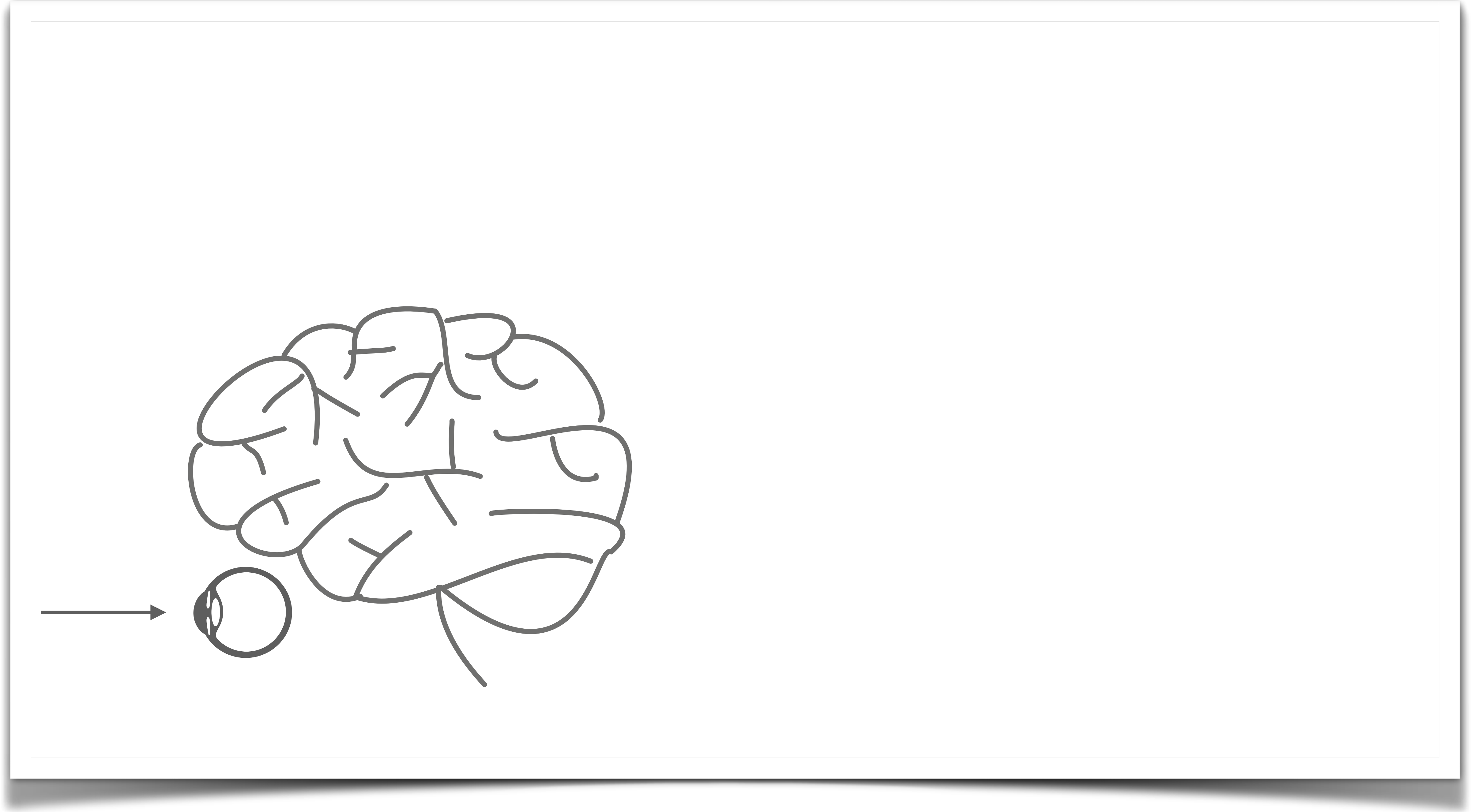
Hubel & Wiesel

How does a single neuron 'encode' visual information?



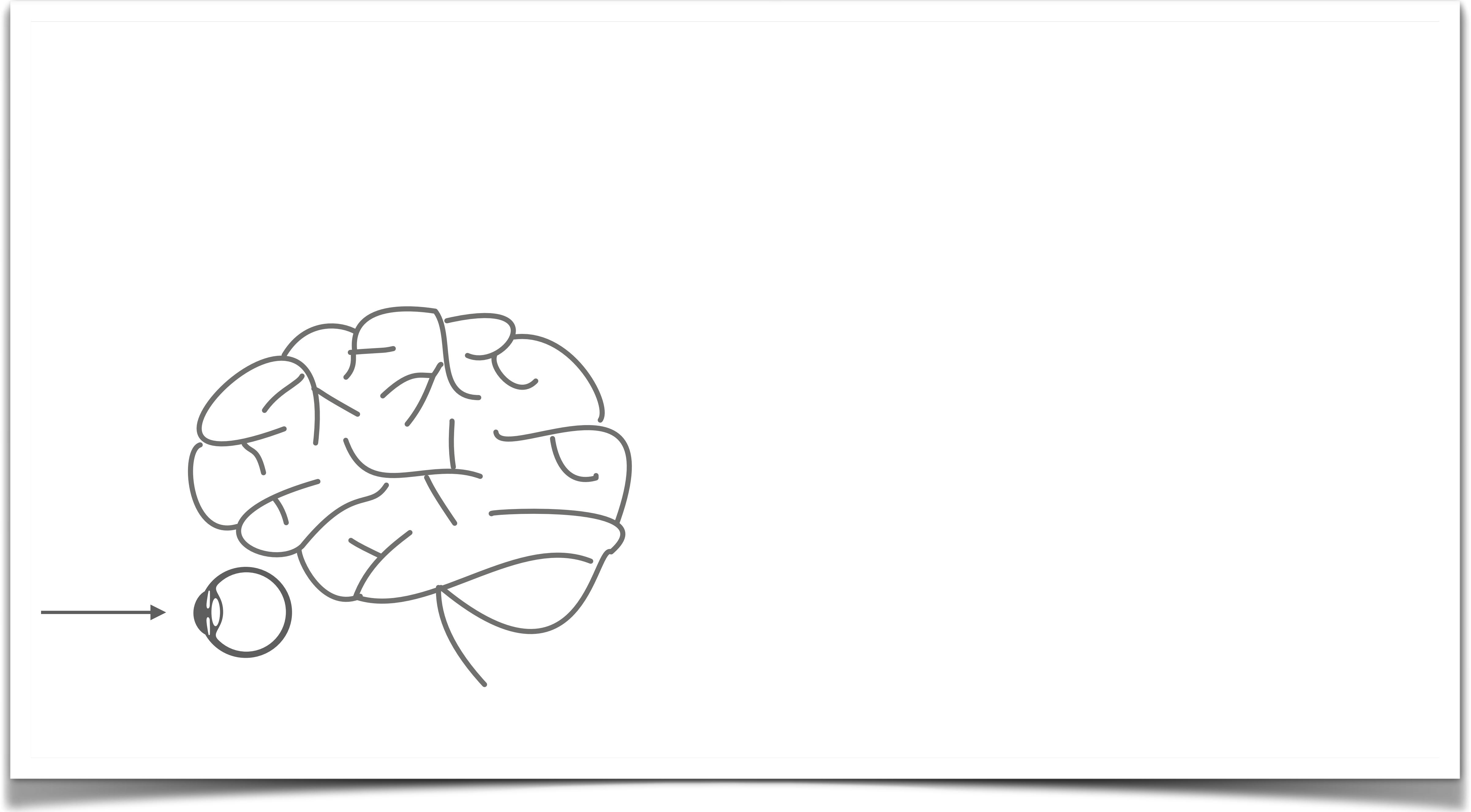


Population Receptive Fields using fMRI



Dumoulin & Wandell, 2008

Population Receptive Fields using fMRI

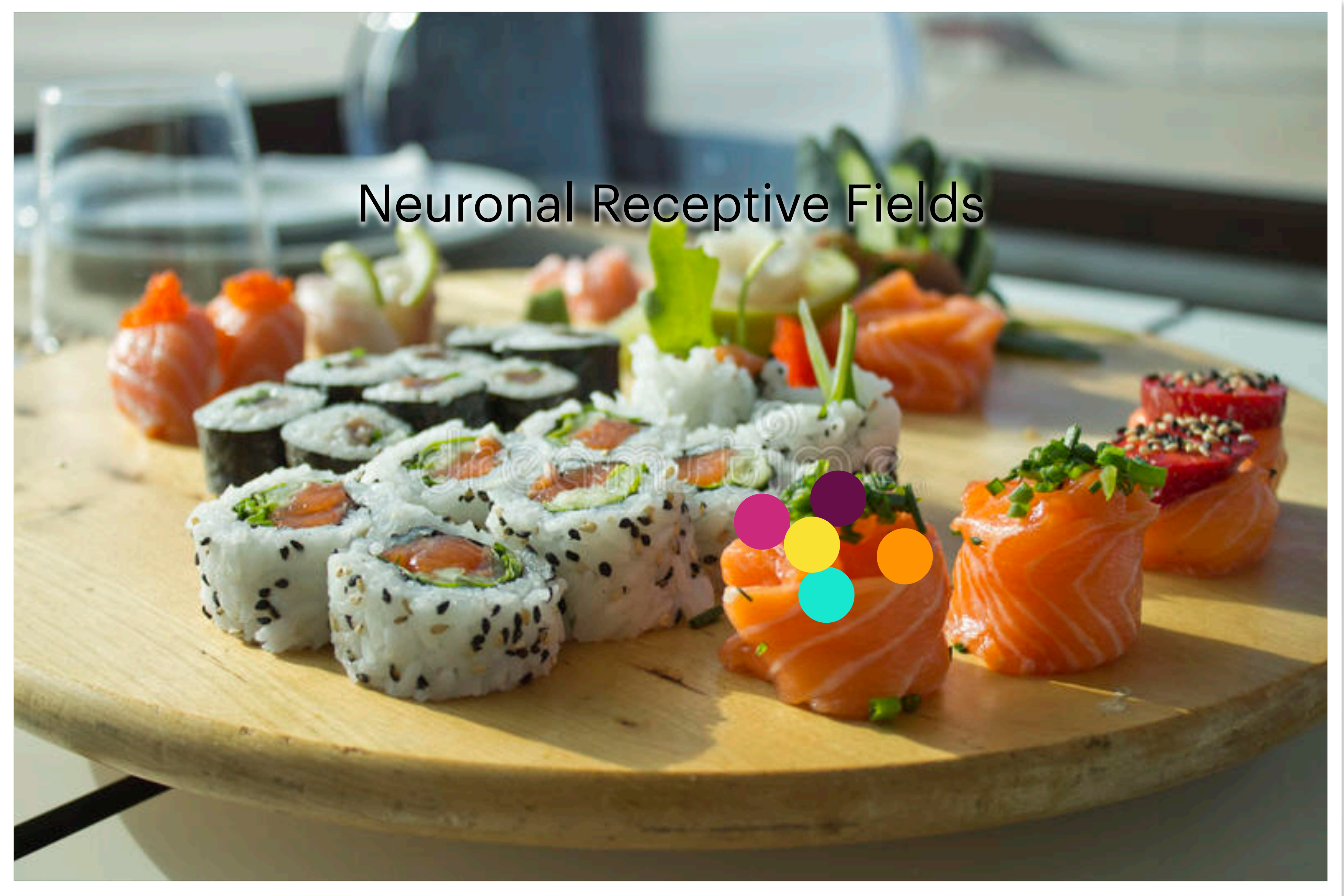


Dumoulin & Wandell, 2008

Population Receptive Fields using fMRI

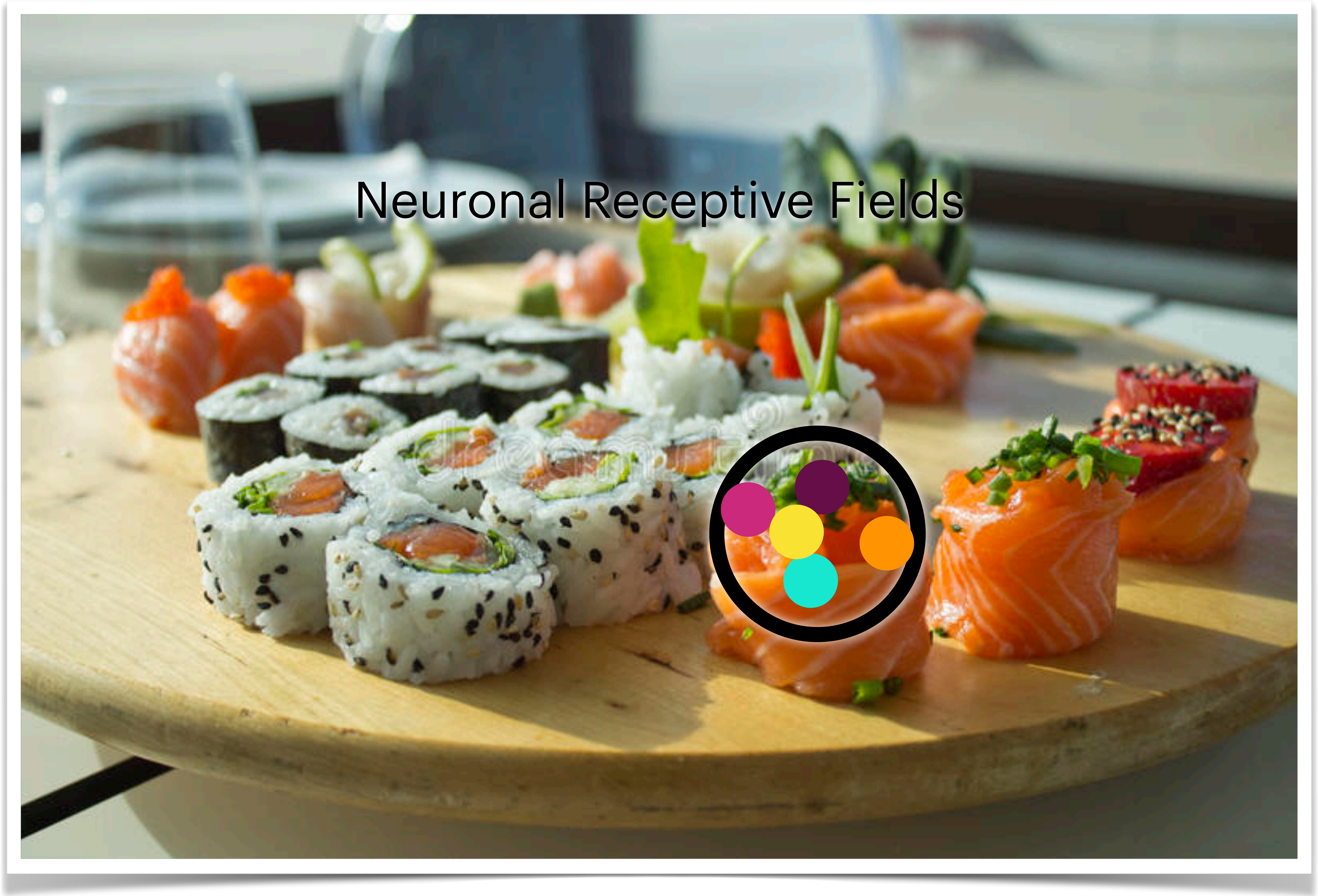
Dumoulin & Wandell, 2008

Population Receptive Fields using fMRI



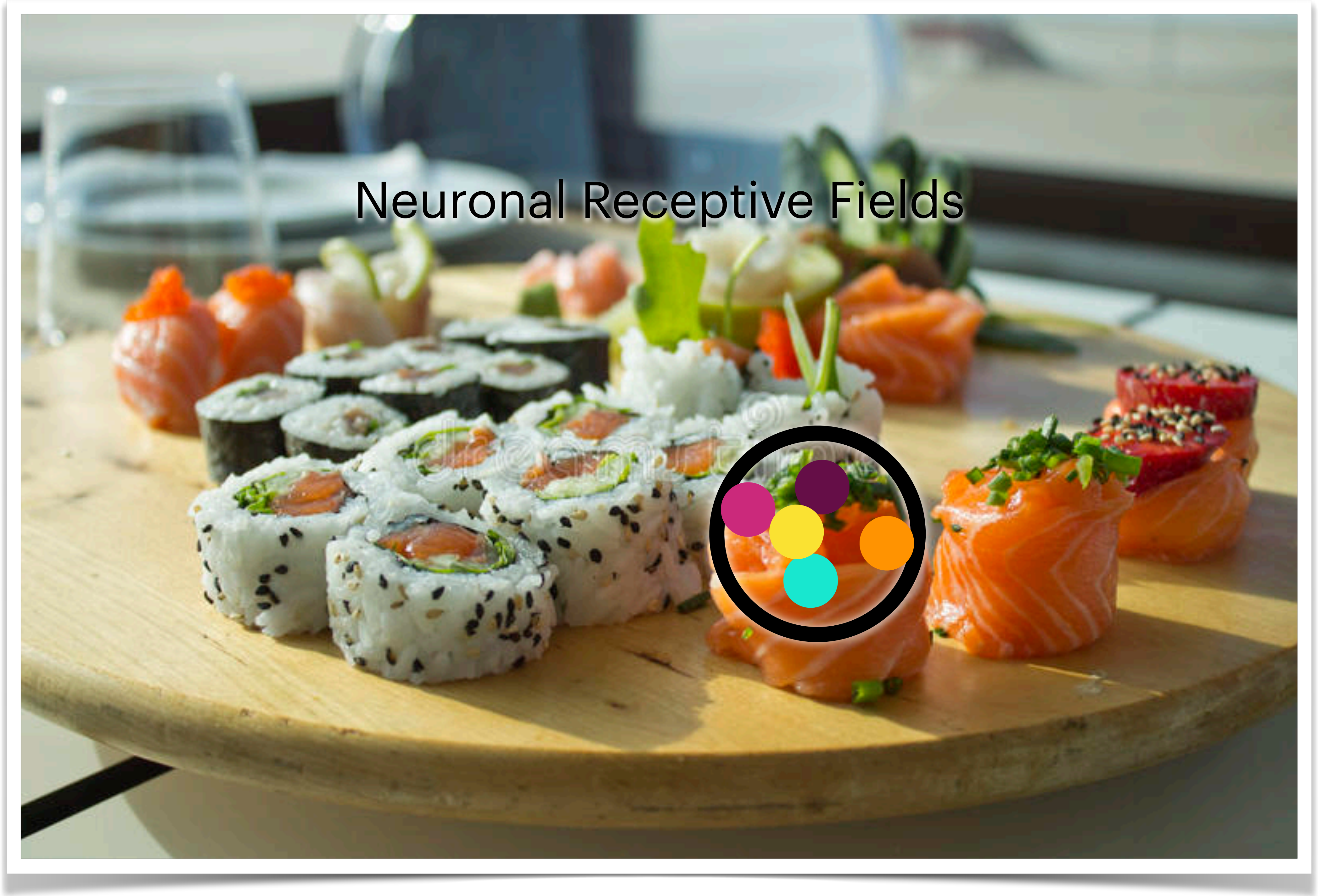
Dumoulin & Wandell, 2008

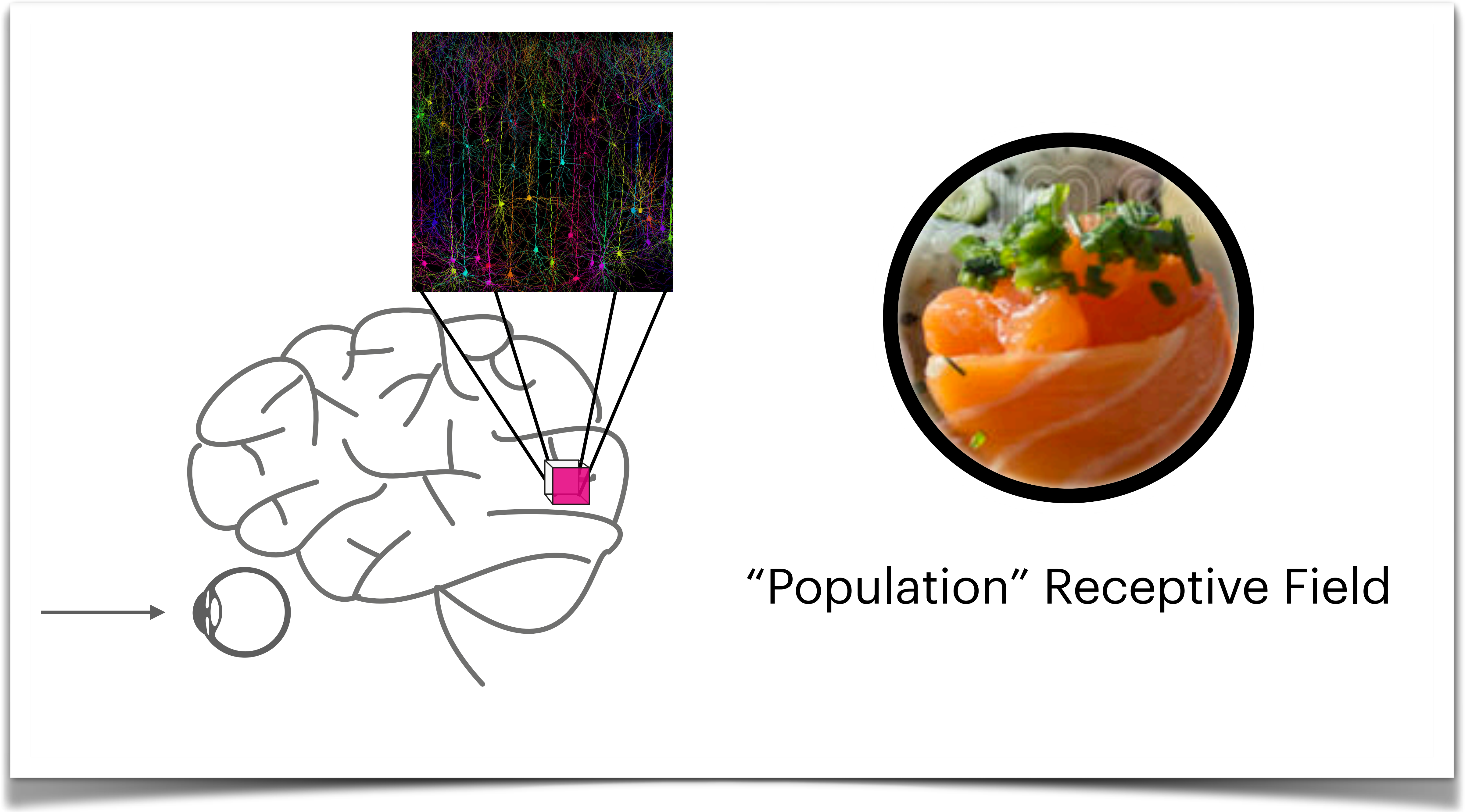
Population Receptive Fields using fMRI



Dumoulin & Wandell, 2008

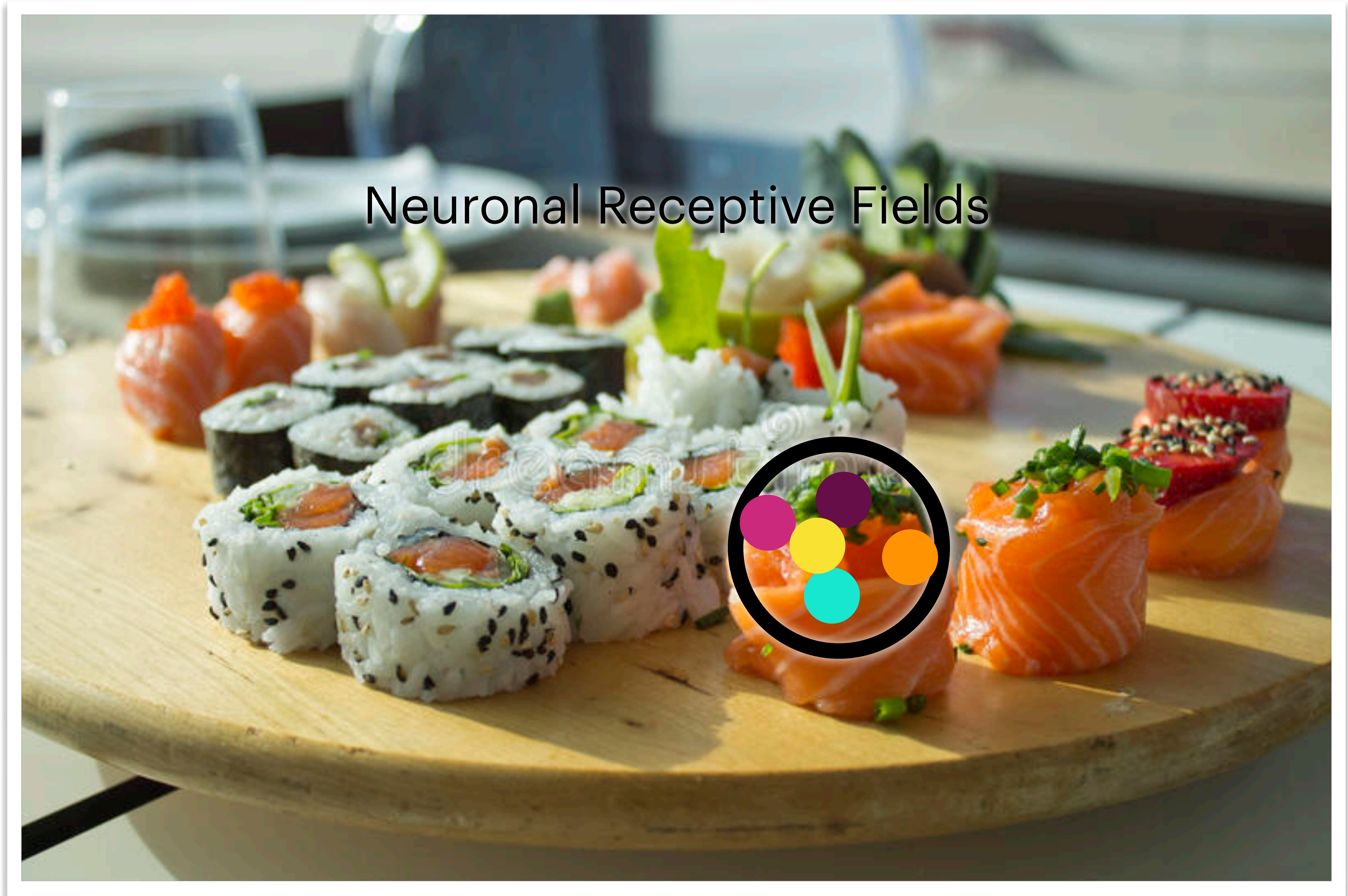
Population Receptive Fields using fMRI

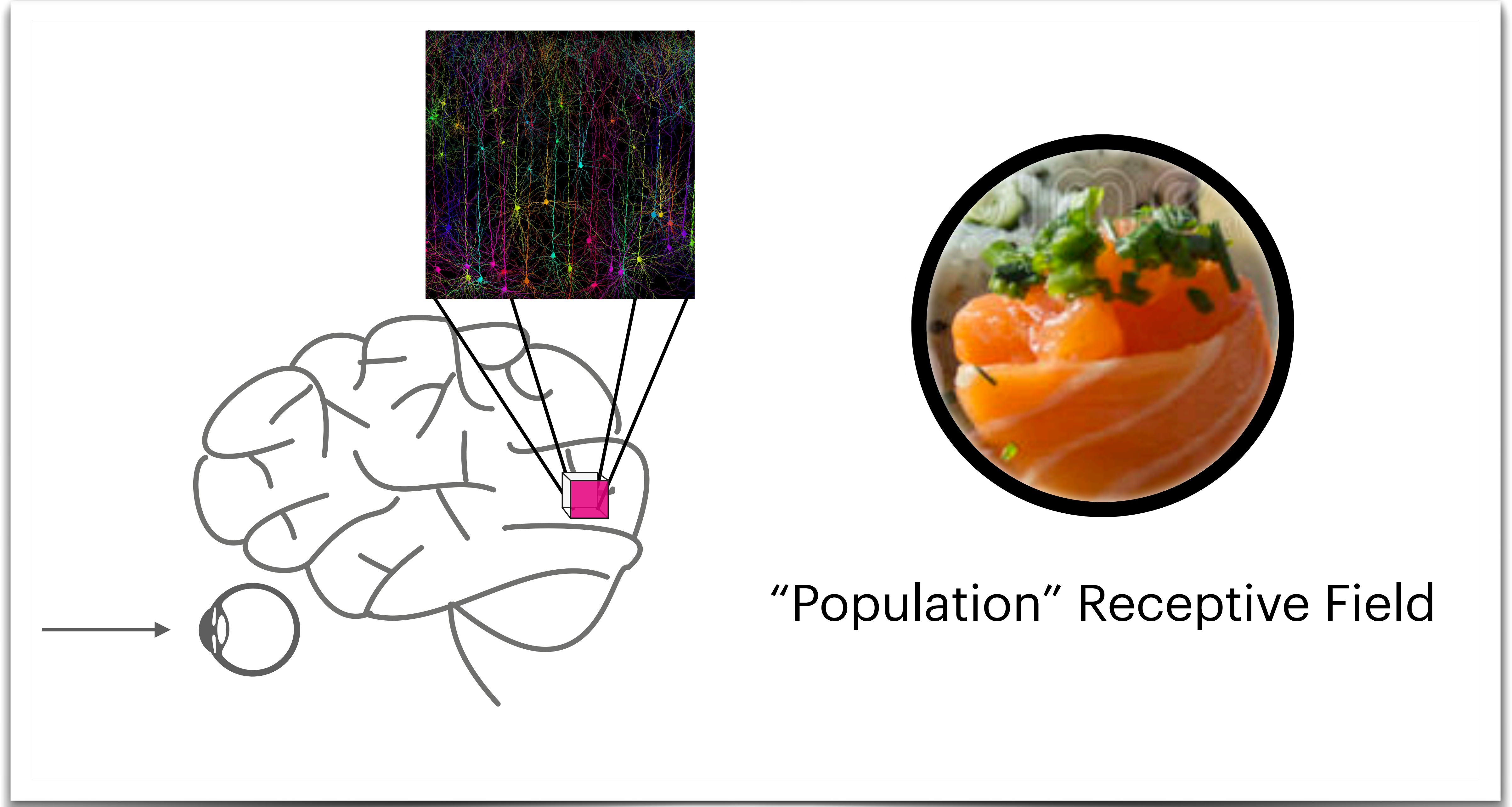




Dumoulin & Wandell, 2008

Population Receptive Fields using fMRI

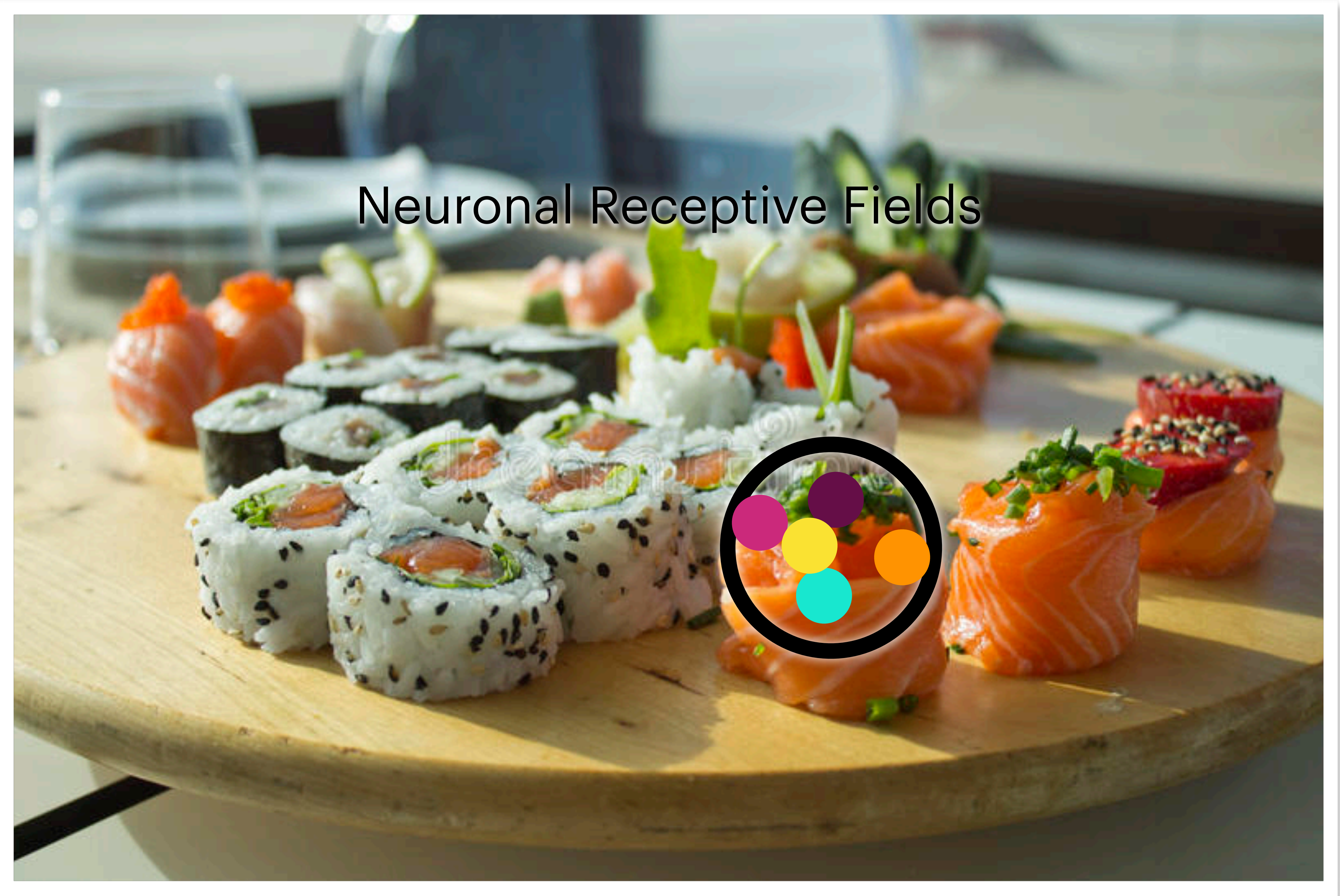




Extend the receptive field computational motif to voxels:
a voxel's neural population (if tuned similarly) has a
population receptive field (pRF)

Dumoulin & Wandell, 2008

Population Receptive Fields using fMRI

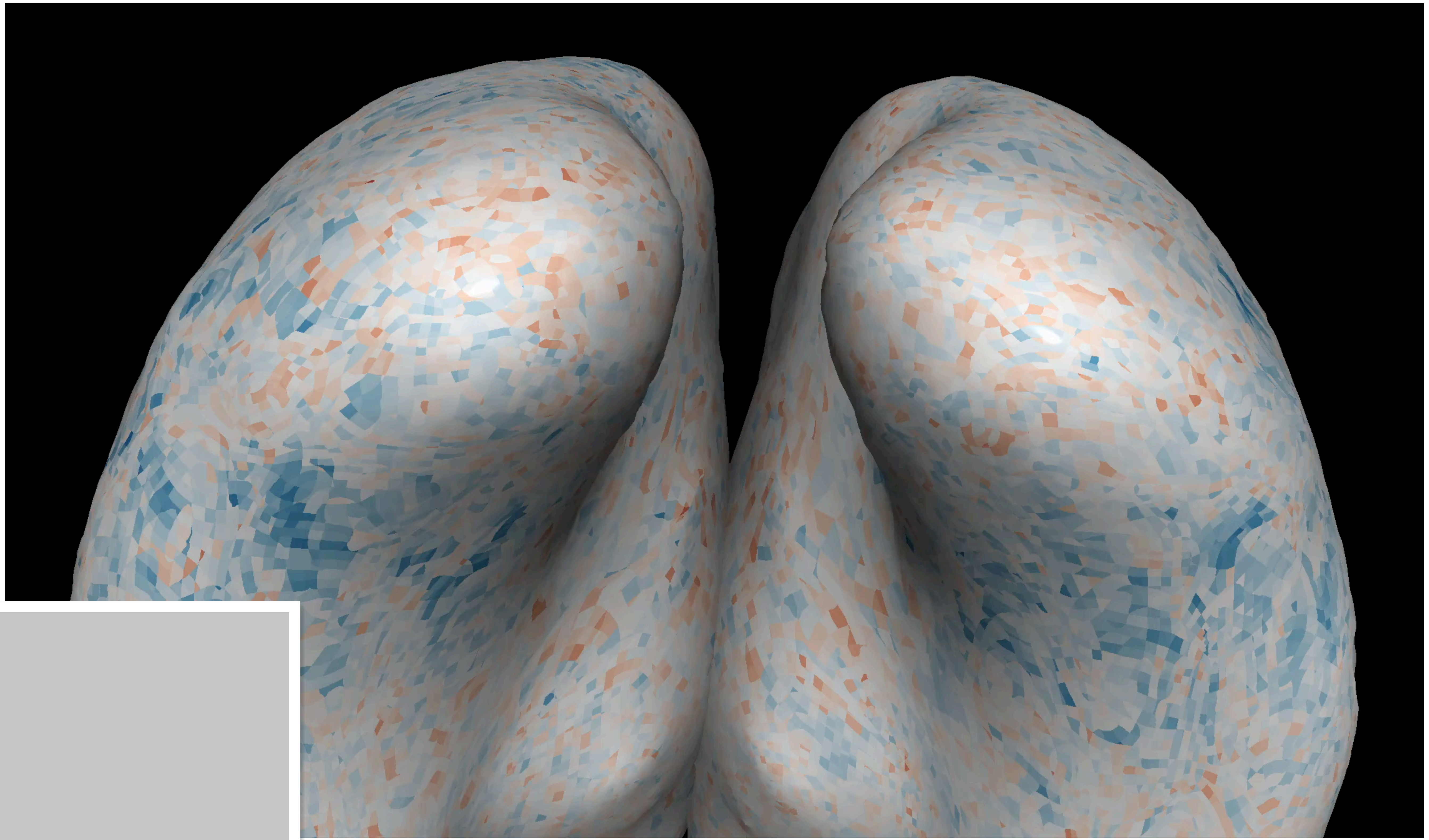


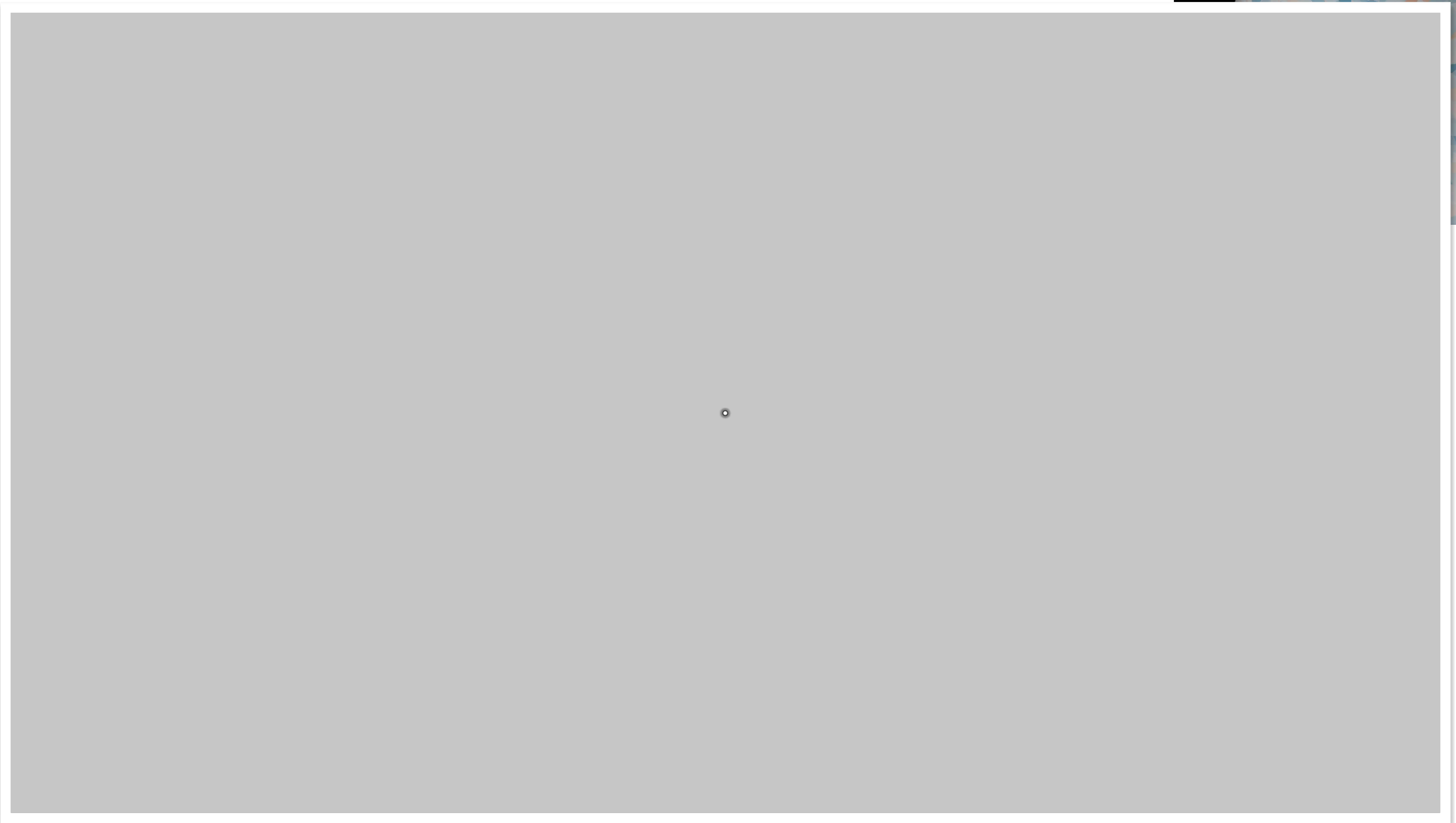
Extend the receptive field computational motif to voxels:
a voxel’s neural population (if tuned similarly) has a
population receptive field (pRF)

Leaves out orientation, focuses on just location!

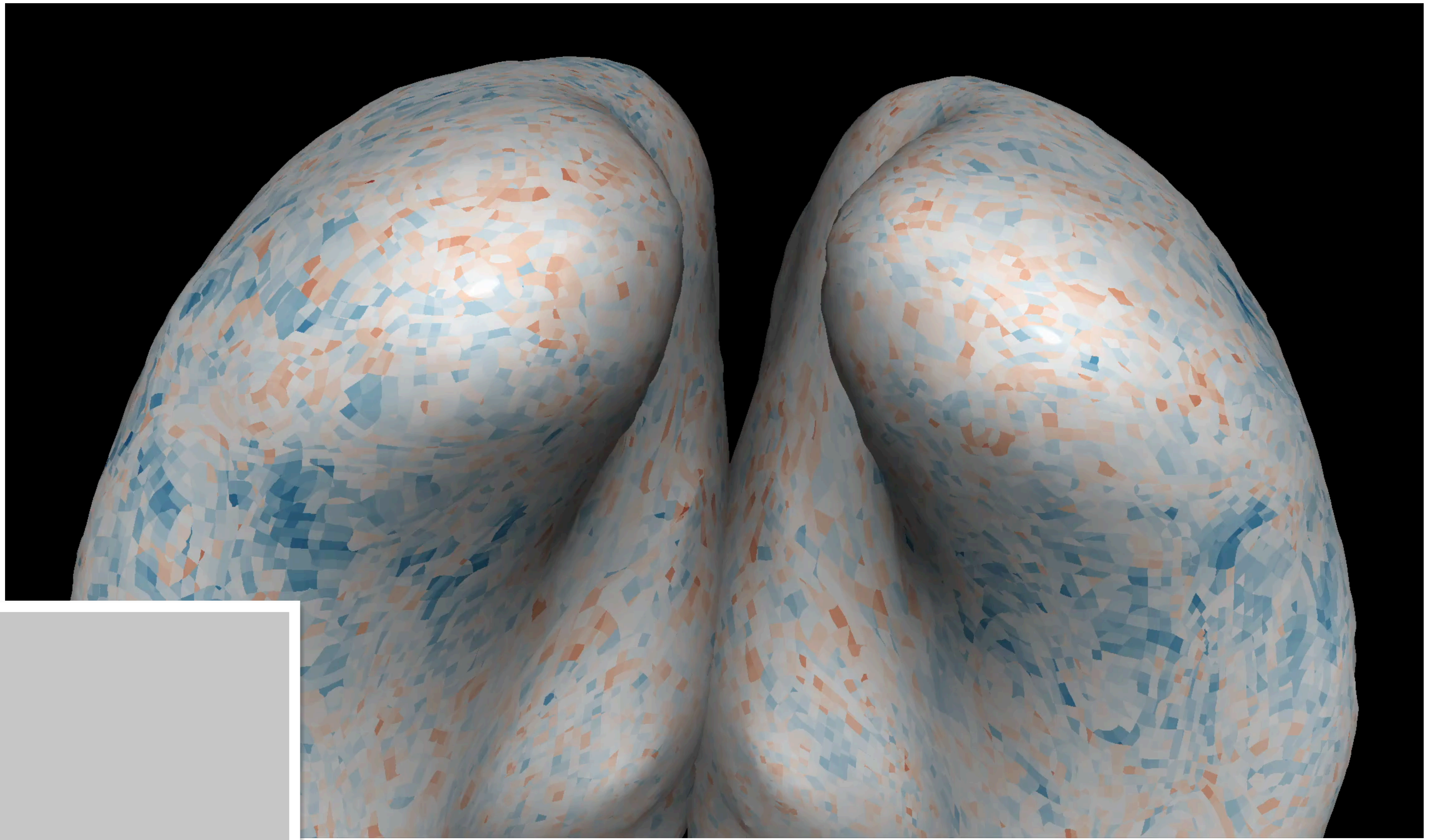
Dumoulin & Wandell, 2008

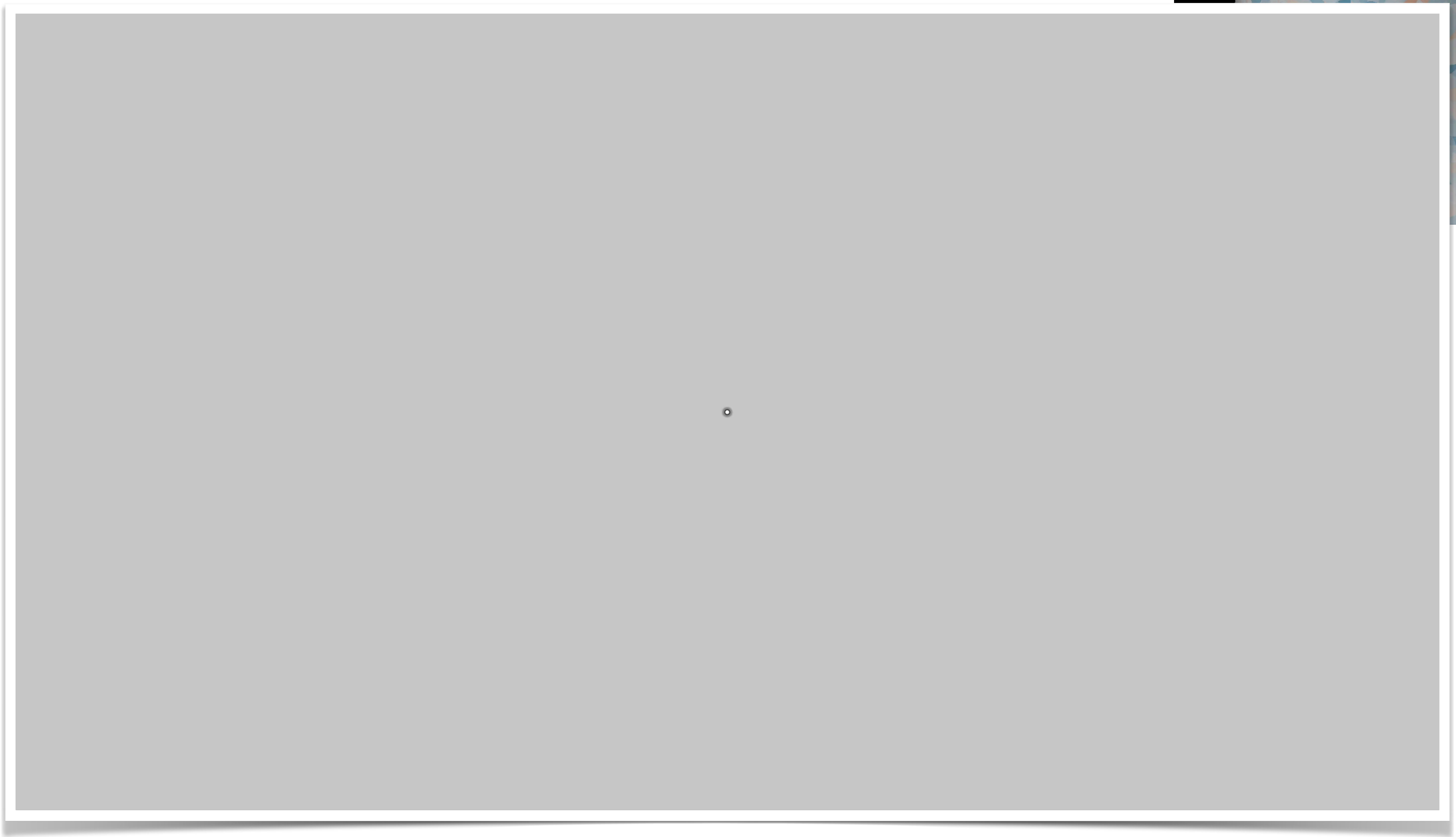
Taking RFs To fMRI



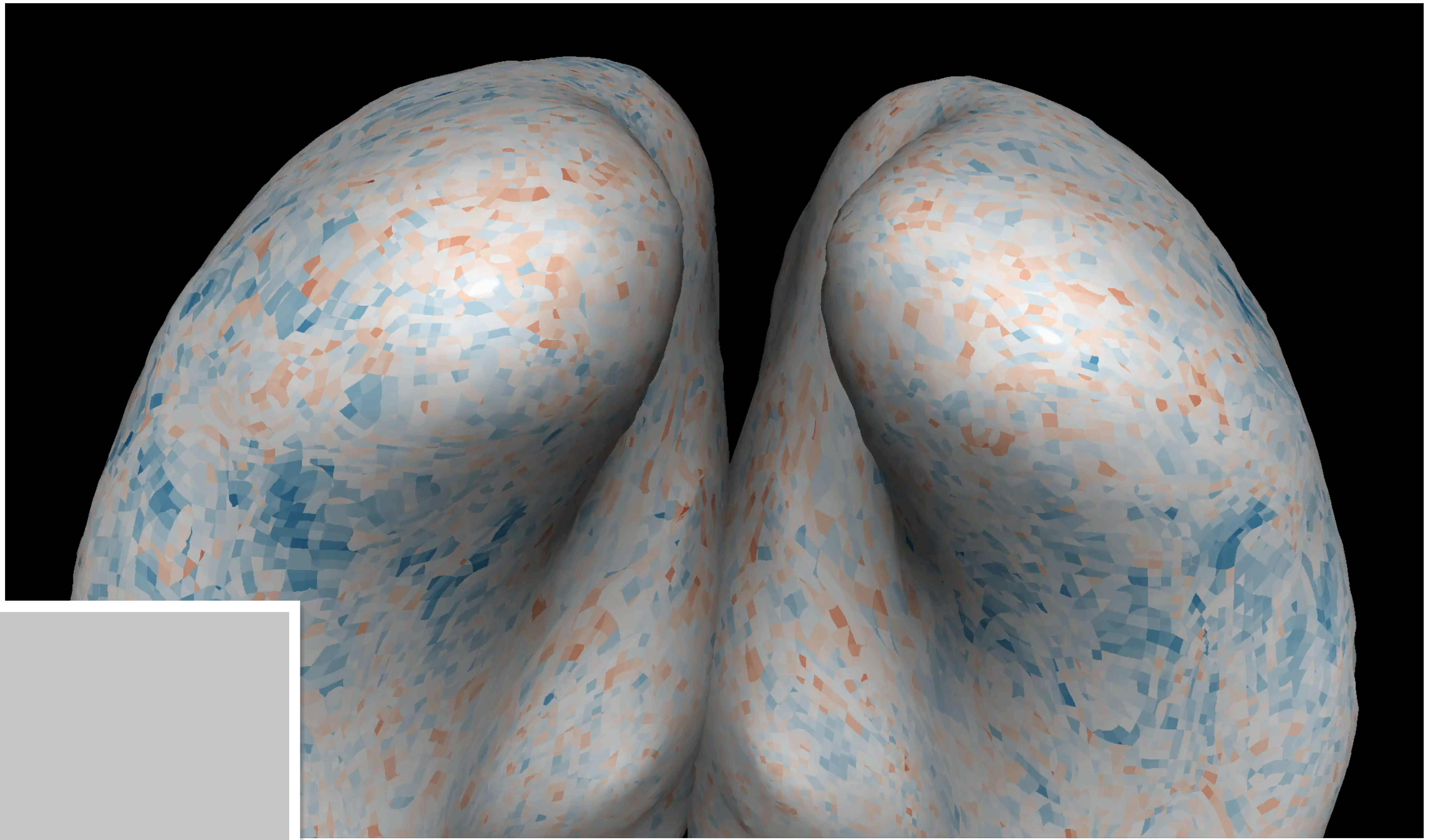


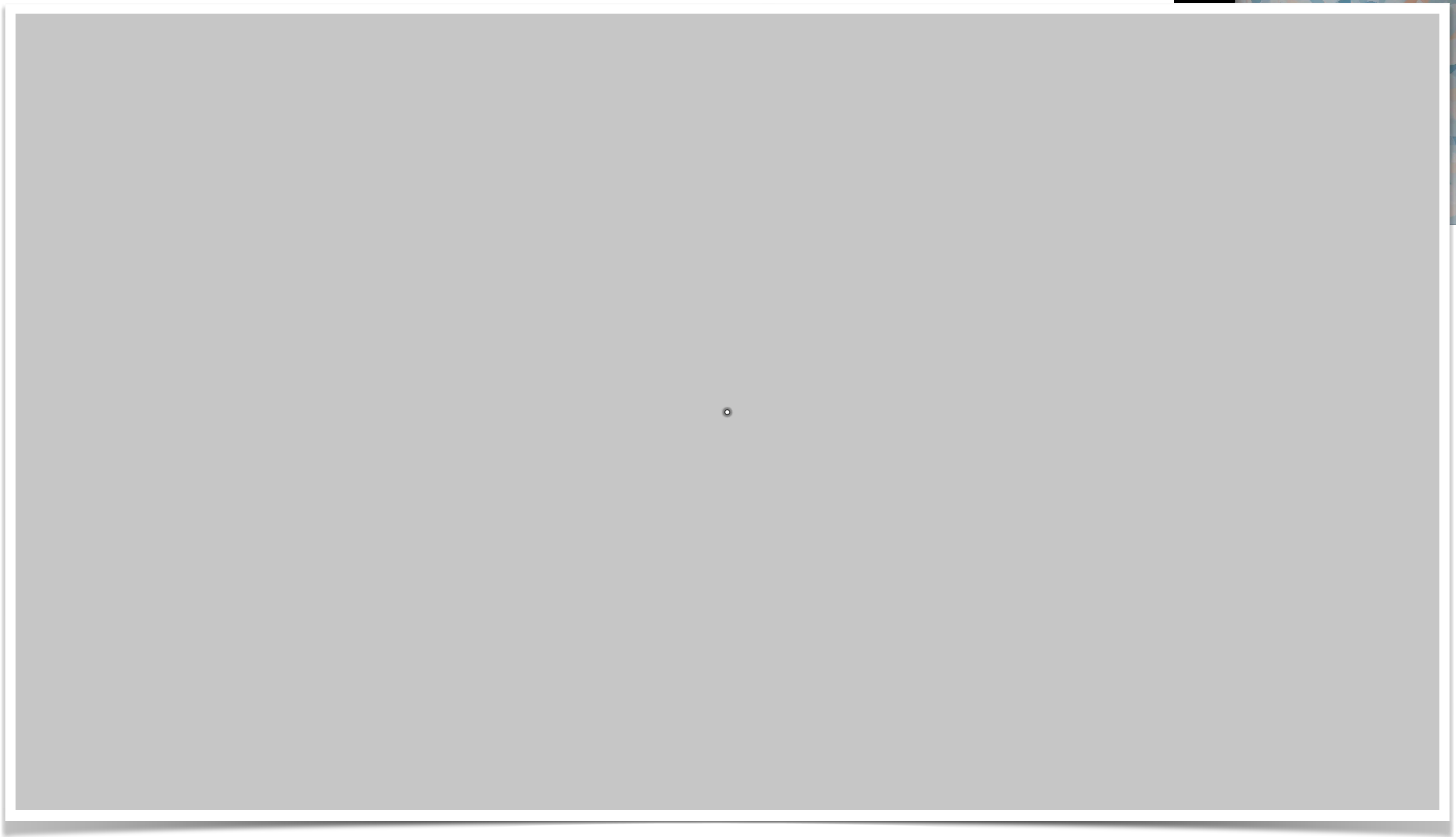
Taking RFs To fMRI





Taking RFs To fMRI





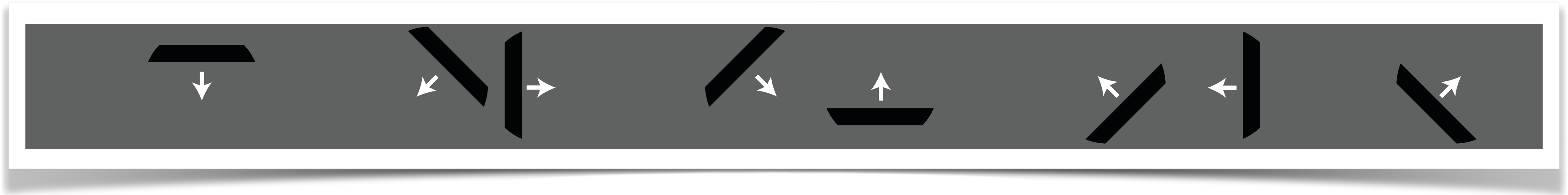
*Structured single-voxel BOLD time course: we can **fit PRF parameters***

Position

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

Size

Estimate best-fitting model parameters given the data



Dumoulin & Wandell, 2008

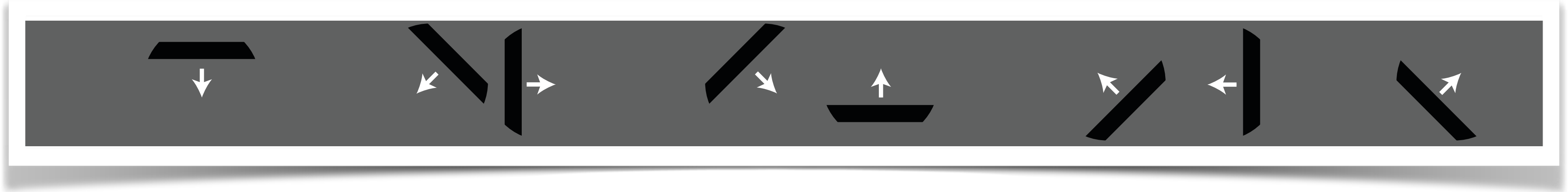
*Structured single-voxel BOLD time course: we can **fit PRF parameters***

Position

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

Size

Estimate best-fitting model parameters given the data



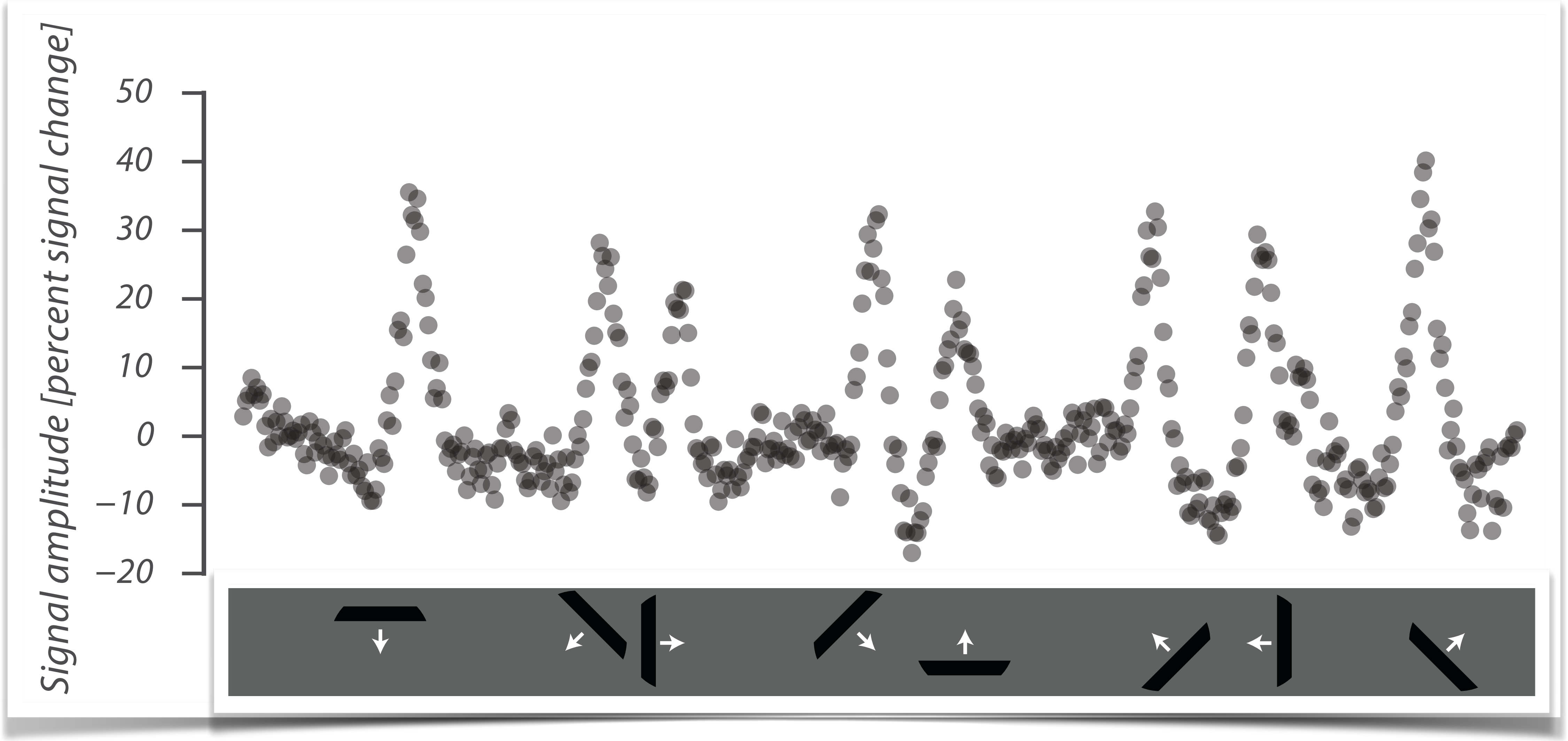
Dumoulin & Wandell, 2008

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

Position

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

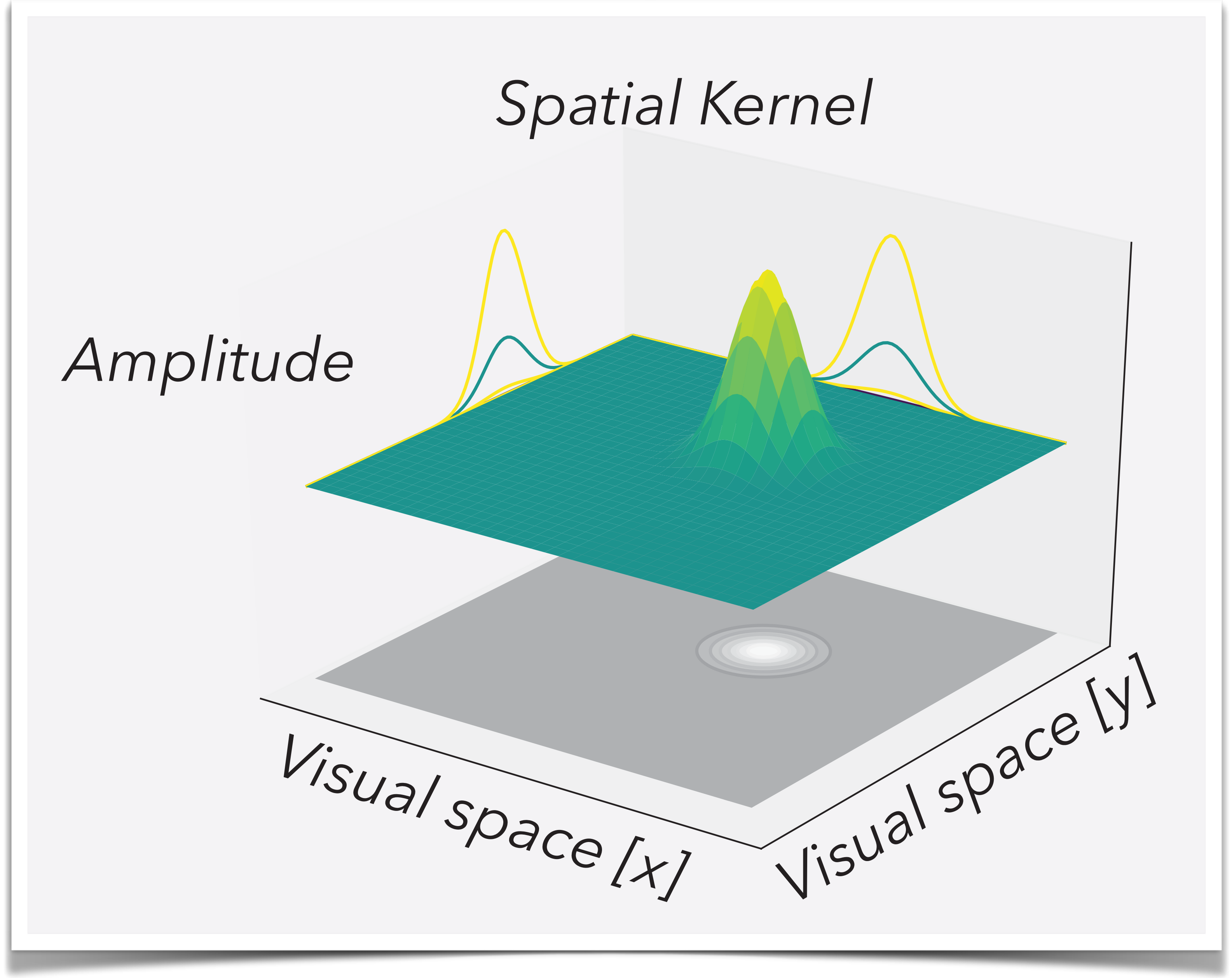
Size



Estimate best-fitting model parameters given the data

Dumoulin & Wandell, 2008

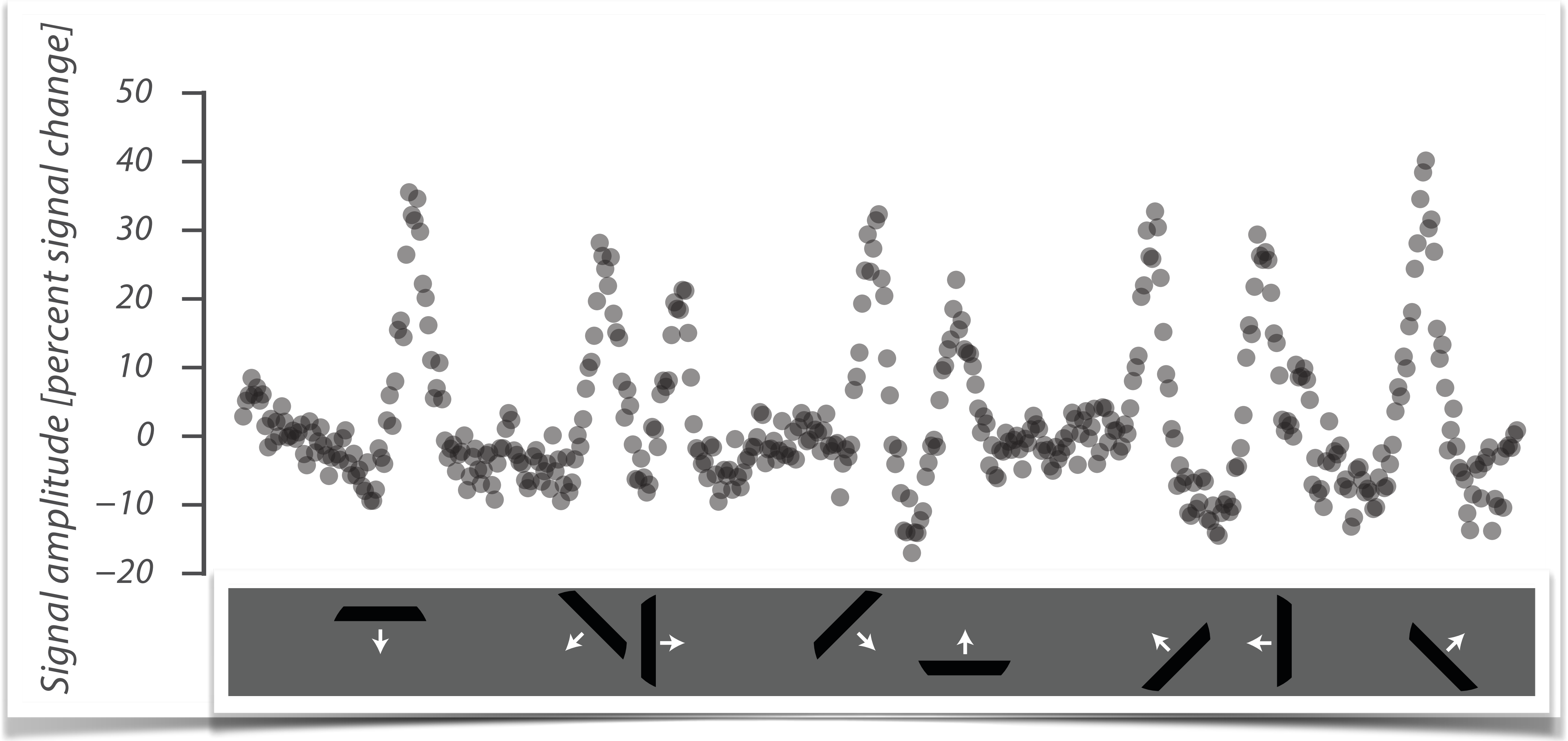
Structured single-voxel BOLD time course: we can ***fit PRF parameters***



Position

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

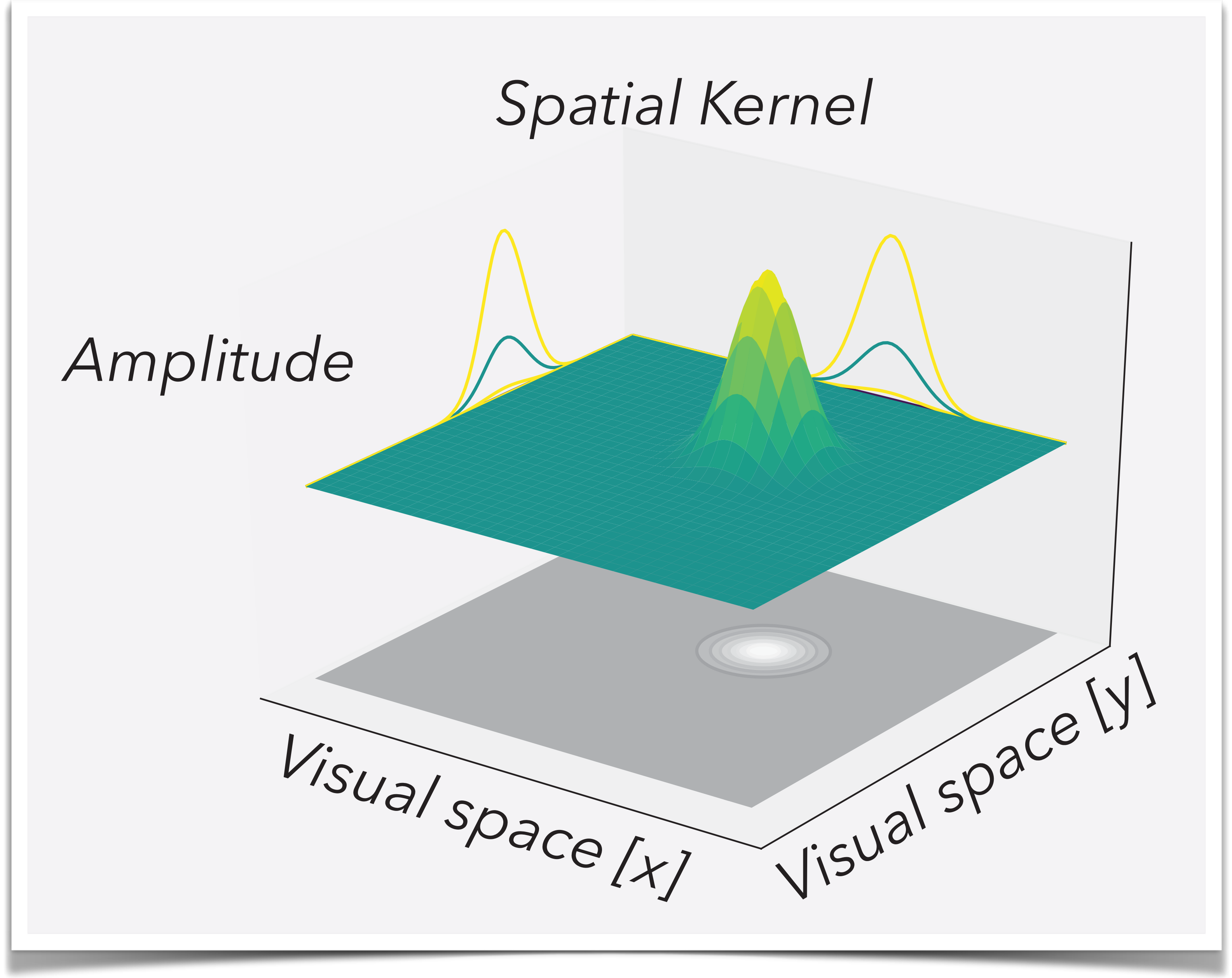
Size



Estimate best-fitting model parameters given the data

Dumoulin & Wandell, 2008

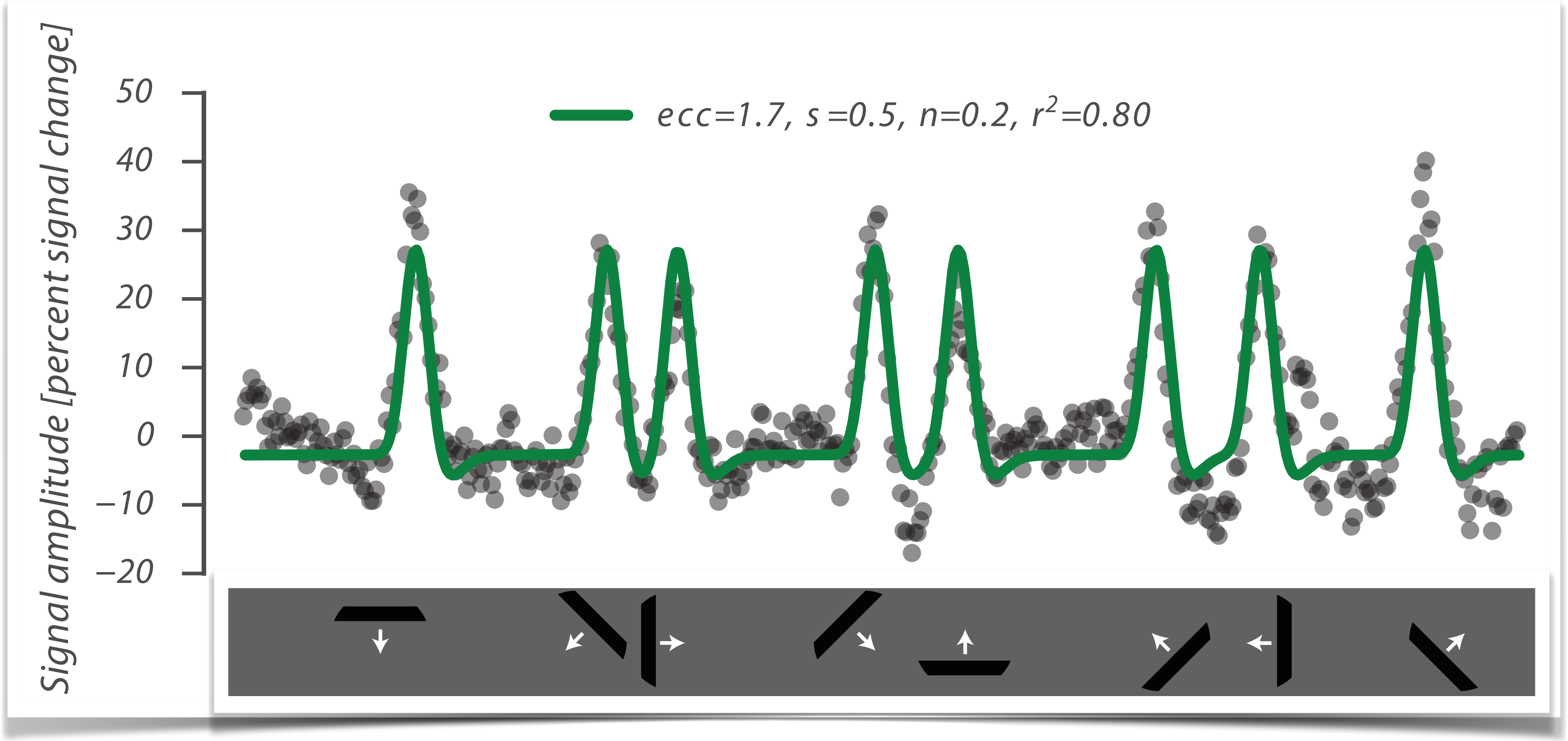
*Structured single-voxel BOLD time course: we can **fit PRF parameters***



Position

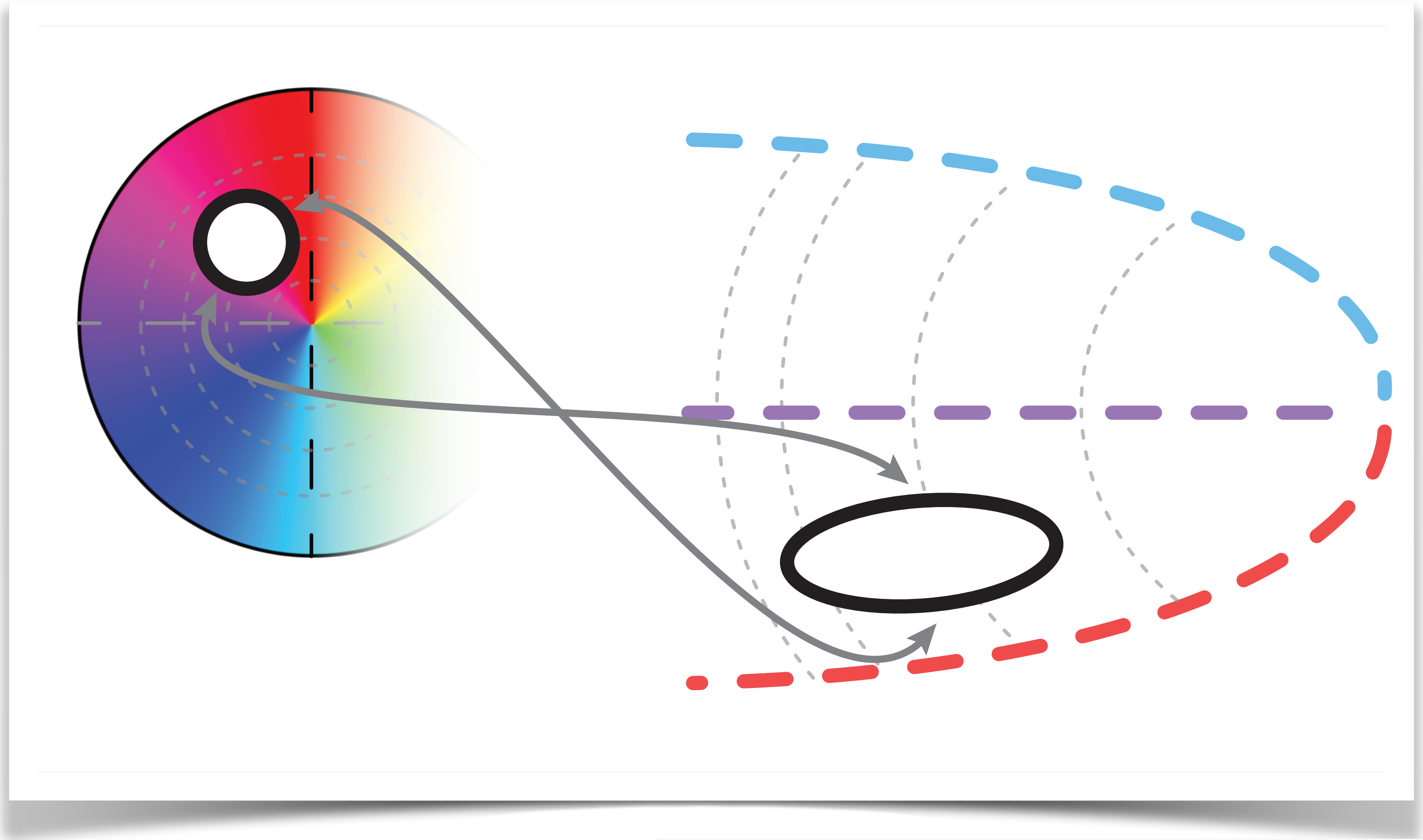
$$g(x_0, y_0, \sigma) = exp(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}),$$

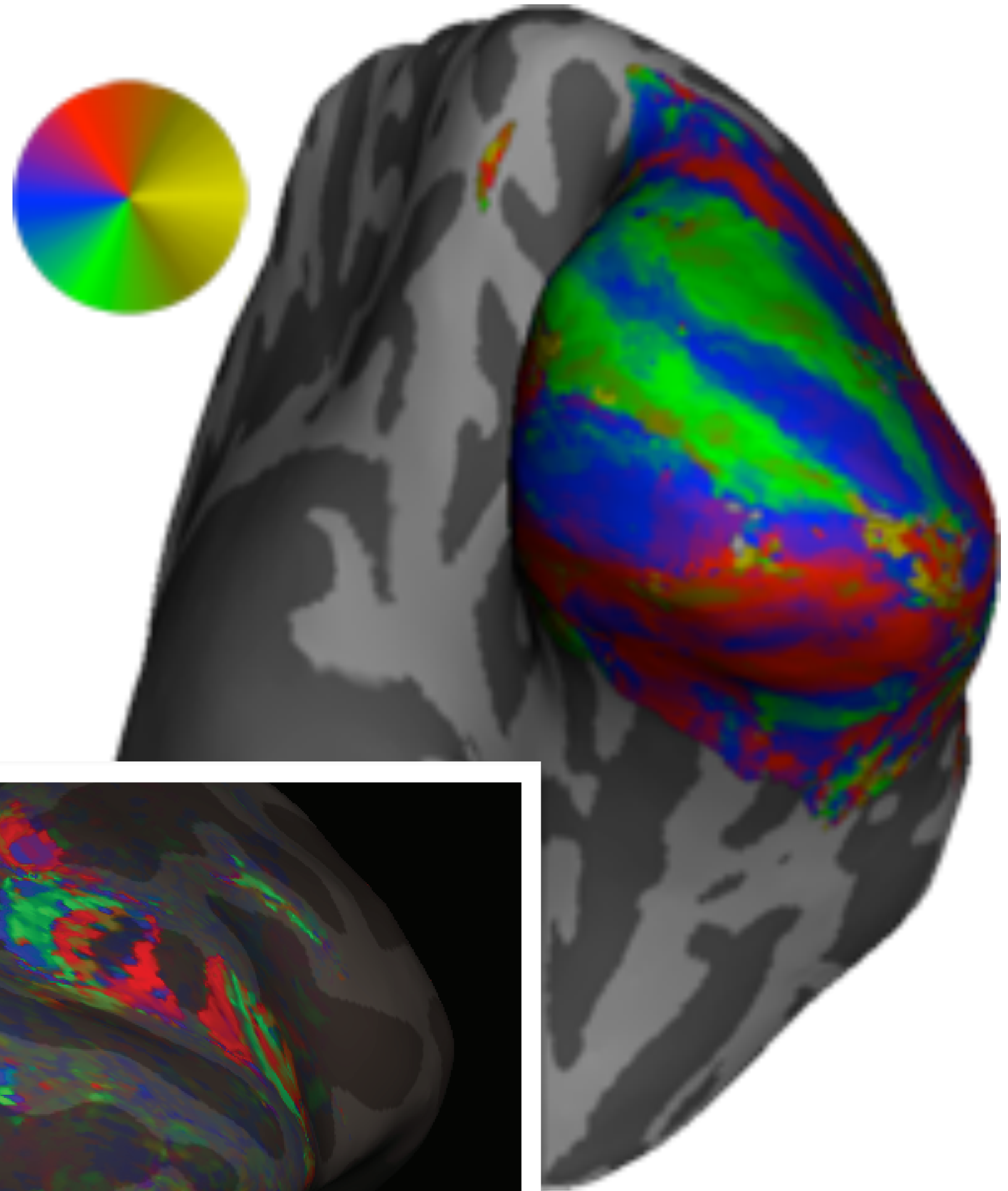
Size

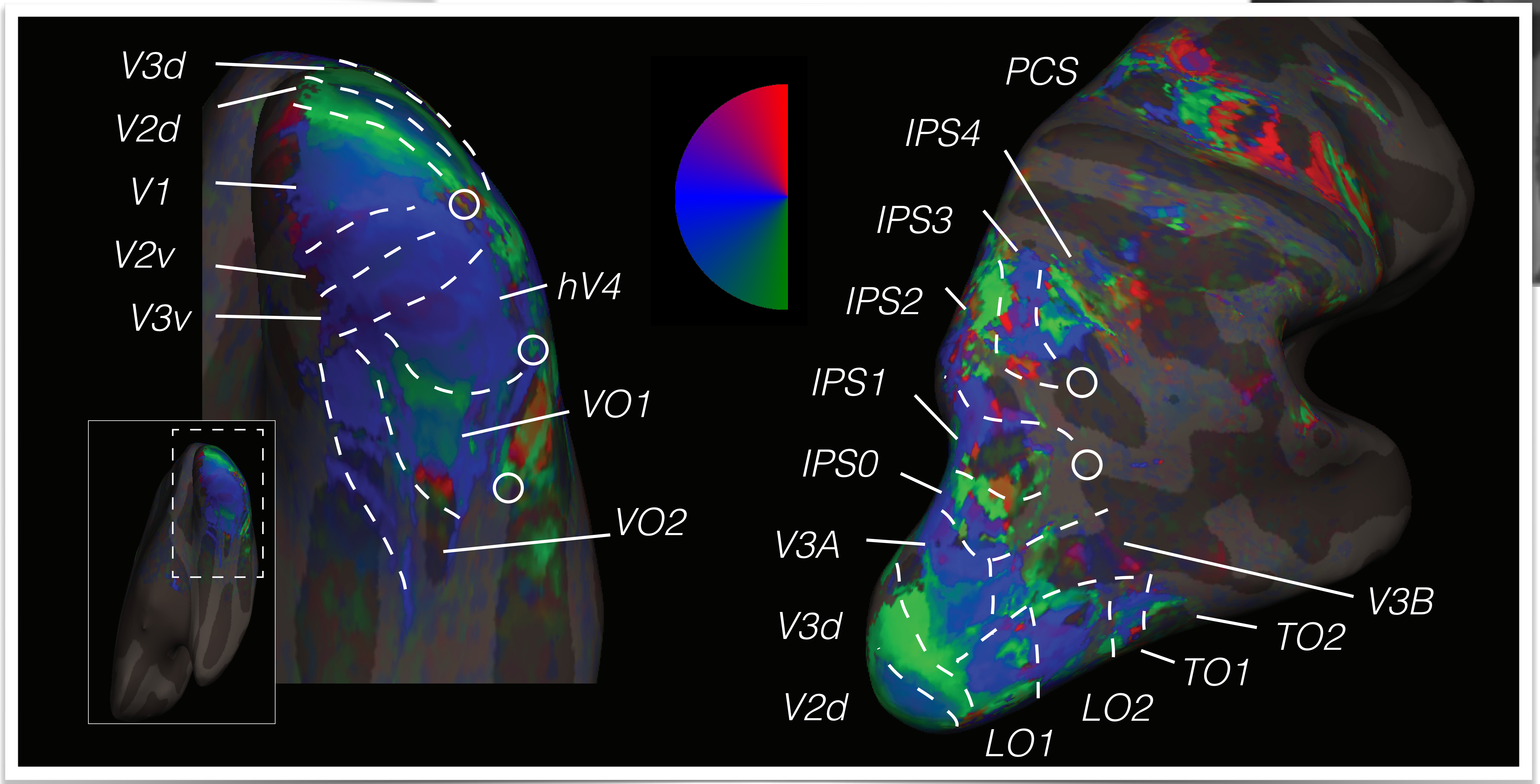


Estimate best-fitting model parameters given the data

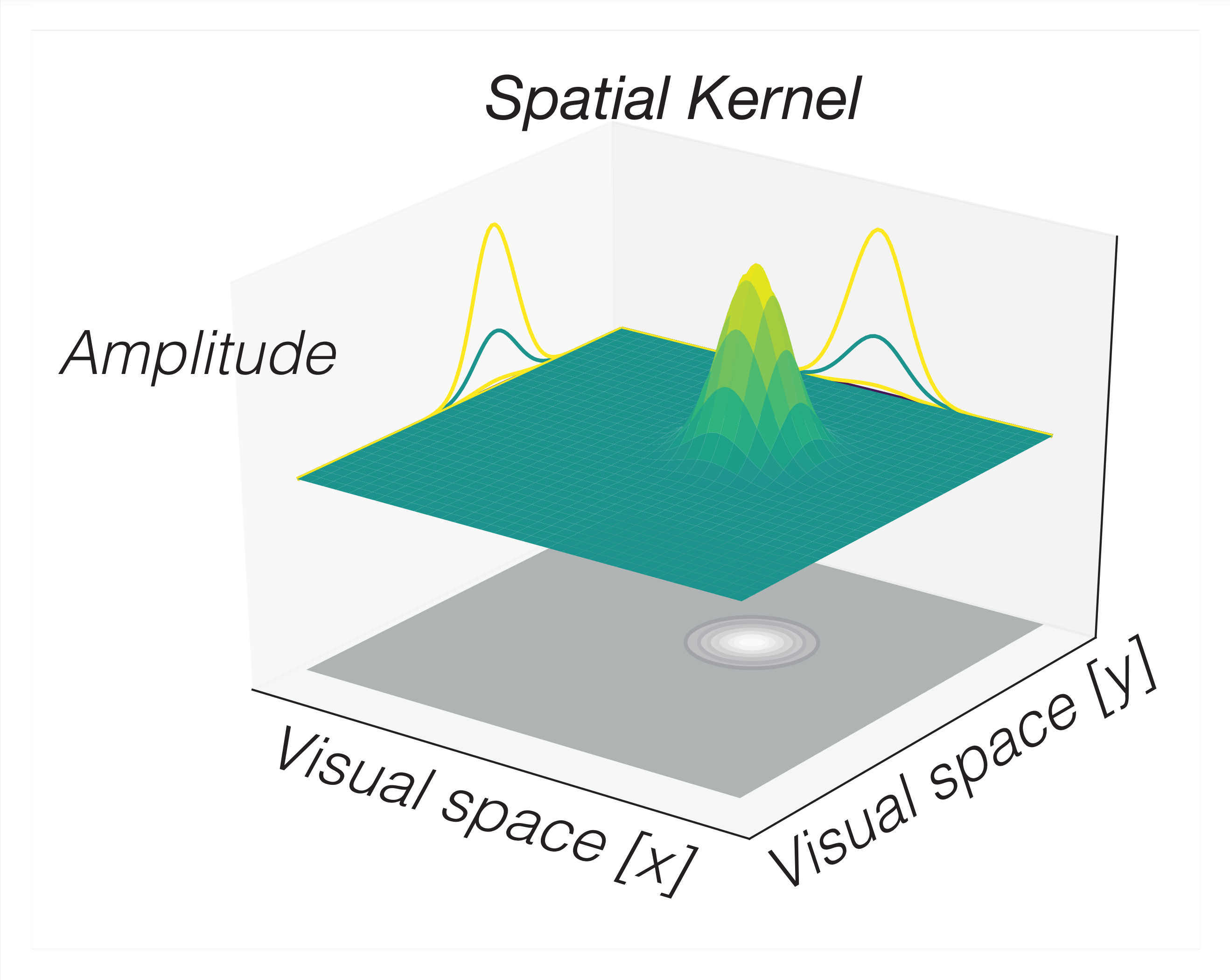
Dumoulin & Wandell, 2008

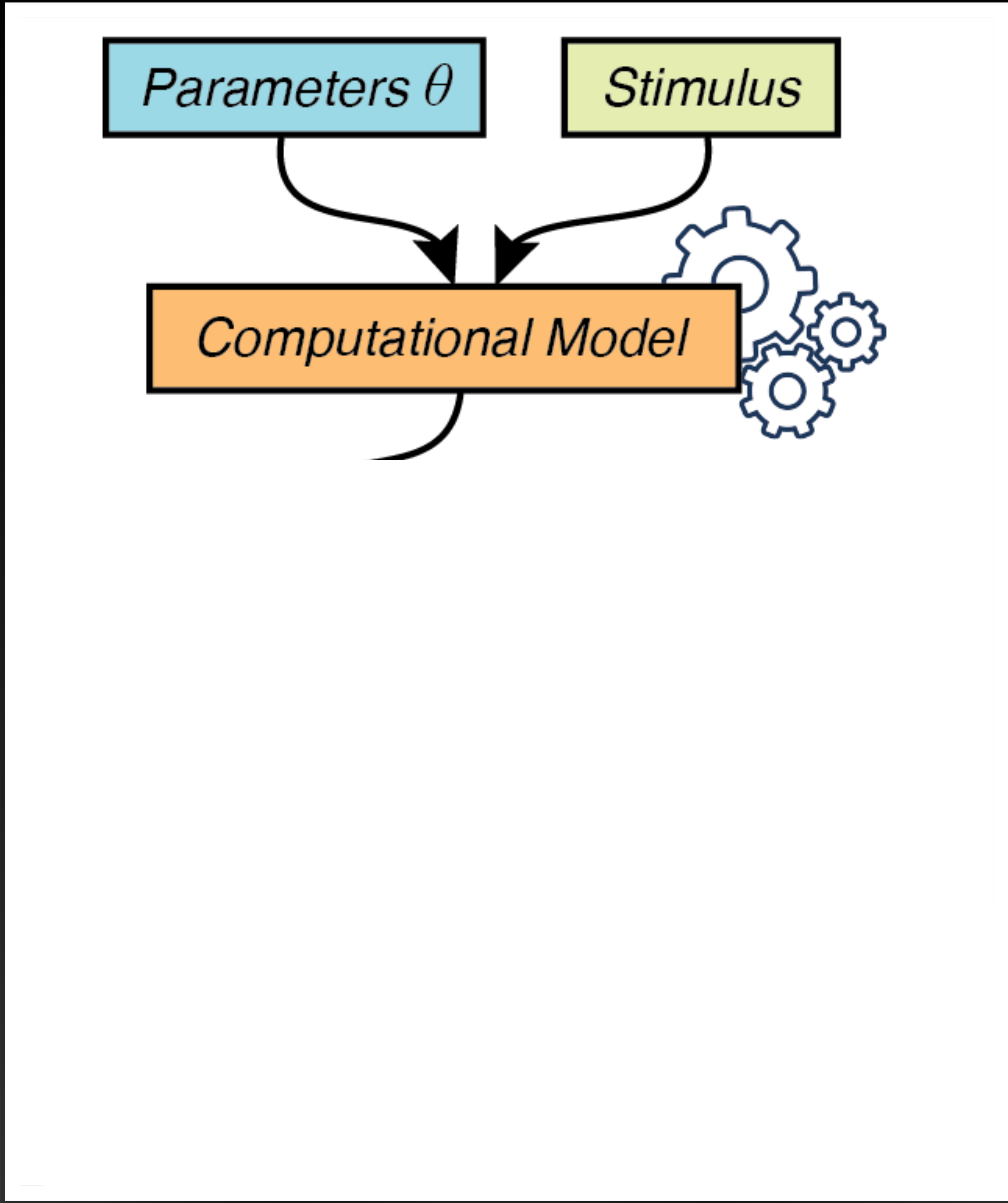




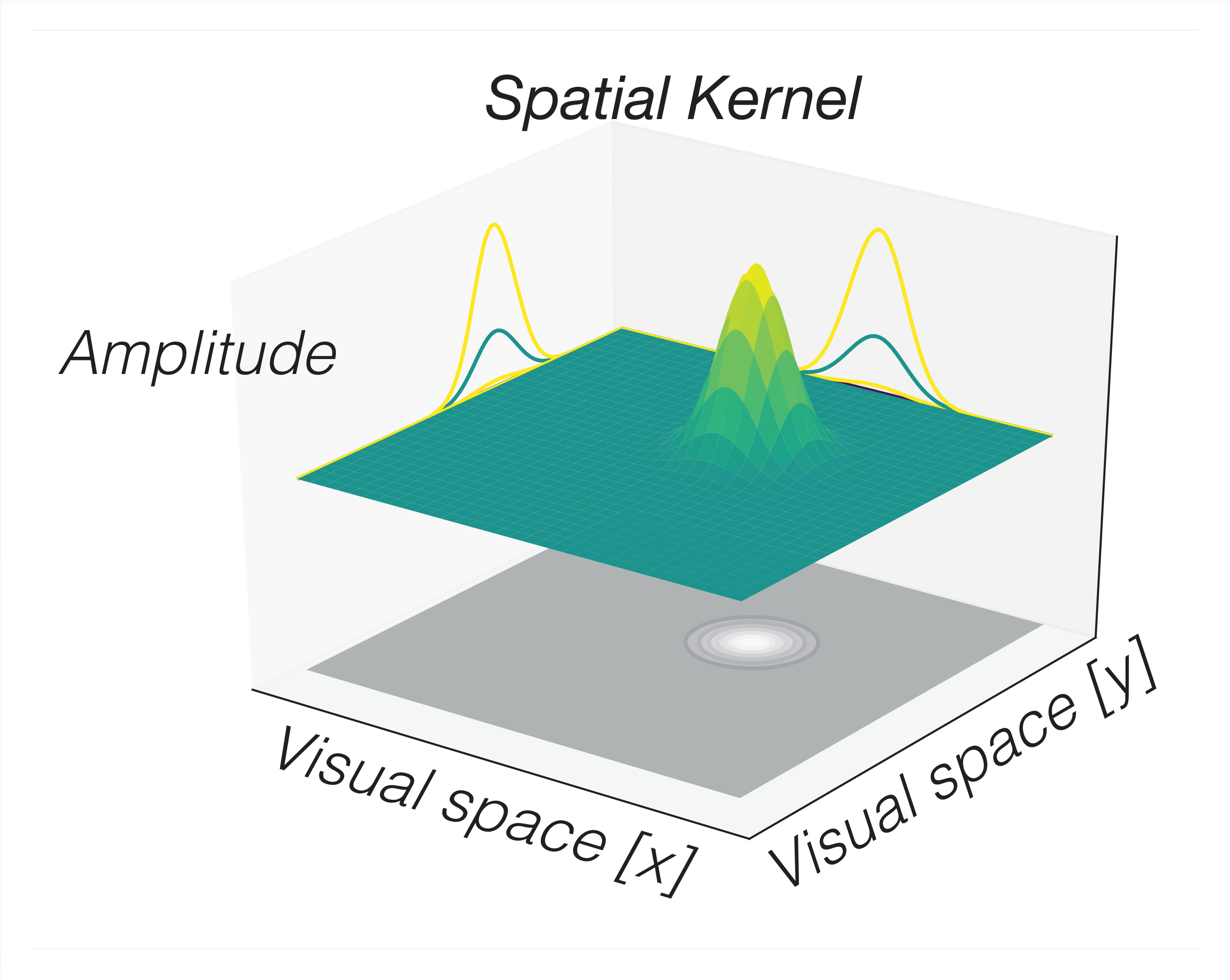

V3d
V2d
V1
V2v
V3v
hV4
VO1
VO2
PCS
IPS4
IPS3
IPS2
IPS1
IPS0
V3A
V3d
V2d
LO1
LO2
TO1
TO2
V3B

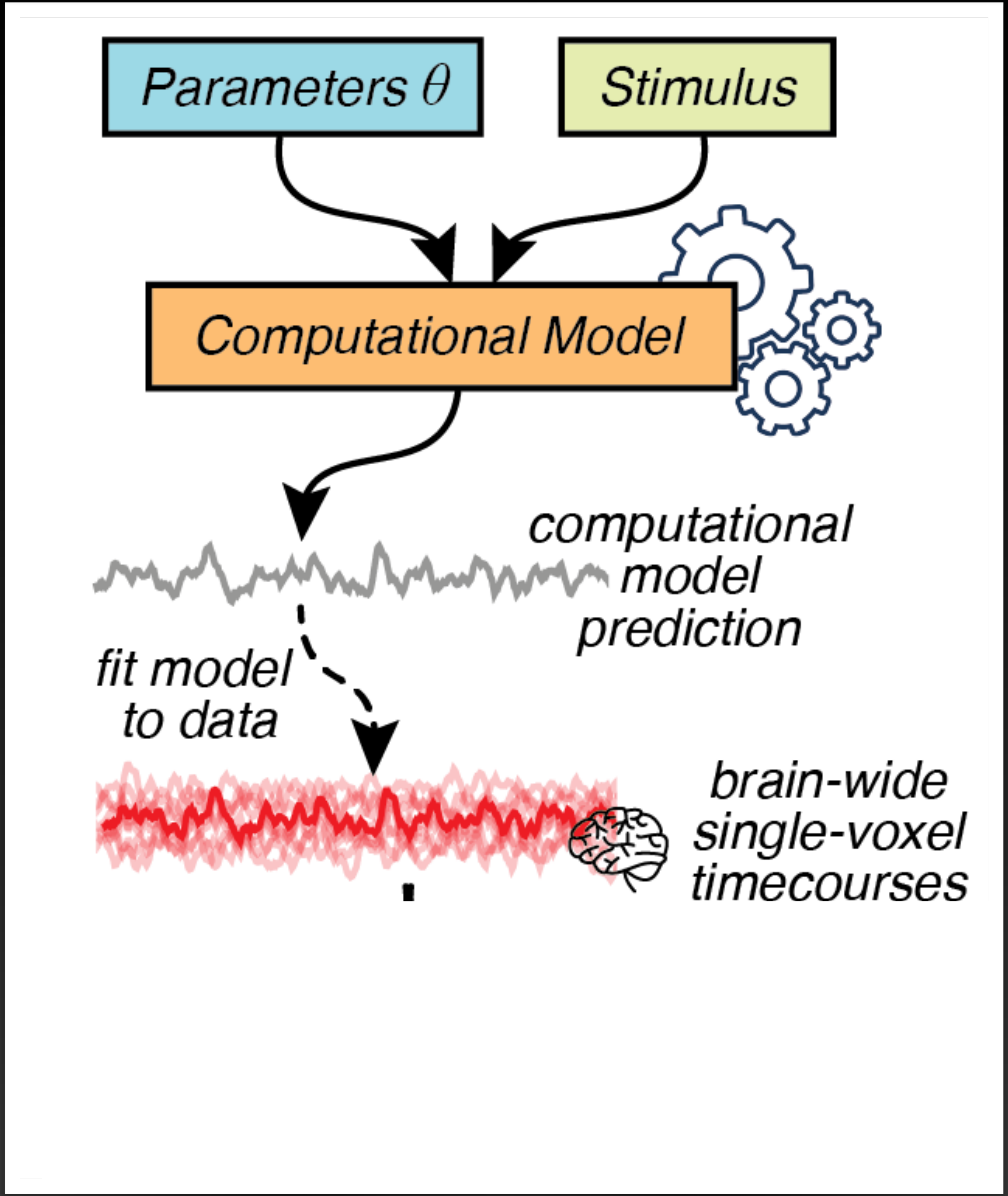
Computational Cognitive Neuroscience



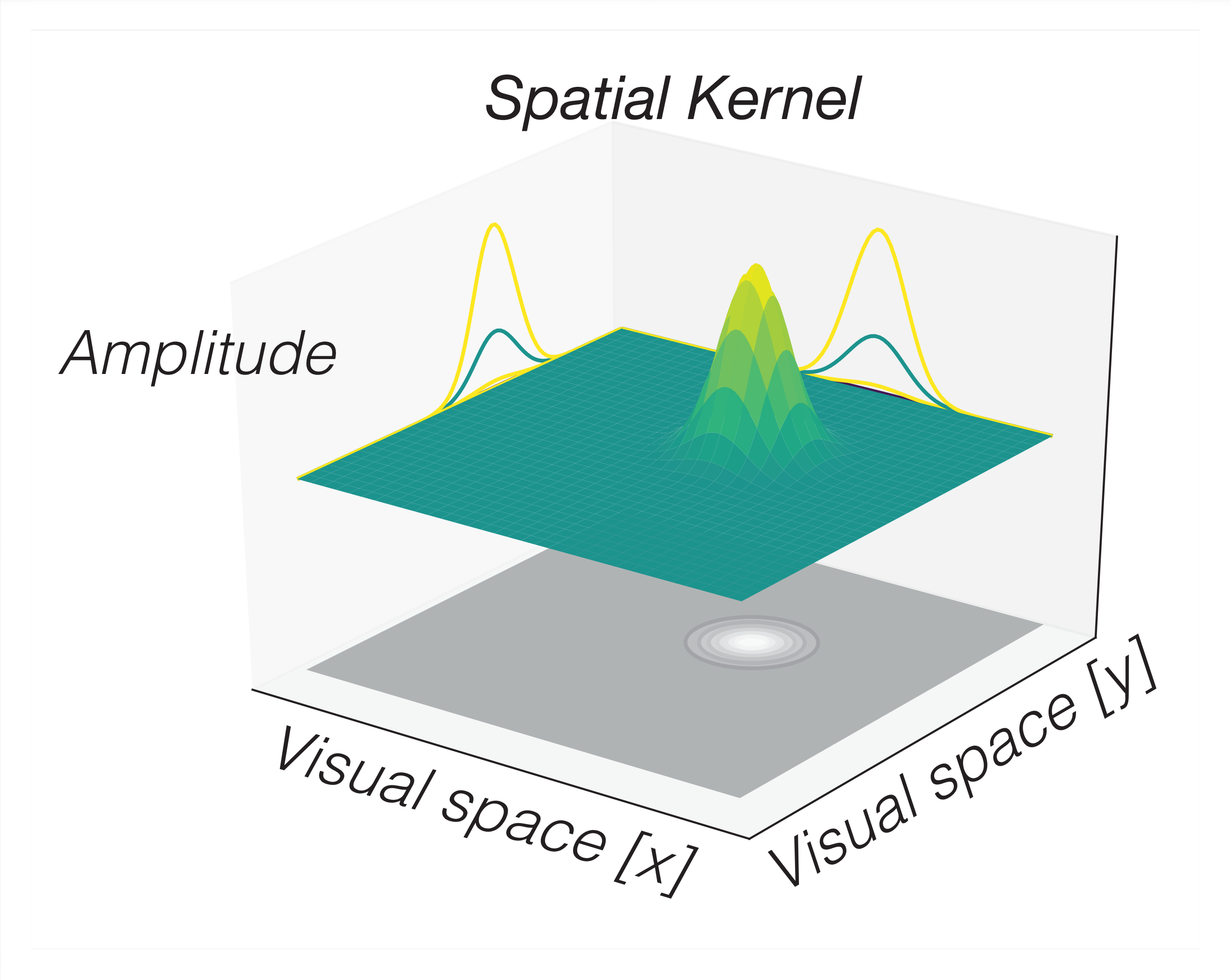


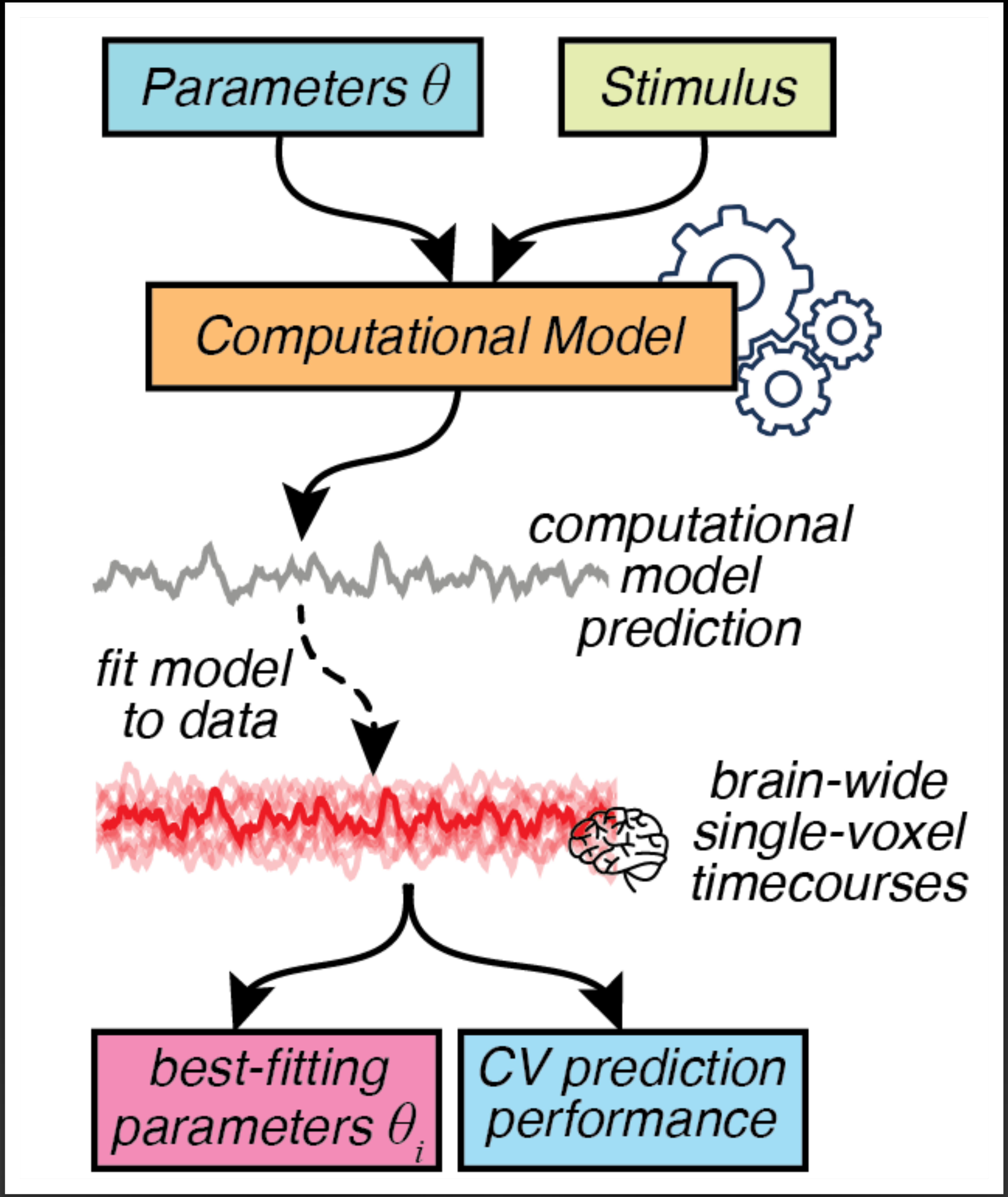
Computational Cognitive Neuroscience



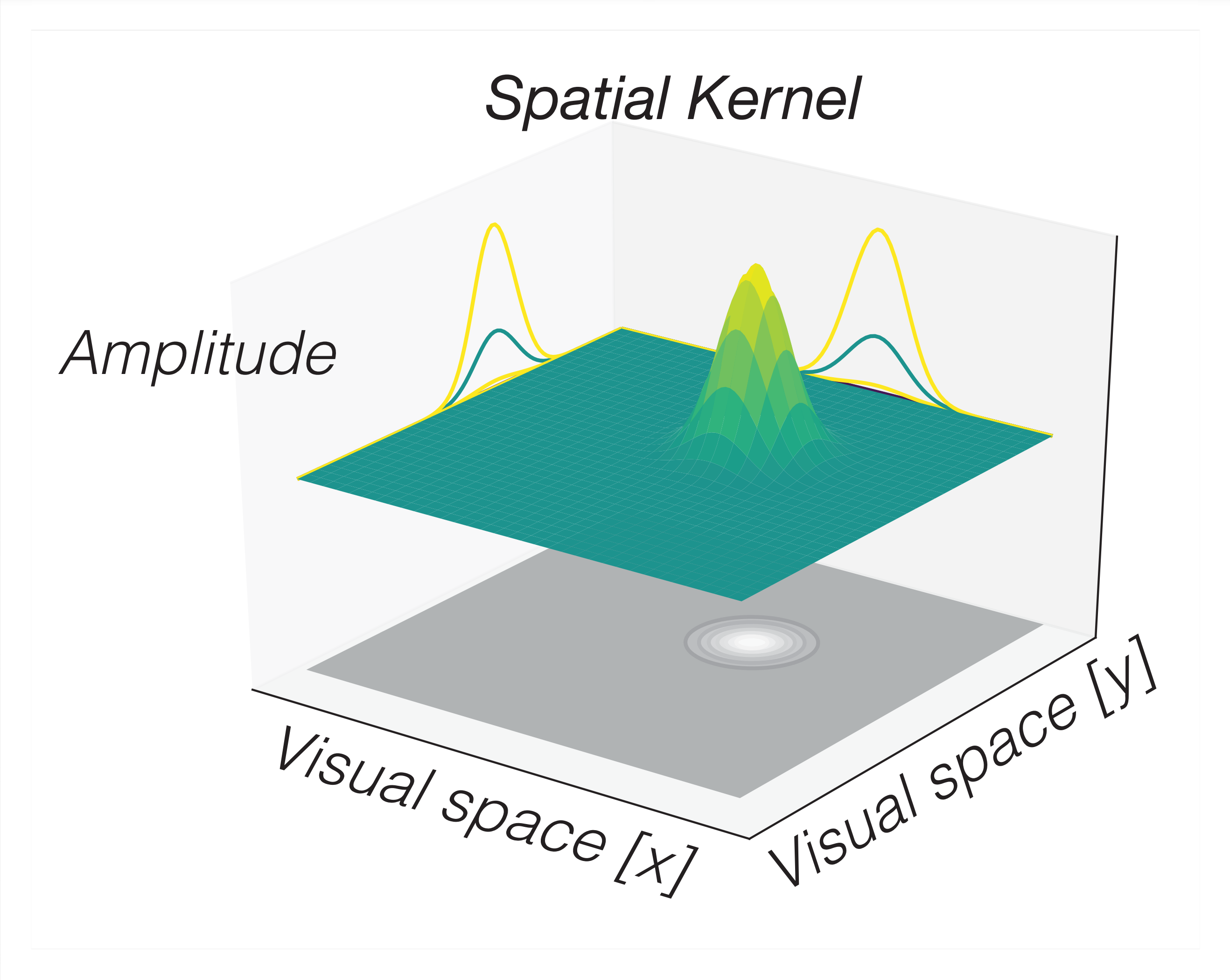


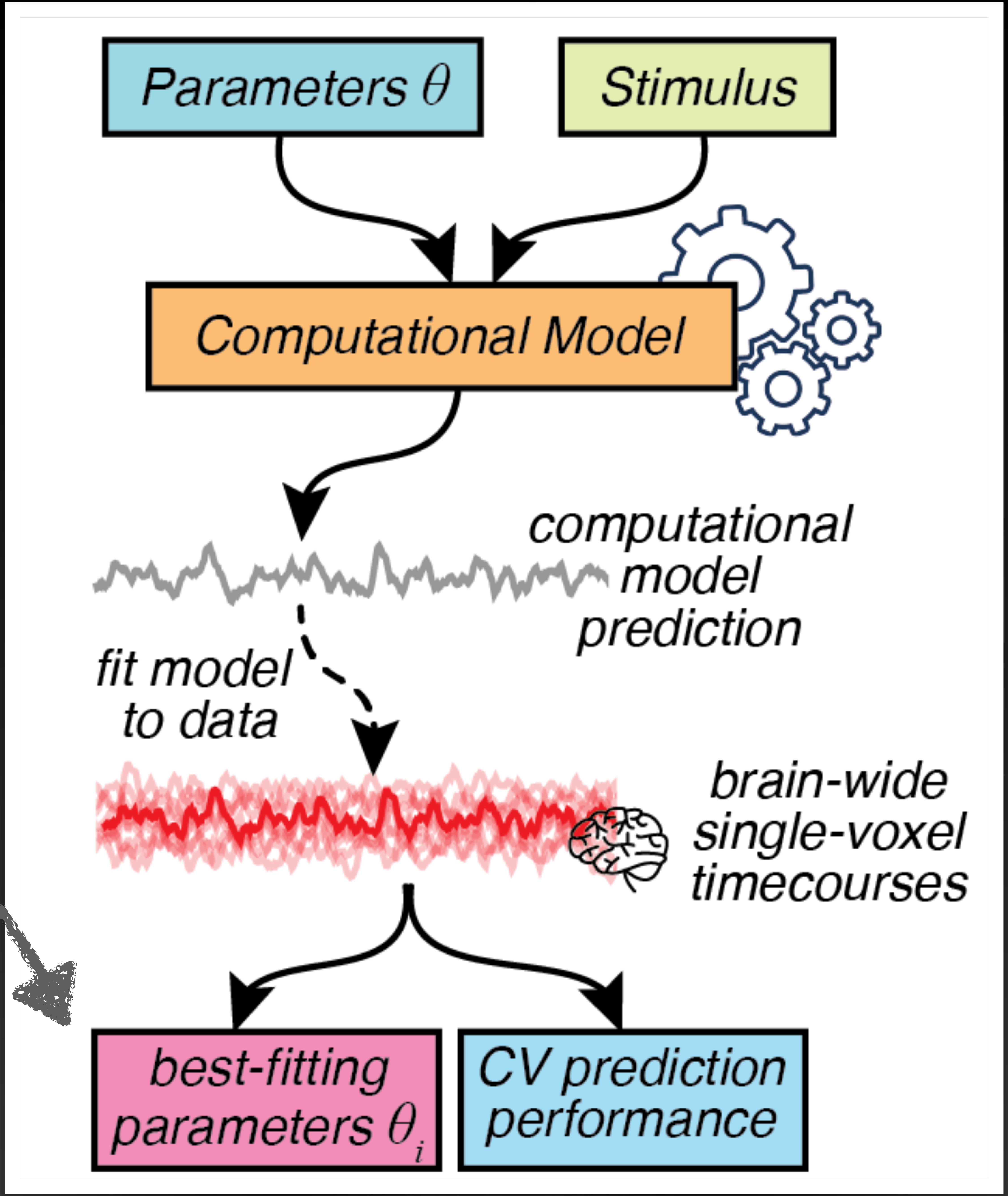
Computational Cognitive Neuroscience





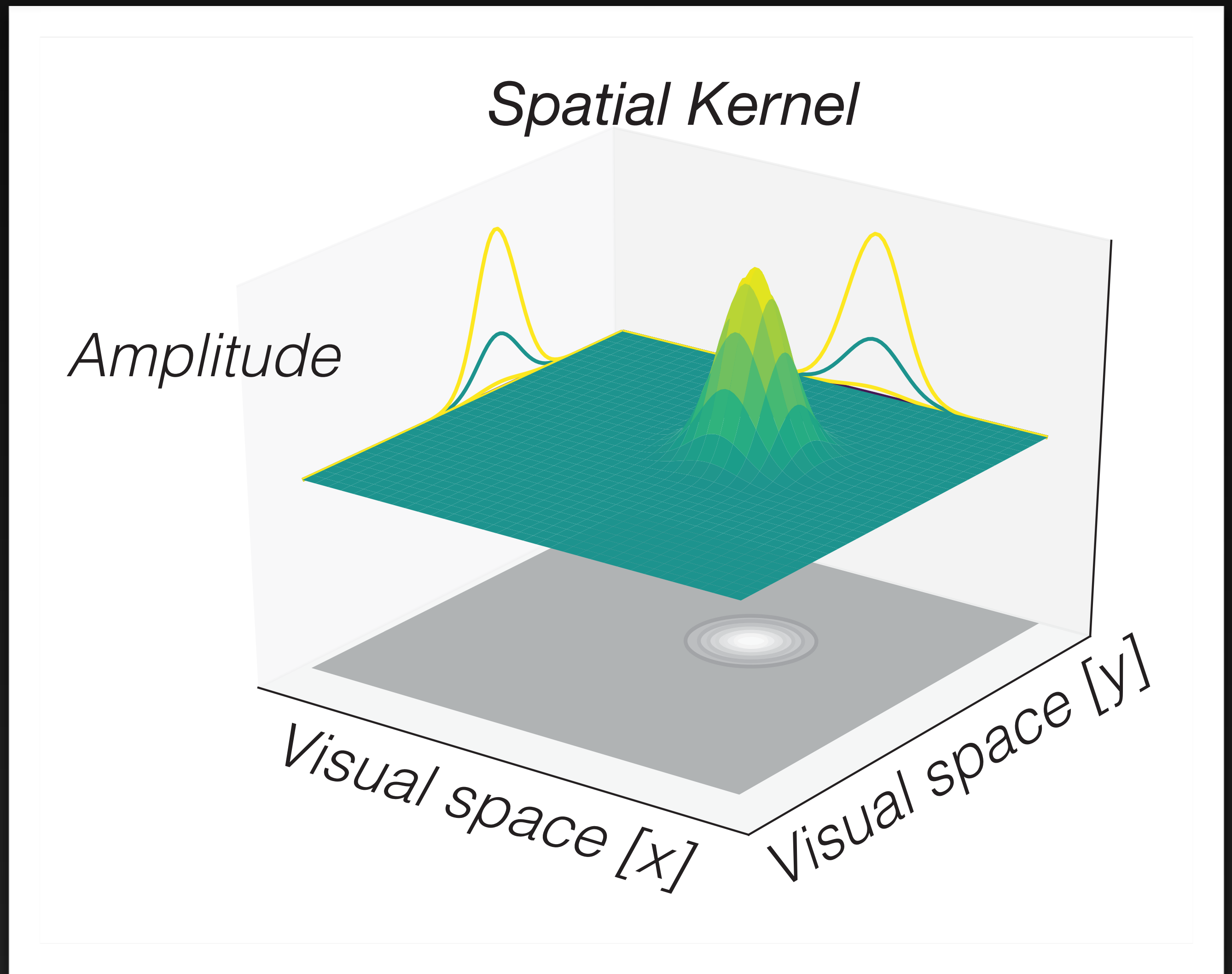
Computational Cognitive Neuroscience

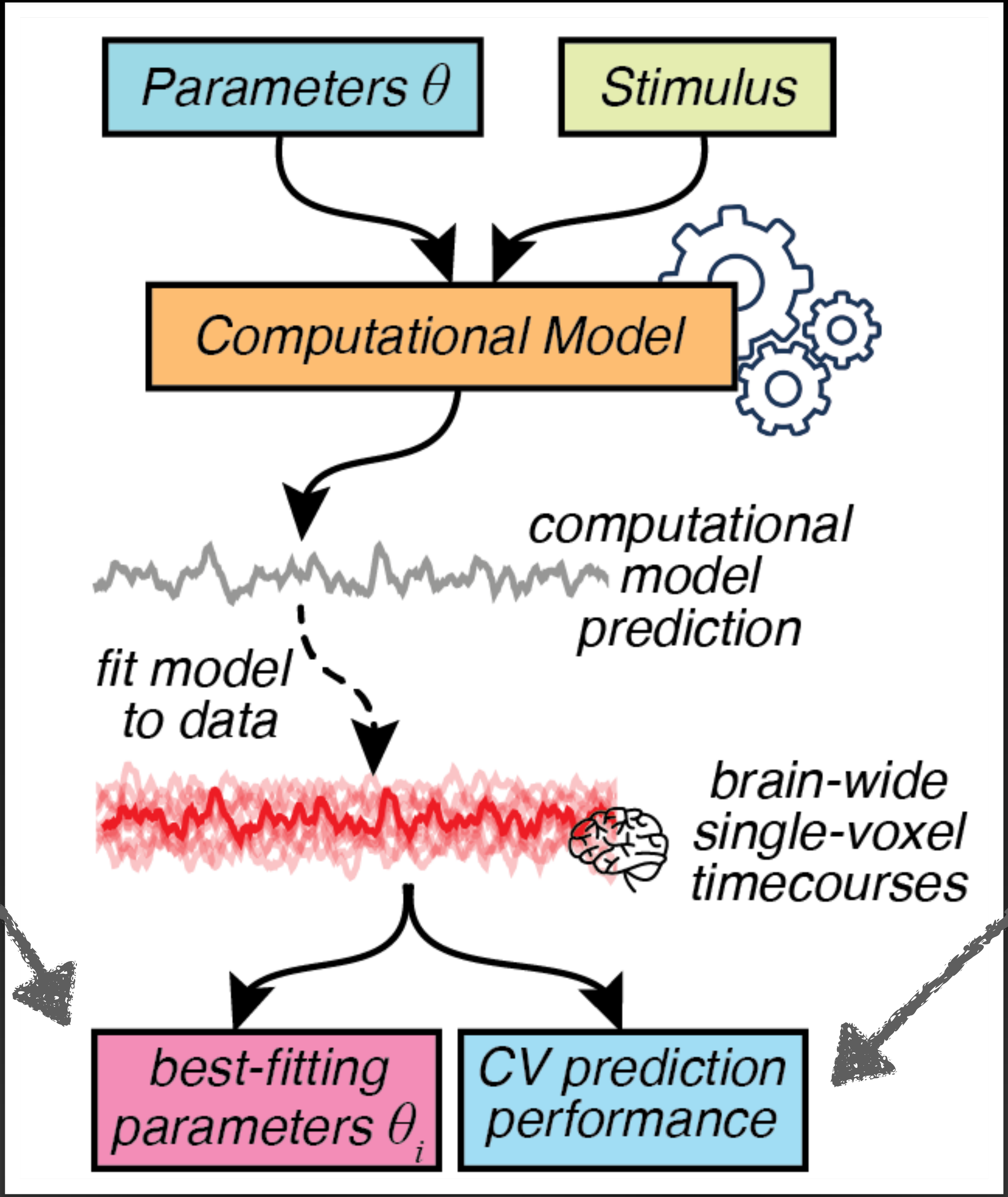




PROBE COMPUTATION, REPRESENTATION

Computational Cognitive Neuroscience





Probe Computation, Representation

Compare Computational Models

Encoding

Kay et al, 2008

Encoding

What is the 'receptive field' of a voxel?

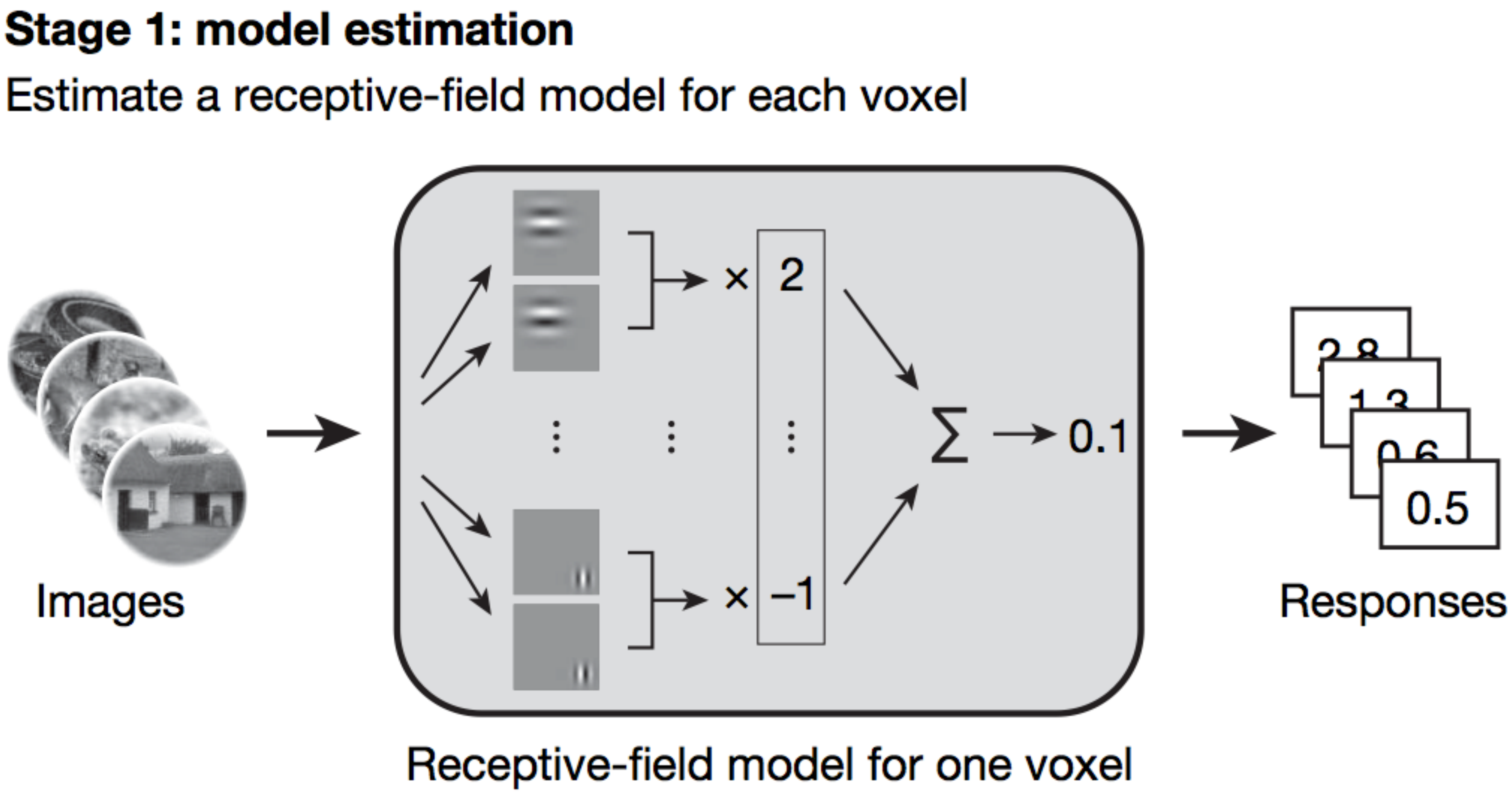
Let's expand to more complex receptive fields

Kay et al, 2008

Encoding

What is the ‘receptive field’ of a voxel?

Let’s expand to more complex receptive fields



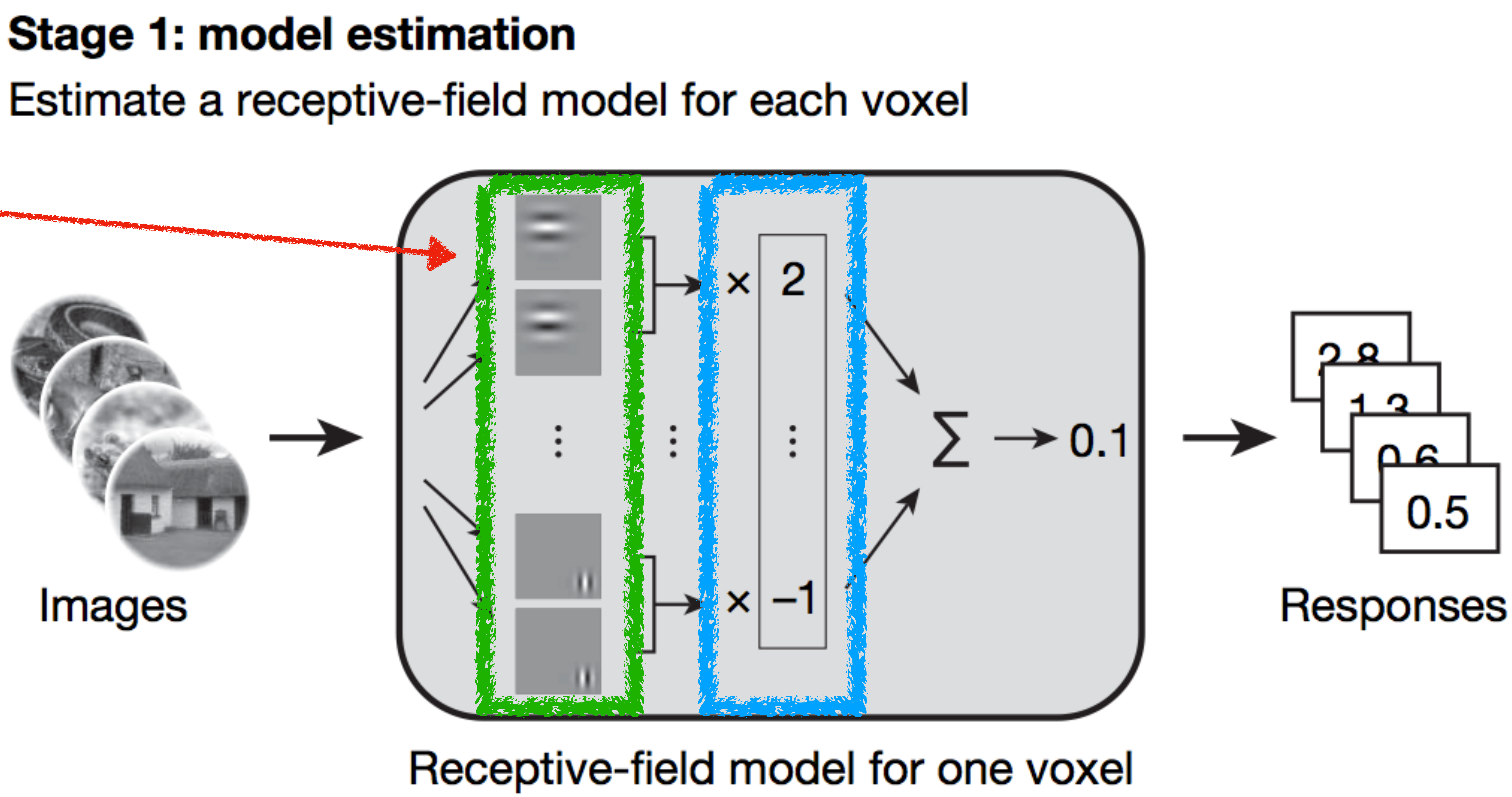
Kay et al, 2008

Encoding

What is the 'receptive field' of a voxel?

Let's expand to more complex receptive fields

Many different models and their predictions



Kay et al, 2008

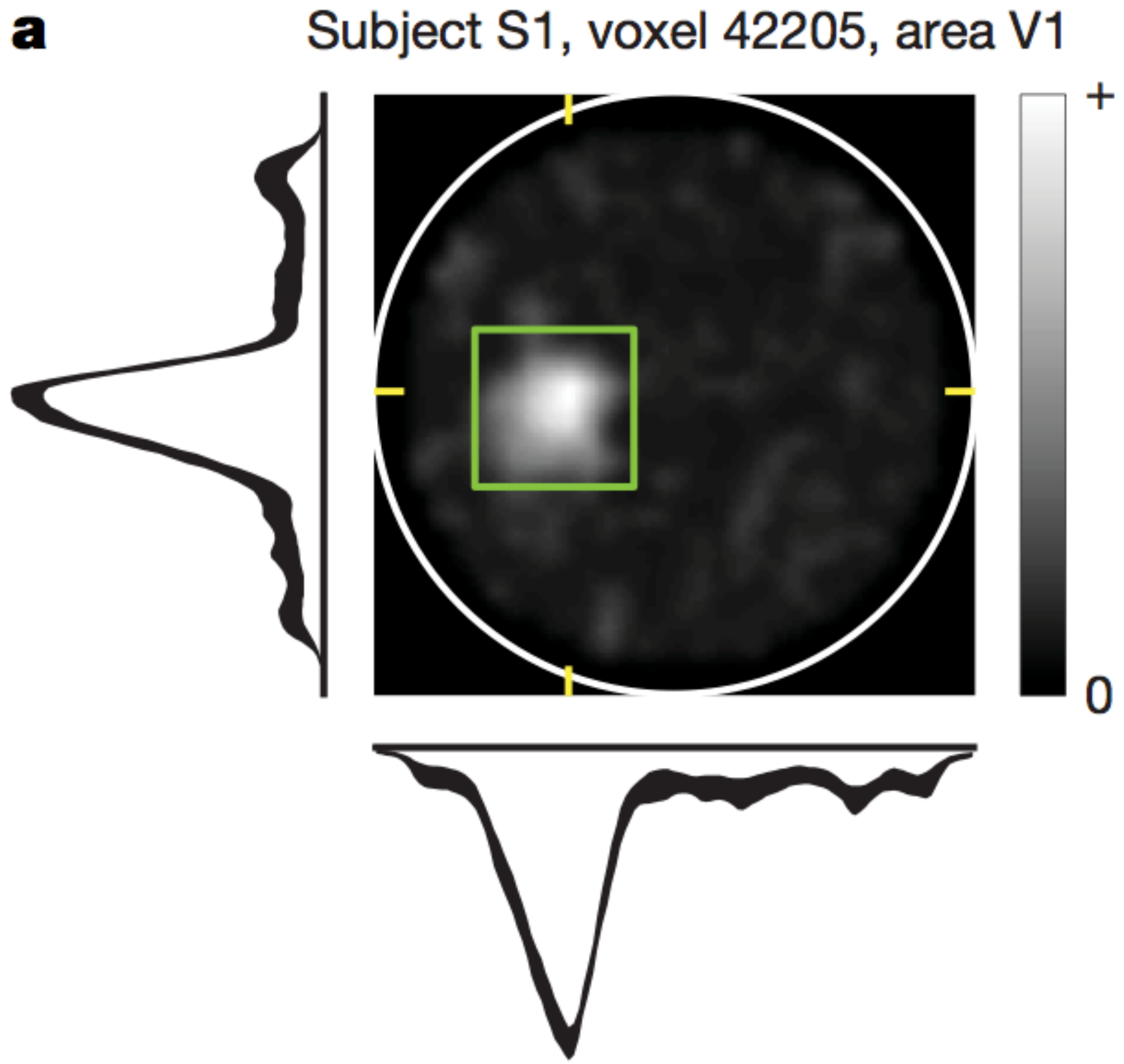
Encoding: adding orientation

What is the ‘receptive field’ of a voxel?

Kay et al, 2008

Encoding: adding orientation

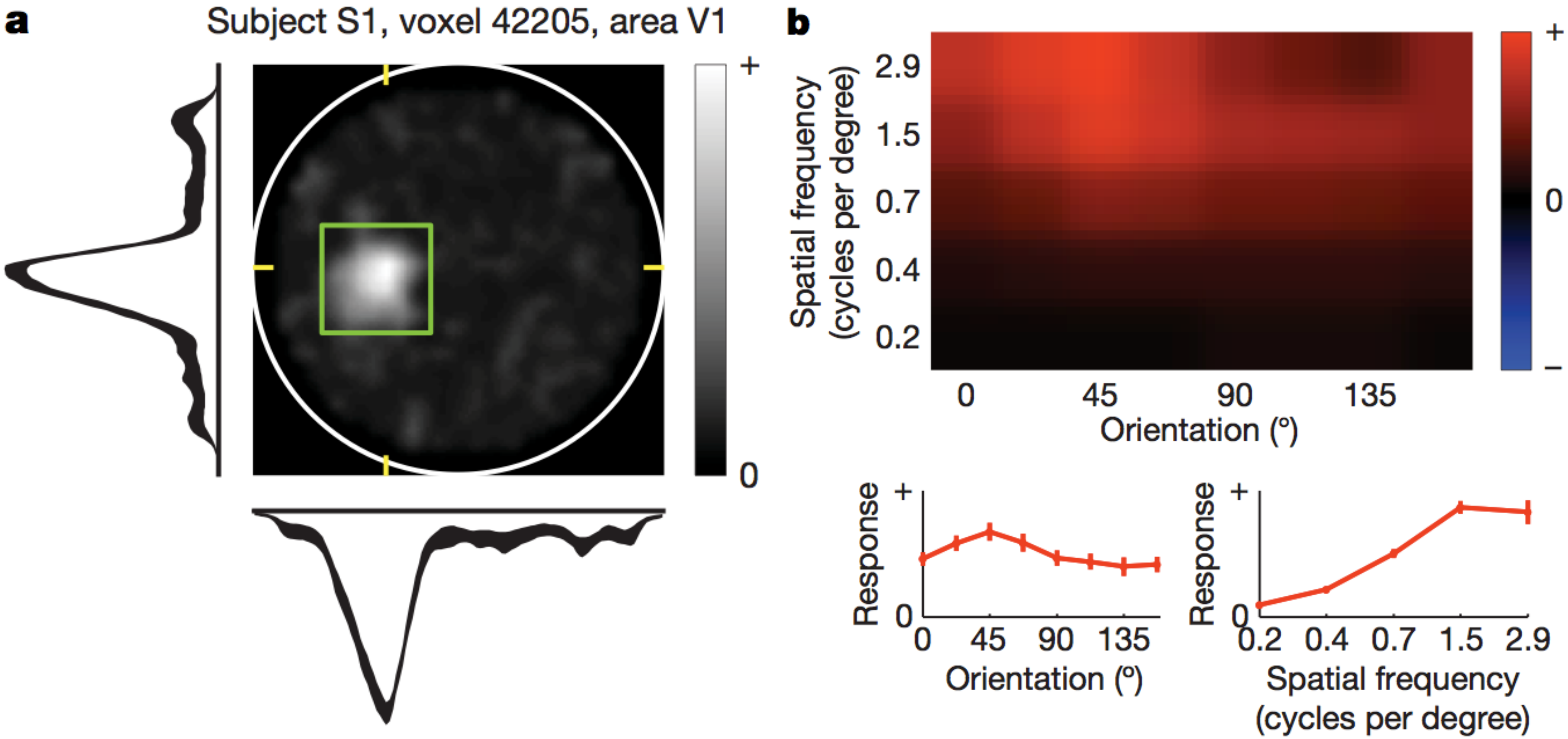
What is the 'receptive field' of a voxel?



Kay et al, 2008

Encoding: adding orientation

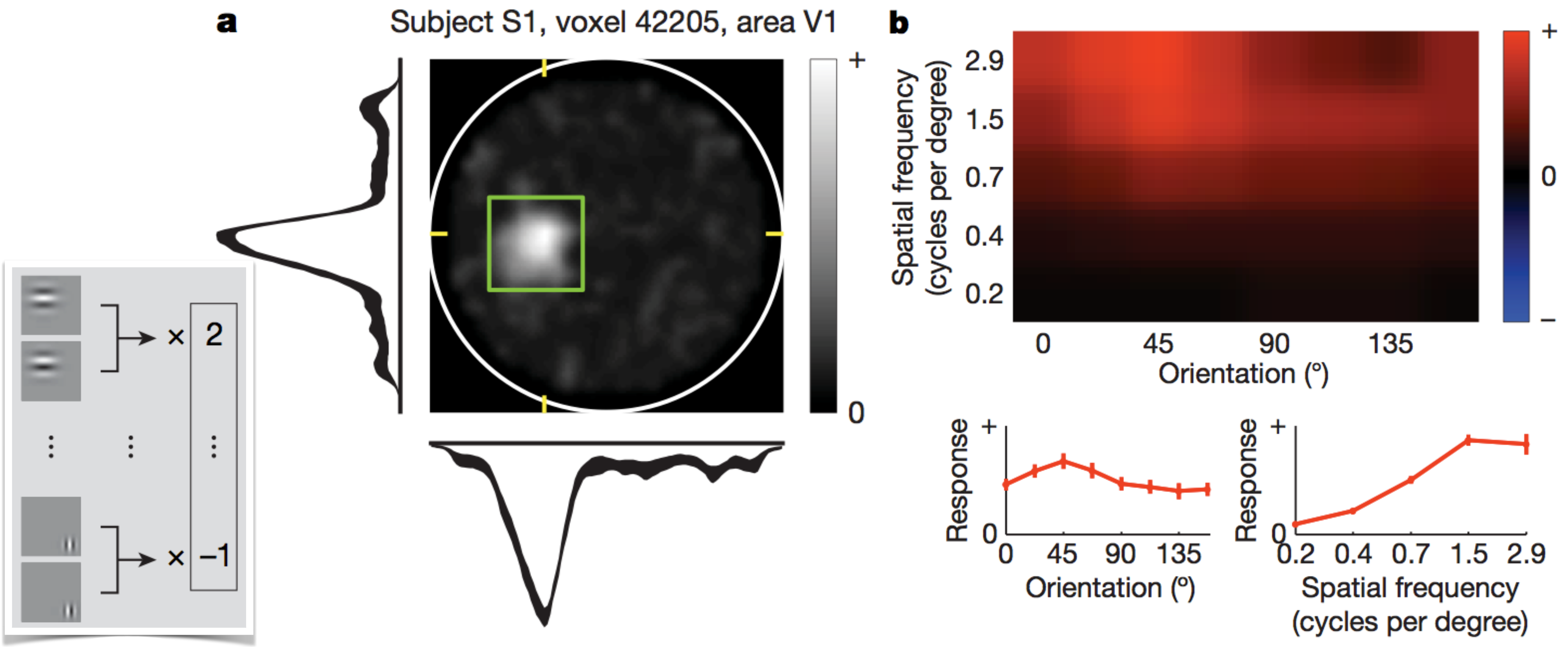
What is the ‘receptive field’ of a voxel?



Kay et al, 2008

Encoding: adding orientation

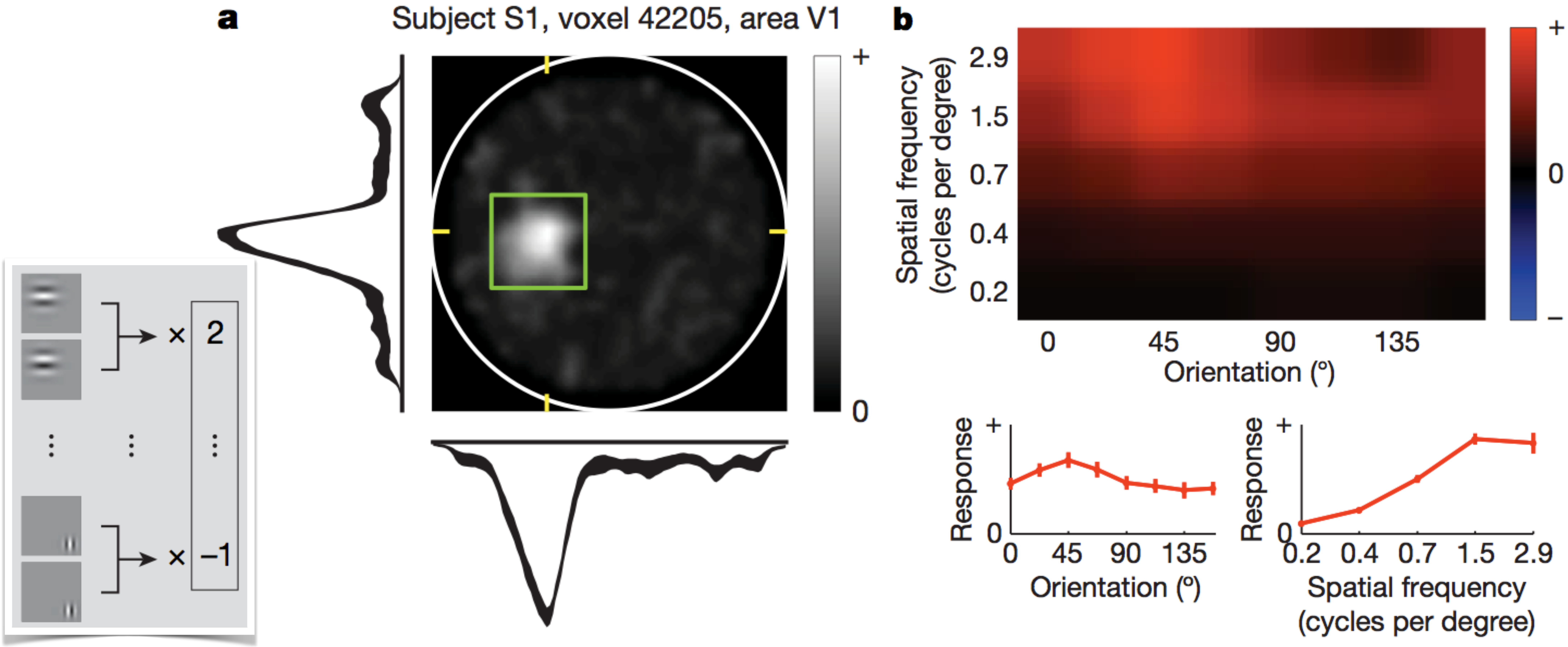
What is the 'receptive field' of a voxel?



Kay et al, 2008

Encoding: adding orientation

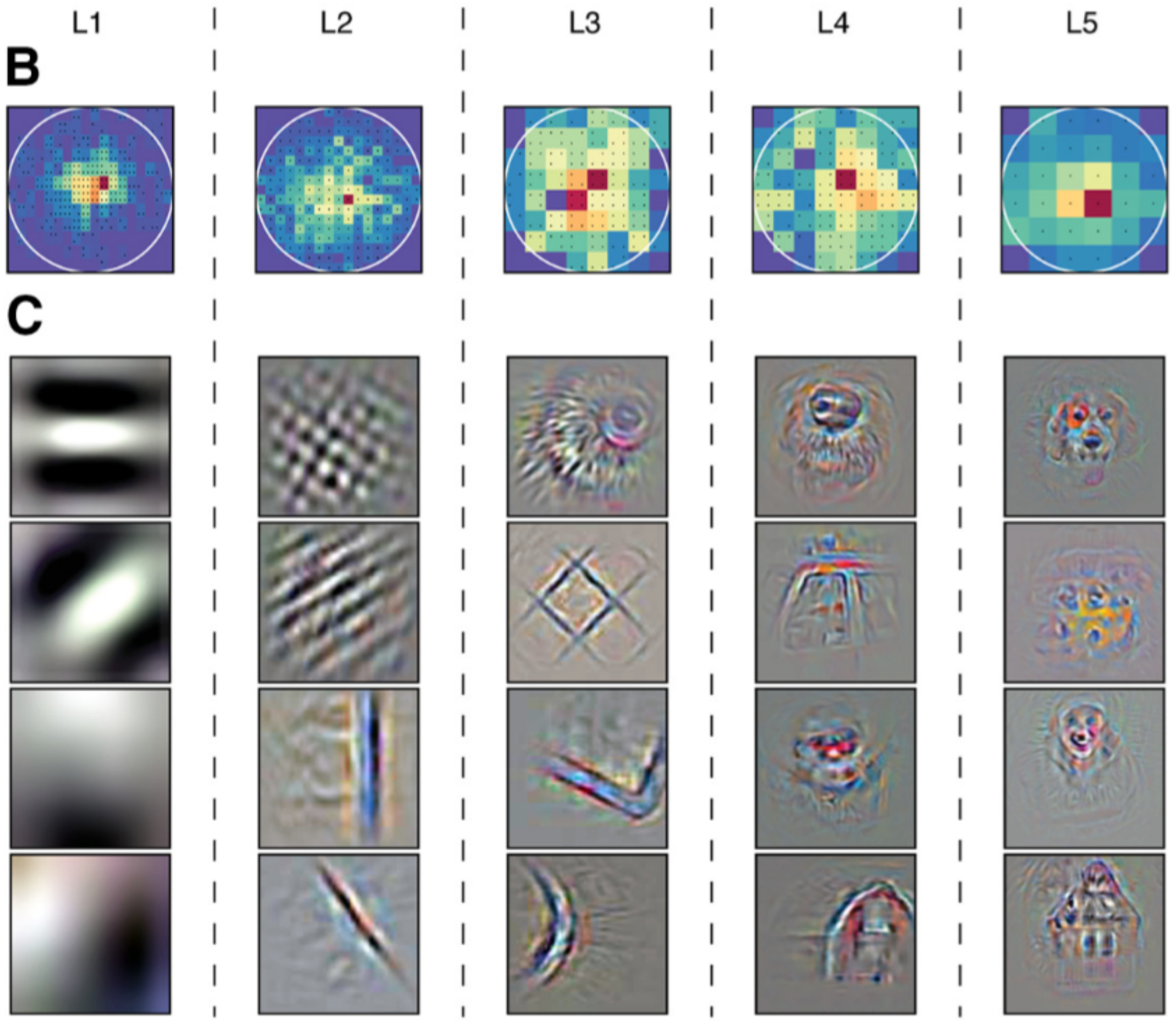
What is the 'receptive field' of a voxel?

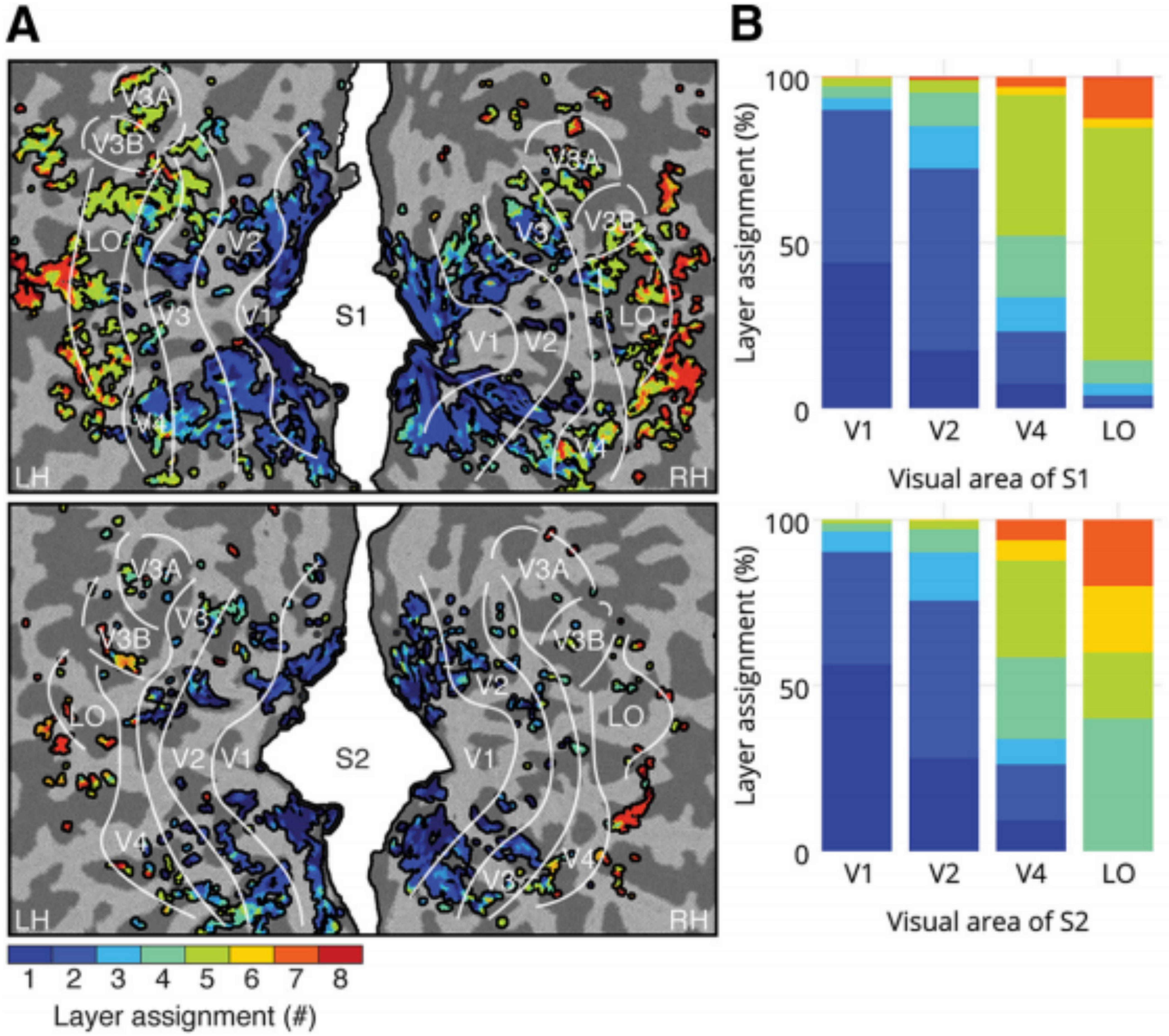


Estimate a huge amount of parameters, using ***penalised regression***

Kay et al, 2008

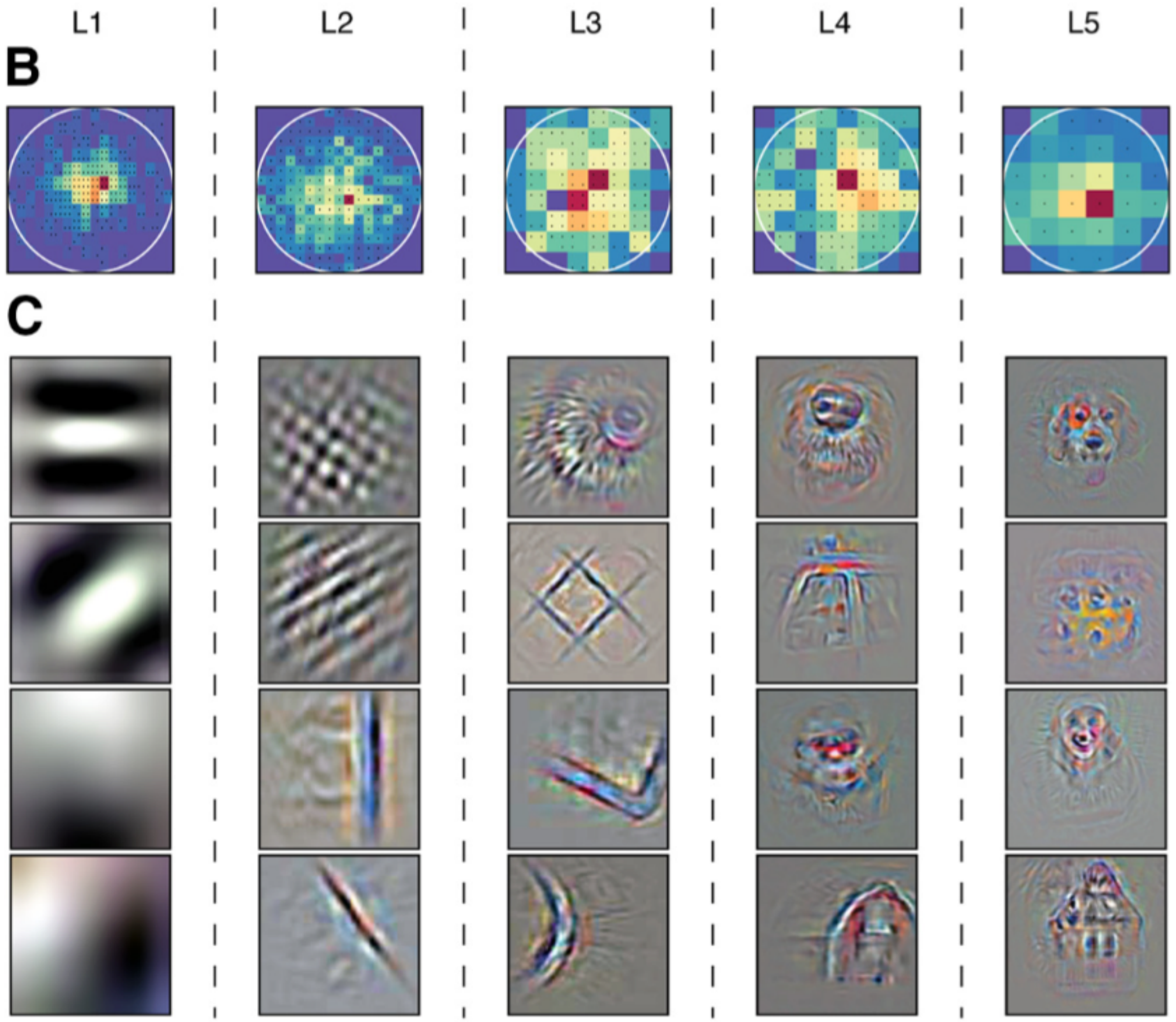
Neural Networks as encoding model

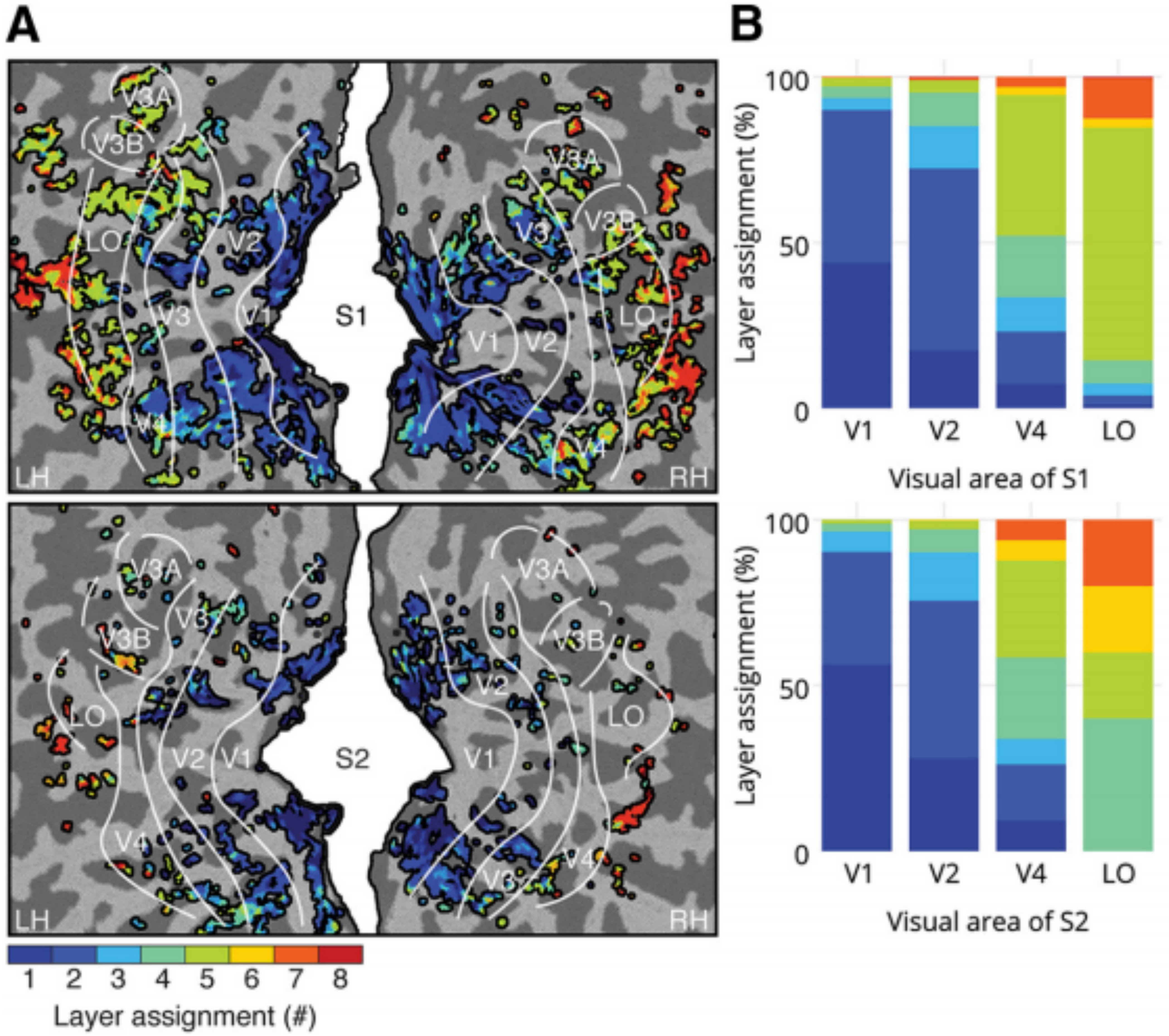




Güçlü U, van Gerven MAJ. 2015

Neural Networks as encoding model





Güçlü U, van Gerven MAJ. 2015

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

Penalised Regression

Ordinary Least Squares

Only stable when # regressors < # timepoints
(when X is *full rank*)

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

When # regressors > # timepoints,
inverting $(X^T X)$ becomes unstable

Penalised Regression

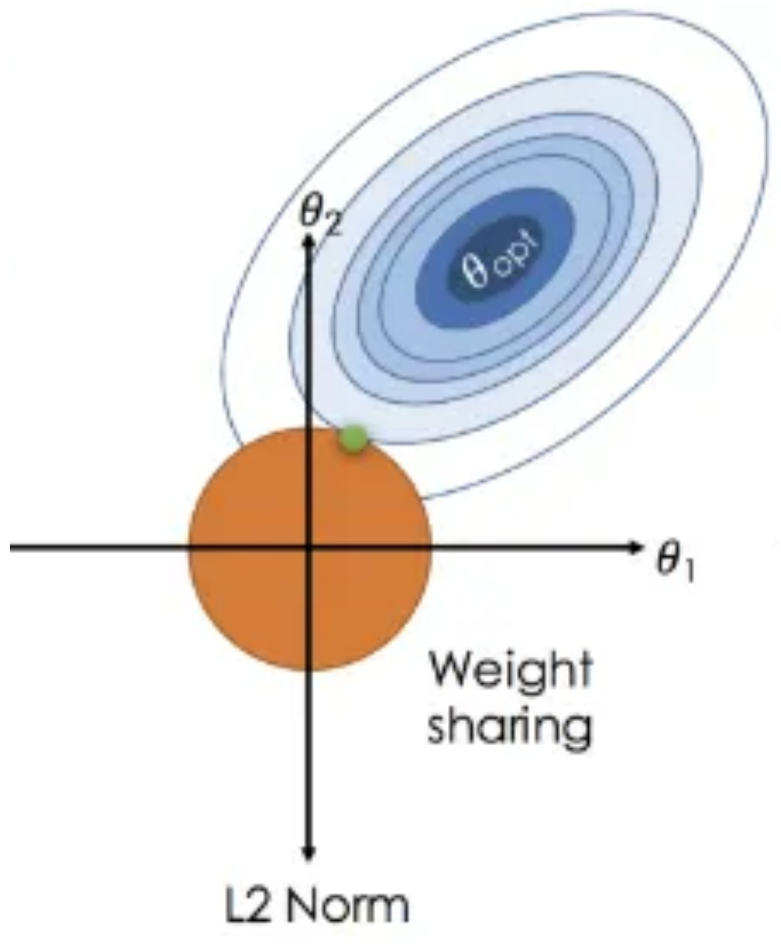
Ordinary Least Squares

Only stable when # regressors < # timepoints
(when X is *full rank*)

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

When # regressors > # timepoints,
inverting $(X^T X)$ becomes unstable

Ridge (Tikhonov or L2 Norm) Regression

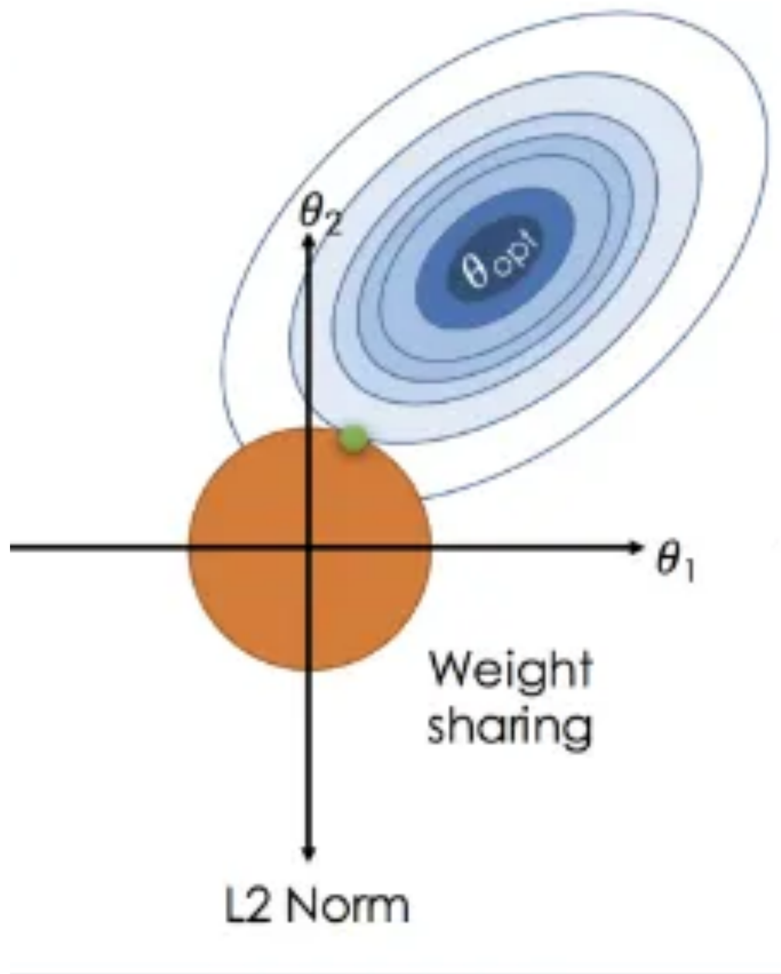


'Penalises' β weight values, forcing them towards 0

$$\hat{\beta}_{ridge} = (X^T X + \lambda I)^{-1} X^T y$$

With $\lambda = 0$, we have OLS.
The higher λ, the stronger the penalisation

Penalised Regression



Ordinary Least Squares

Only stable when # regressors < # timepoints
(when X is *full rank*)

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

When # regressors > # timepoints,
inverting $(X^T X)$ becomes unstable

Ridge (Tikhonov or L2 Norm) Regression

'Penalises' β weight values, forcing them towards 0

$$\hat{\beta}_{ridge} = (X^T X + \lambda I)^{-1} X^T y$$

With $\lambda = 0$, we have OLS.
The higher λ, the stronger the penalisation

Need to cross-validate to find λ

Recap: how to fit encoding models?

Two ways:

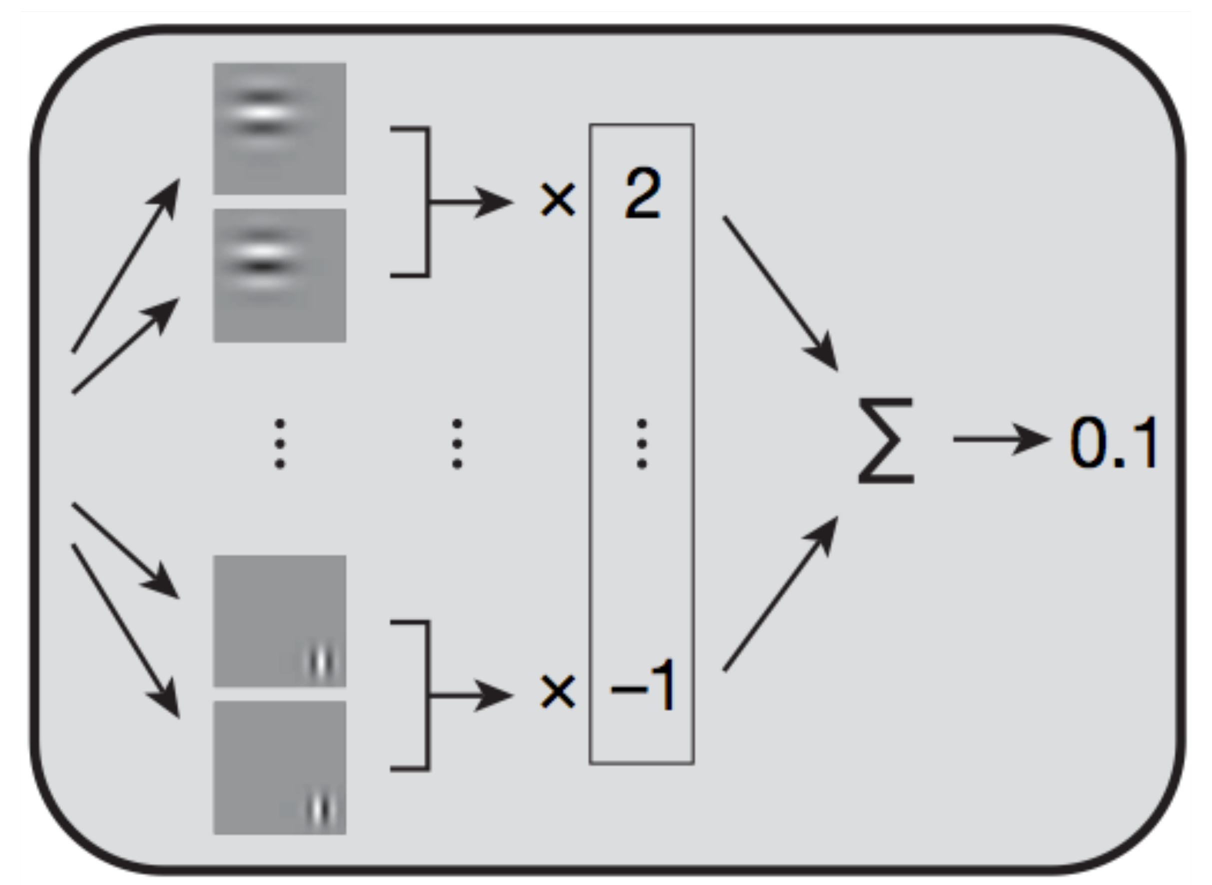
Recap: how to fit encoding models?

Two ways:

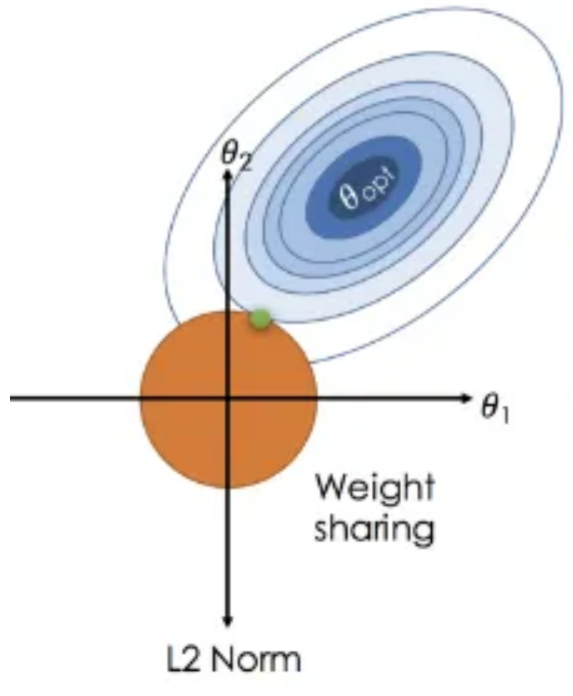
1. **Linear (Penalised) regression, with a very large design matrix**

   More data-driven, less constrained

2048 regressors



Receptive-field model for one voxel



Recap: how to fit encoding models?

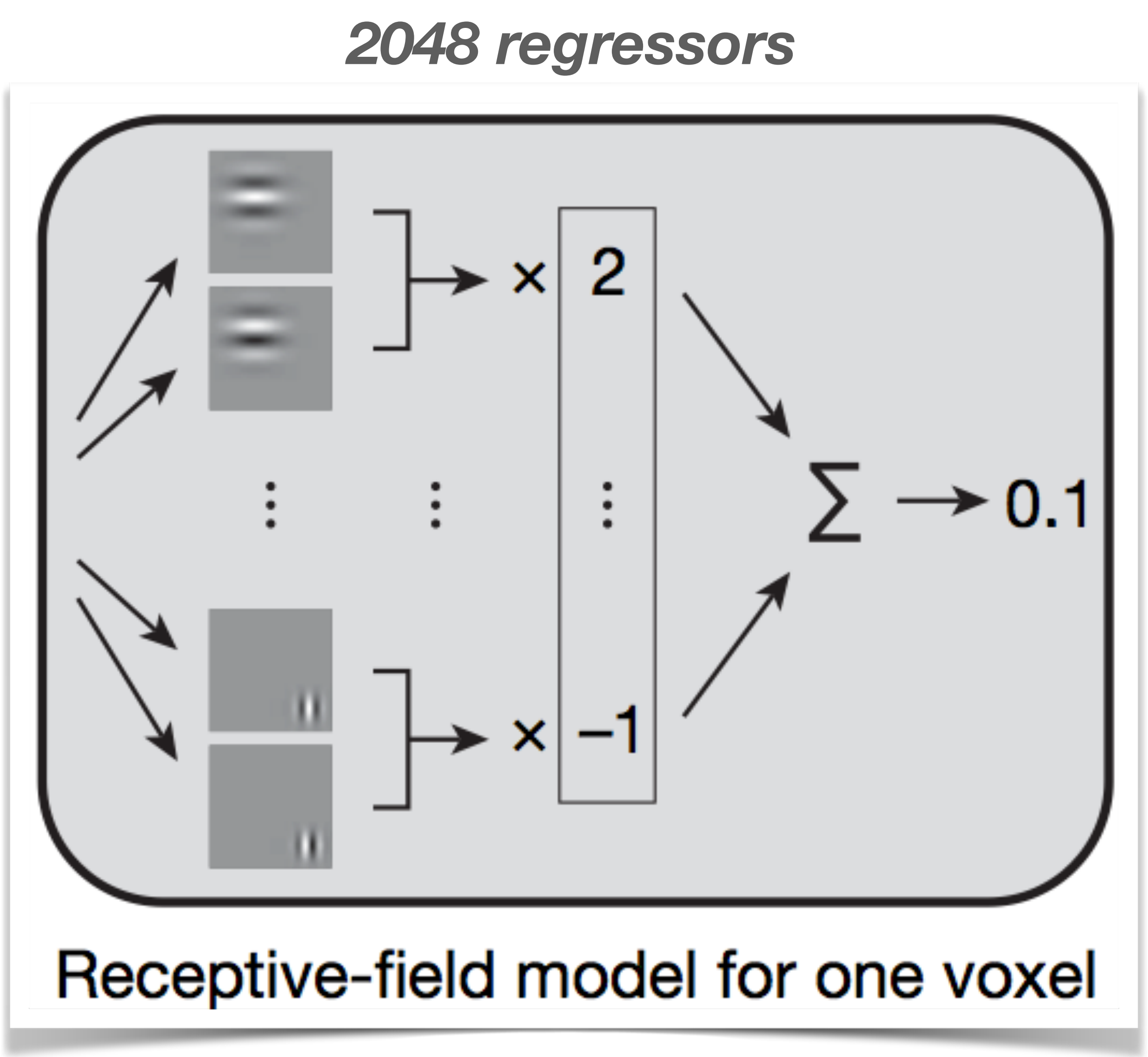
Two ways:

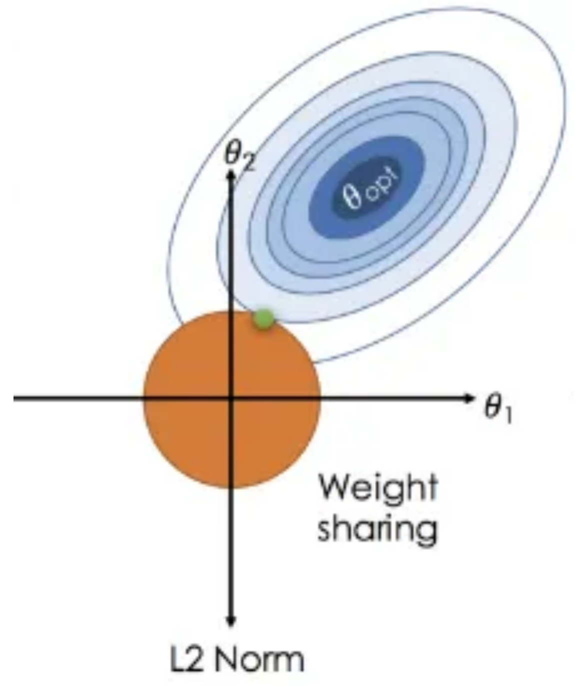
1. **Linear (Penalised) regression, with a very large design matrix**

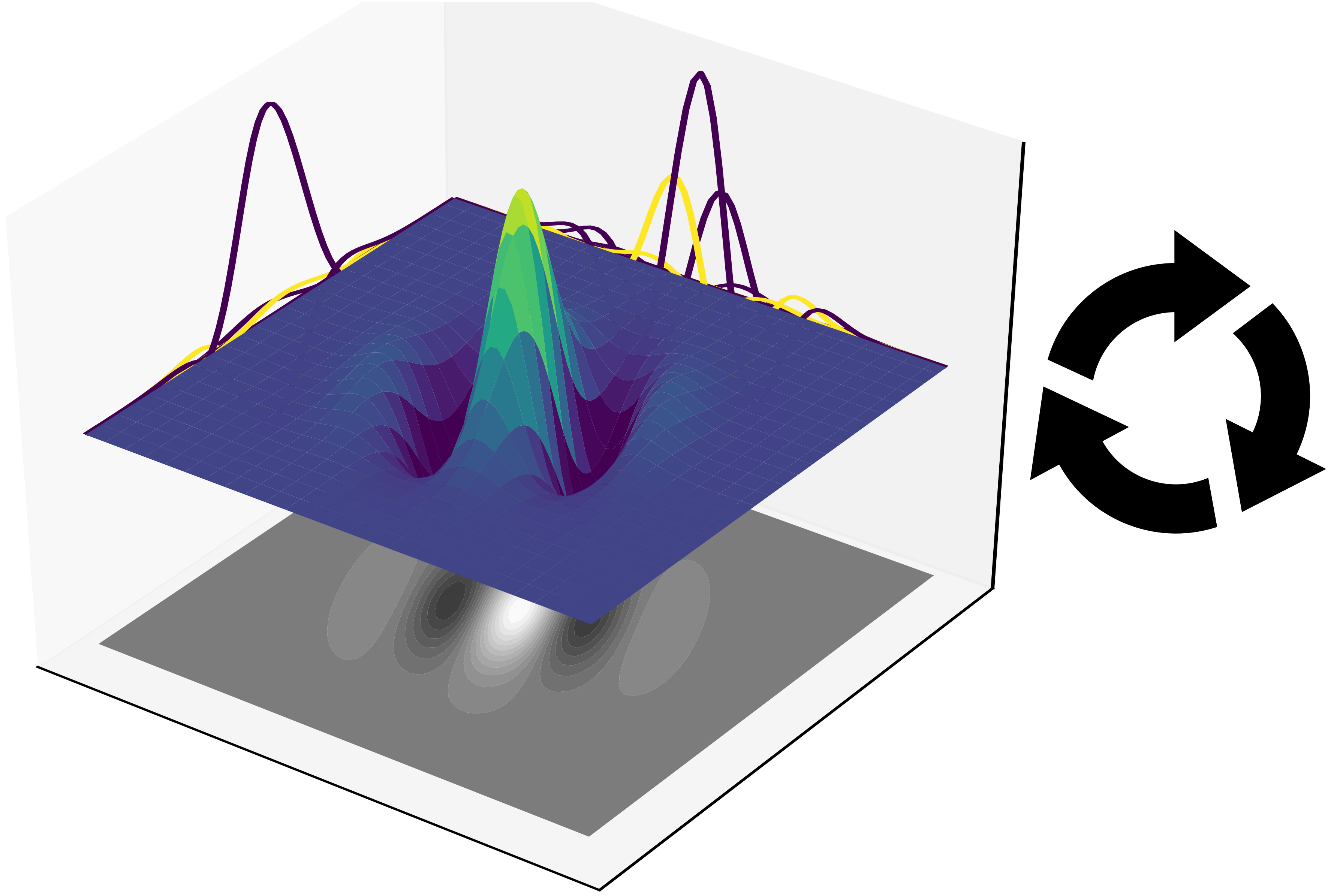
   More data-driven, less constrained

2. **Find best-fitting parameters of a mathematical model**

   More model-driven, only specific, controlled situations







Recap: how to fit encoding models?

Two ways:

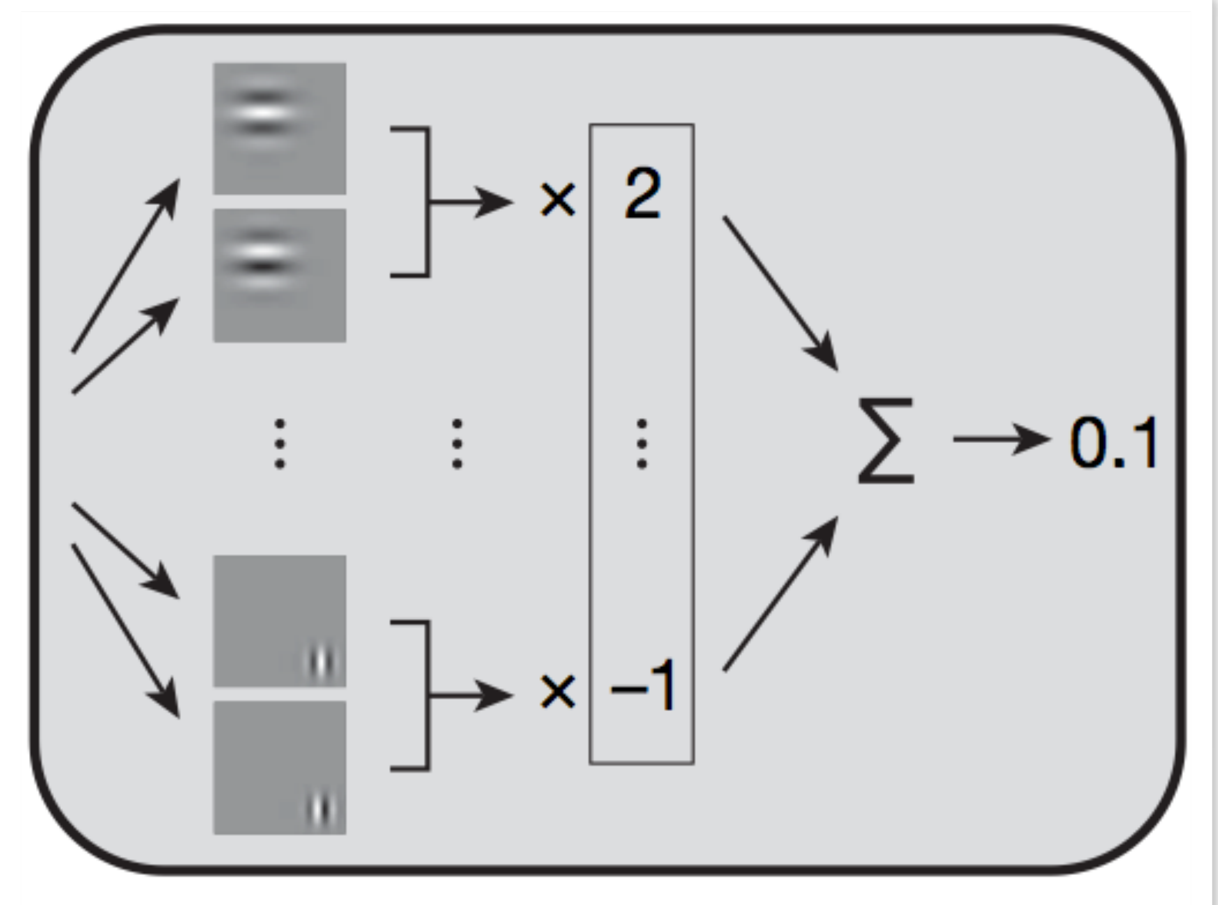
1. **Linear (Penalised) regression, with a very large design matrix**

   More data-driven, less constrained

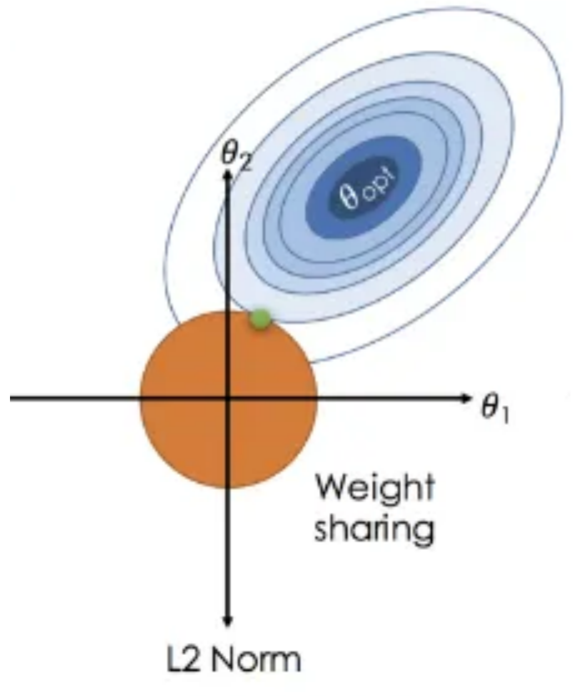
2. **Find best-fitting parameters of a mathematical model**

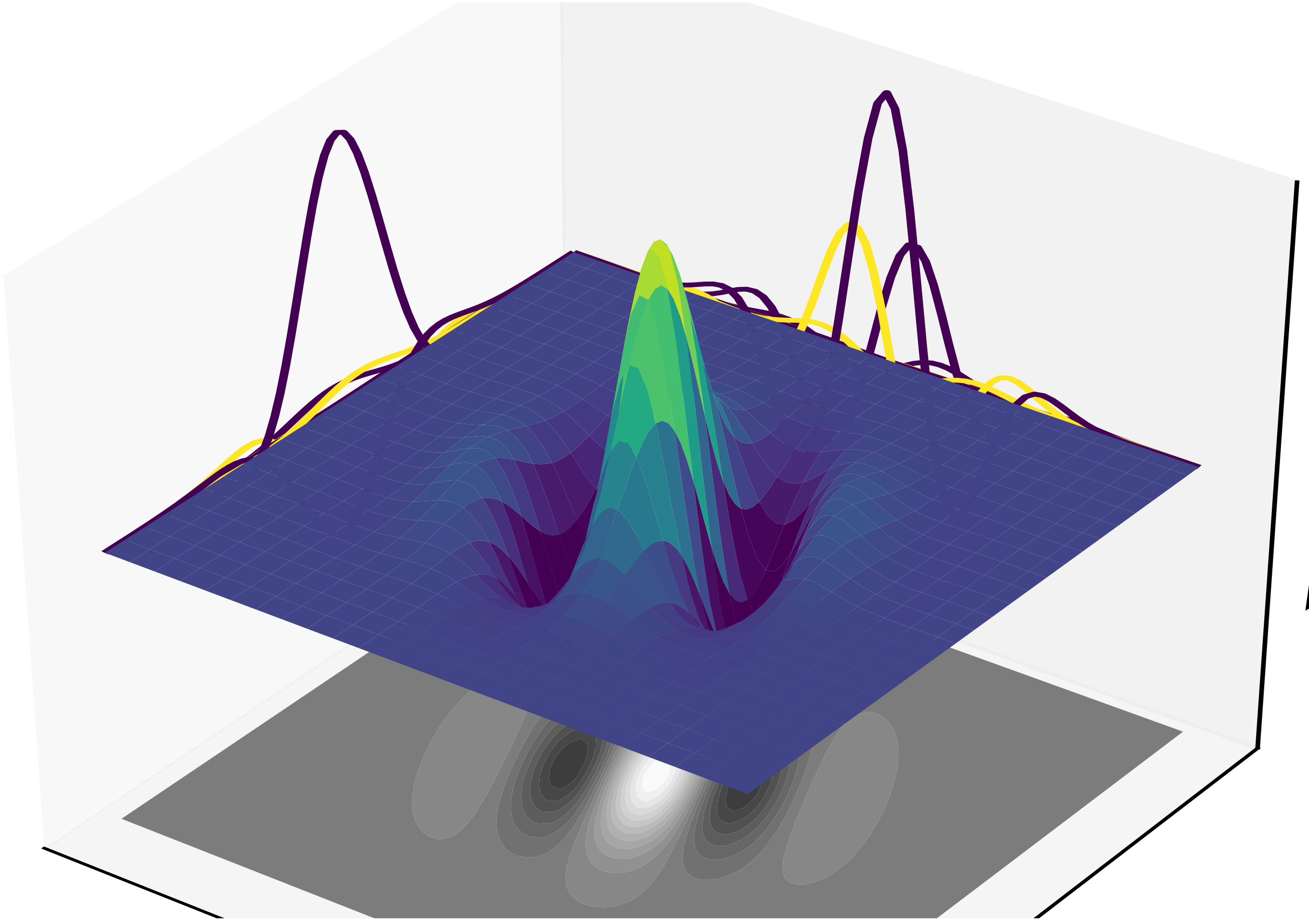
   More model-driven, only specific, controlled situations

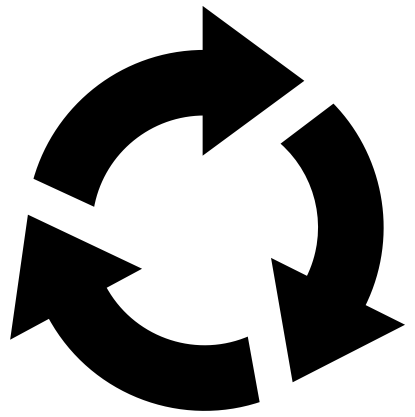
2048 regressors



Receptive-field model for one voxel



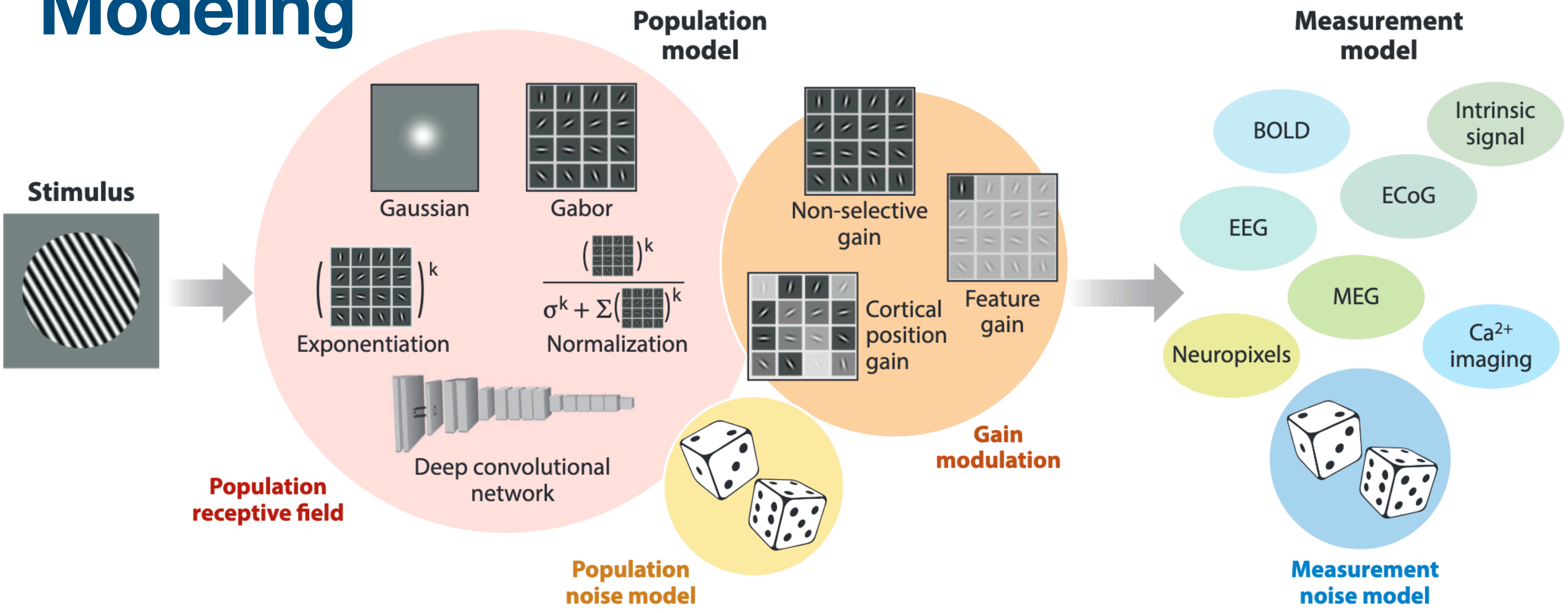




7 parameters

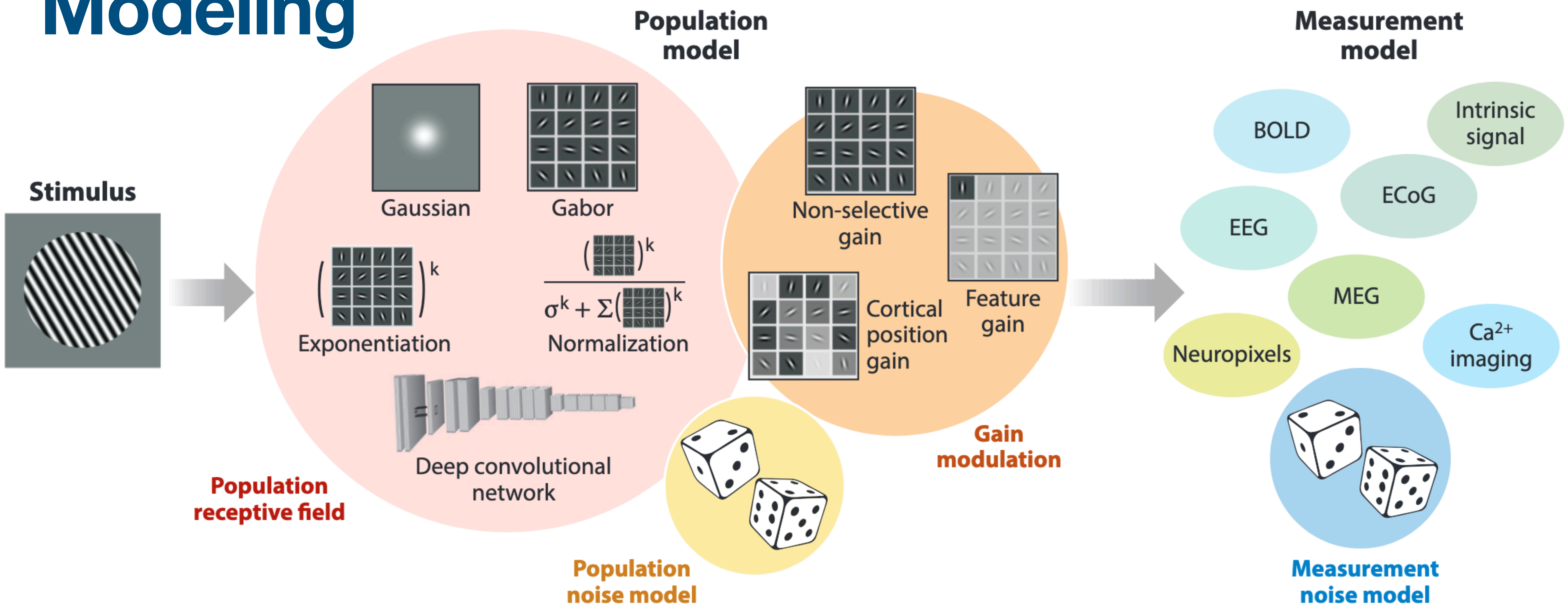
$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right)$$

Modeling



From Gardner & Merriam, 2022

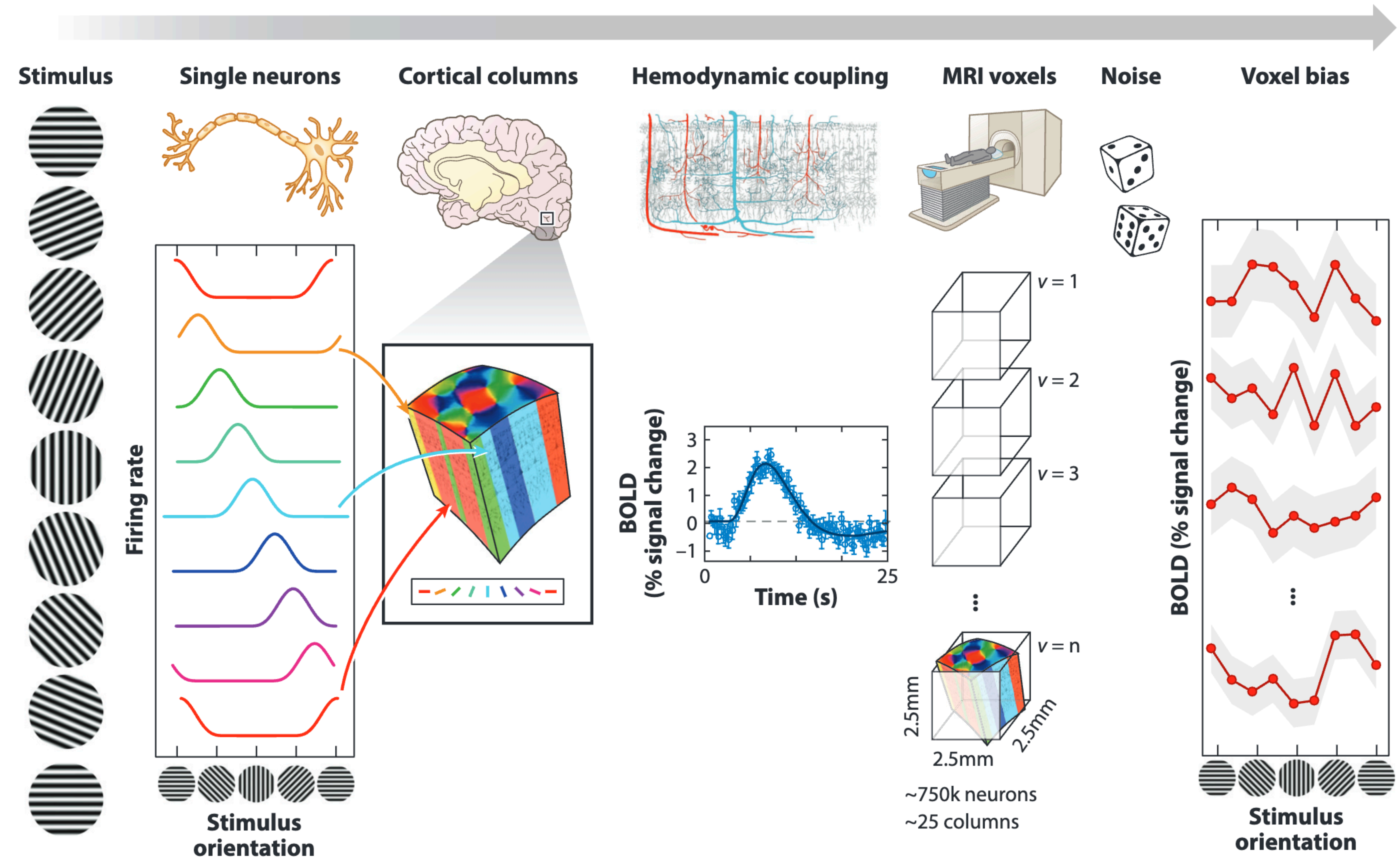
Modeling



Very explicit what our underlying model!

From Gardner & Merriam, 2022

Modeling



Very explicit what our underlying model!

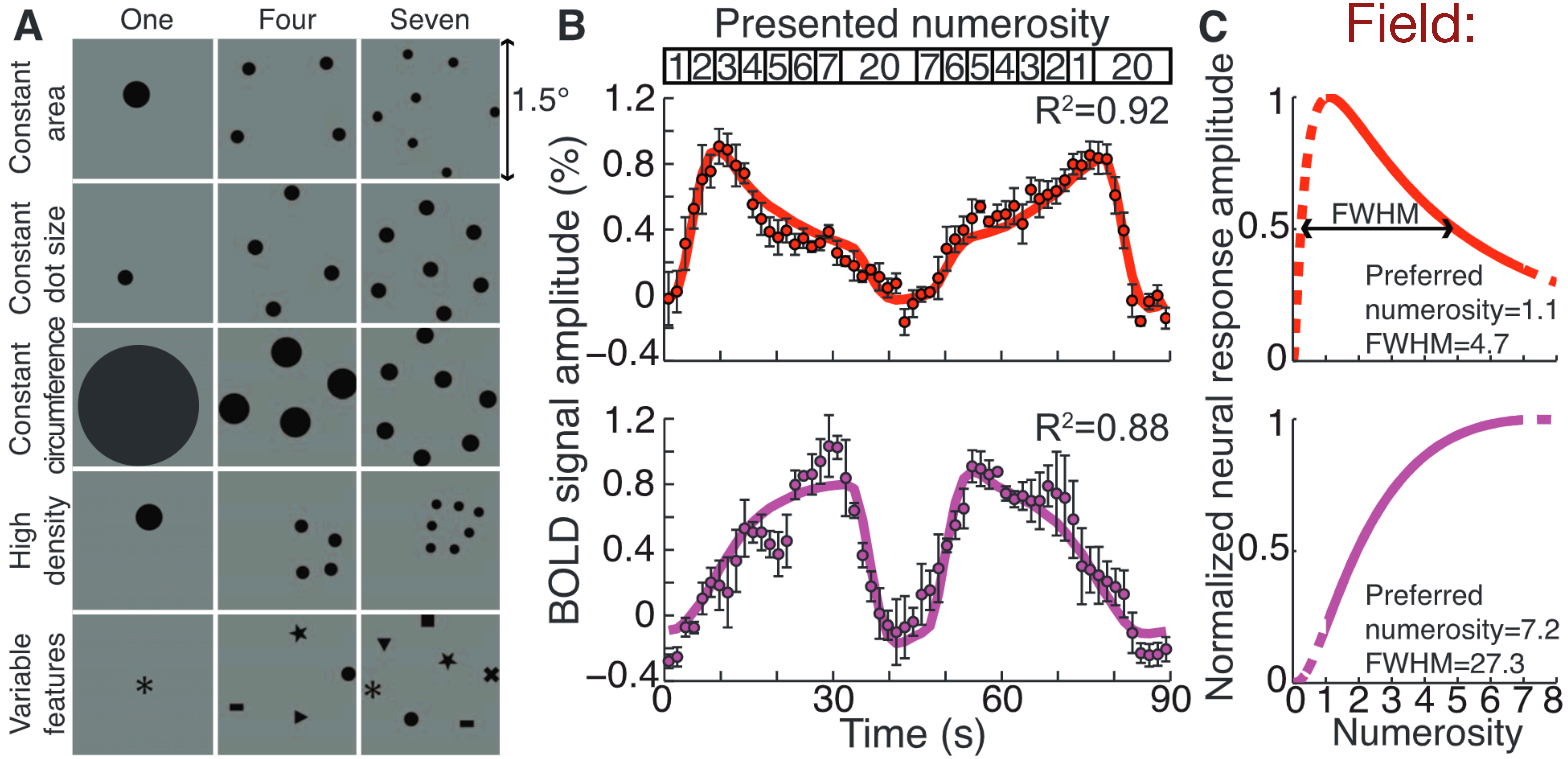
From Gardner & Merriam, 2022

Does it have to be visual?

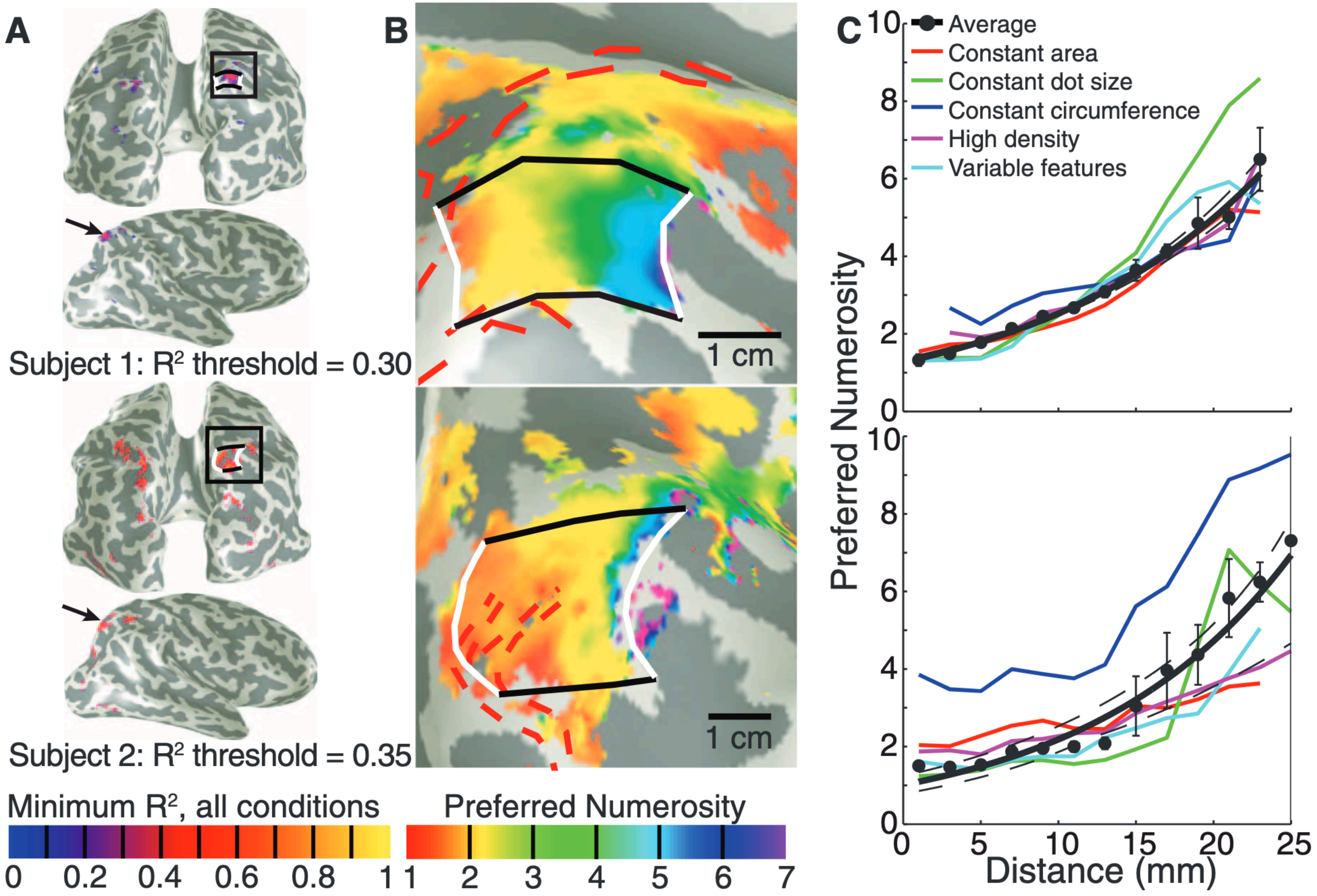
But more…

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

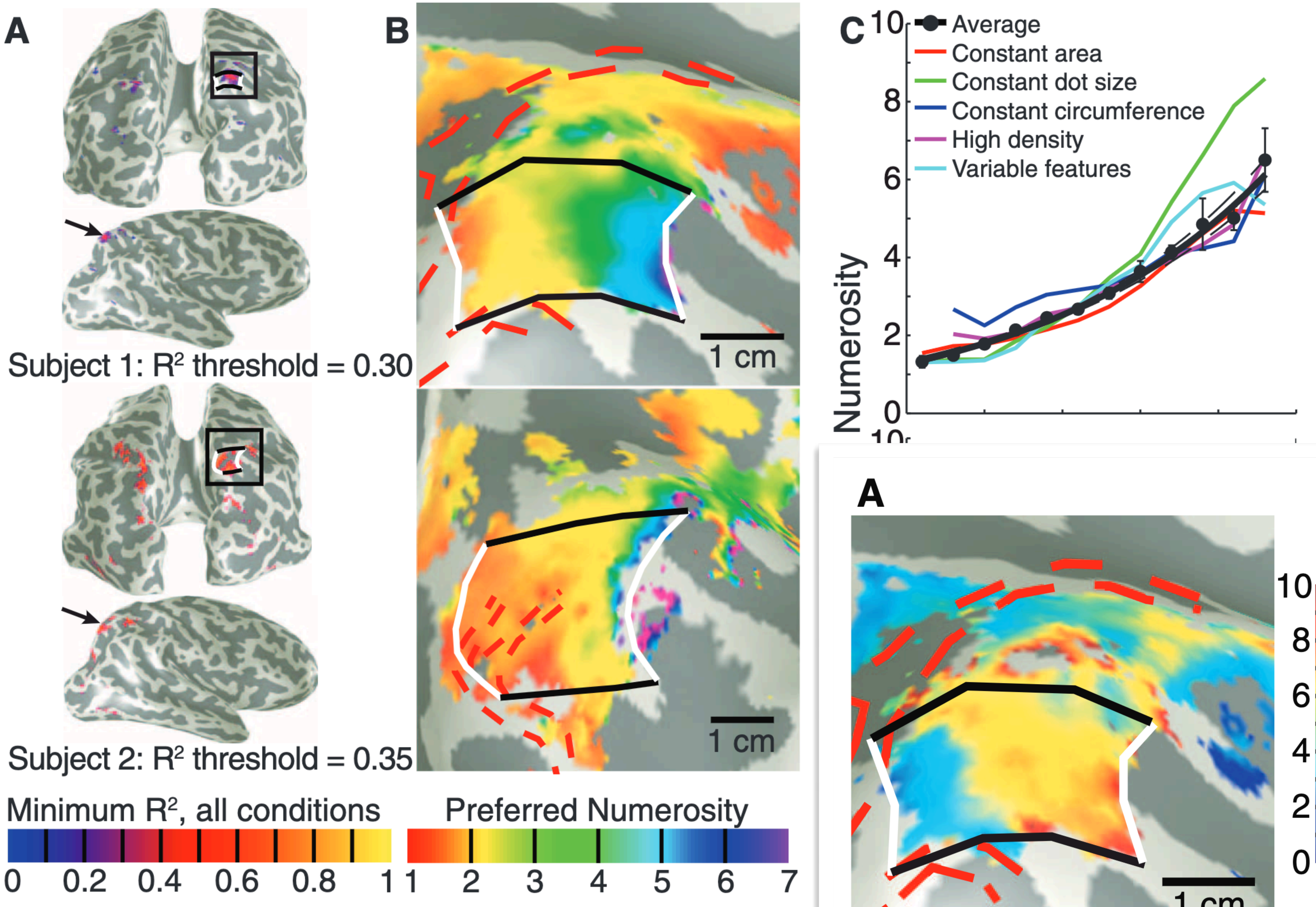
Receptive Field:

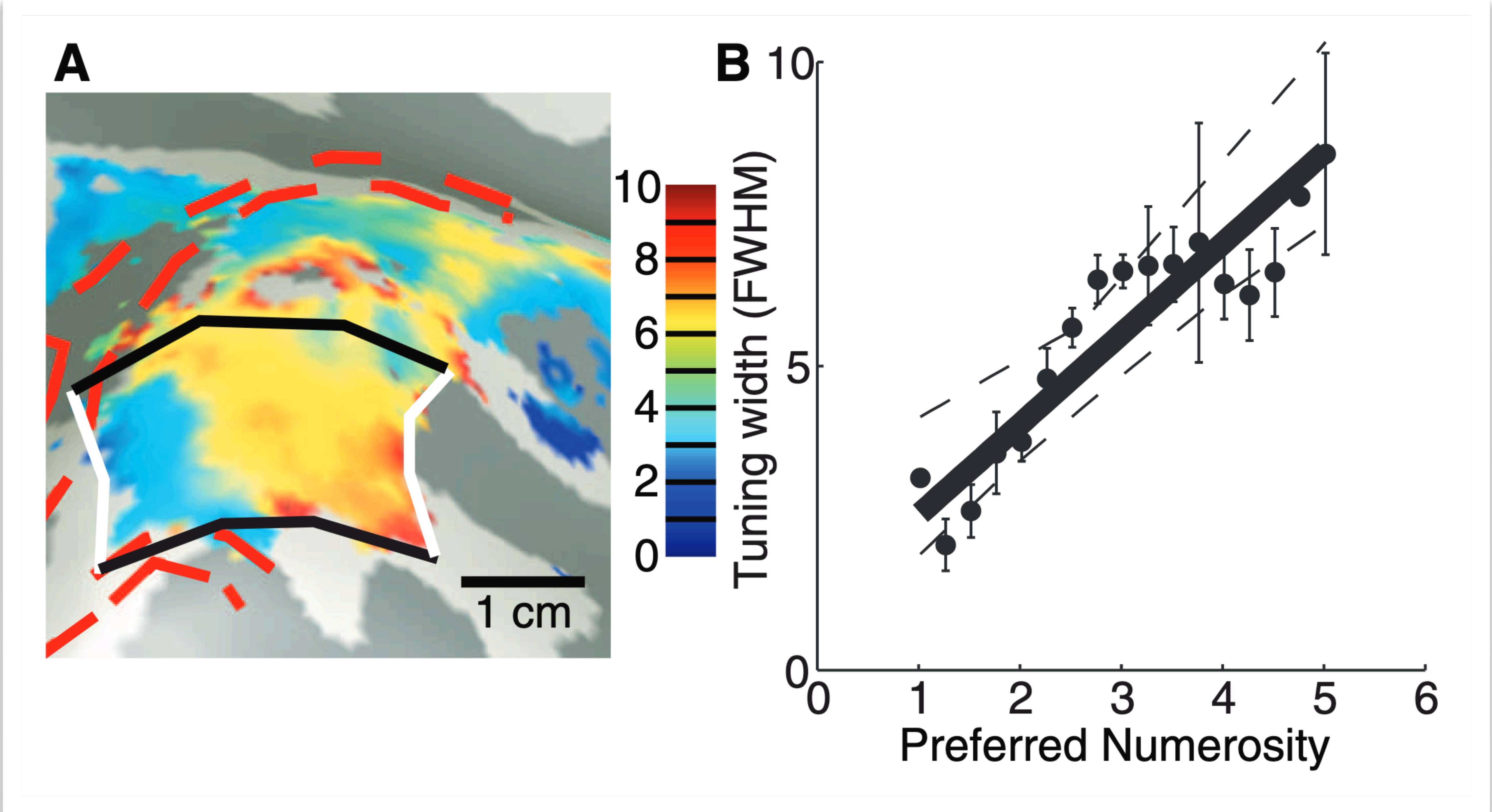


These receptive fields are layed out on maps…



These receptive fields are layed out on maps…





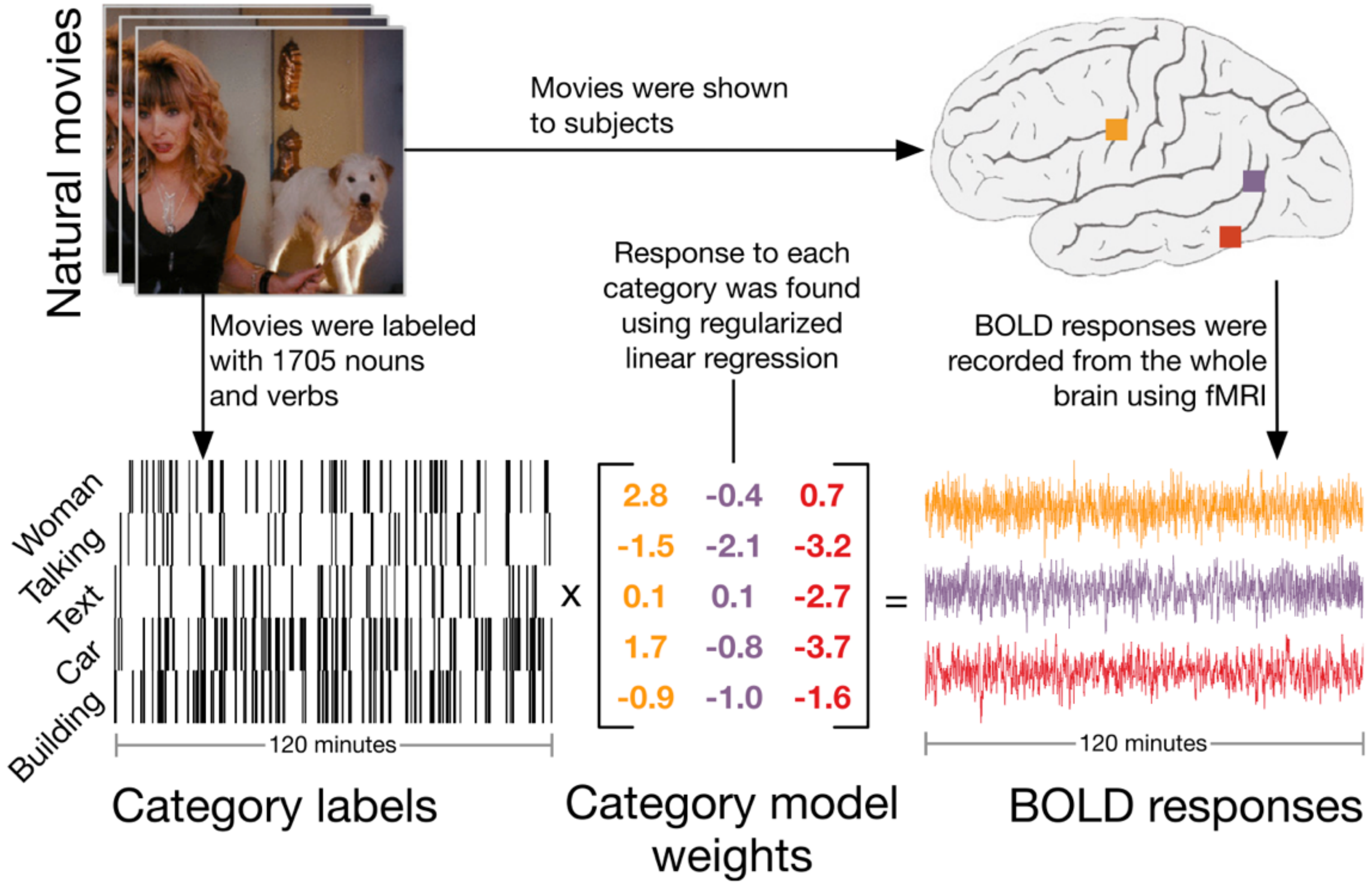
Encoding in higher spaces

Encoding in higher spaces

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

Encoding in higher spaces

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

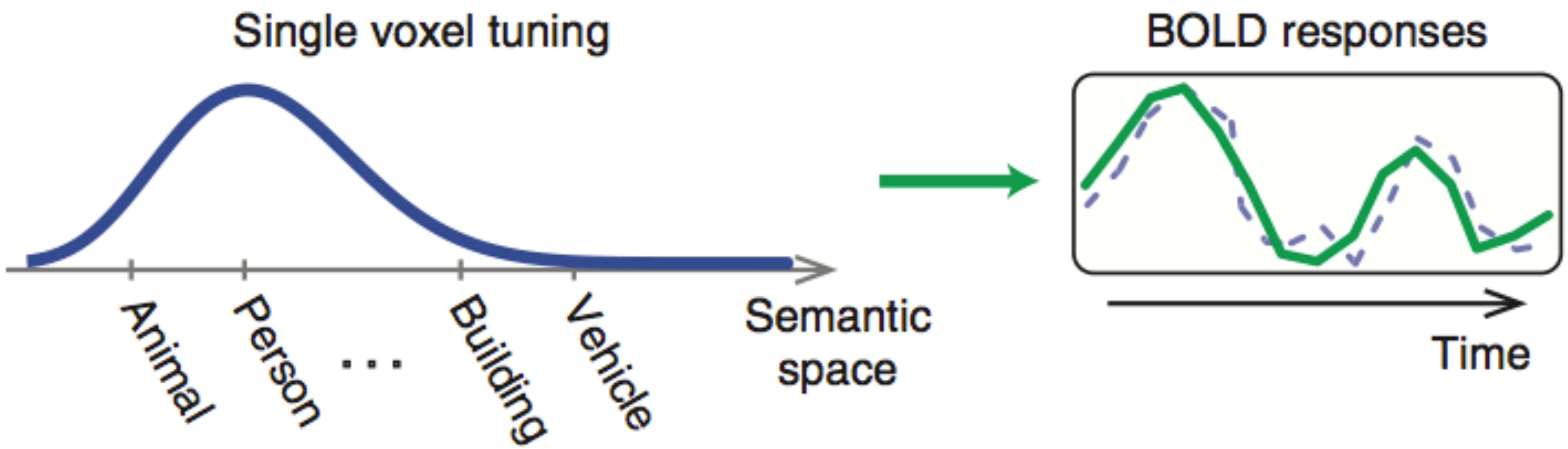


Encoding in higher spaces

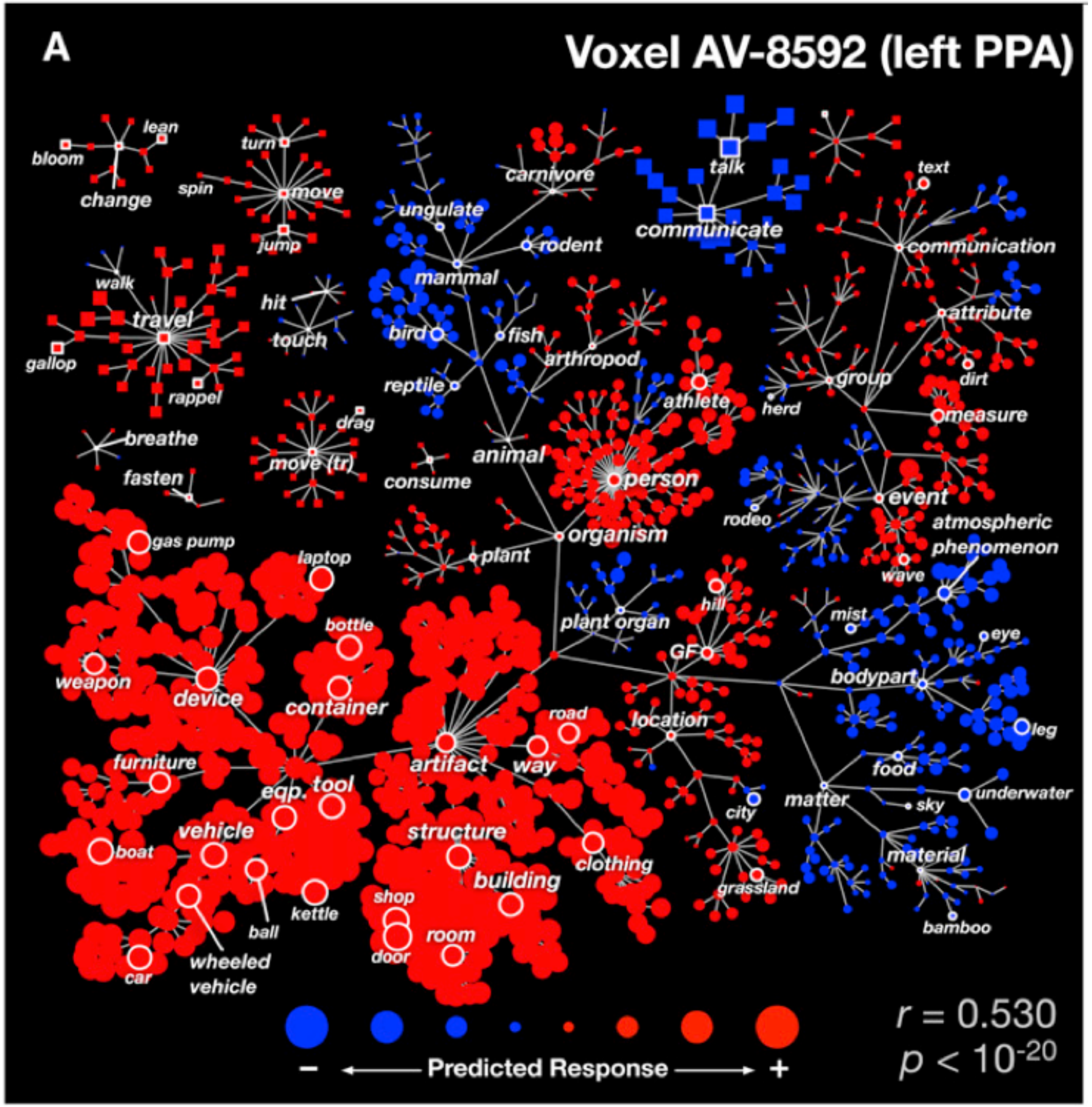
A voxel's receptive field doesn't have to be retinotopic, but can live in a more abstract (semantic) space.

Encoding in higher spaces

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



Single-voxel
tuning curve
in a
semantic space



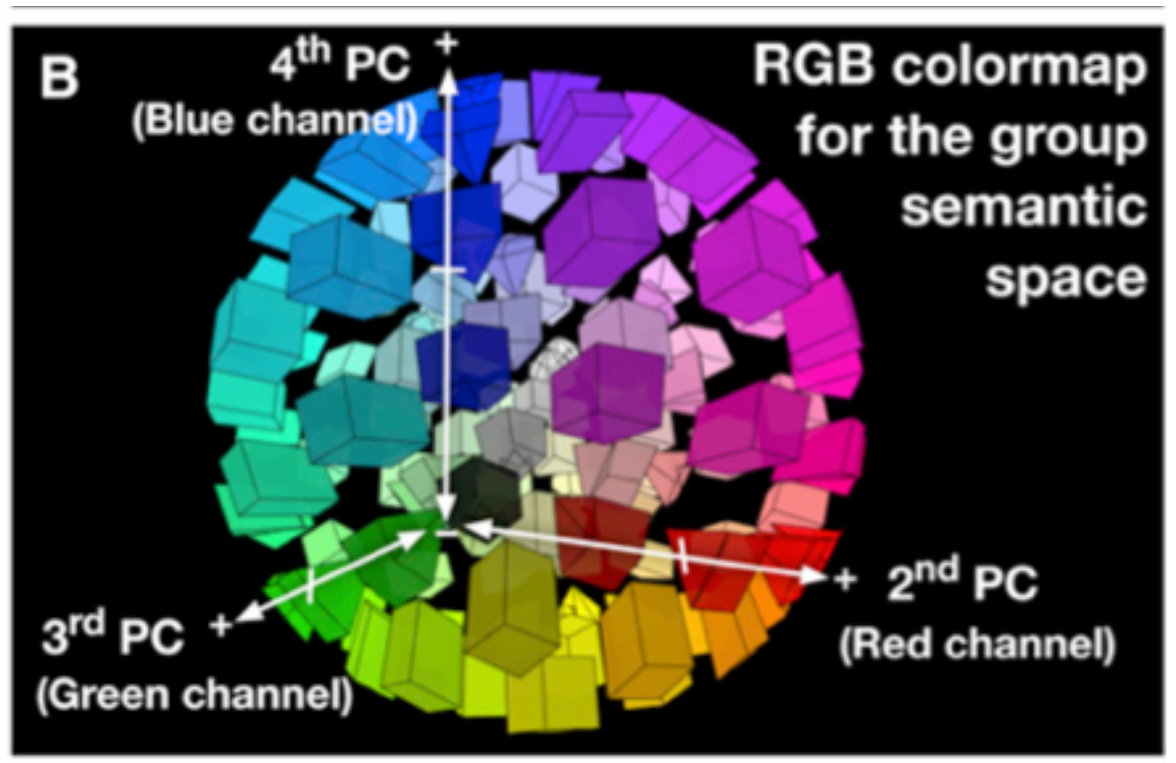
Making sense of these results…

Making sense of these results…

- Problem: doing this for all voxels gives us a very-high-dimensional space of voxels by regressors’ beta weights.

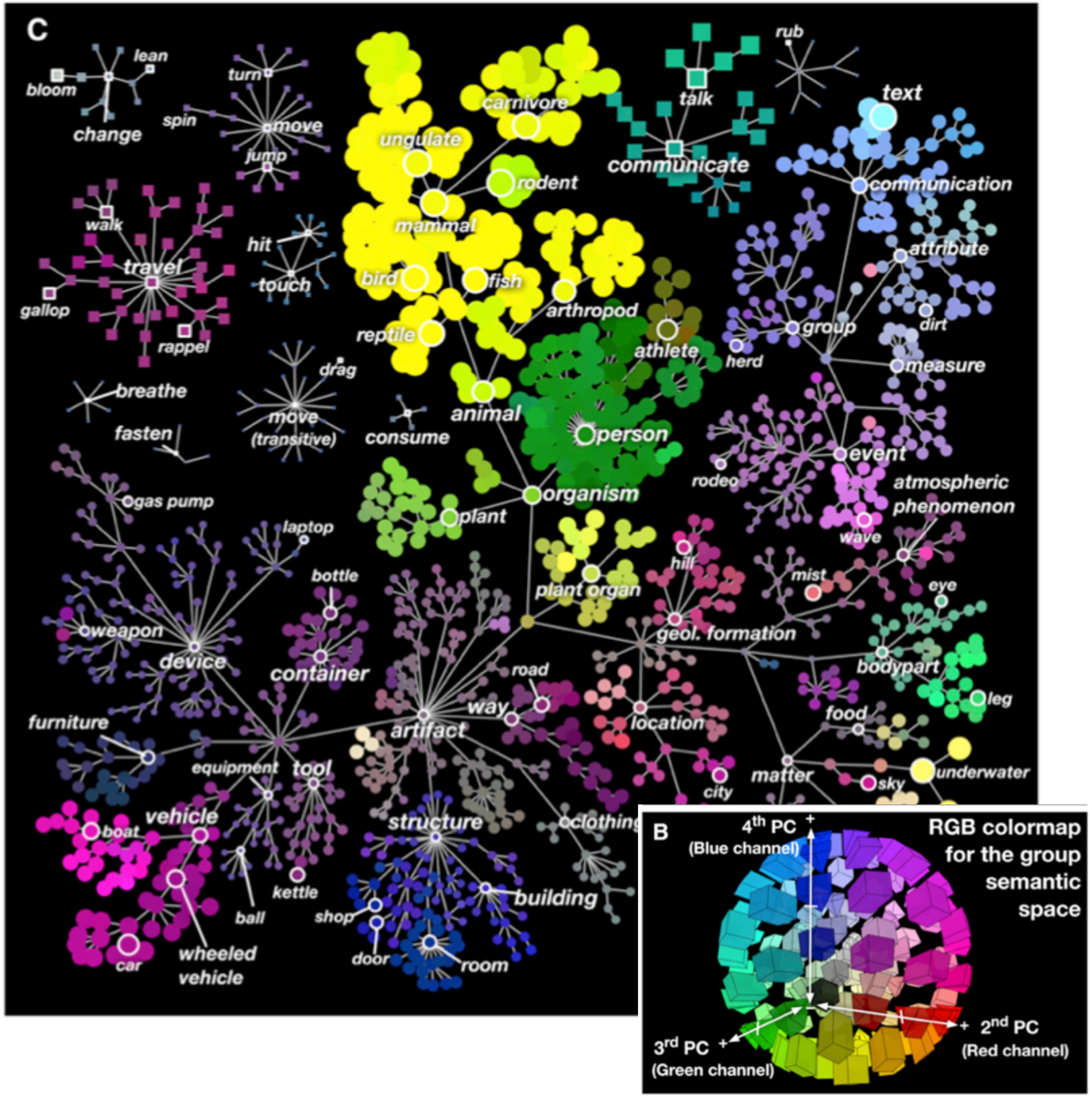
Making sense of these results…

- Problem: doing this for all voxels gives us a very-high-dimensional space of voxels by regressors' beta weights.
- Dimensionality reduction (Factor analysis/SVD/PCA) yields a couple of components



Making sense of these results…

- Problem: doing this for all voxels gives us a very-high-dimensional space of voxels by regressors’ beta weights.
- Dimensionality reduction (Factor analysis/SVD/PCA) yields a couple of components
- Projecting onto semantic structure:



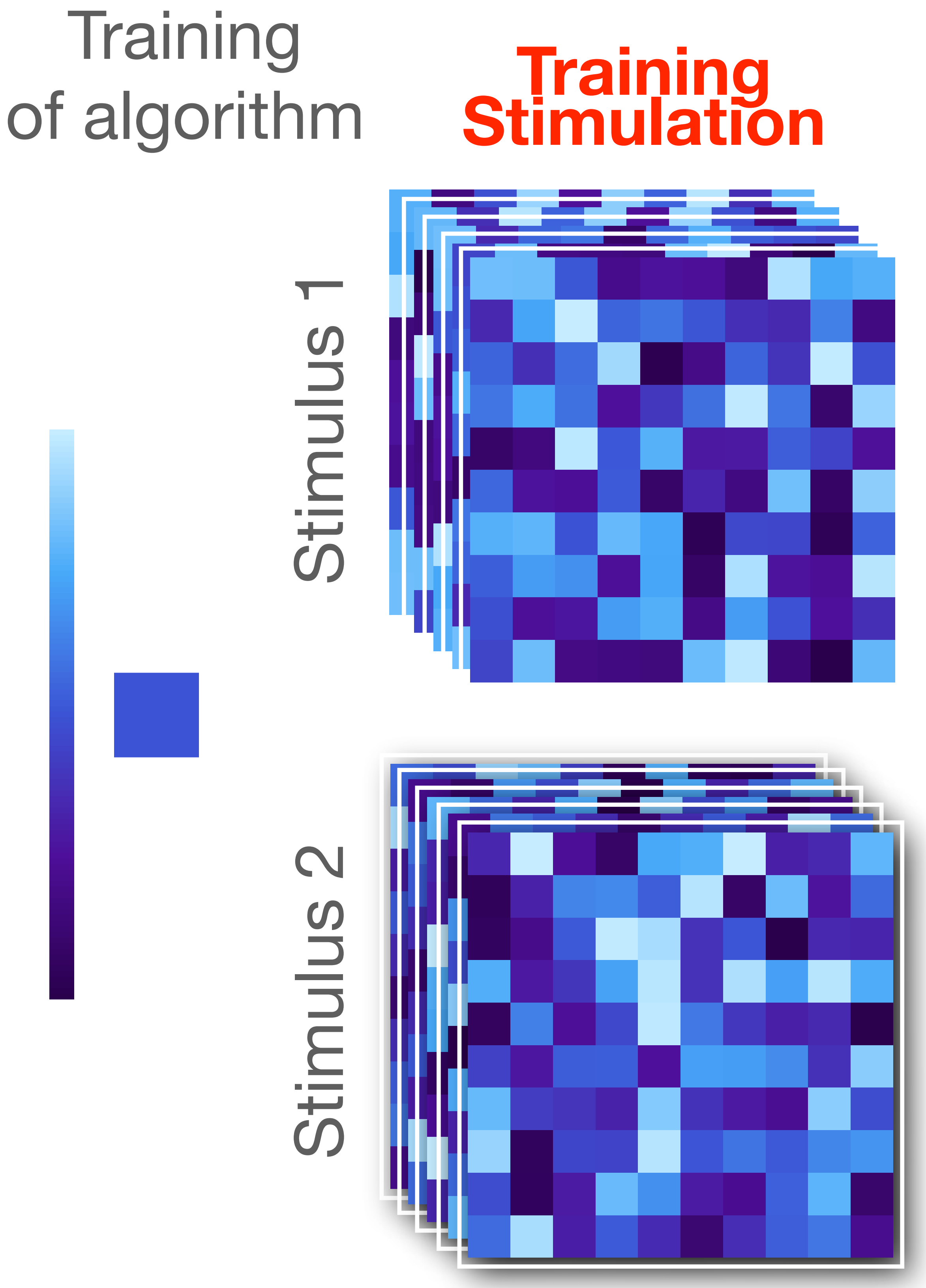
roll
reach
talk
gesture
hit
condition
abstraction
player
move
eat
male
information
tree
herb
plant
mustache
cloud
machine
rock
appliance
object
piece
causal_agent
glass
medium
point
face
leg
cabinet
art
conduit
fluid
gas
implement
picture
game_equipment
structure
field
building
shop
motor vehicle
car
S4
S3a
S3b
S1
S2
V2
V1
V3
LBelt
PBelt
A1
MBelt
V1
V2
V3
S4
S3b
S2
S1
S3a
PBelt
Belt
A1
MBelt

Decoding

Decoding

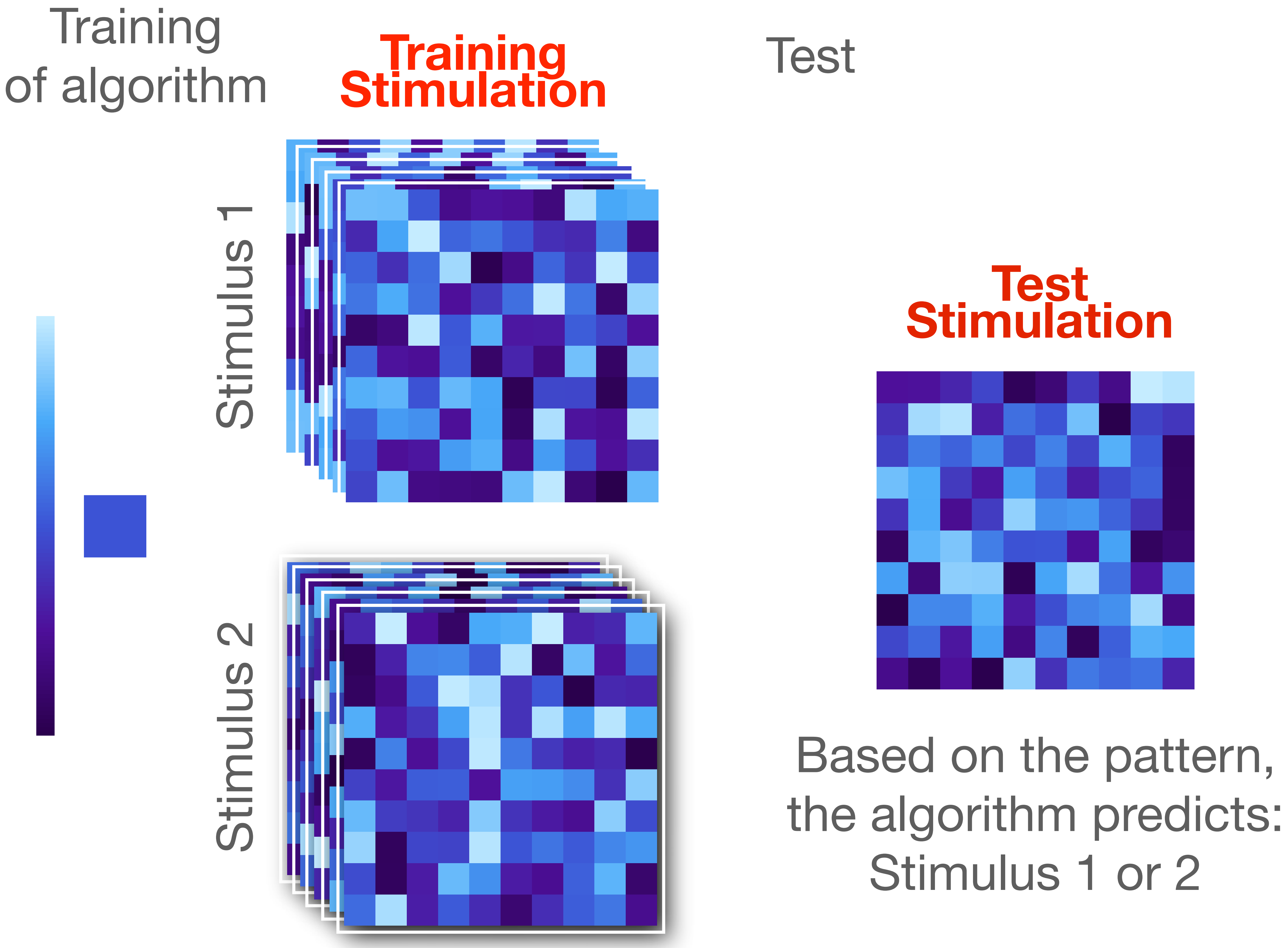
Decoding

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



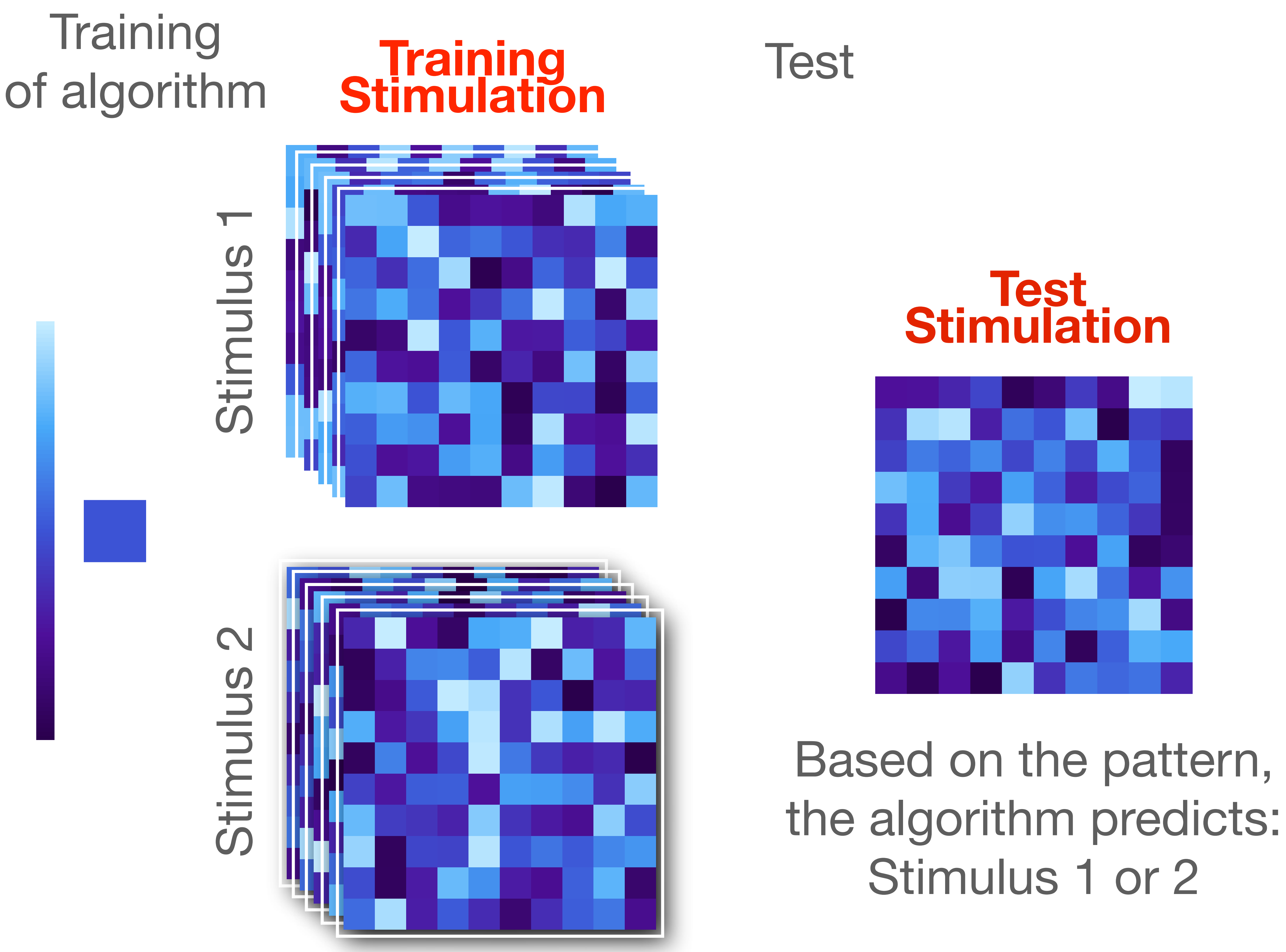
Decoding

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



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: percentage correct

Gauge on information content in a specific region of interest

Decoding

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.*

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?

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

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.*

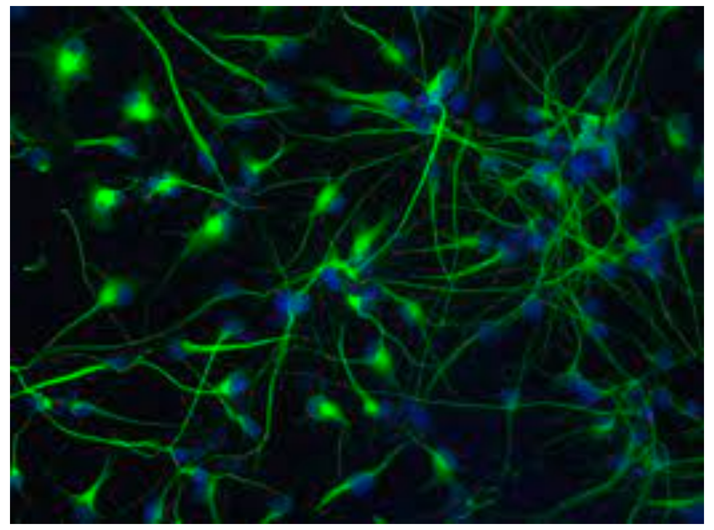
Strong debates:

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?*

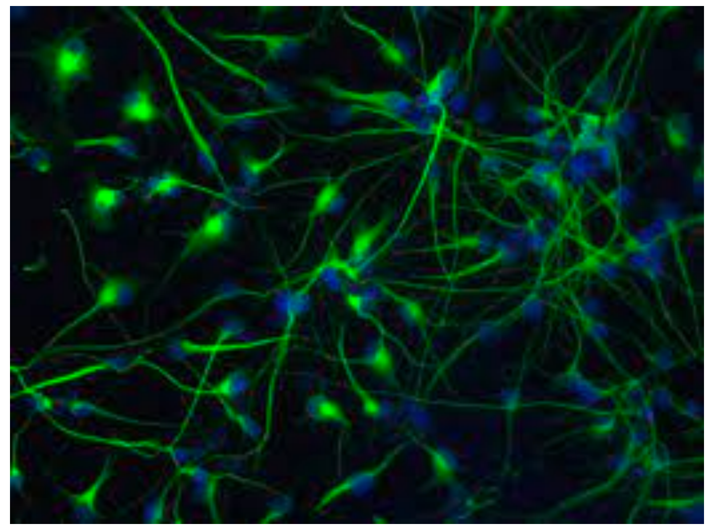
We don’t know, because the information-processing and measurement mechanism remains implicit.

Encoding



Encoding

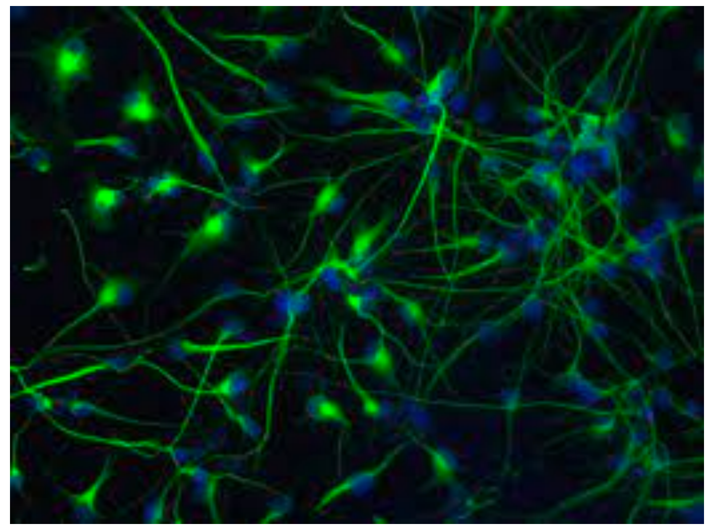
*If we can find what a neuron/sensor/voxel responds to, we will have a model of how this unit encodes information: **an encoding model***



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***

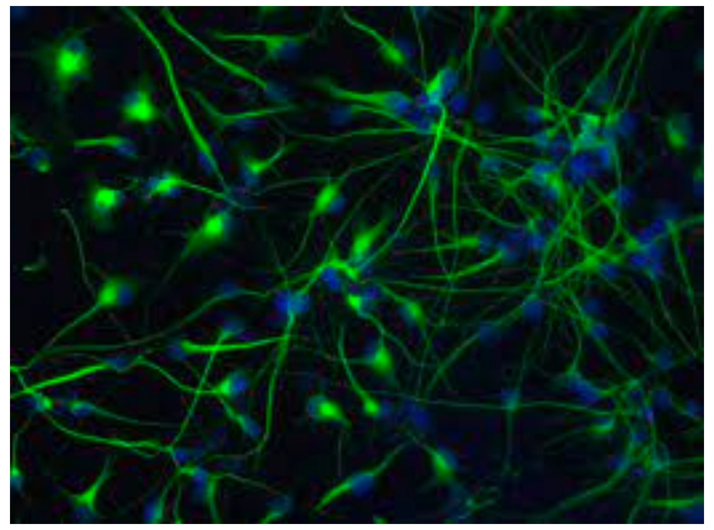
*The paradigm: see **voxels as a population of neurons**:*
*the encoding model describes **neural population responses***



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***

The paradigm: see ***voxels as a population of neurons****:*
the encoding model describes ***neural population responses***



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***

Bayesian decoding from an encoding model

Bayesian decoding from an encoding model

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

Bayesian decoding from an encoding model

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

s = stimulus, b = BOLD response pattern

Bayesian decoding from an encoding model

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

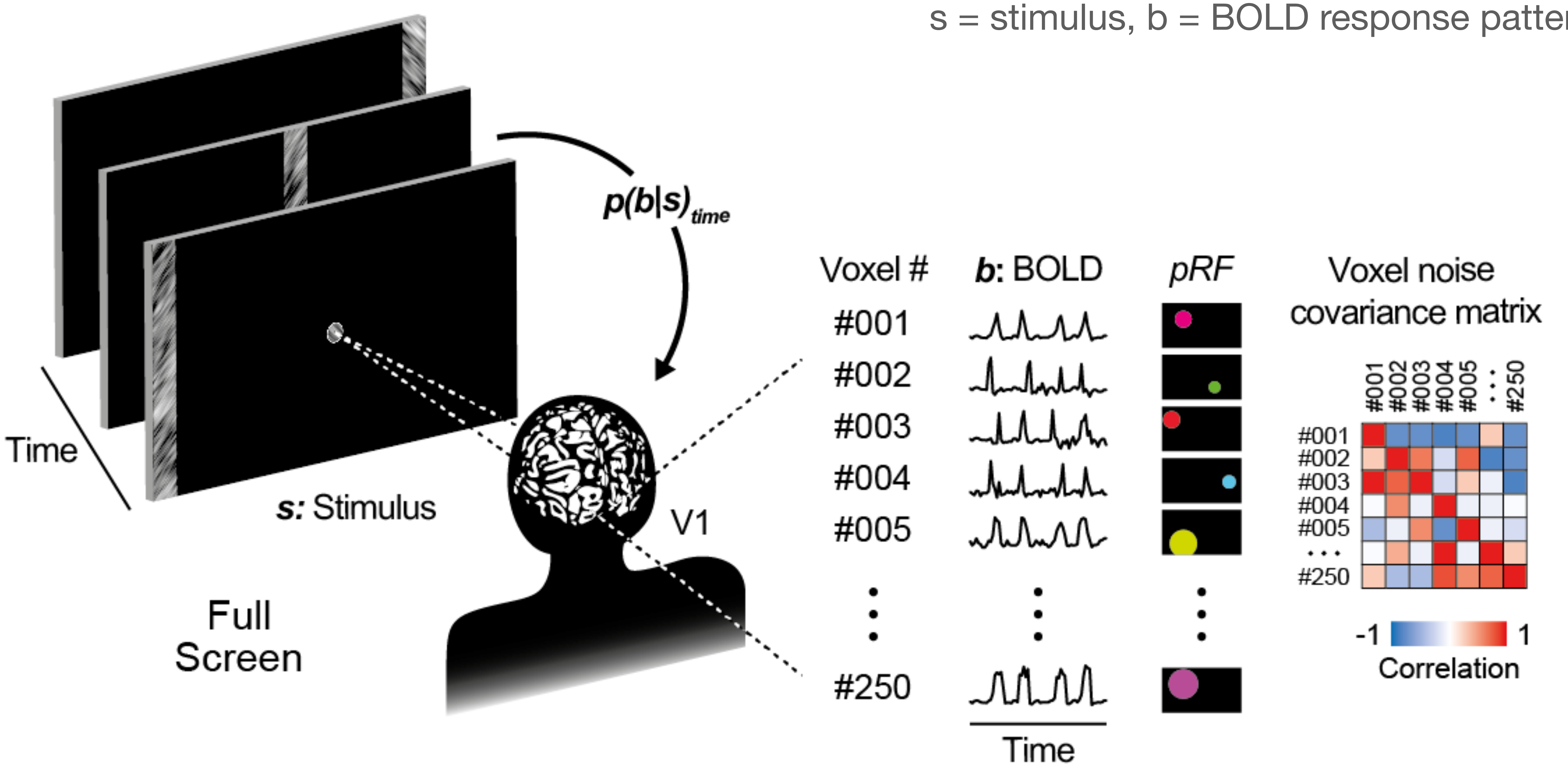
s = stimulus, b = BOLD response pattern

Bayesian decoding from an encoding model

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

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

s = stimulus, b = BOLD response pattern



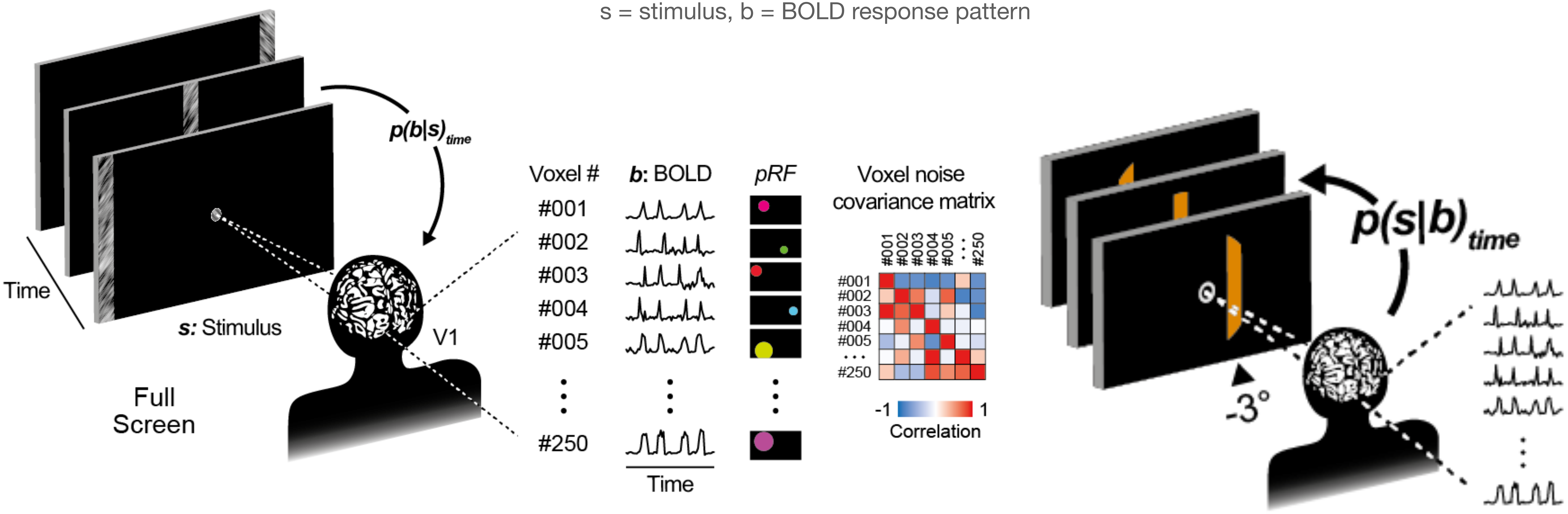
Bayesian decoding from an encoding model

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

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

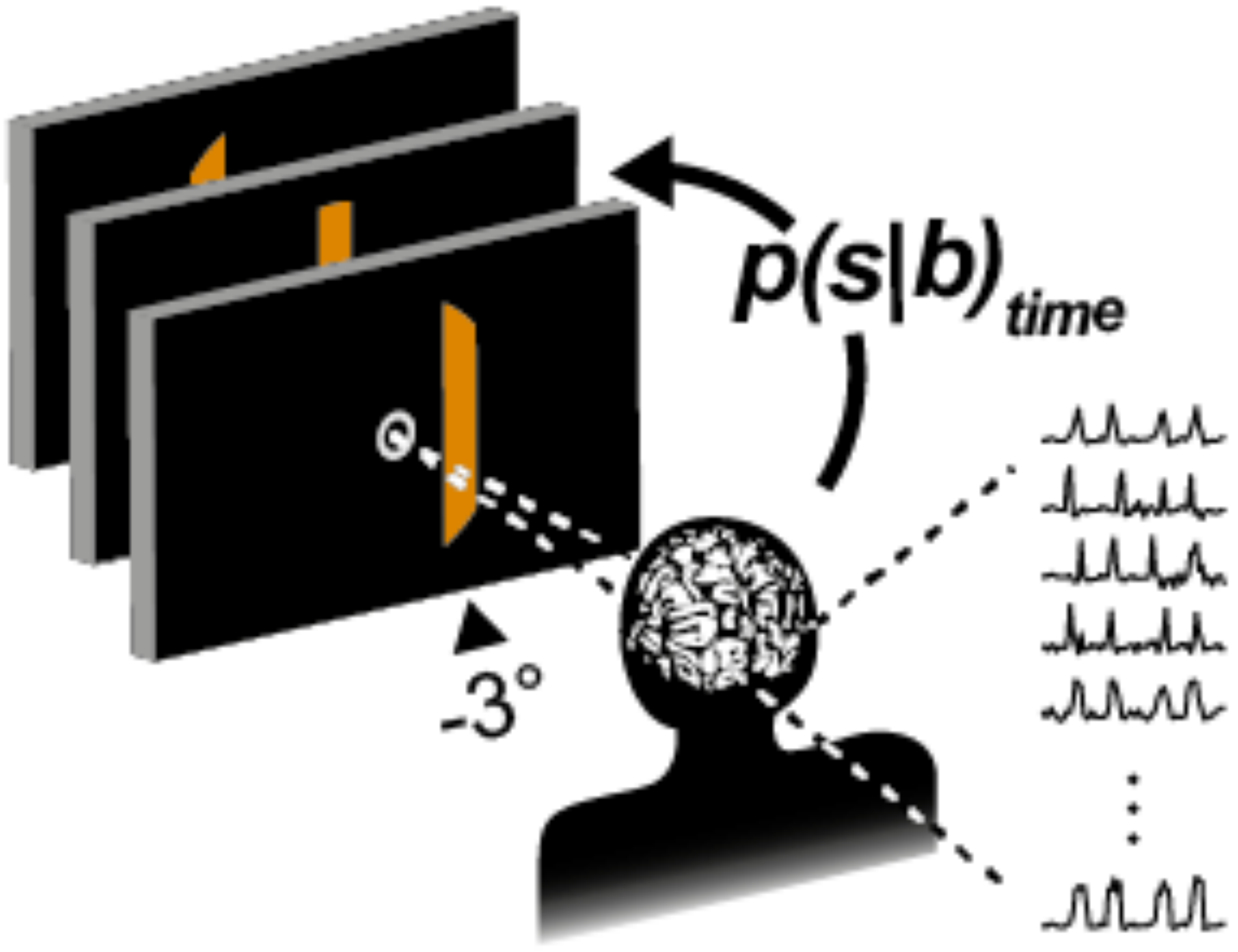
We can model/calculate
the probability of a BOLD response pattern
given a stimulus

s = stimulus, b = BOLD response pattern

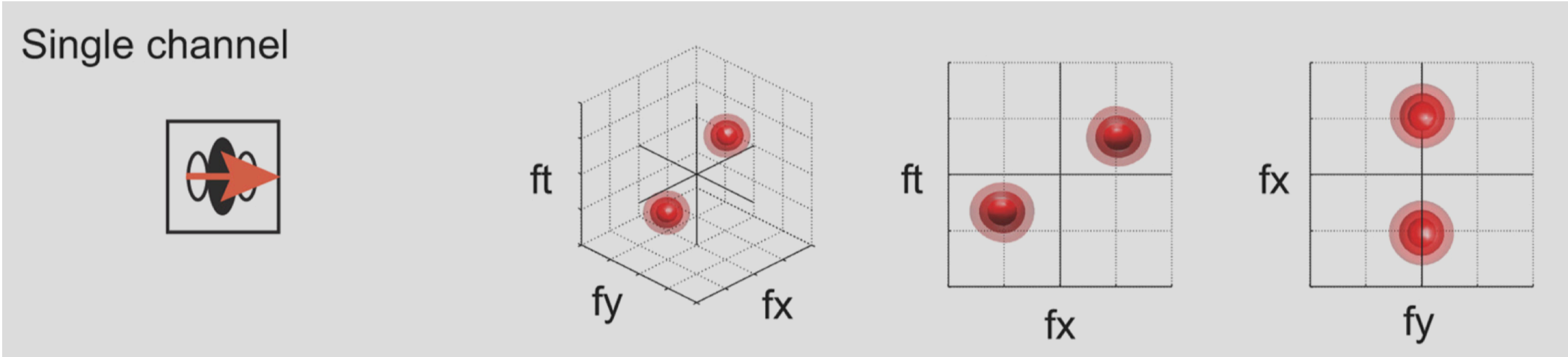


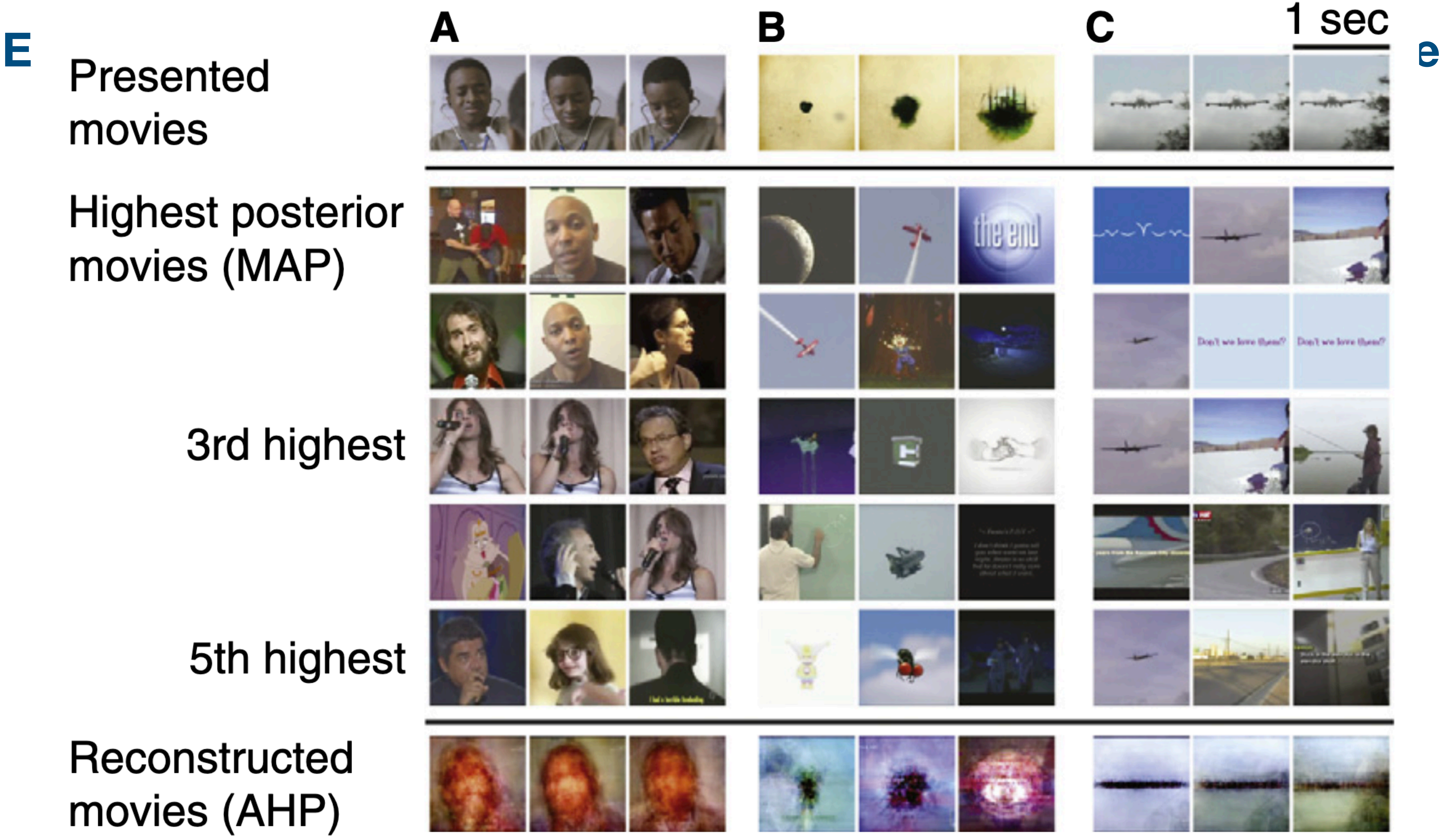
Bayesian decoding from an encoding model

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



Extend this to movies: Add spatiotemporal receptive field structure



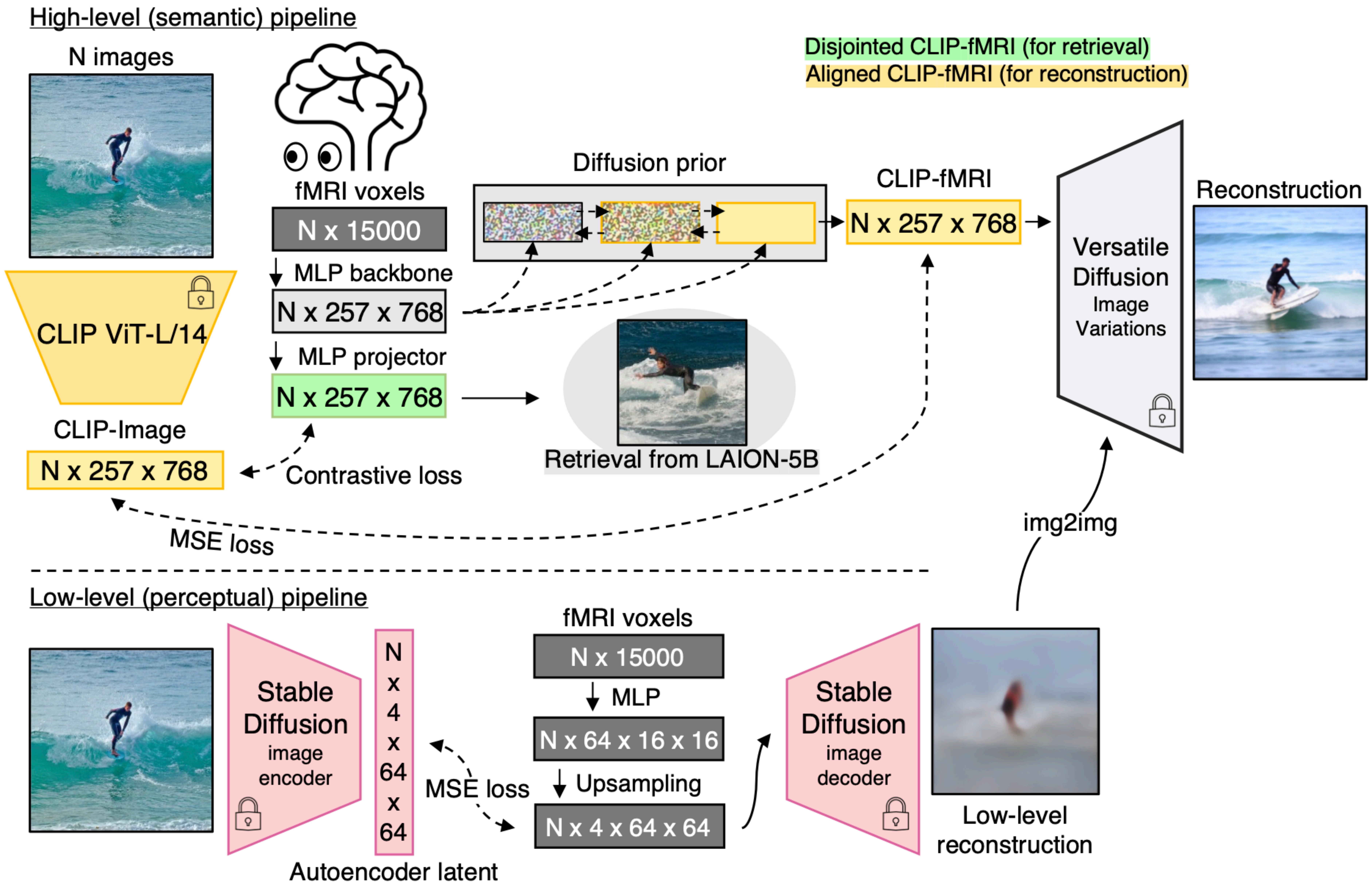

E
A
B
C
1 sec
e
Presented movies
Highest posterior movies (MAP)
3rd highest
5th highest
Reconstructed movies (AHP)

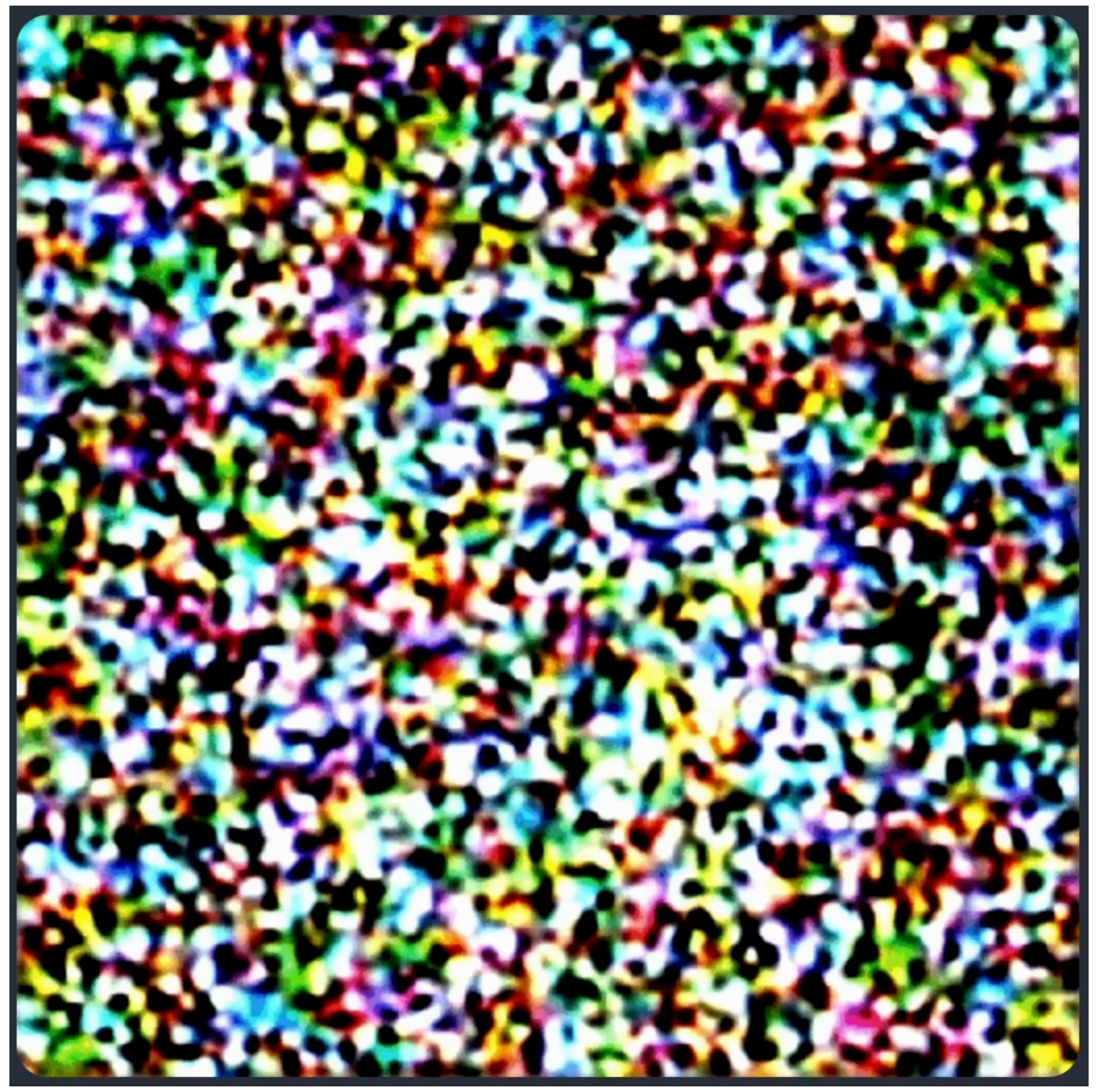
Presented clip

Clip reconstructed from brain activity

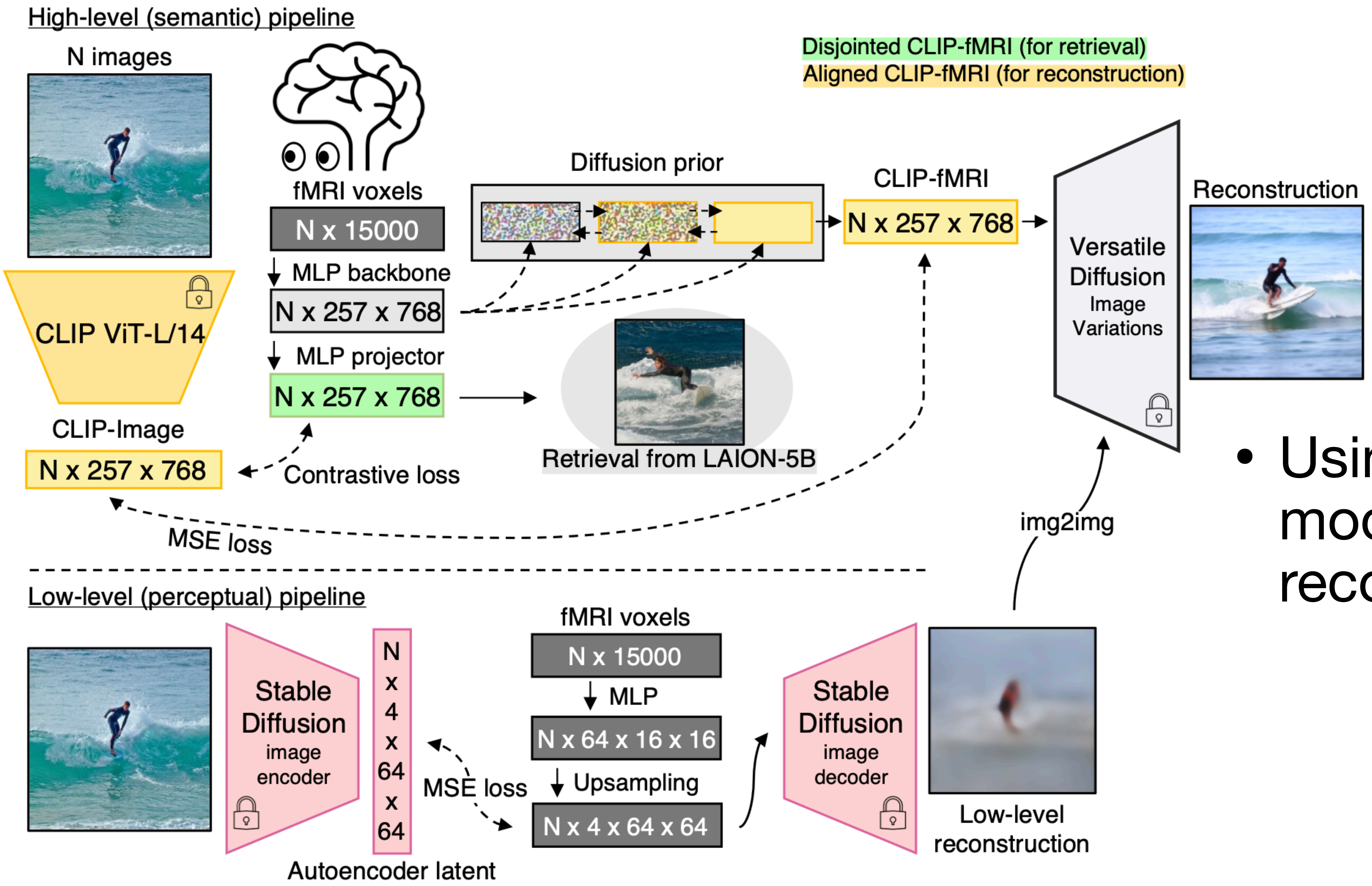
Stimulus reconstruction using AI

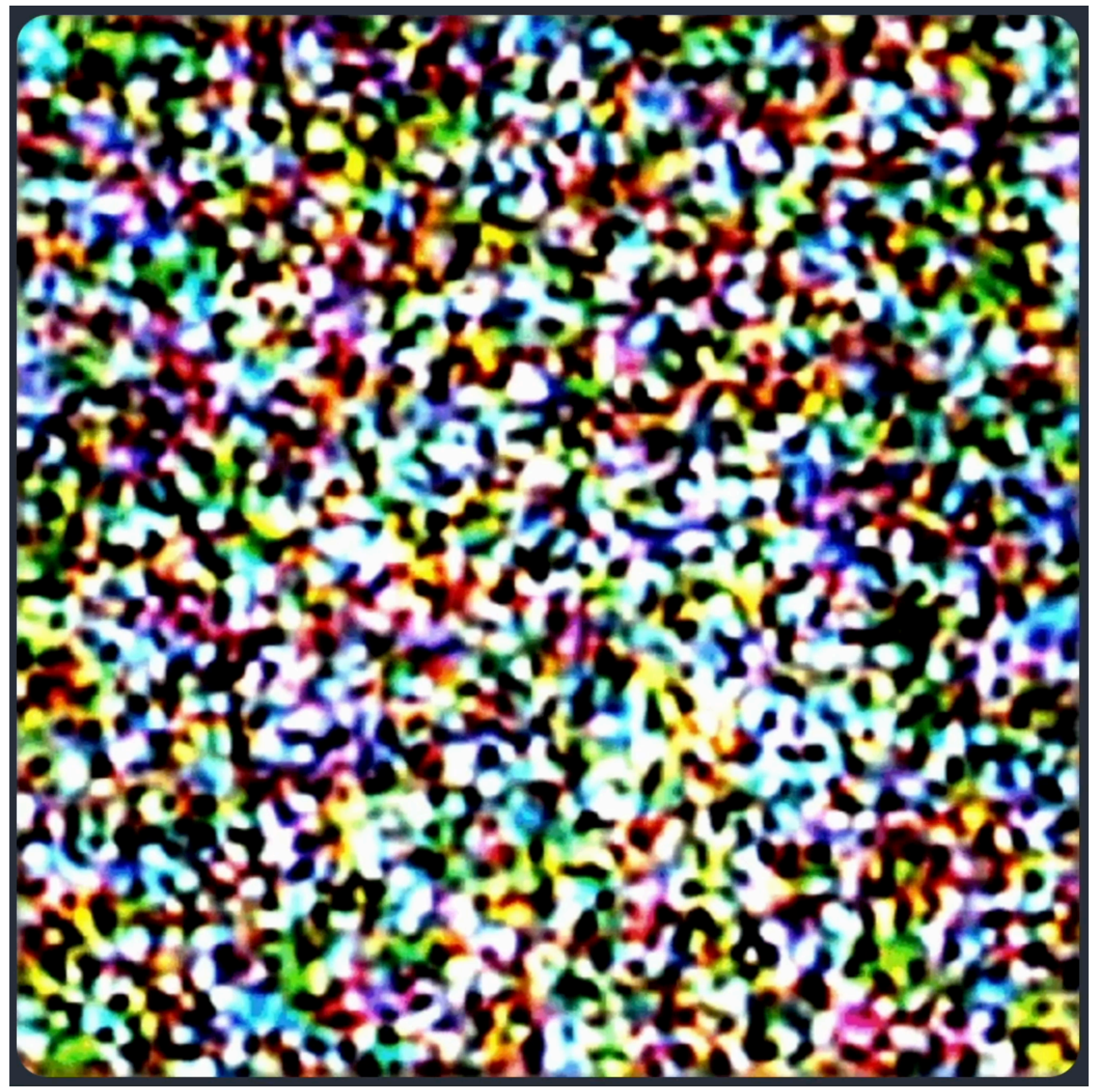
Stimulus reconstruction using AI





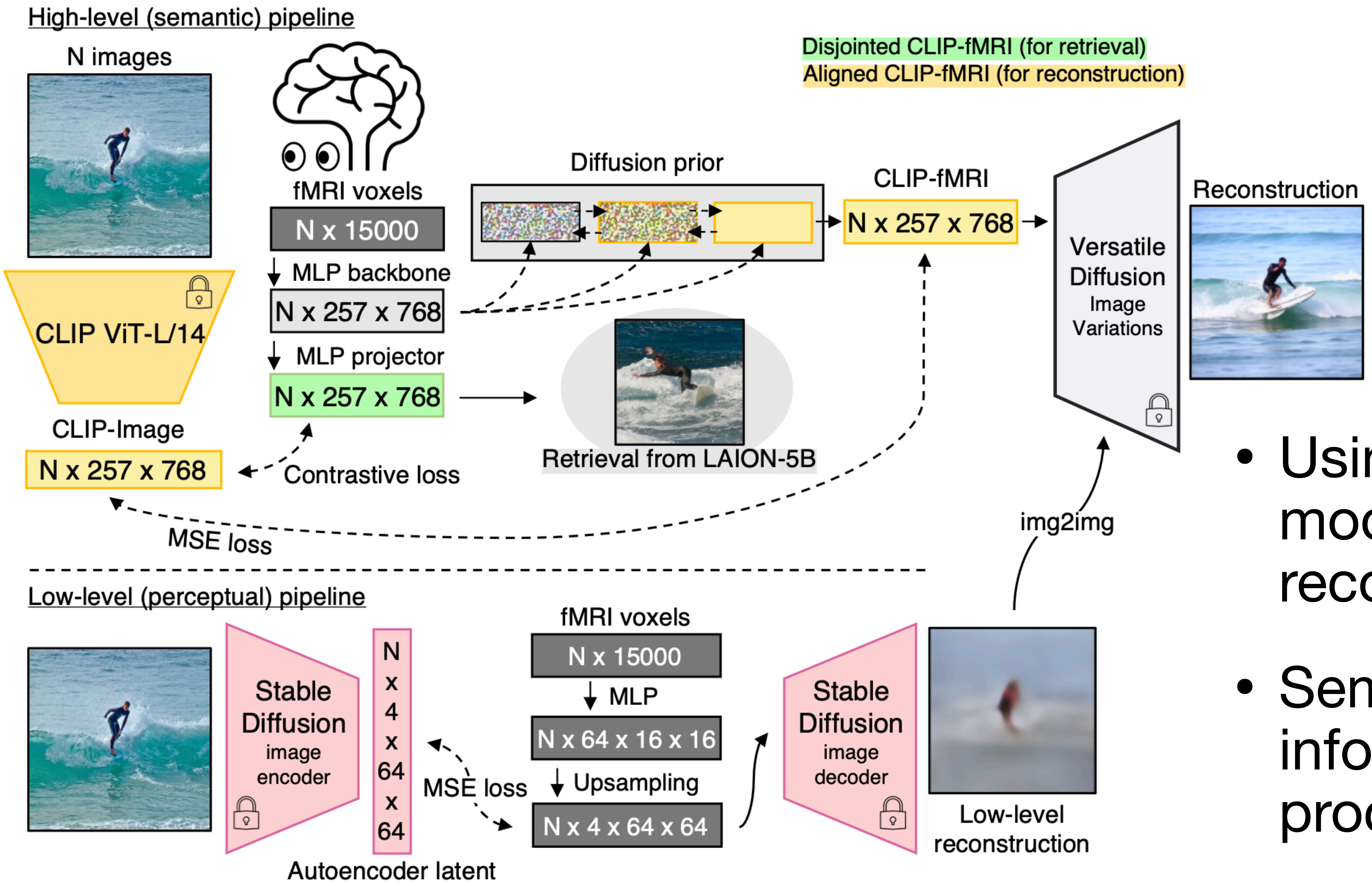
Stimulus reconstruction using AI

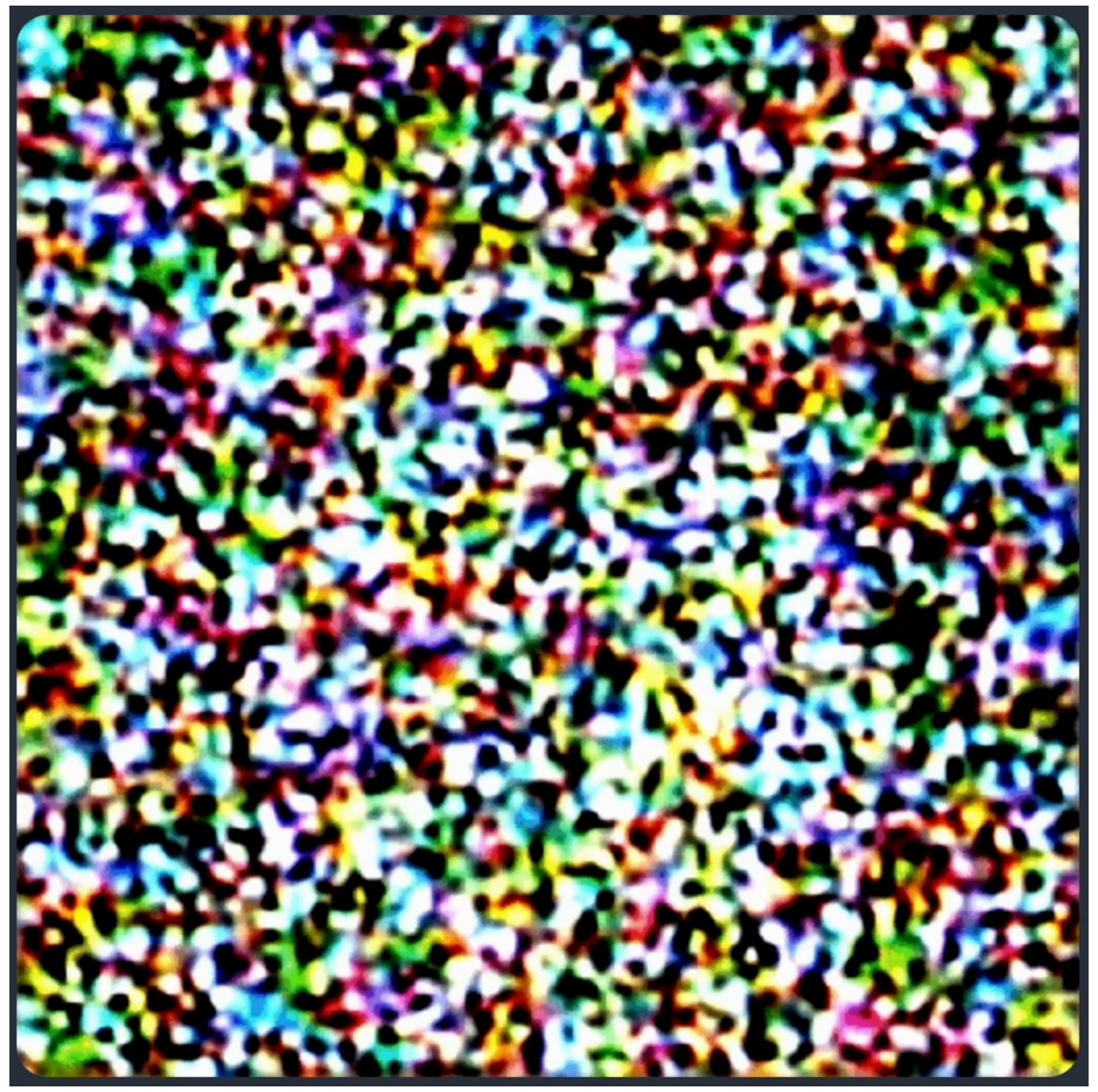




- Using diffusion models for reconstruction

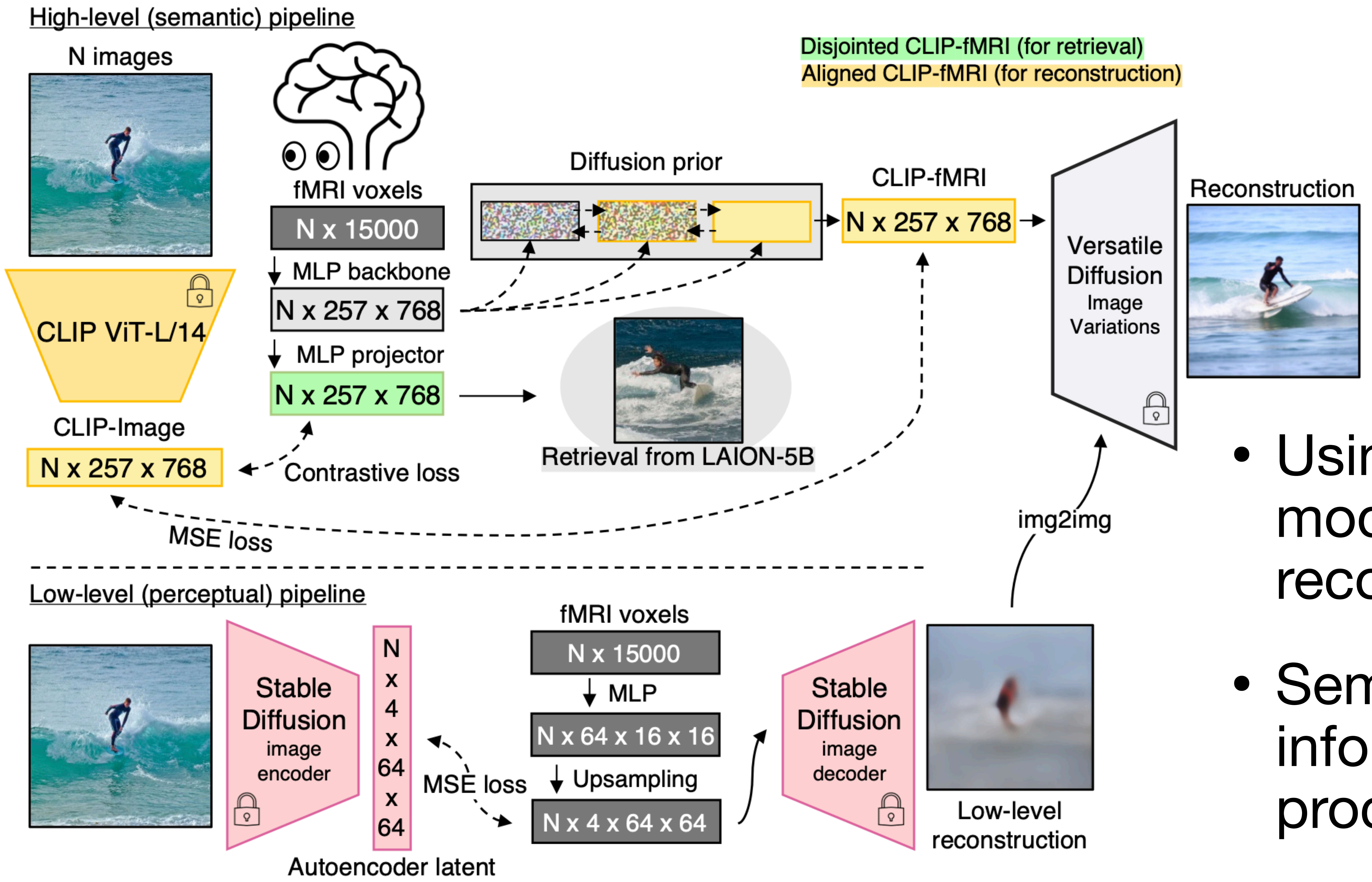
Stimulus reconstruction using AI

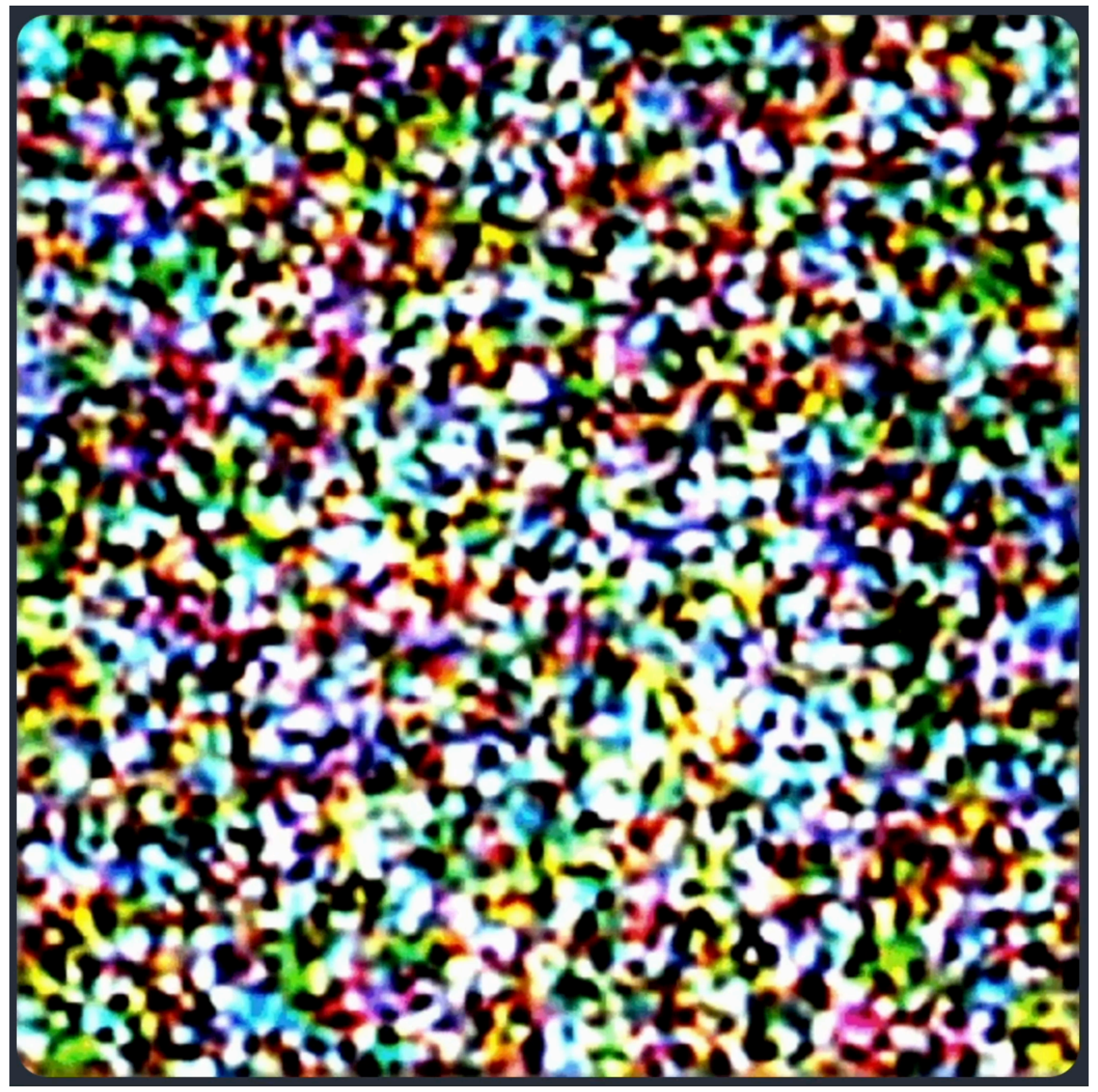




- Using diffusion models for reconstruction
- Semantic & Visual information processing streams

Stimulus reconstruction using AI





- Using diffusion models for reconstruction
- Semantic & Visual information processing streams

Take-away points

- Encoding/pRF models make neuron-voxel relationships explicit, and leverage the receptive field analogy
- Encoding/pRF models are explicit models of information processing:
    - ***Powerful!***
    - ***Hard!***
- Decoding based on single-voxel encoding model fits makes the patterns of MV**P**A explicit
    - ***Powerful!***
    - ***Hard!***