brain+cognitive sciences

CENTER FOR Brains Minds+ Machines

Computational modeling of human vision: Part II

Decoding

Kohitij Kar
Postdoctoral Associate
DiCarlo Lab

Recommended reading

Annu. Rev. Neurosci. 1998. 21:227–77

SENSE AND THE SINGLE NEURON: Probing the Physiology of Perception

A. J. Parker
University Laboratory of Physiology, University of Oxford, Parks Road, Oxford, OX1 3PT, United Kingdom

W. T. Newsome
Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305

Neuron
Perspective

How Does the Brain Solve Visual Object Recognition?

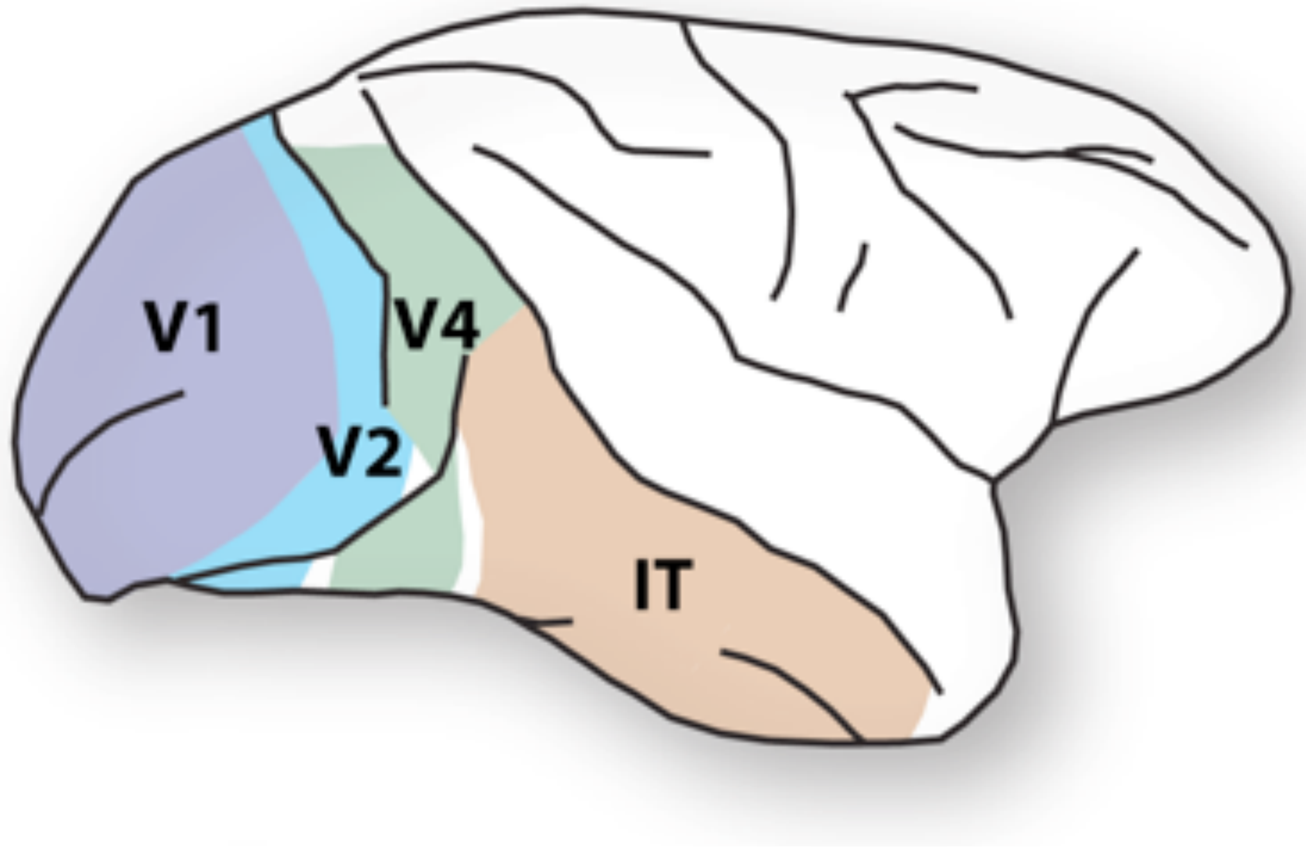
James J. DiCarlo,[1,*] Davide Zoccolan,[2] and Nicole C. Rust[3]
[1]Department of Brain and Cognitive Sciences and McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Cognitive Neuroscience and Neurobiology Sectors, International School for Advanced Studies (SISSA), Trieste, 34136, Italy
[3]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA
*Correspondence: dicarlo@mit.edu
DOI 10.1016/j.neuron.2012.01.010

How you might start to think about models of vision?

Visual Stimulus

(e.g. Image: glass of water)

Encoding models



Decoding models

Behavior

(e.g. pick up the glass of water)

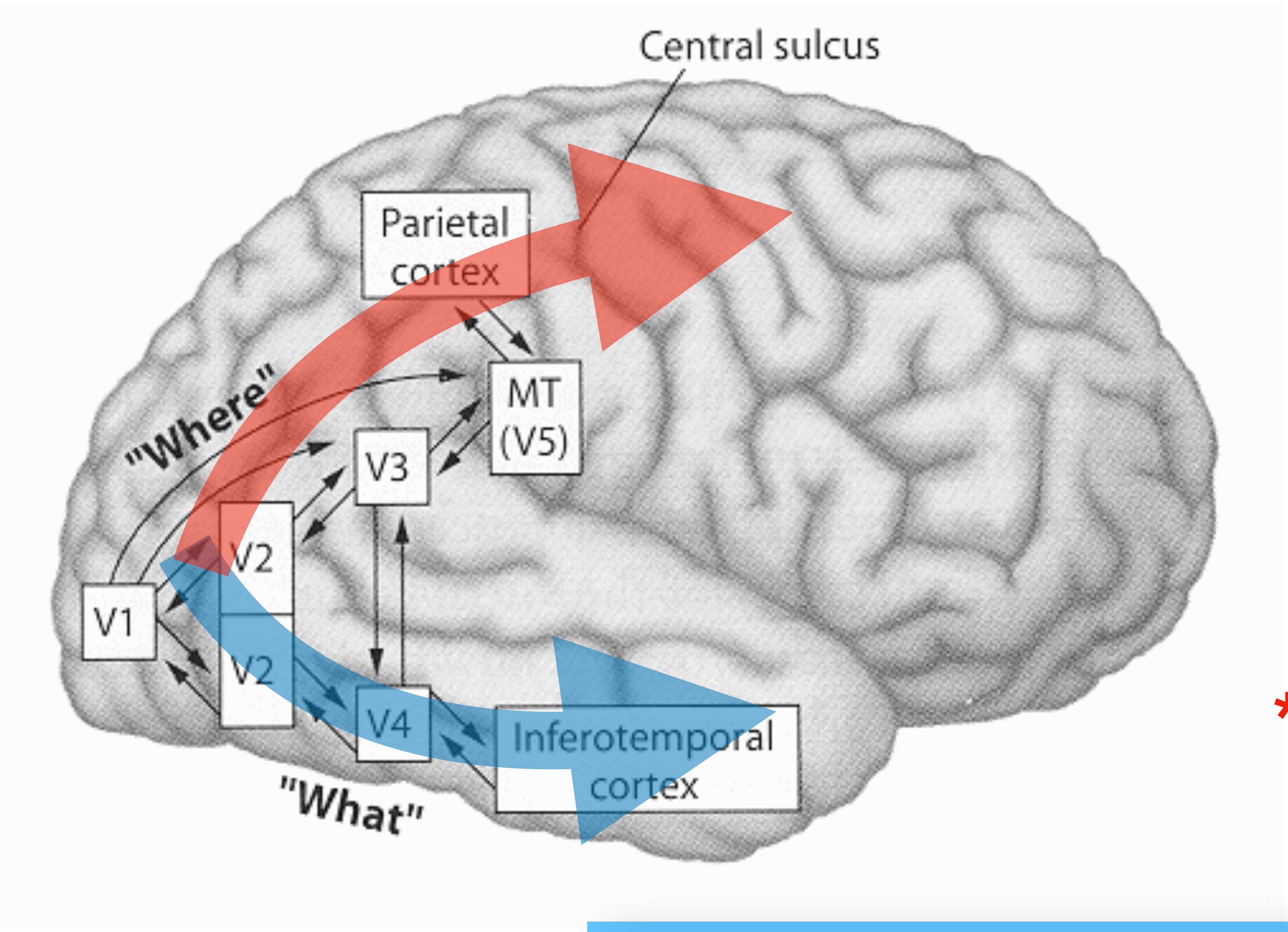
Encoding: How is the visual input represented in the brain?

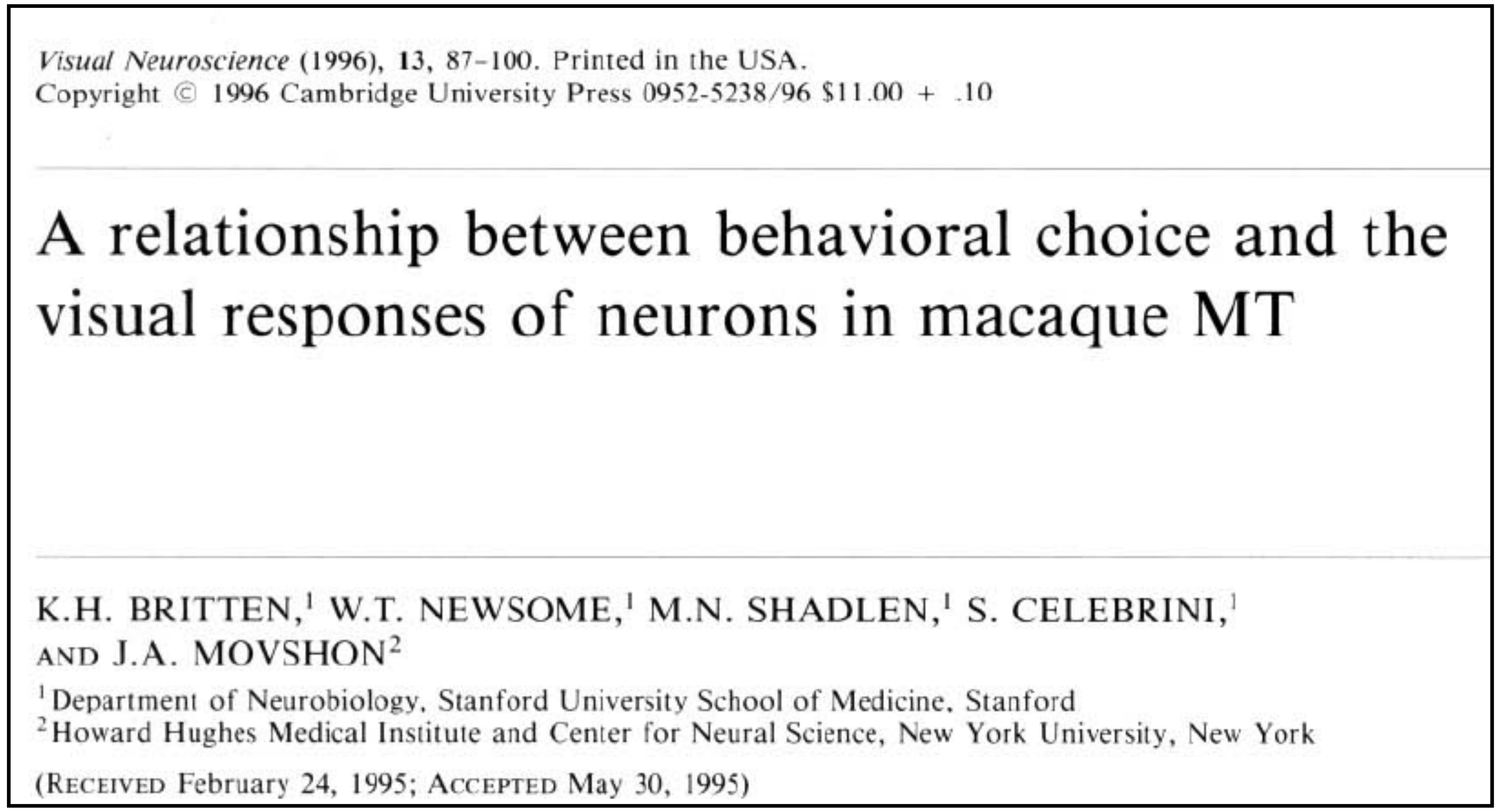
Decoding: How is the representation used by the brain to carry out behavioral tasks?

This needs to be defined quantitatively to evaluate how good the decoding models are

What model is implemented in the brain?

Let's talk today about behaviors associated with the ventral stream





Visual Neuroscience (1996), 13, 87–100. Printed in the USA.
Copyright © 1996 Cambridge University Press 0952-5238/96 $11.00 + .10

A relationship between behavioral choice and the visual responses of neurons in macaque MT

K.H. BRITTEN,[1] W.T. NEWSOME,[1] M.N. SHADLEN,[1] S. CELEBRINI,[1] AND J.A. MOVSHON[2]

[1]Department of Neurobiology, Stanford University School of Medicine, Stanford
[2]Howard Hughes Medical Institute and Center for Neural Science, New York University, New York

(RECEIVED February 24, 1995; ACCEPTED May 30, 1995)

* more on this during the Psychophysics and data analysis tutorial

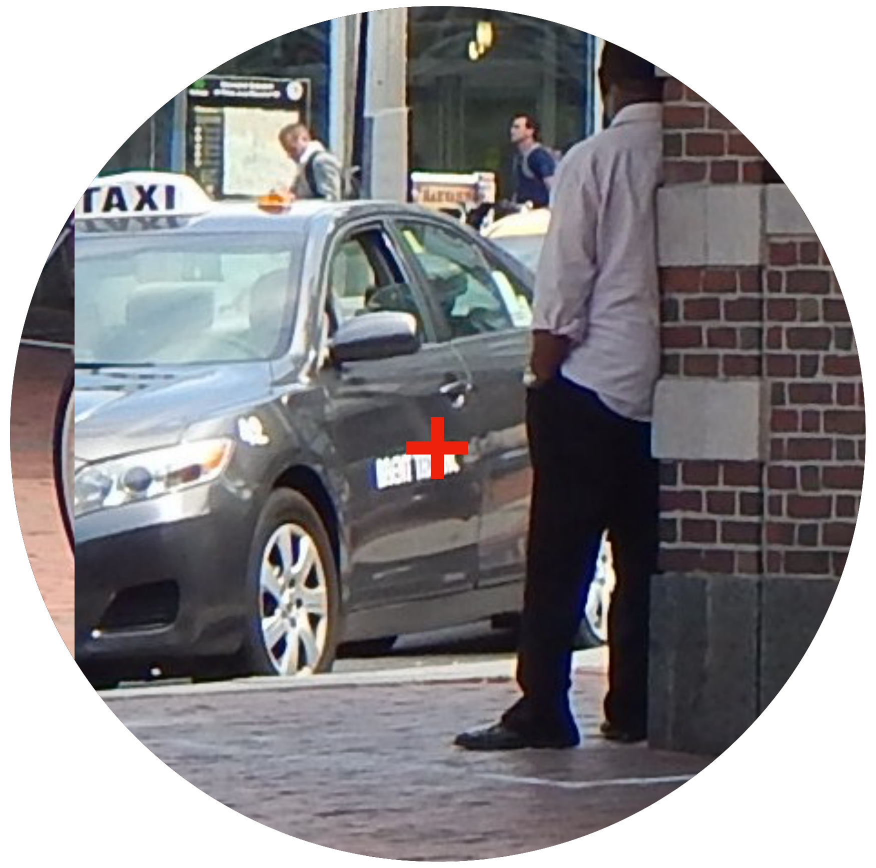
Aug 15: 8-9 pm

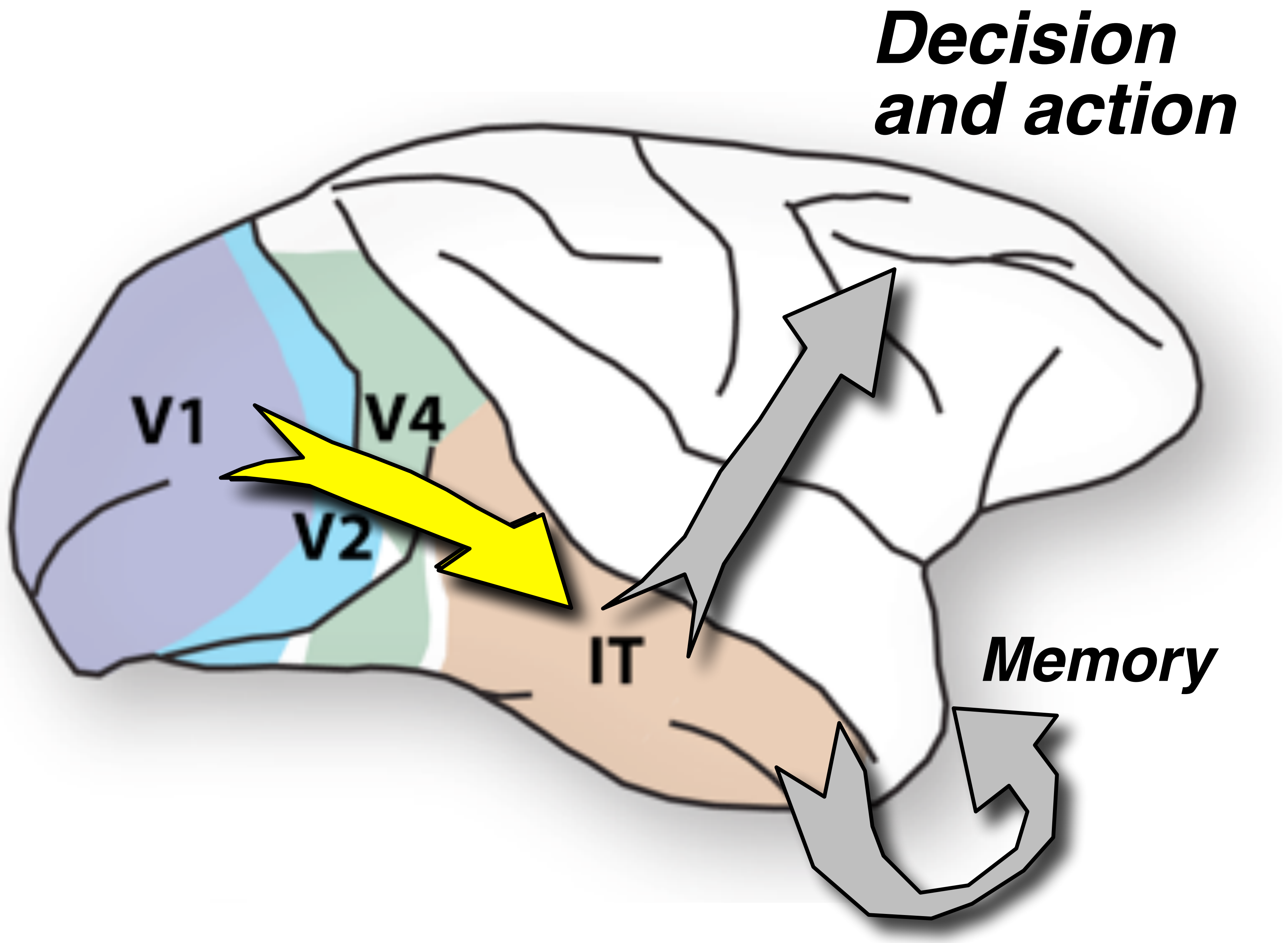
Lets mainly talk about visual object recognition today

"Understanding how biological visual systems recognize objects is one of the ultimate goals in computational neuroscience….." Riesenhuber and Poggio, 2000

Rapid object identity inference

Ventral Stream model of core object recognition





Behavioral Task

8 deg image at center of gaze, 100 msec viewing time

1

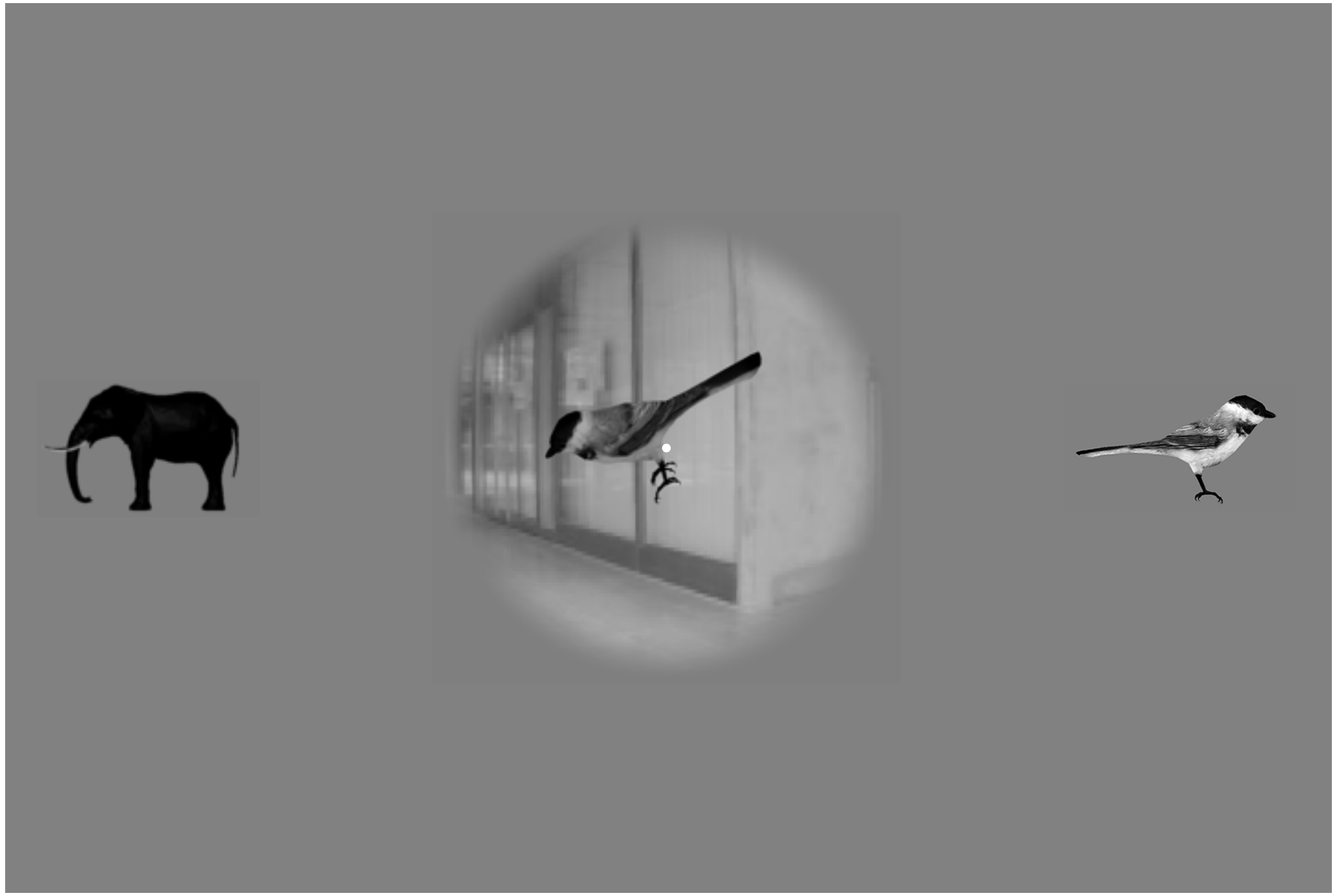
Define & operationalize a behavior

Behavioral Task

8 deg image at center of gaze, 100 msec viewing time

1

Define & operationalize a behavior

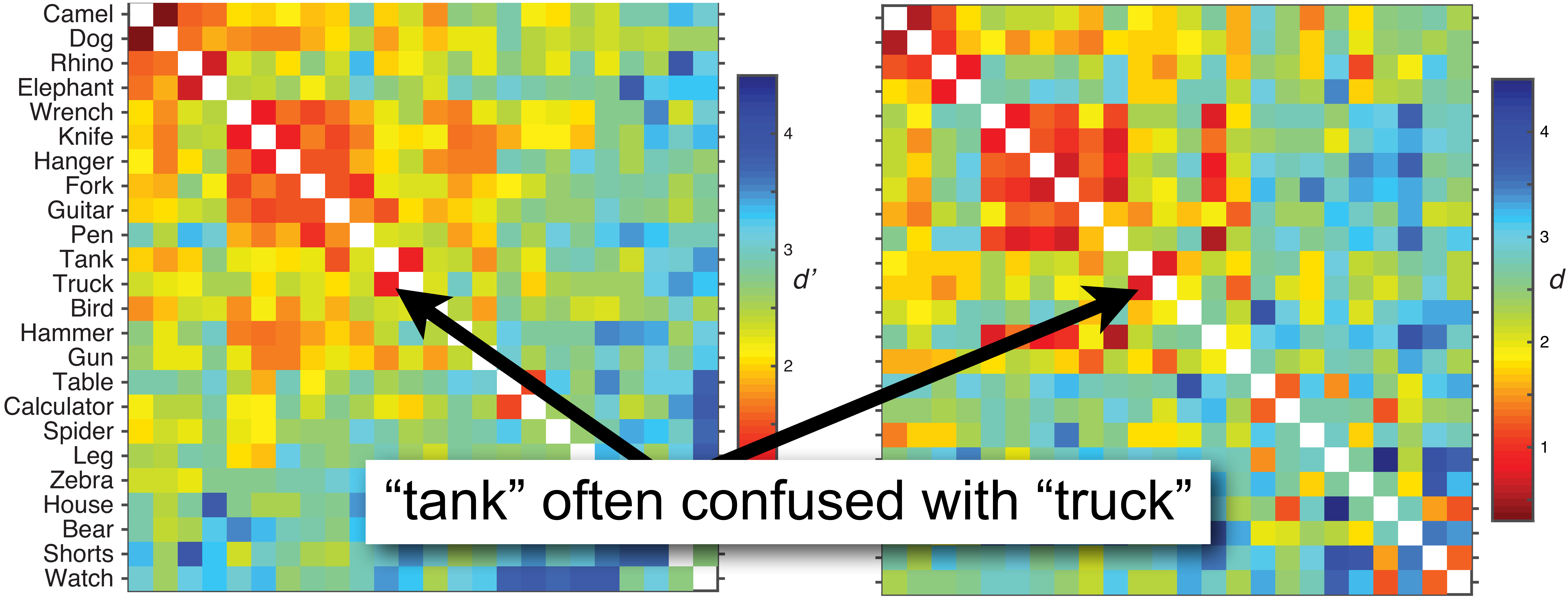


What animal model shall we choose to study the behavior?

2

Which animal model shall we choose?

Behavior of primate species A | **Behavior of primate species B**



Summary: human behavior = monkey behavior

Now go get more internal component measurements!

Rajalingham, Schmidt, & DiCarlo, ***J. Neuroscience*** *(2015)*
Rajalingham, Issa, Bashivan, Kar, Schmidt, & DiCarlo, ***J. Neuroscience*** *(2018)*

measure

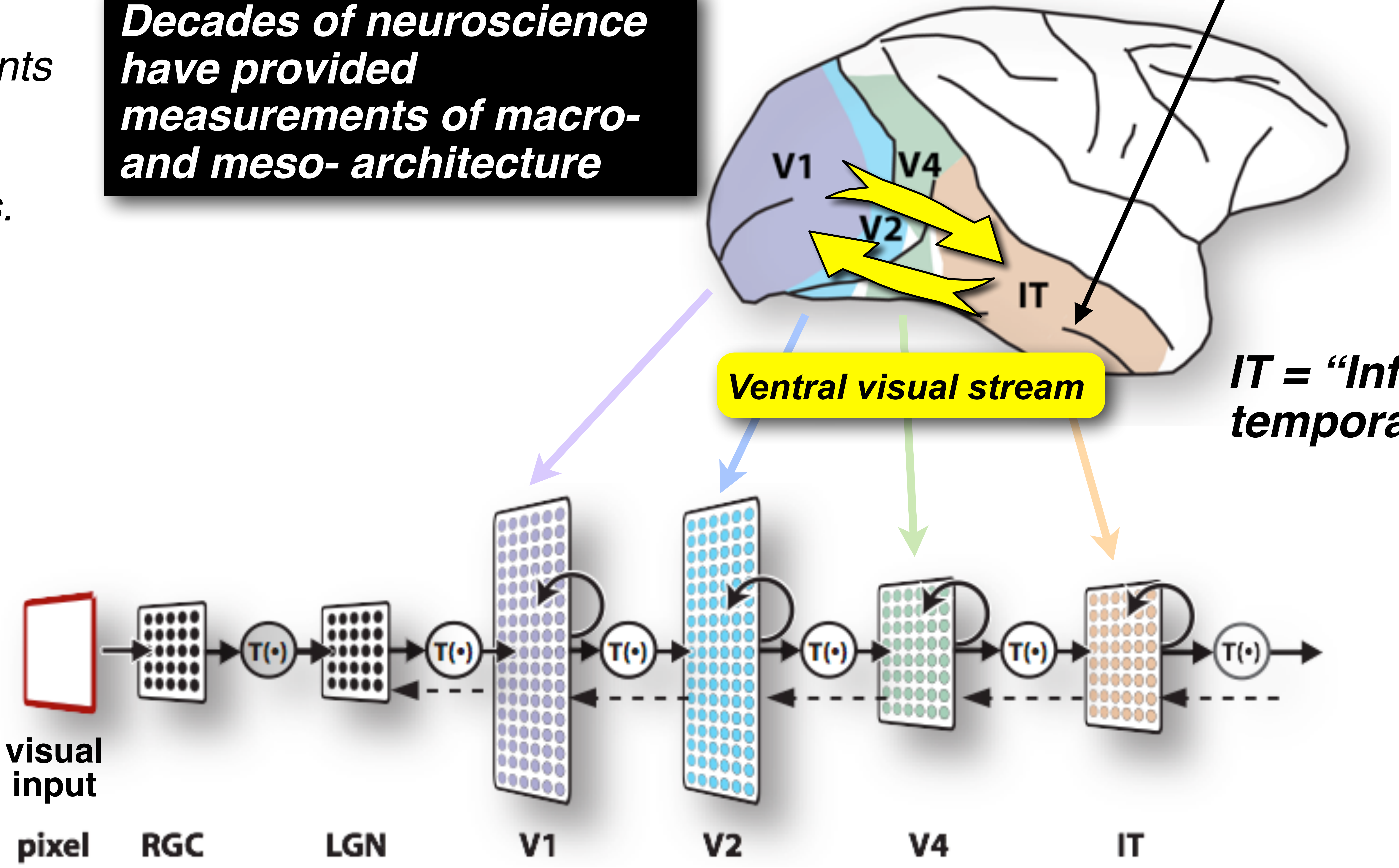
3

Get measurements of internal system components.

Where to look in the monkey brain?

Decades of neuroscience have provided measurements of macro- and meso- architecture

Lesions here result in deficits in visual recognition.



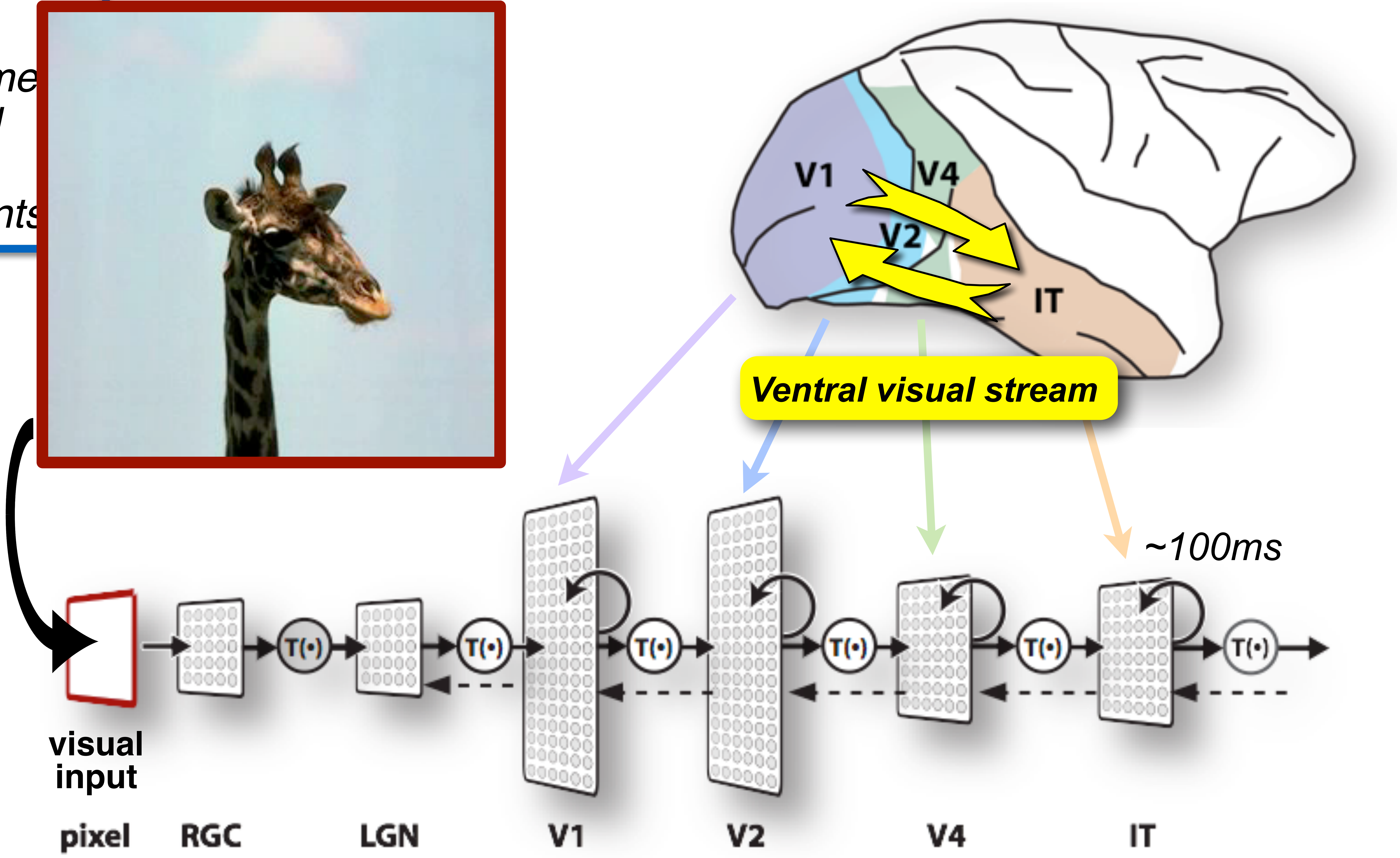
IT = "Inferior temporal cortex"

measure

3

Get measureme[nts] of internal system component[s]

Decades of neuroscience physiology:

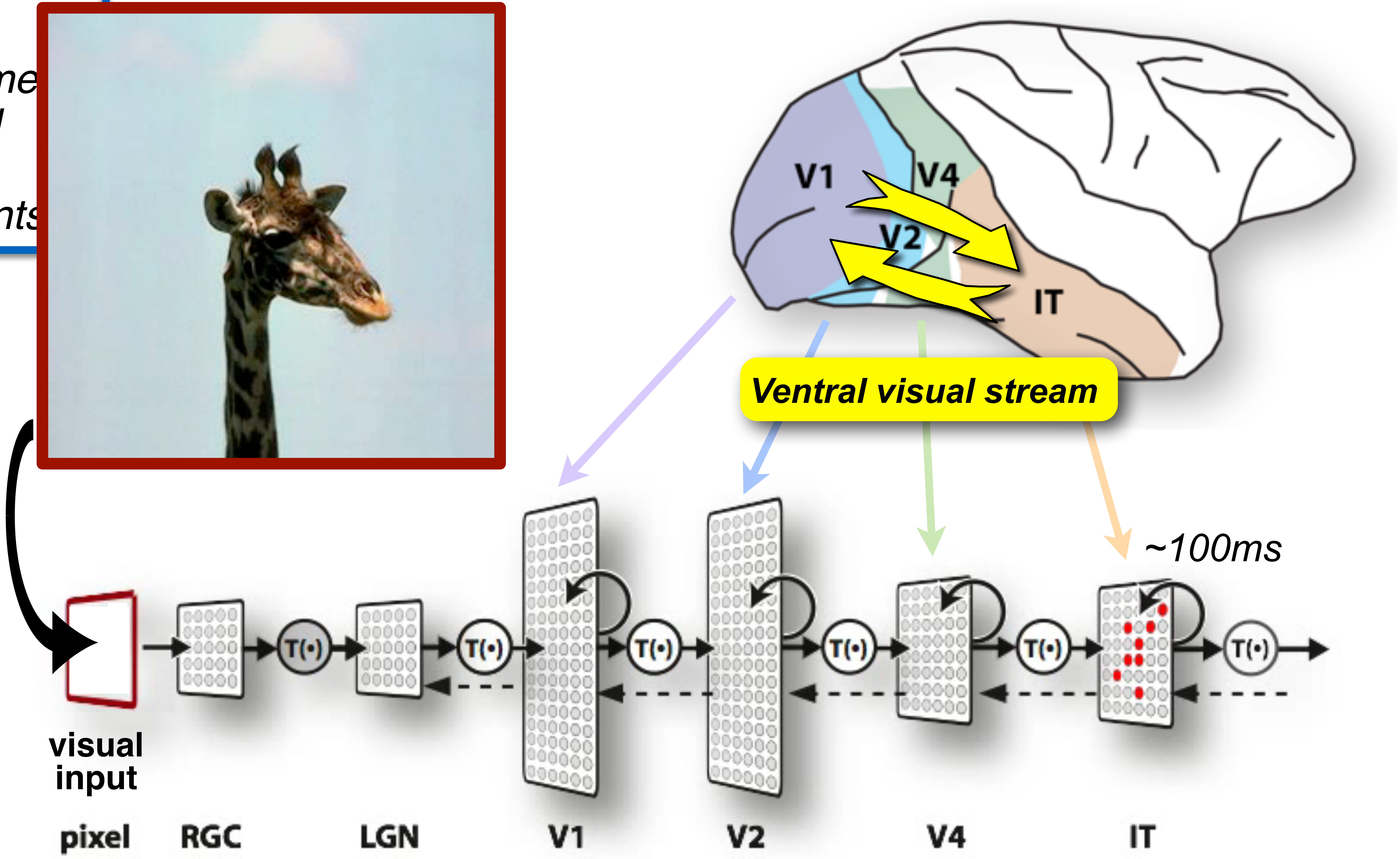


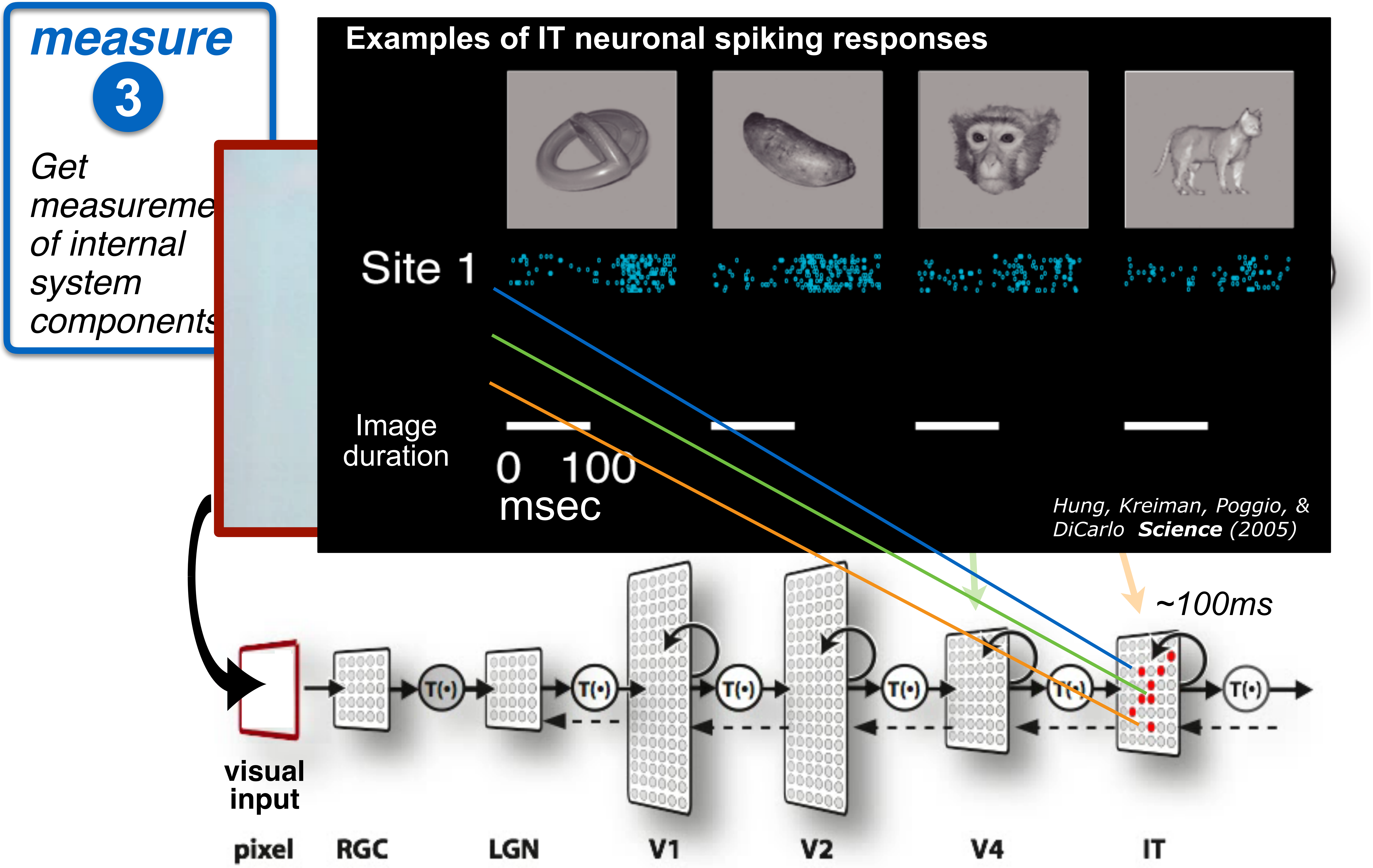
measure

3

Get measureme of internal system components

Decades of neuroscience physiology:



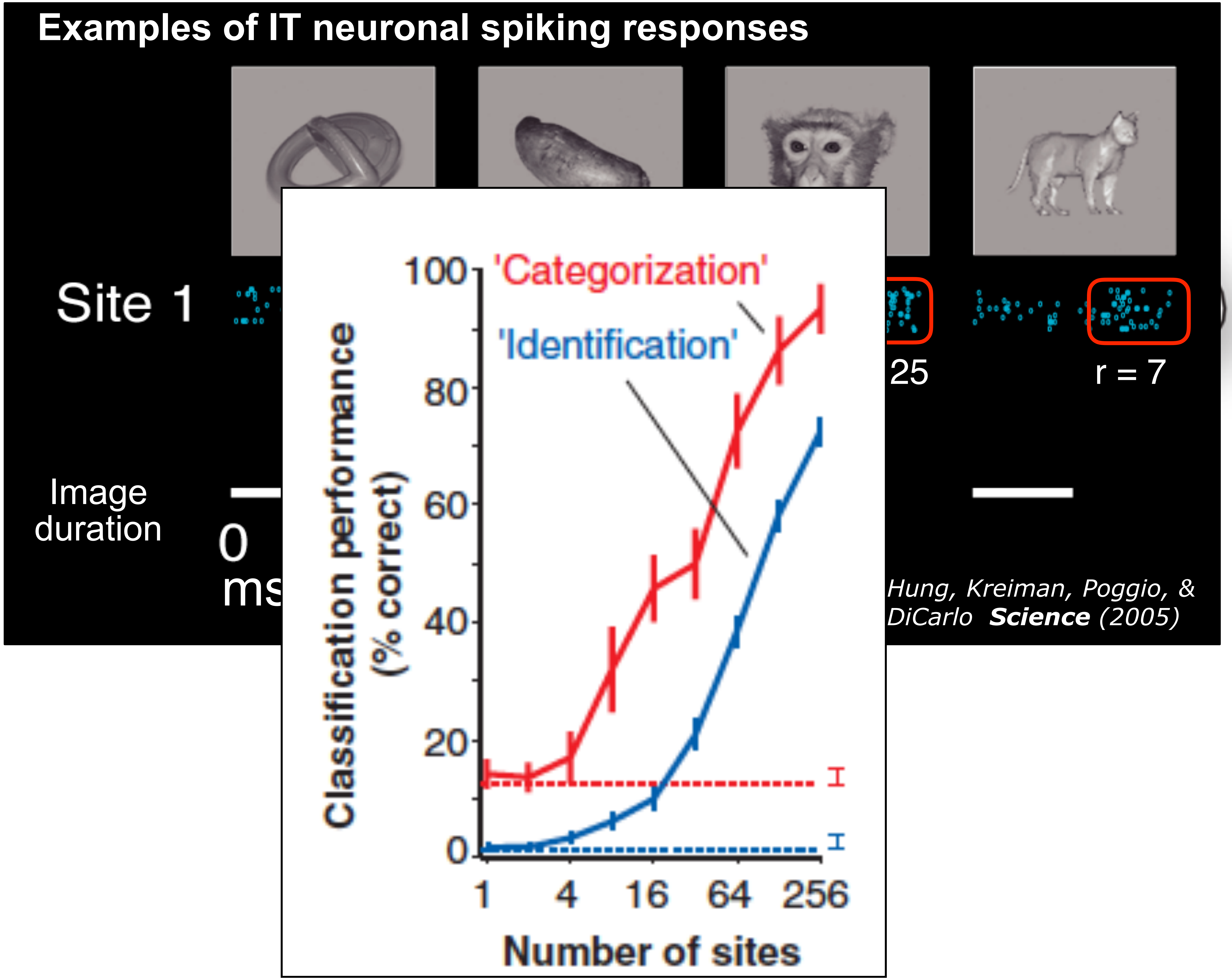


measure
3
Get measureme
of internal
system
components
Examples of IT neuronal spiking responses
Site 1
Image duration
0 100
msec
Hung, Kreiman, Poggio, & DiCarlo Science (2005)
~100ms
visual input
pixel
RGC
LGN
V1
V2
V4
IT
T(•)

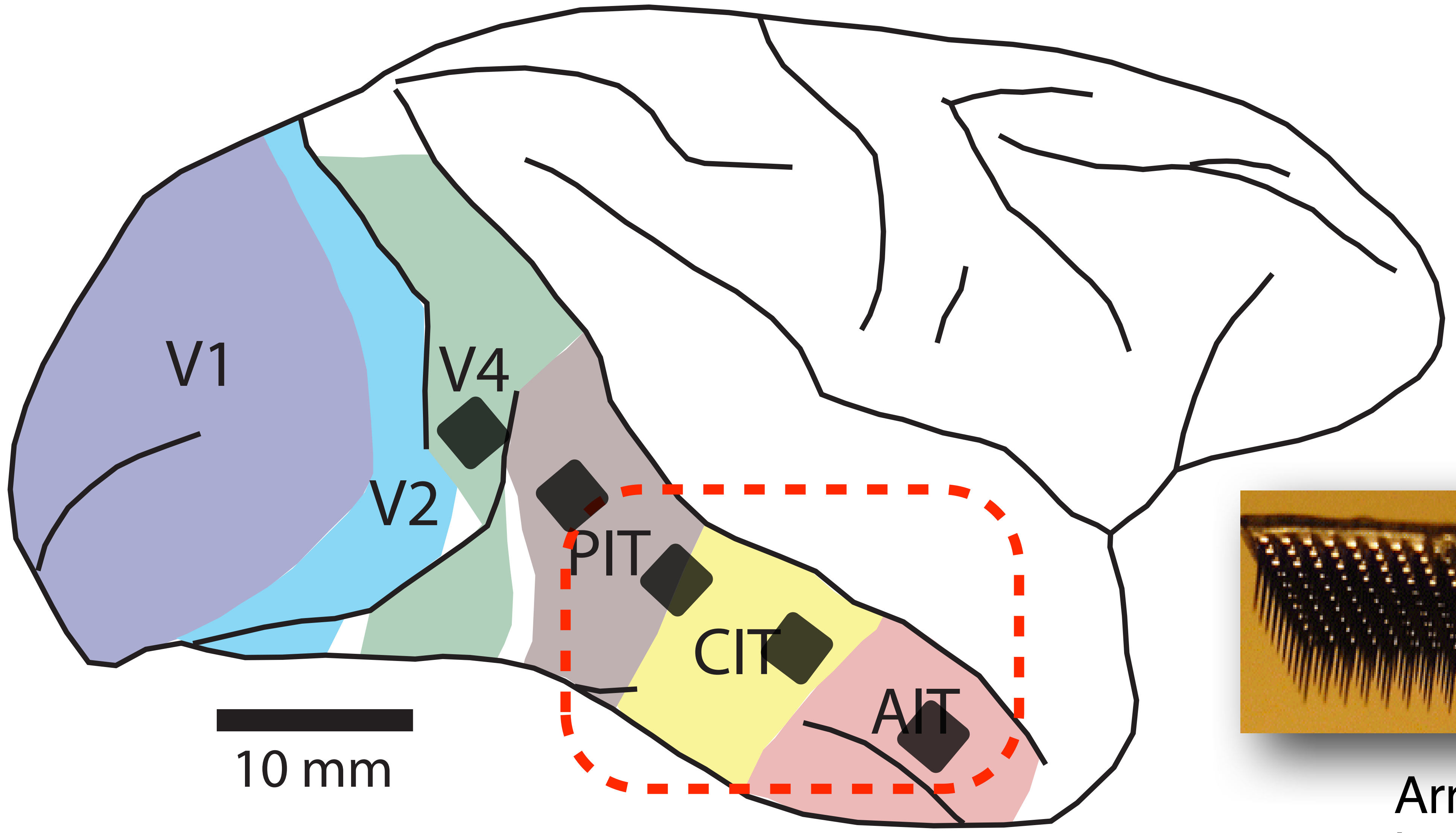
measure

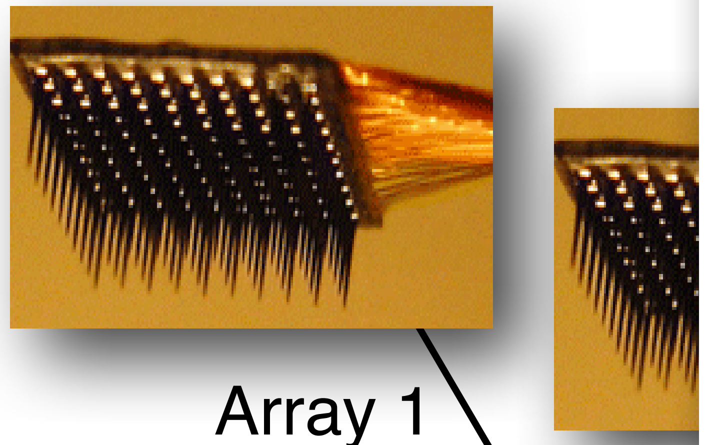
3

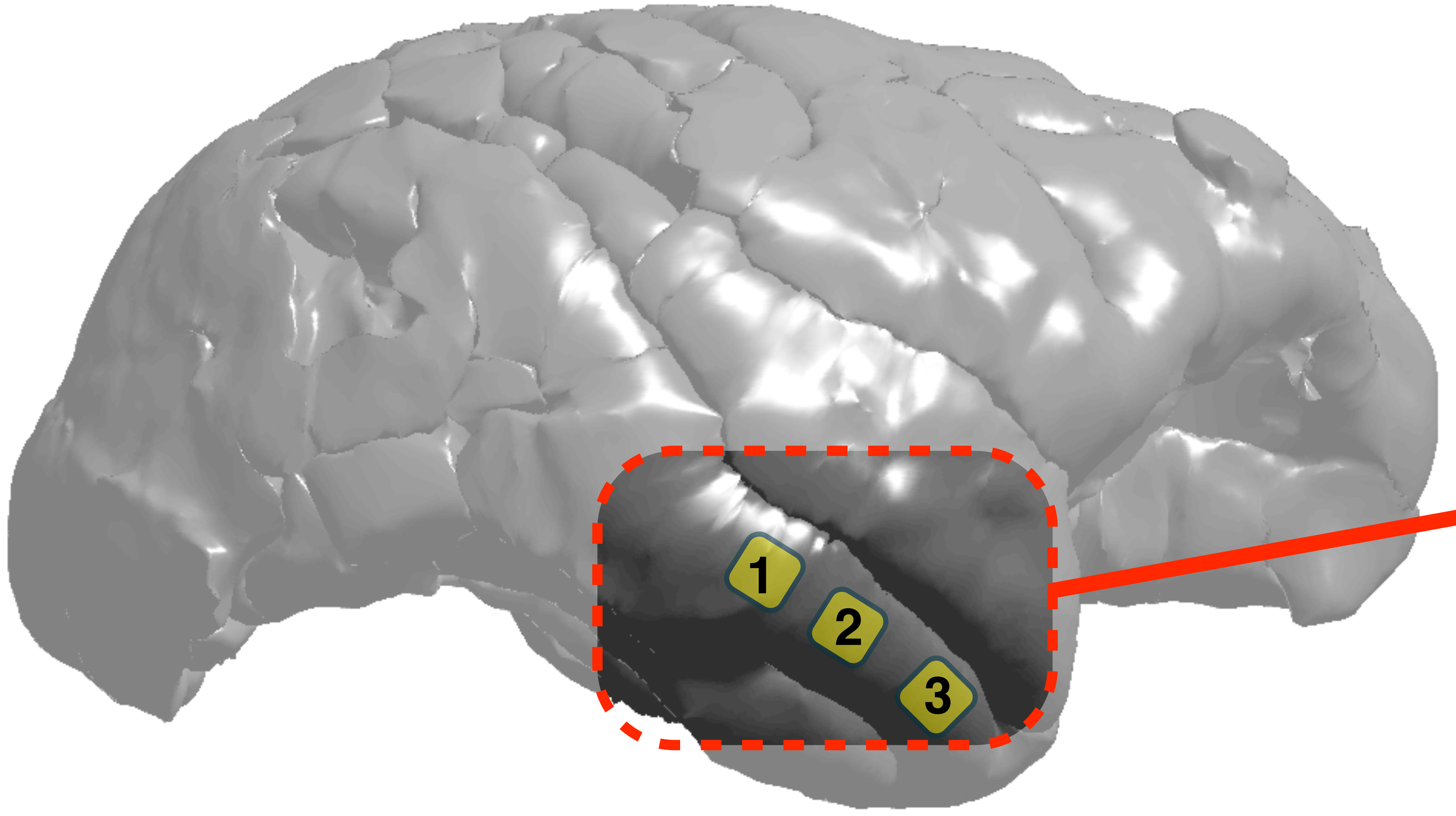
Get measurements of internal system components.

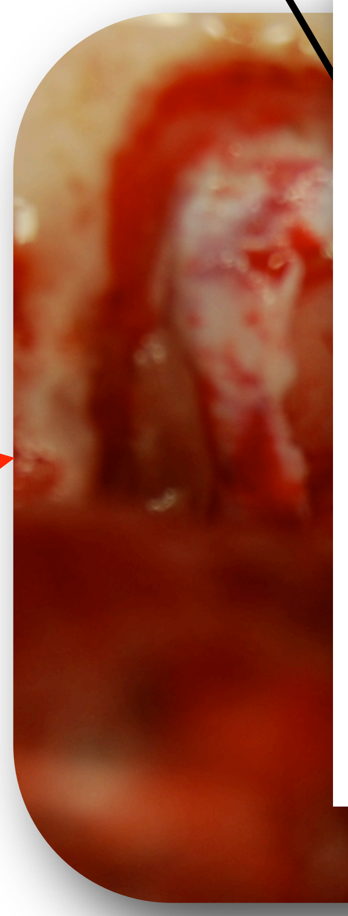


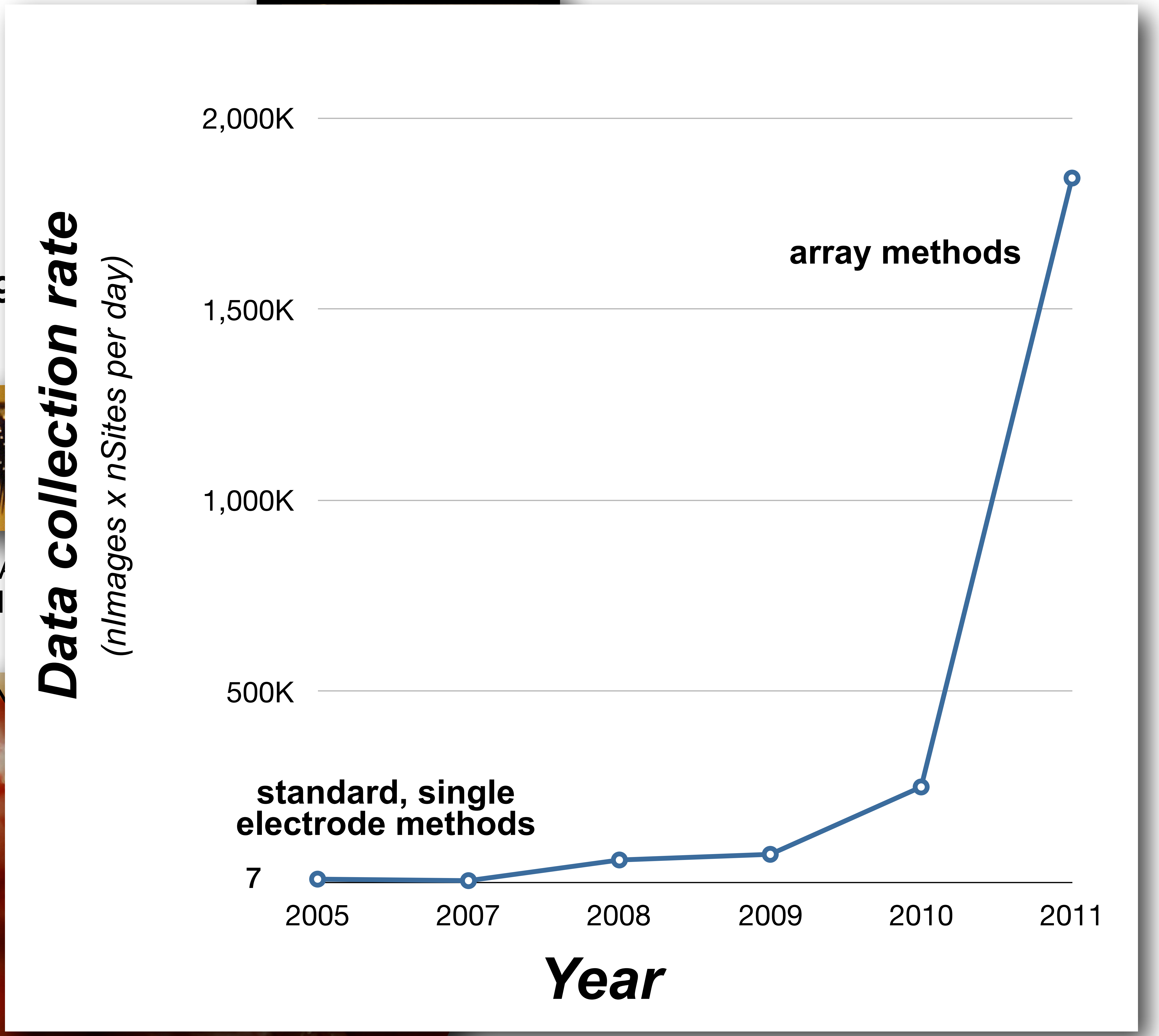
The data collection over the years have scaled up

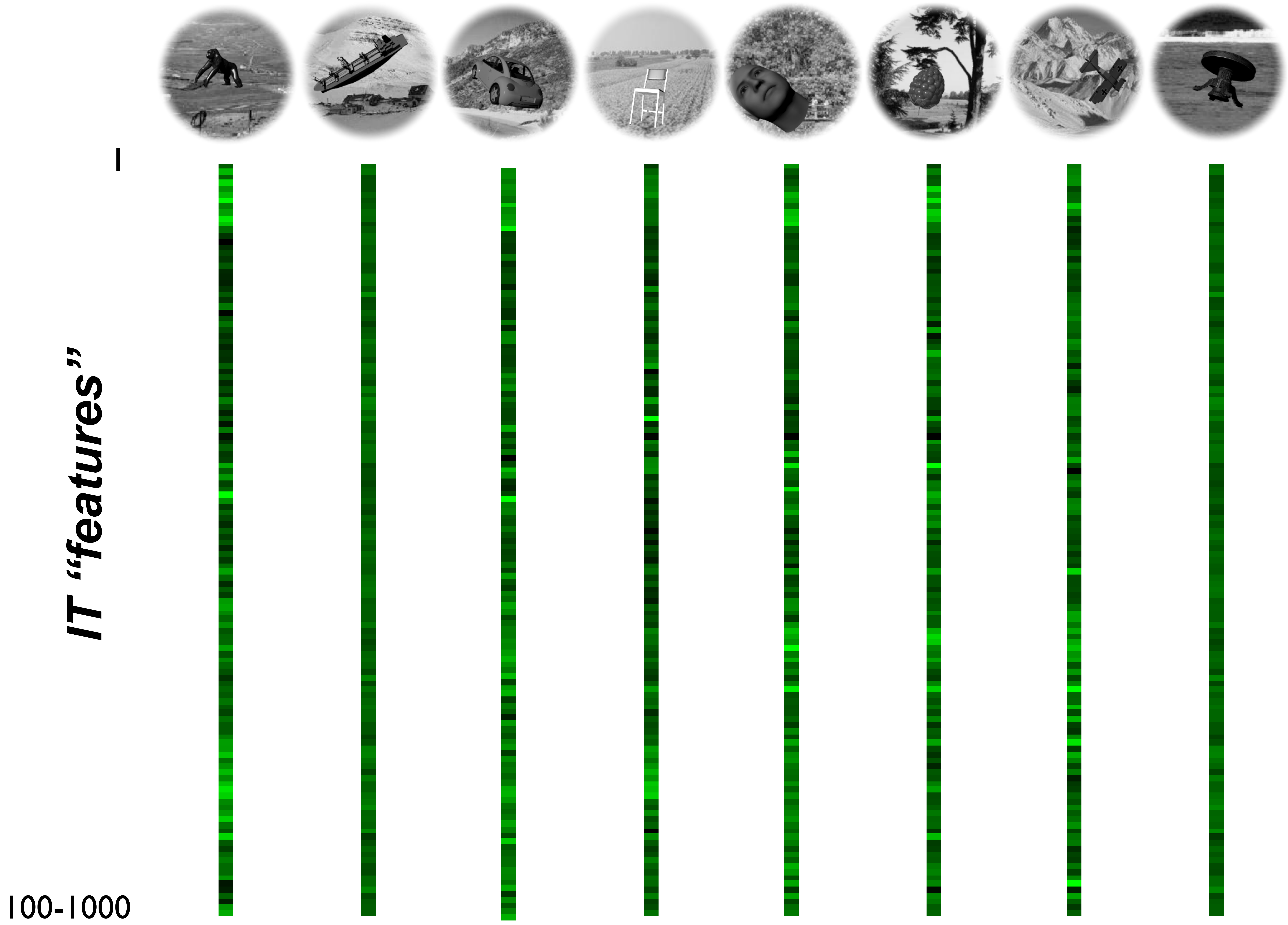




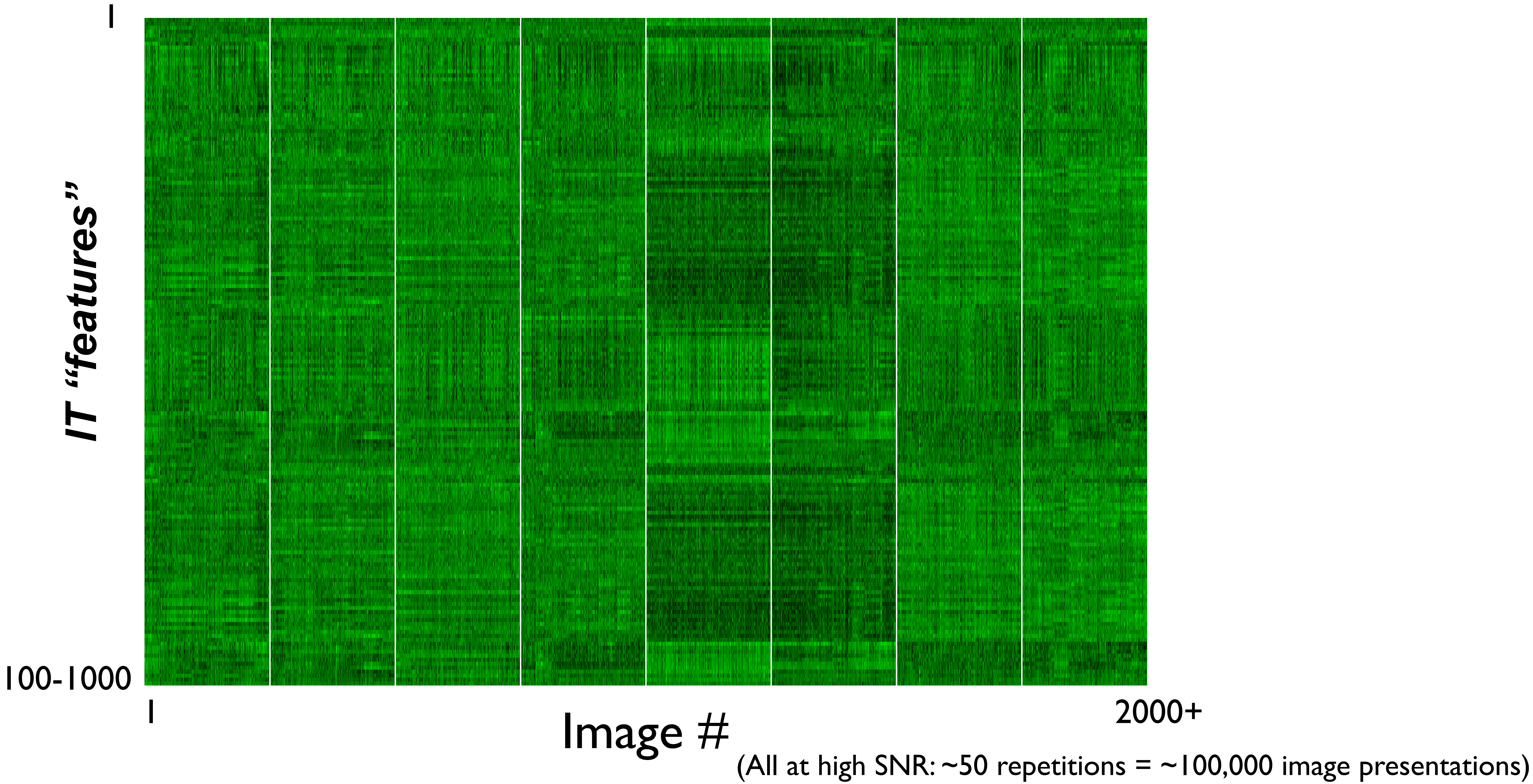








1
IT "features"
100-1000



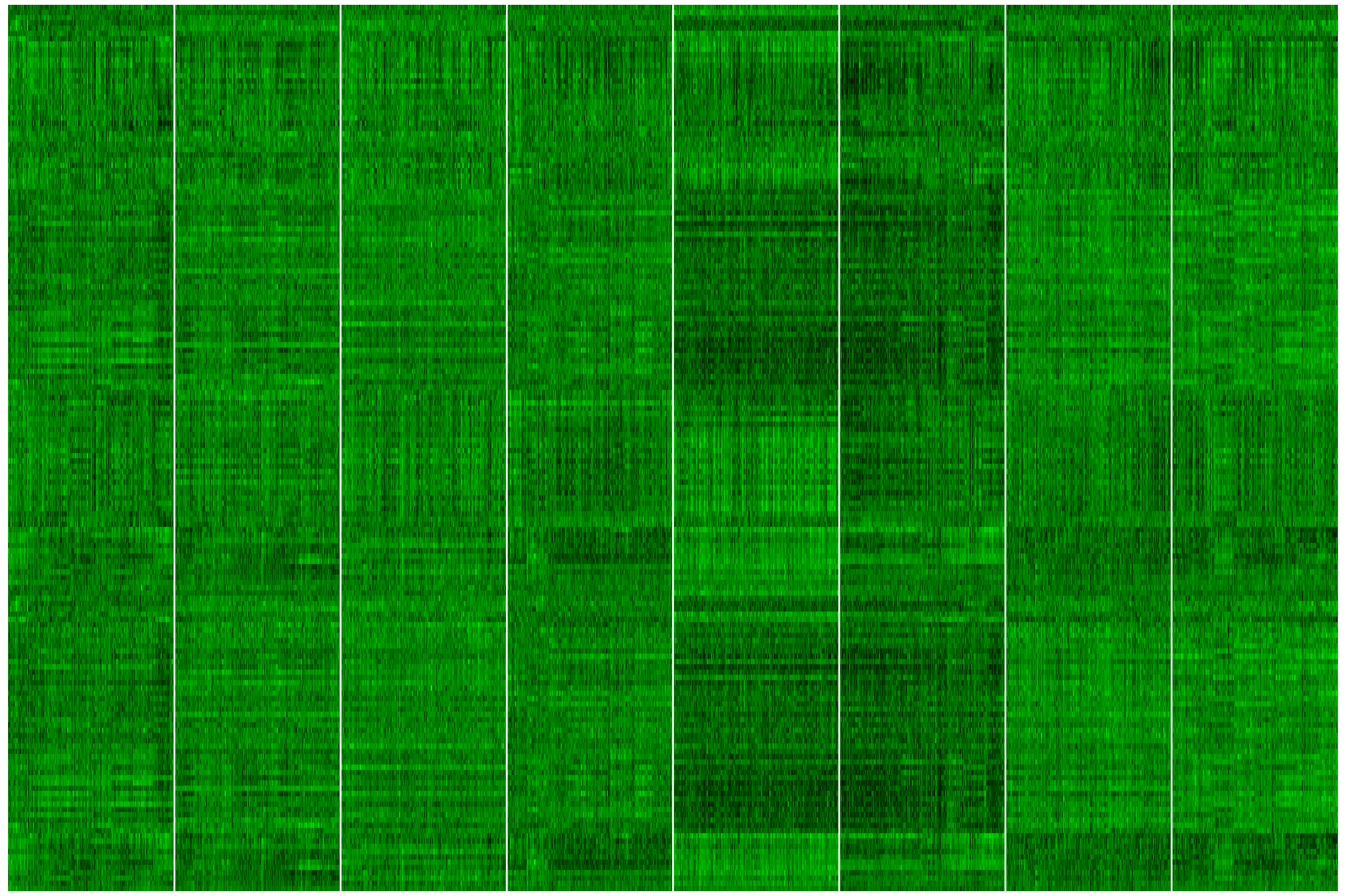
1
100-1000
IT “features”
1
2000+
Image #
(All at high SNR: ~50 repetitions = ~100,000 image presentations)

measure

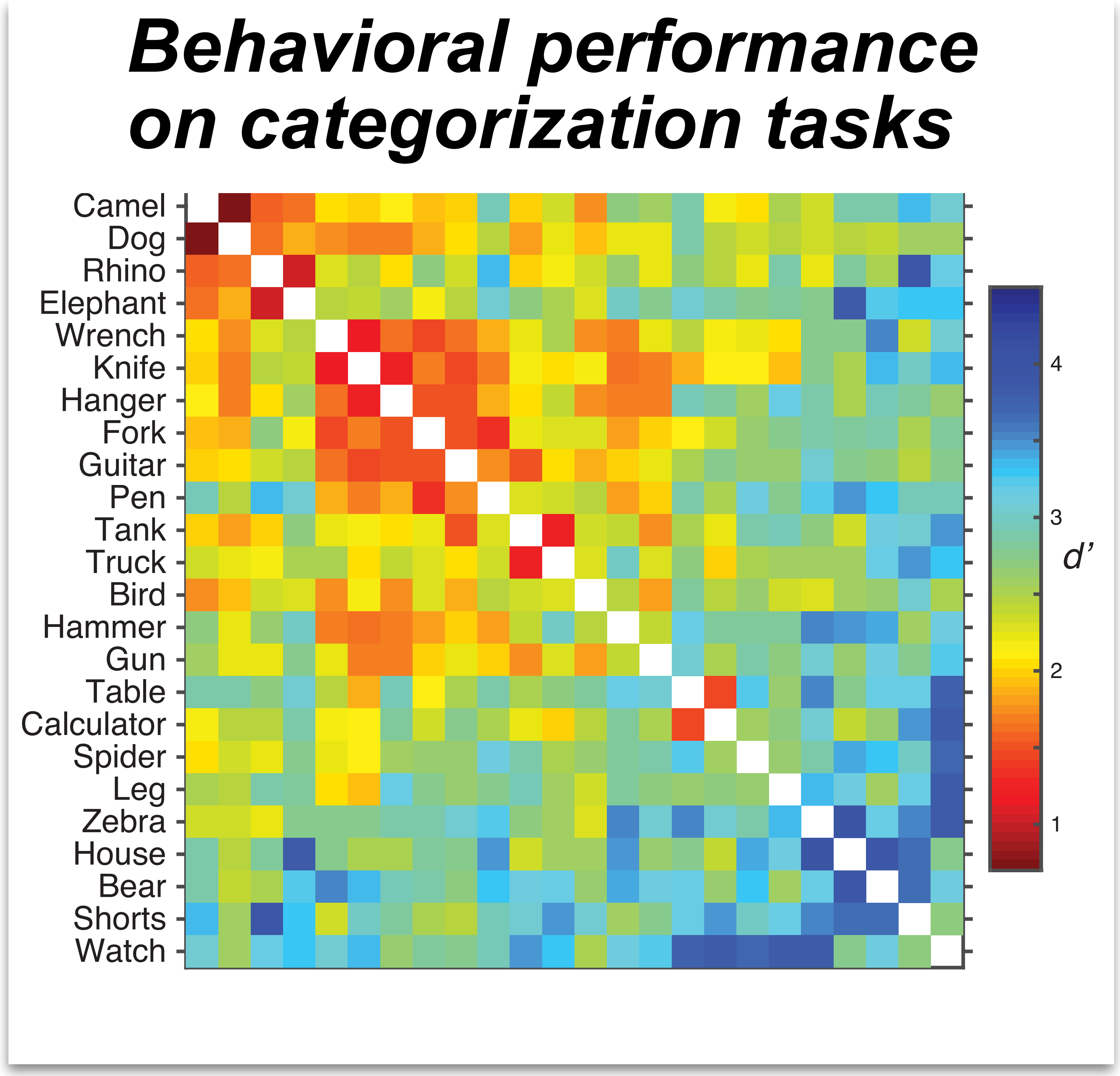
4

Test population decoding models that can fully explain behavior

IT "features"

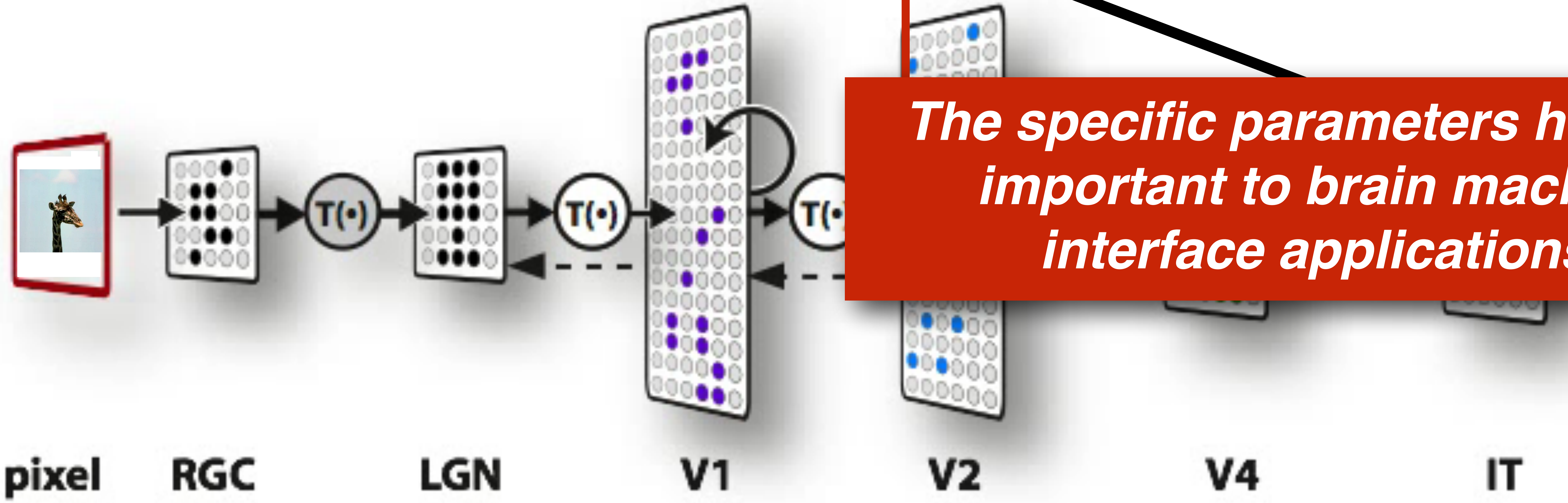


Linking neurons to behavior



*Majaj, Hong, Soloman, & DiCarlo, **J Neuro** (2015)*

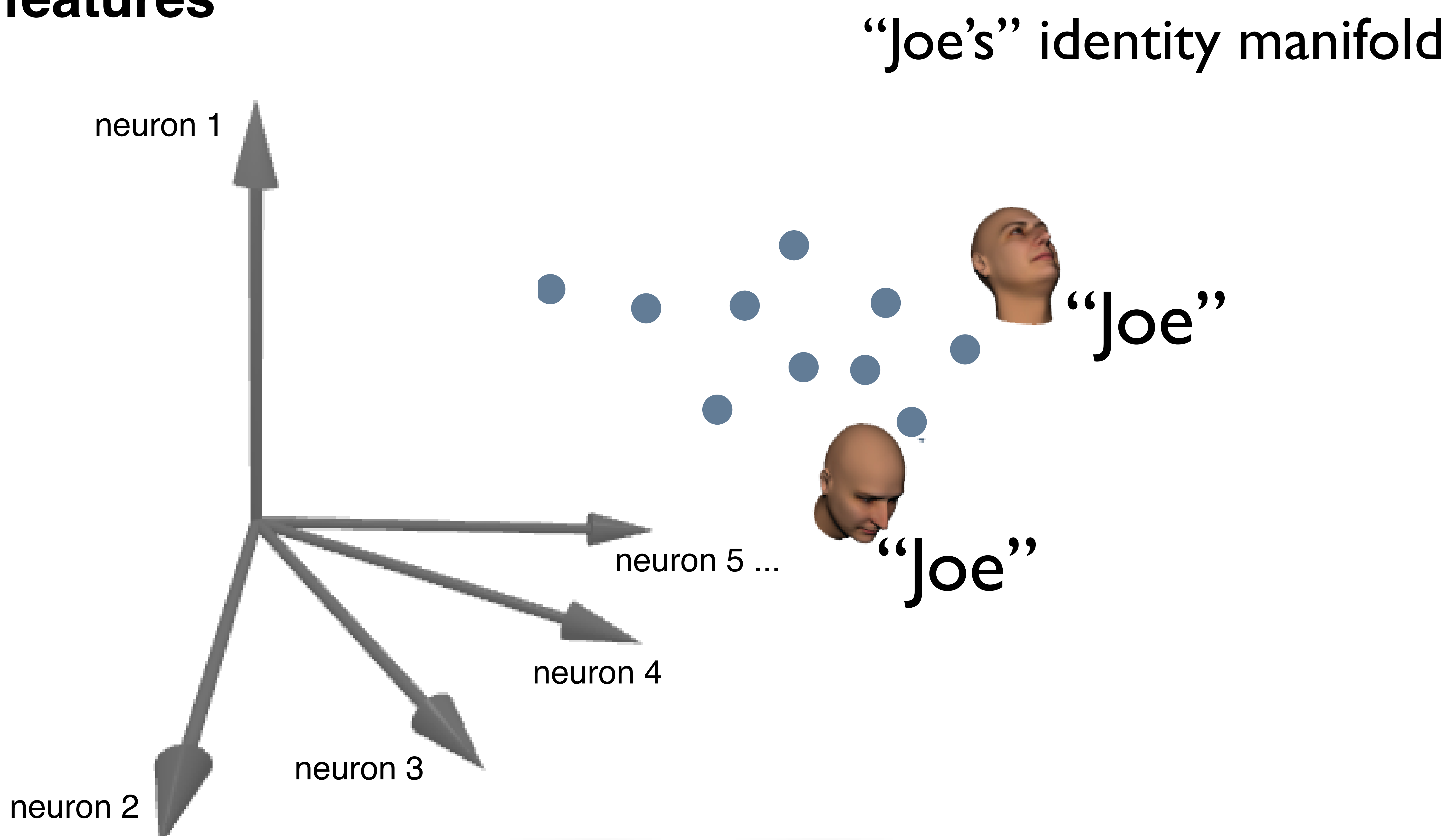
Hong, Yamins*, Majaj & DiCarlo. **Nat. Neuro.** (2016)*

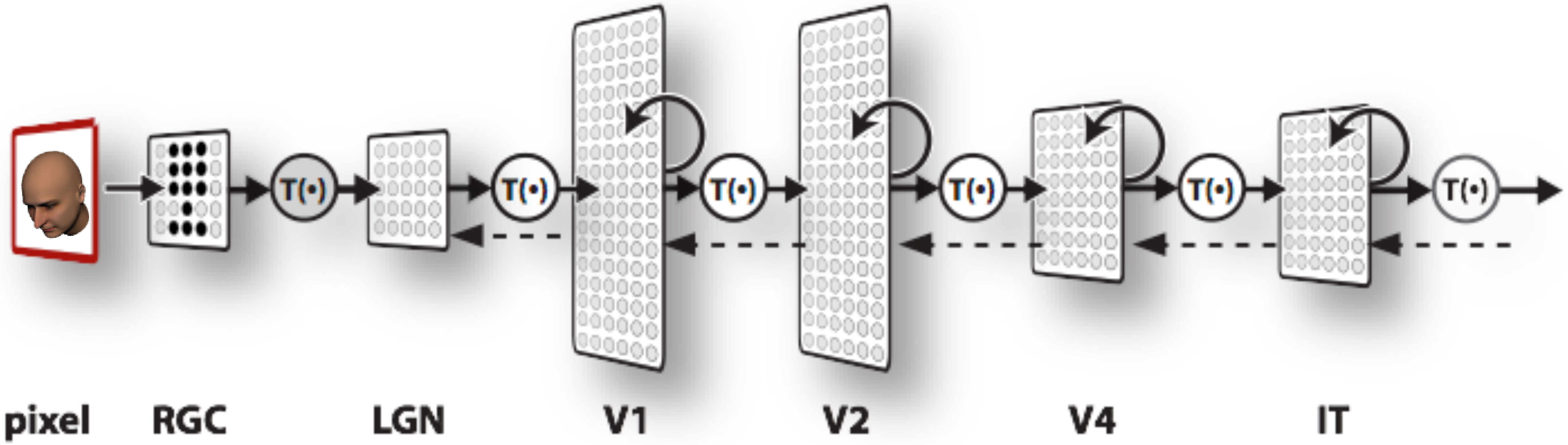


The specific parameters here are important to brain machine interface applications.

Consistency
(Spearman correlation)
0.95
0.9
0.8
0.6
0.4
0.2
Human-to-human consistency
Pixel
V1-like
PHOG
SIFT
SLF
L3
Computervision
algorithm hypotheses
V4-based
hypotheses
Max
Corr
SVM
LAWs of RAD
IT-based
hypotheses
Max
Corr
SVM
LAWs of RAD
Behavioral performance
on categorization tasks
Camel
Dog
Rhino
Elephant
Wrench
Knife
Hanger
Fork
Guitar
Pen
Tank
Truck
Bird
Hammer
Gun
Table
Calculator
Spider
Leg
Zebra
House
Bear
Shorts
Watch
d'
king
ns to
vior
euro (2015)
Neuro. (2016)
e are
Image index
1
2
3
4
5
6
5760
168
IT
1
70
170
pixel
RGC

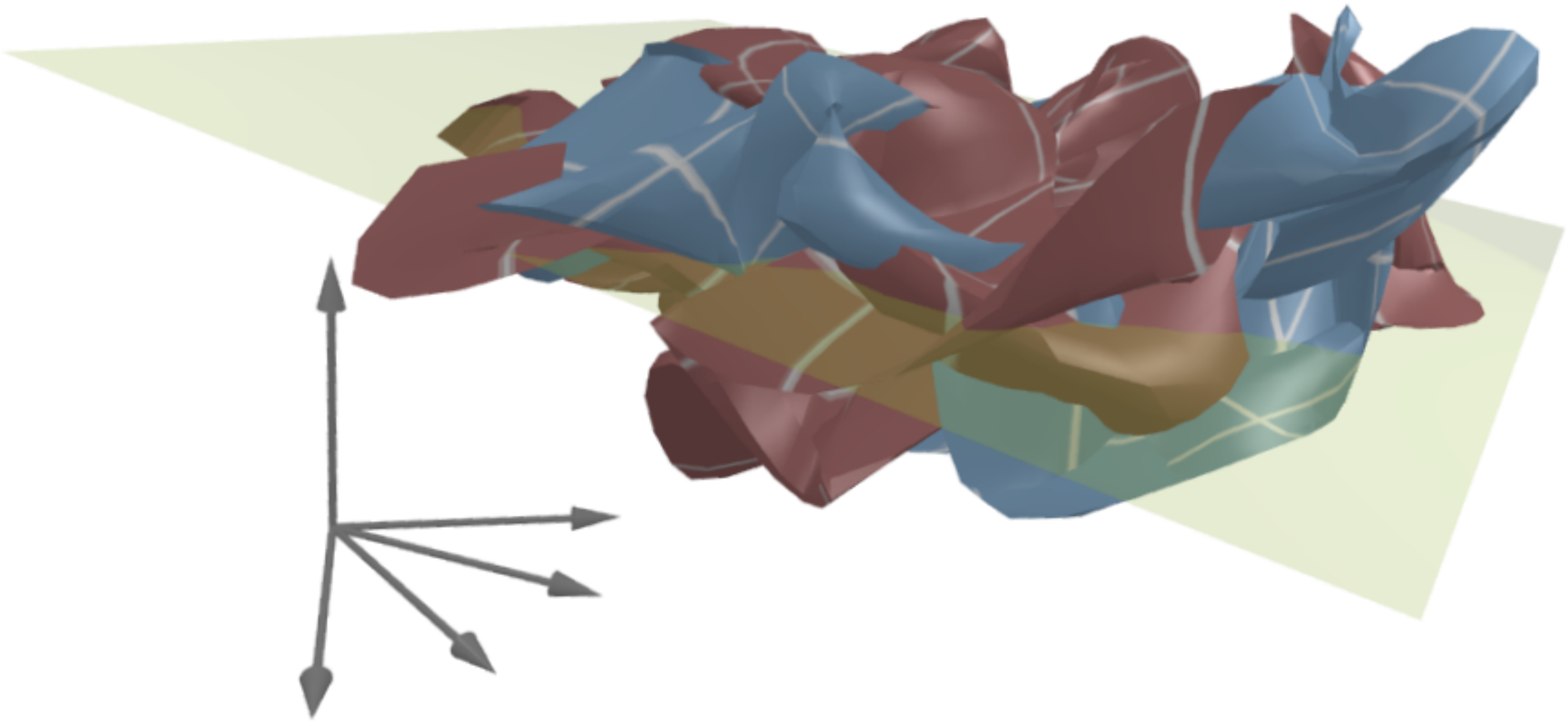
The visual brain represents the image as populations of visually-evoked "features"





Object manifolds get untangled along the ventral stream

V1-like population representation



individual 2

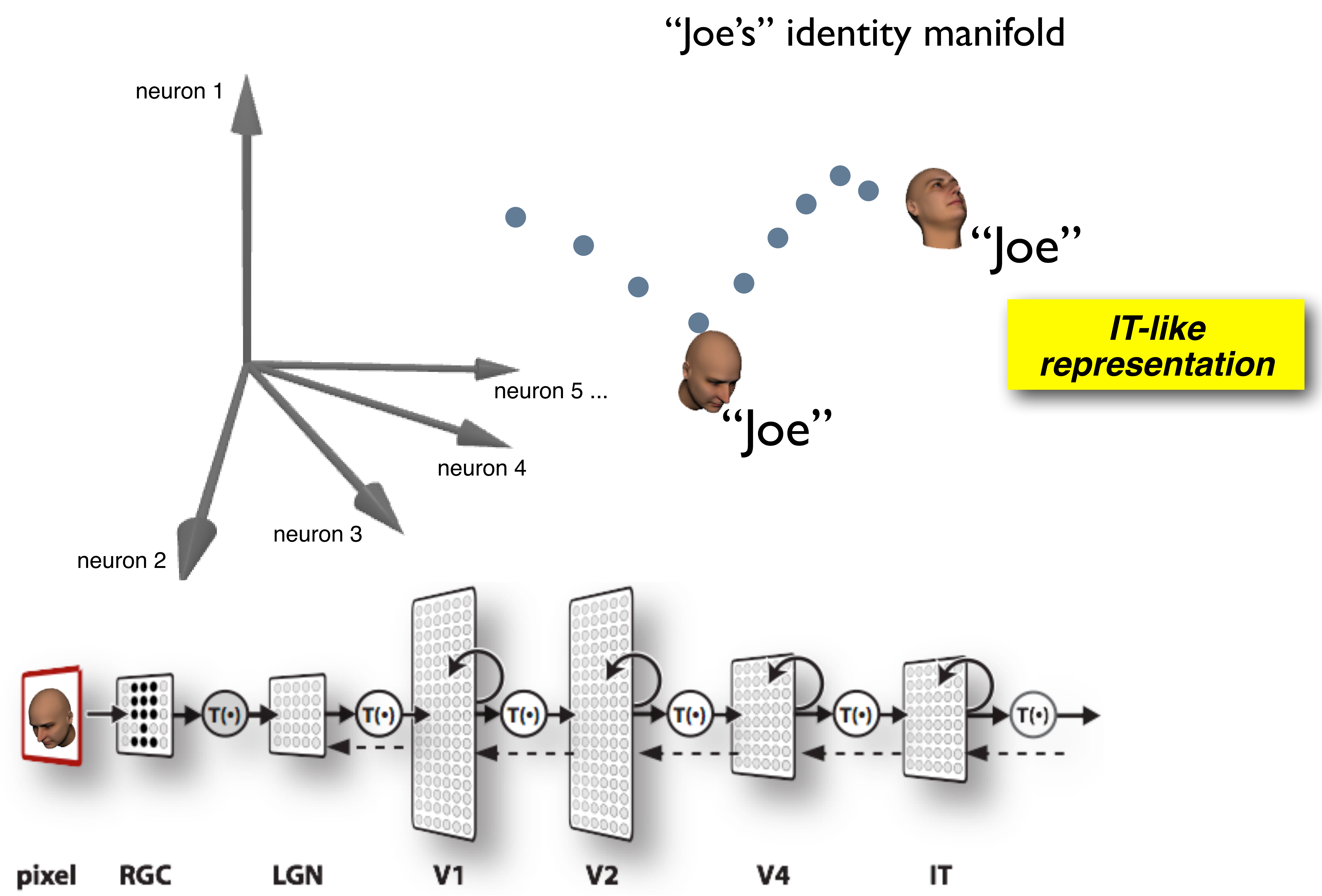
ineffective separating hyperplane

individual 1

object manifolds are "tangled"

(Due to identity-preserving image variation.)

*DiCarlo and Cox, **TICS** (2007); Pinto, Cox, and DiCarlo, **PLoS Comp Bio** (2008)*

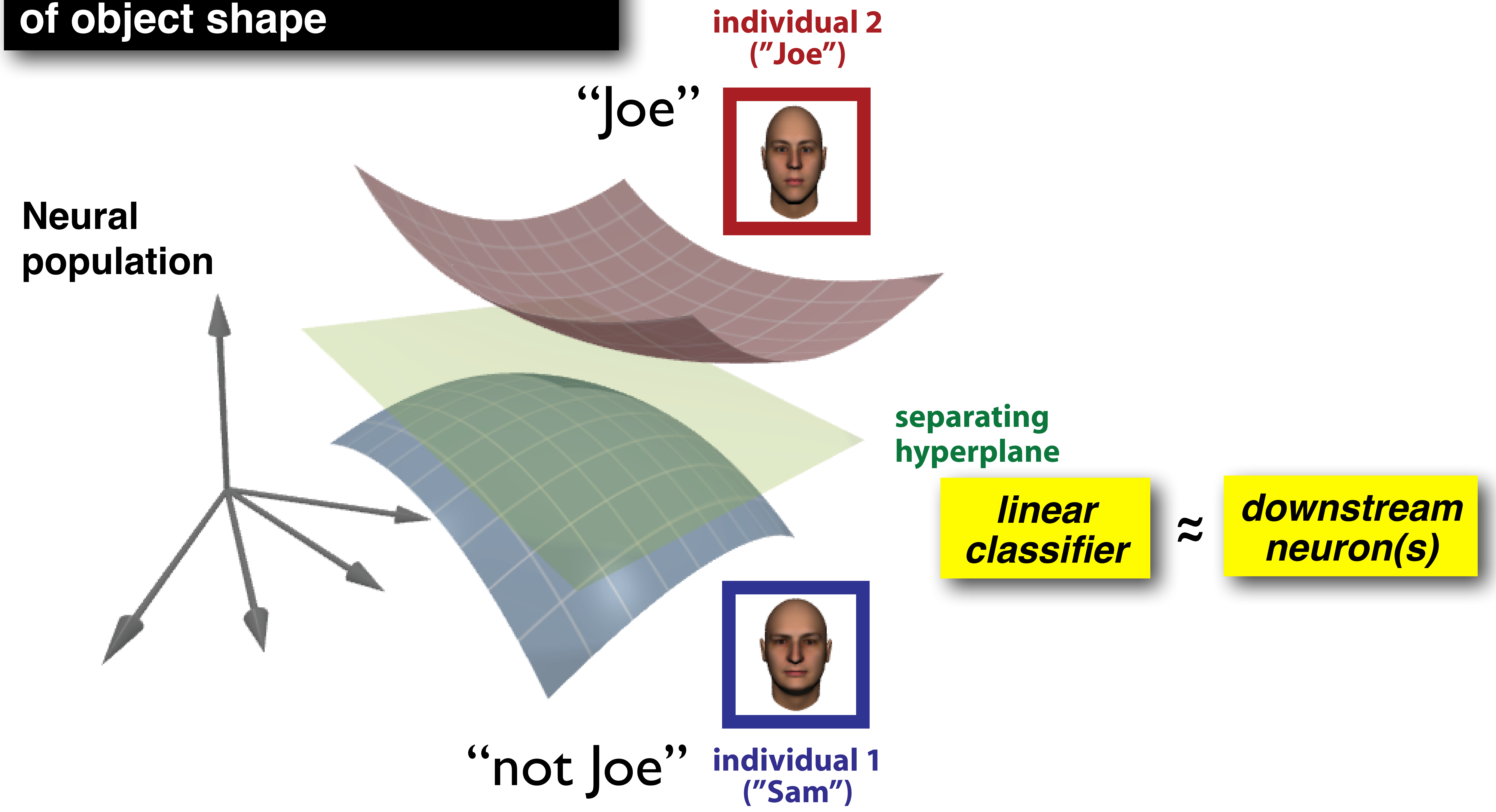


“Joe’s” identity manifold
neuron 1
neuron 2
neuron 3
neuron 4
neuron 5 ...
“Joe”
“Joe”
IT-like representation
T(•)
pixel
RGC
LGN
V1
V2
V4
IT

The computational crux of object and face recognition

A “good” set of visual features: e.g. IT

== “Explicit” representation of object shape

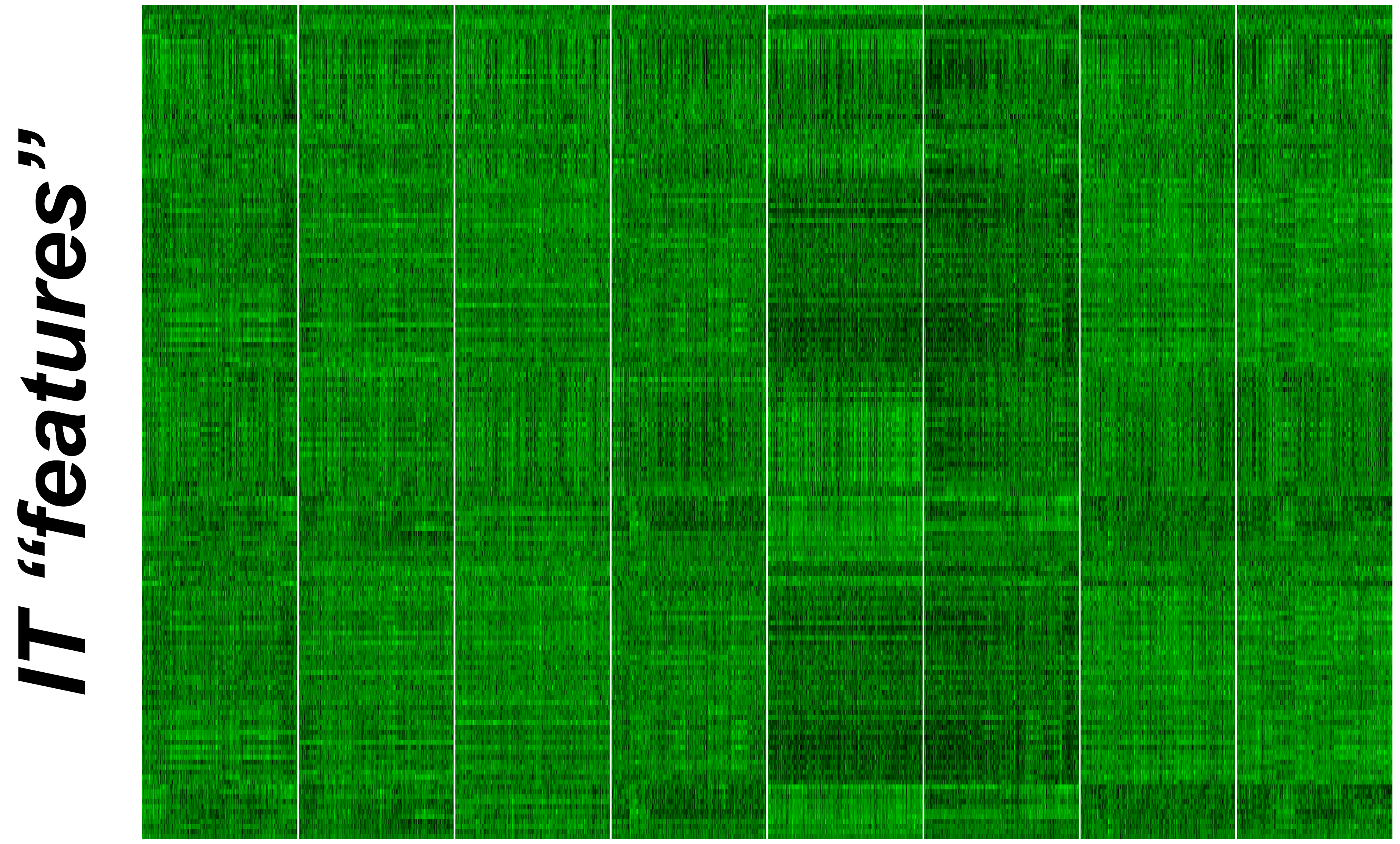


*DiCarlo and Cox, **TICS** (2007)*

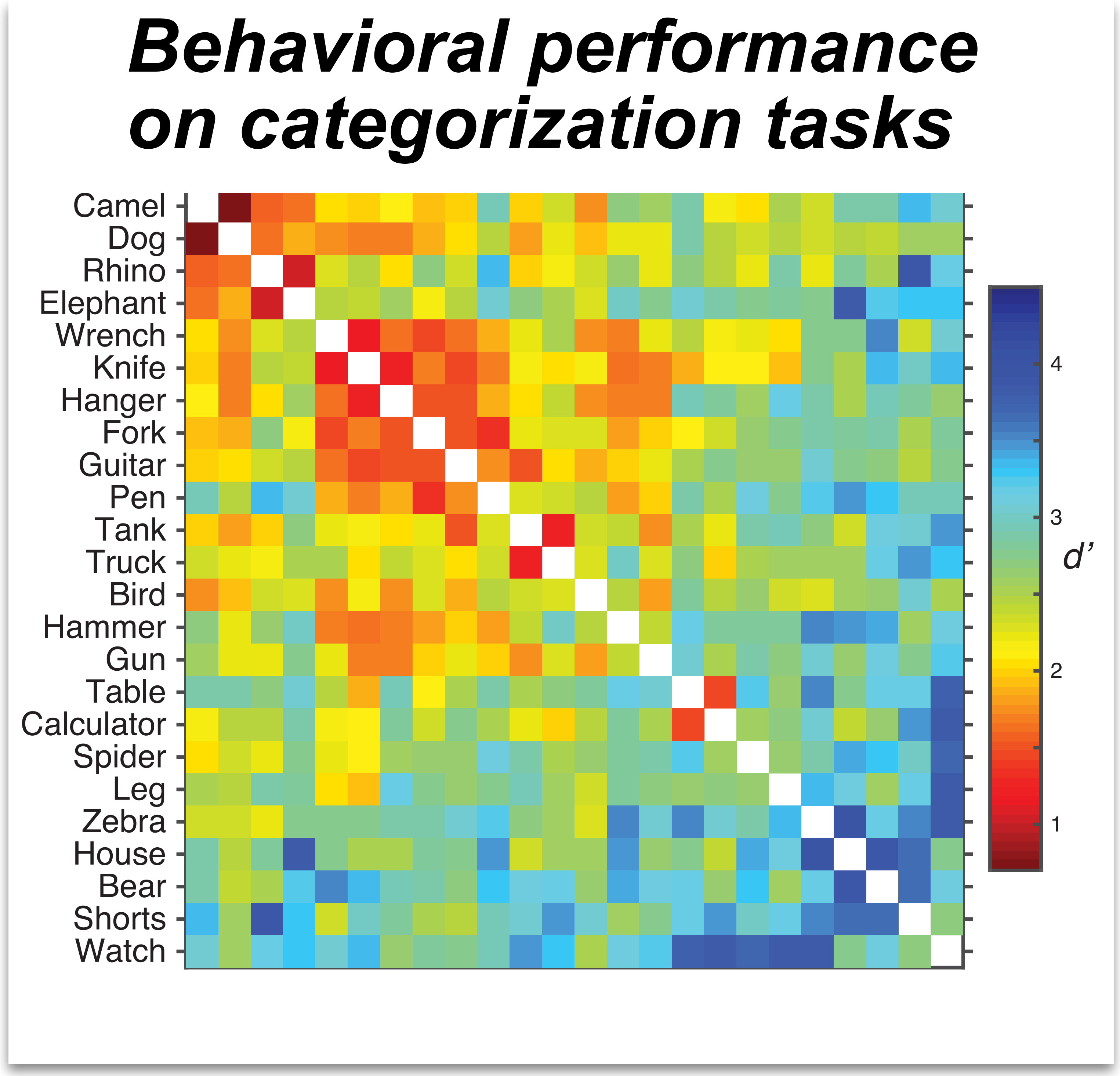
measure

4

Test population decoding models that can fully explain behavior

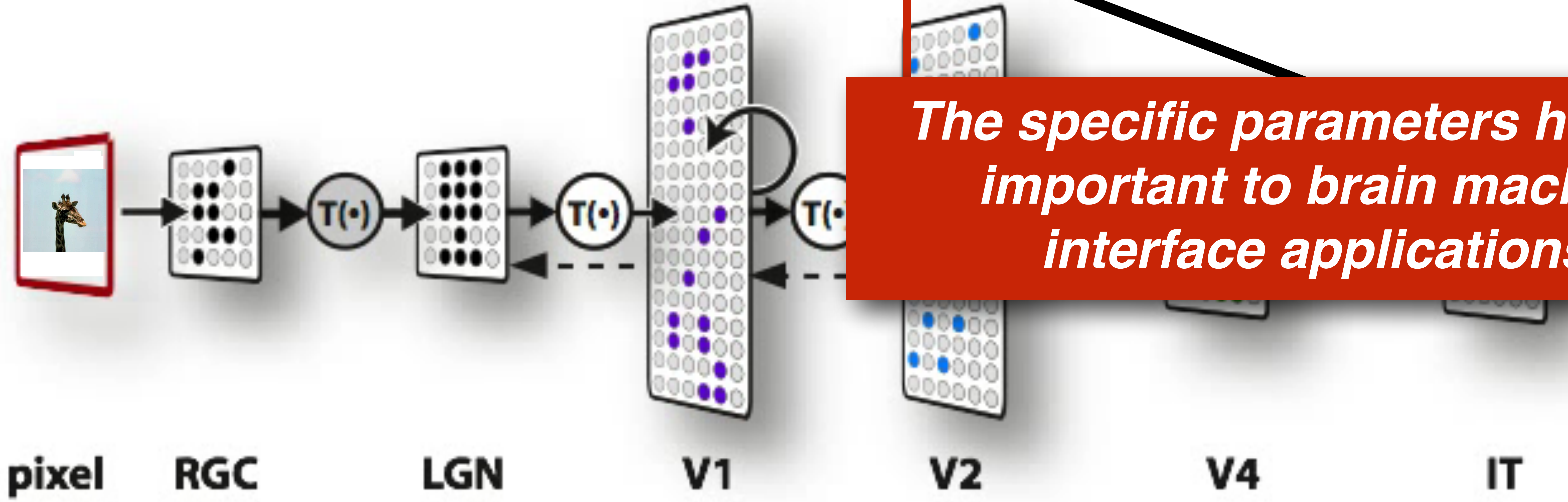


Linear decoder accurately predicts!



Majaj, Hong, Soloman, & DiCarlo, ***J Neuro*** *(2015)*

Hong, Yamins*, Majaj & DiCarlo.* ***Nat. Neuro.*** *(2016)*



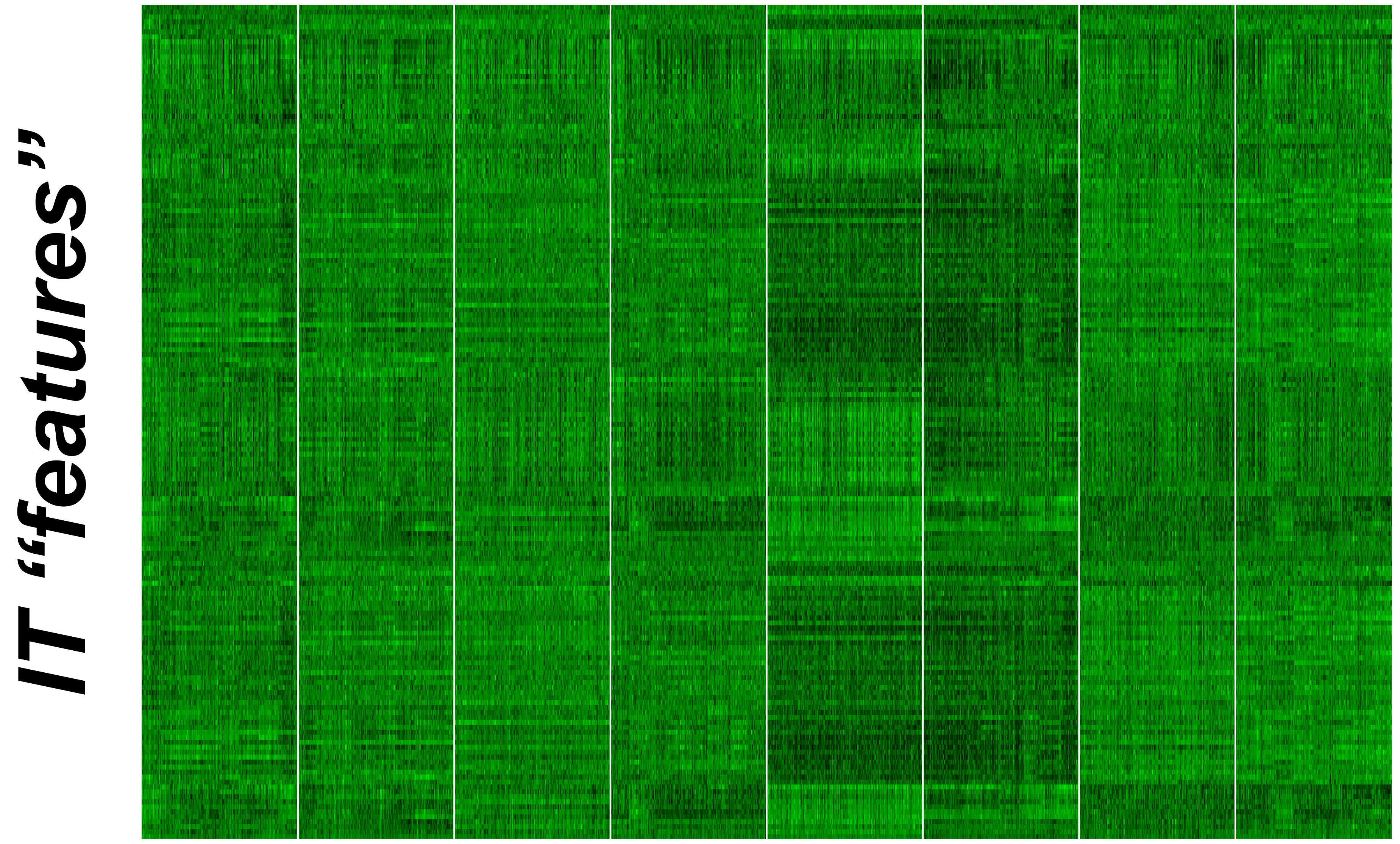
The specific parameters here are important to brain machine interface applications.


measure
4
Linear
Behavioral performance on categorization tasks
Camel
Dog
d'
LaWS of RAD IT linking hypothesis
16 sites
64 sites
128 sites
LaWS of RAD V4 linking hypothesis
16 sites
64 sites
128 sites
n = 64 tasks
Predicted d' on tasks
Human d' on tasks
Human-like
IT
168 sites
128
64
4
Consistency
Performance
pixel
RGC
LGN
V1
V2
V4
IT

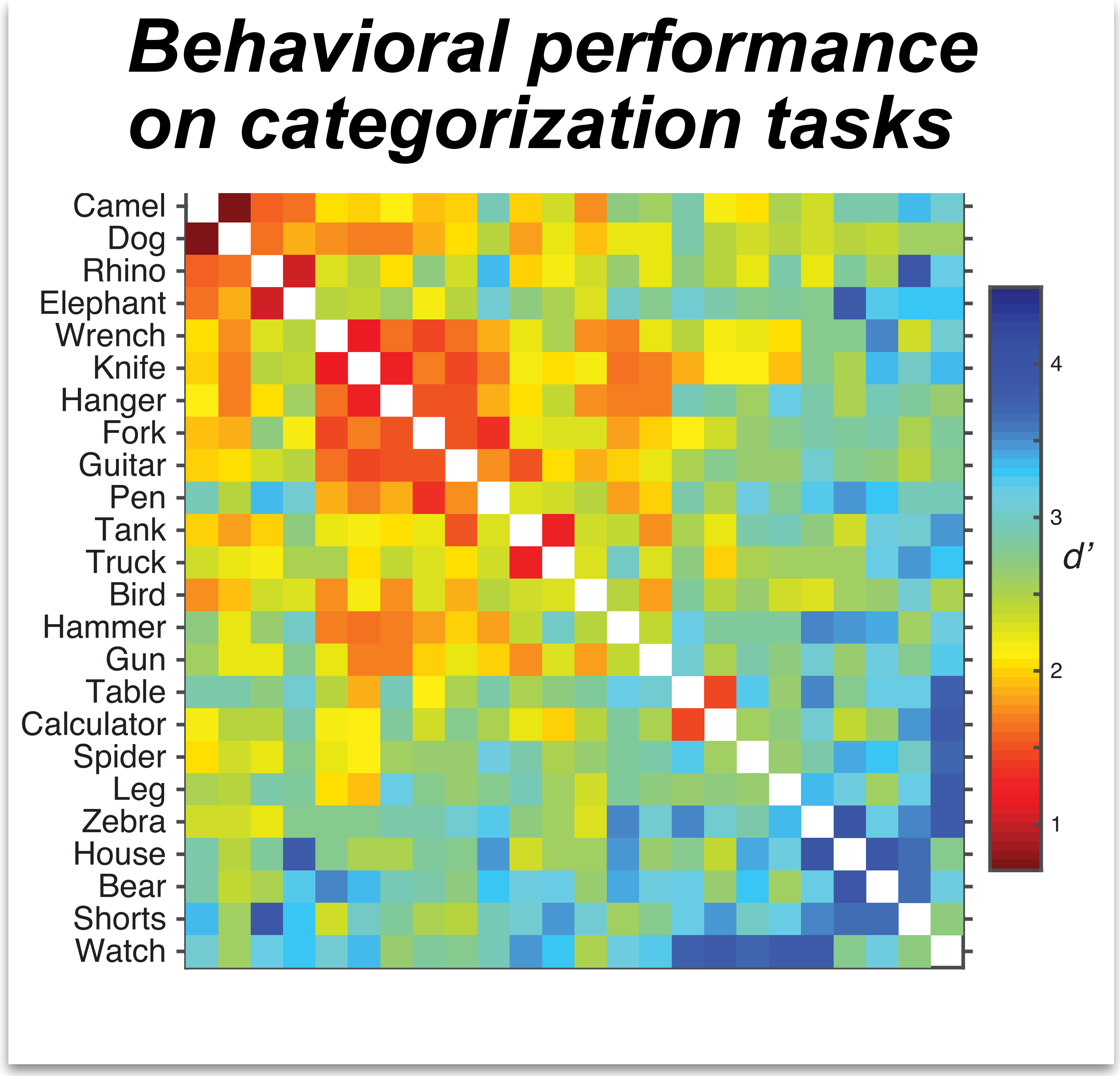
Test against finer grain behavior

5

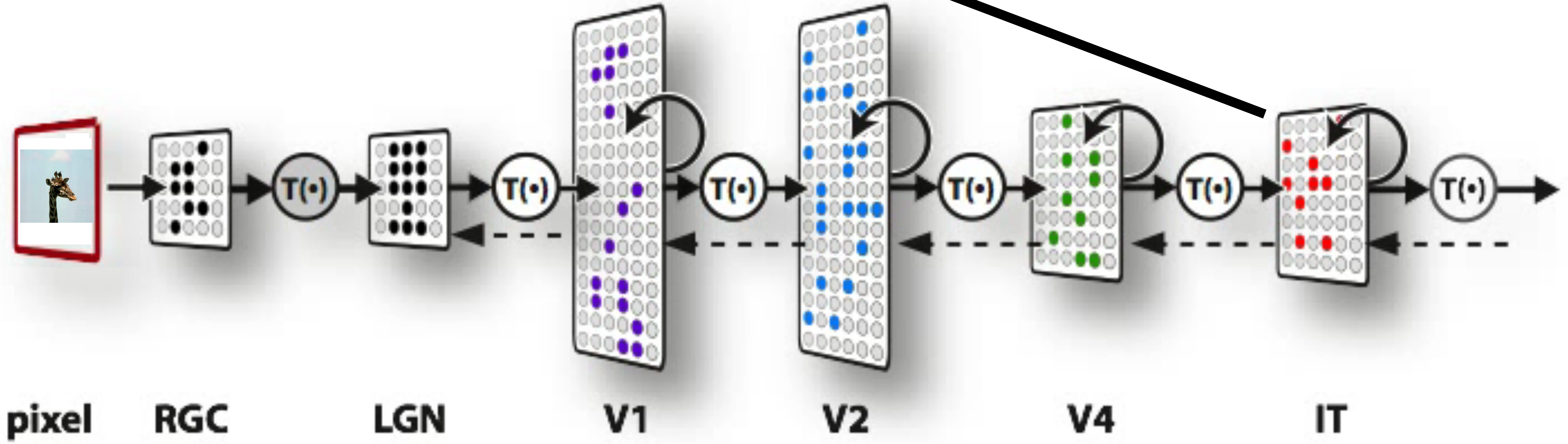
Test whether the same population decoding models can explain finer grain behavioral measurements



Linear decoder accurately predicts!



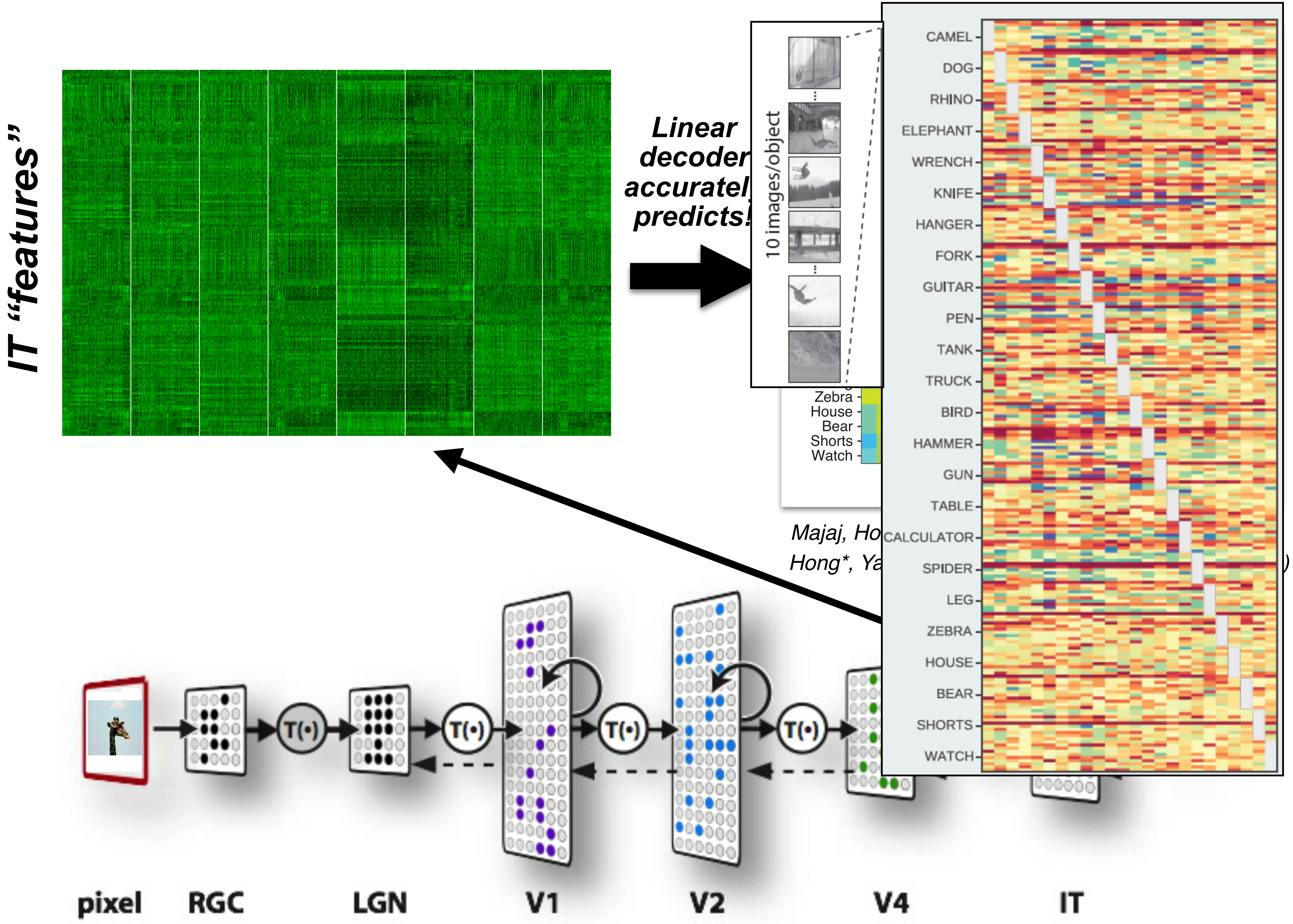
*Majaj, Hong, Soloman, & DiCarlo, **J Neuro** (2015)*
Hong, Yamins*, Majaj & DiCarlo. **Nat. Neuro.** (2016)*



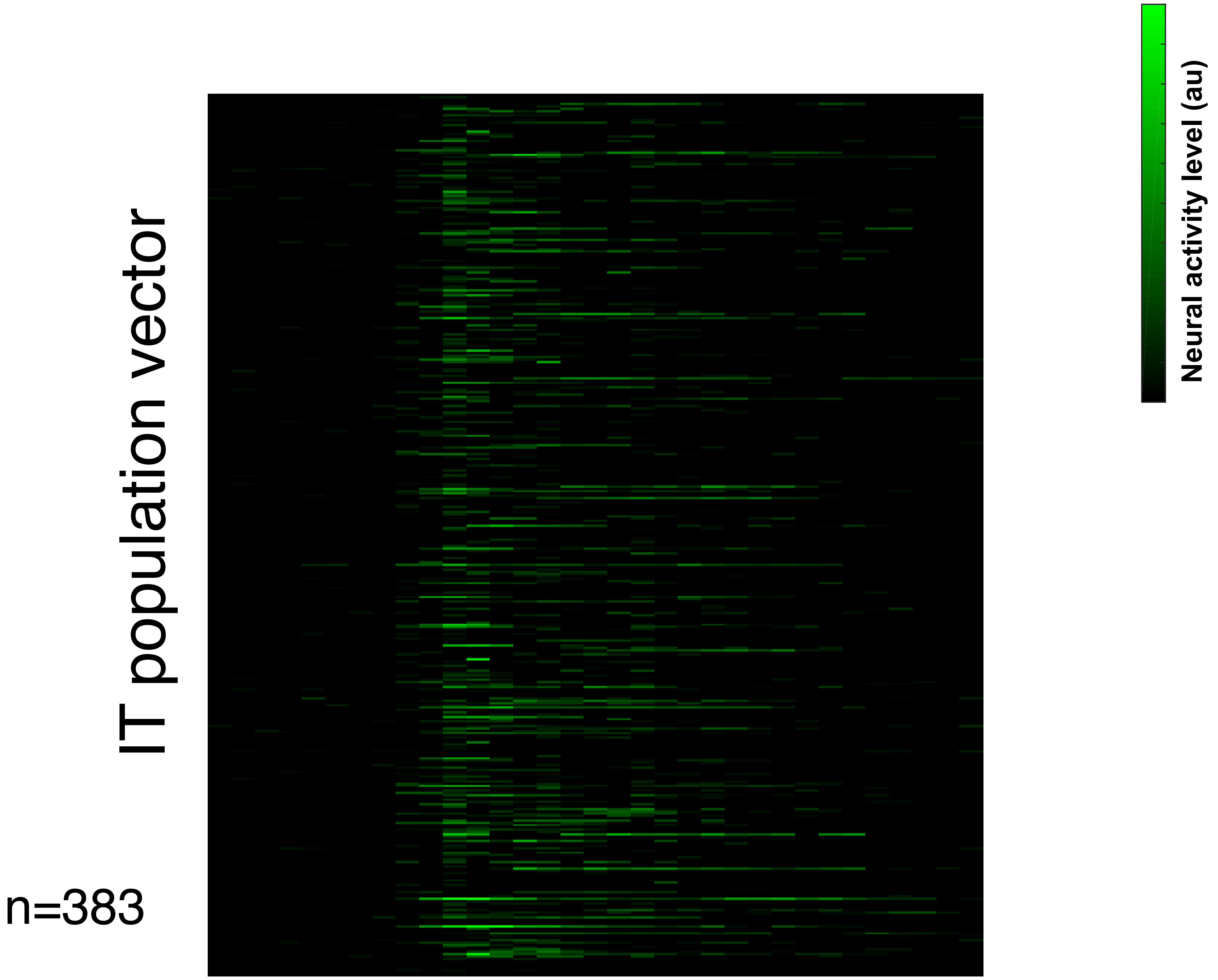
Test against finer grain behavior

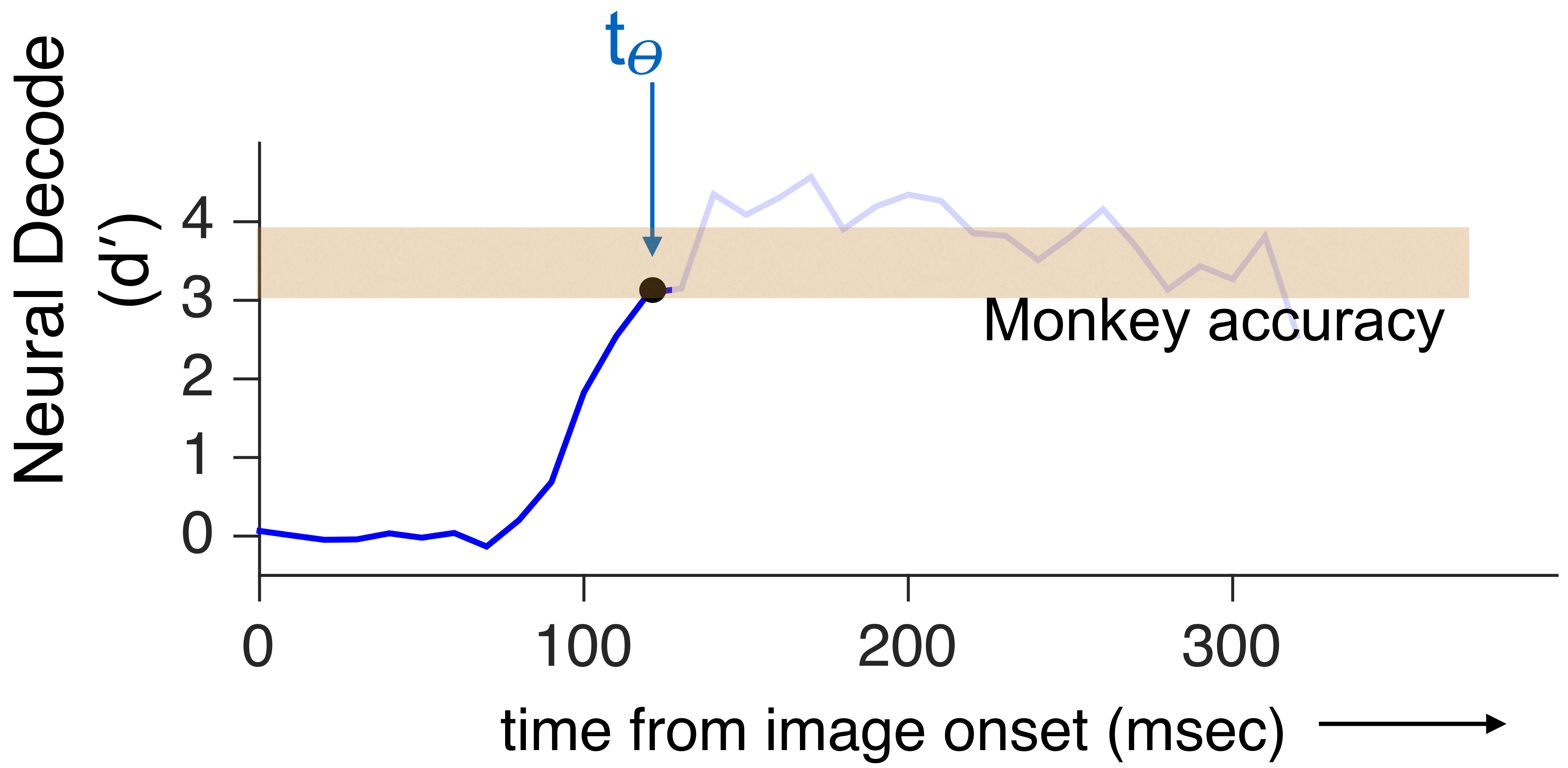
5

Test whether the same population decoding models can explain finer grain behavioral measurements



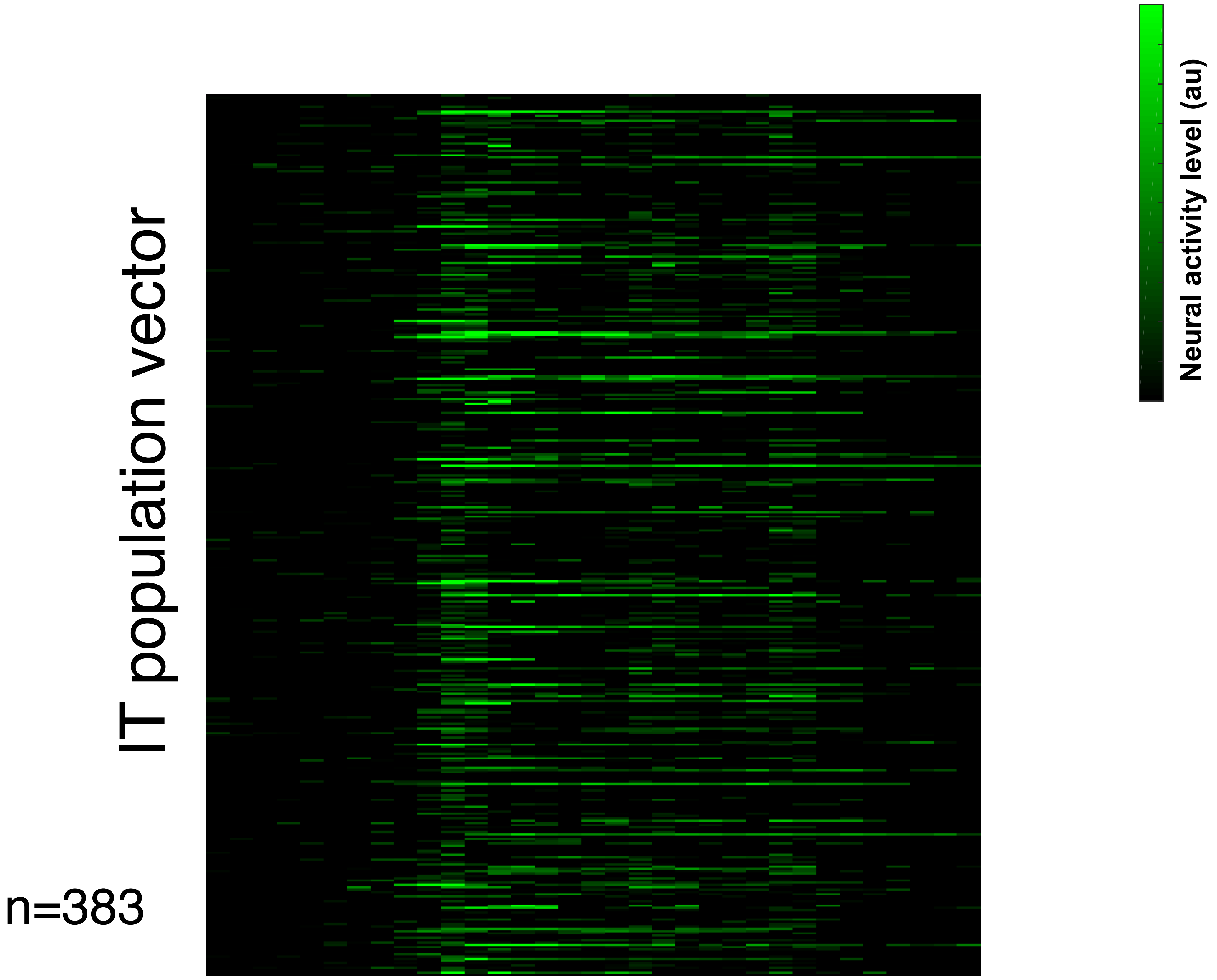
FACE

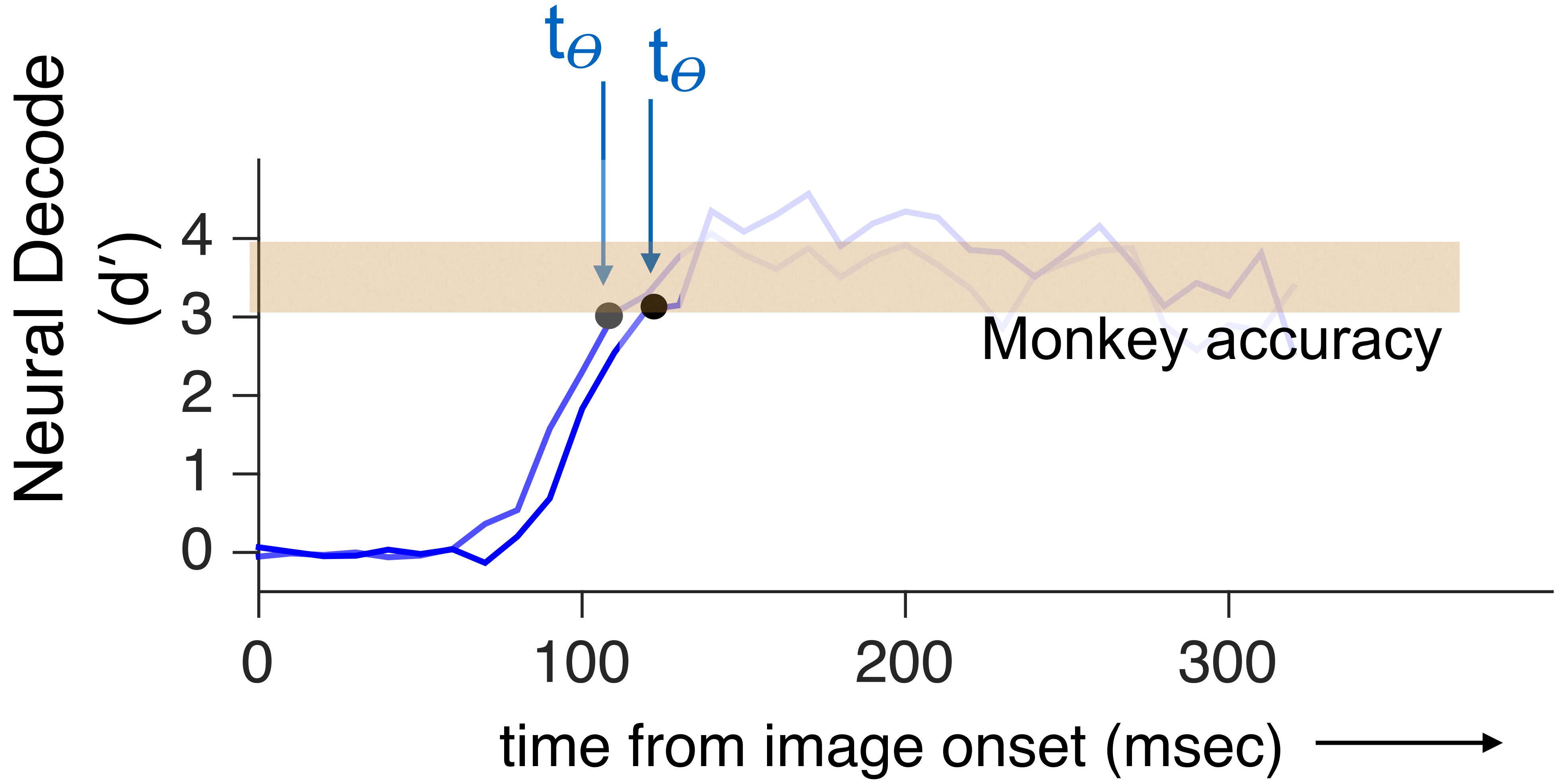




Kar, Kubilius, Issa, Schmidt, DiCarlo. BioRxiv (2018)*

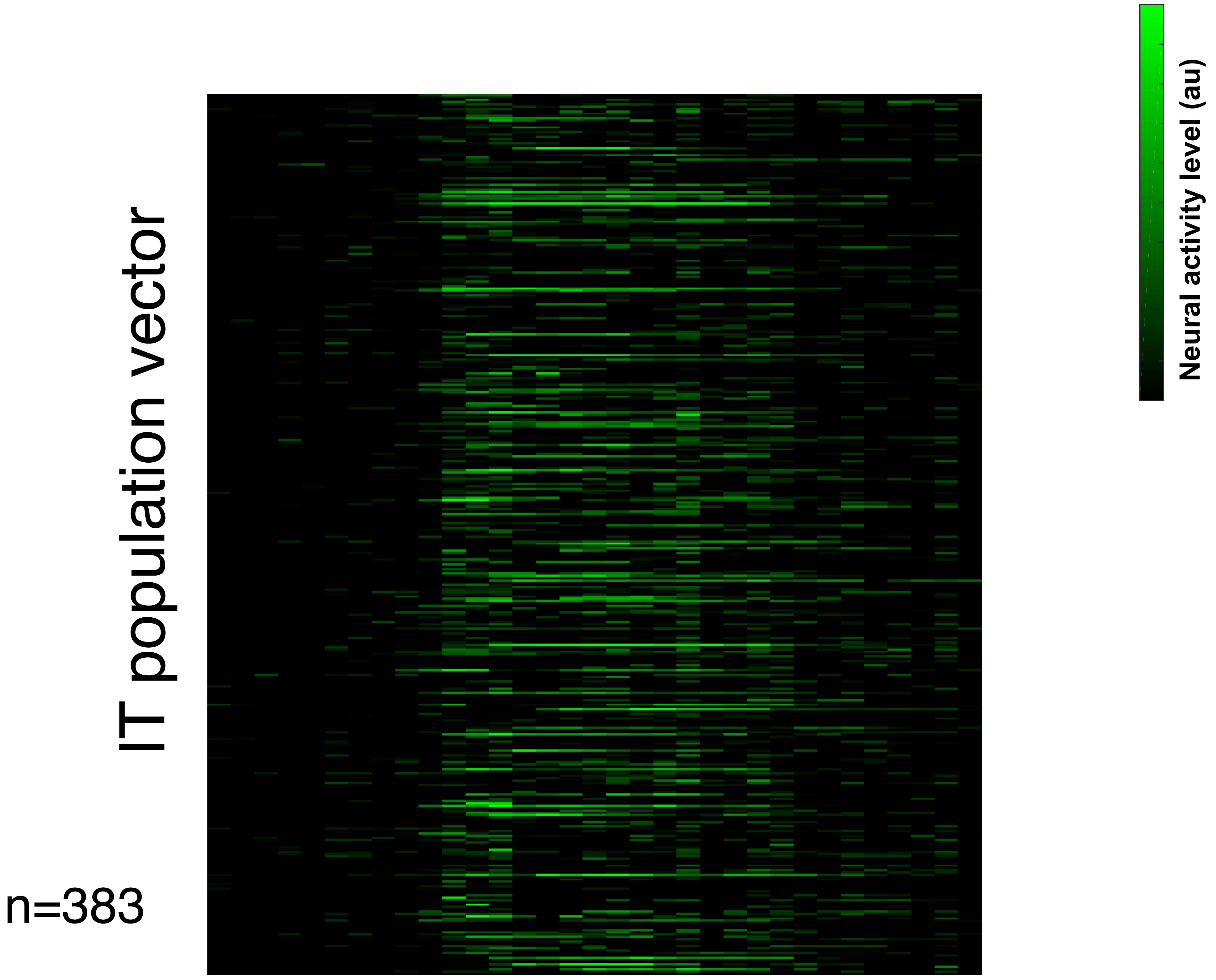
ZEBRA

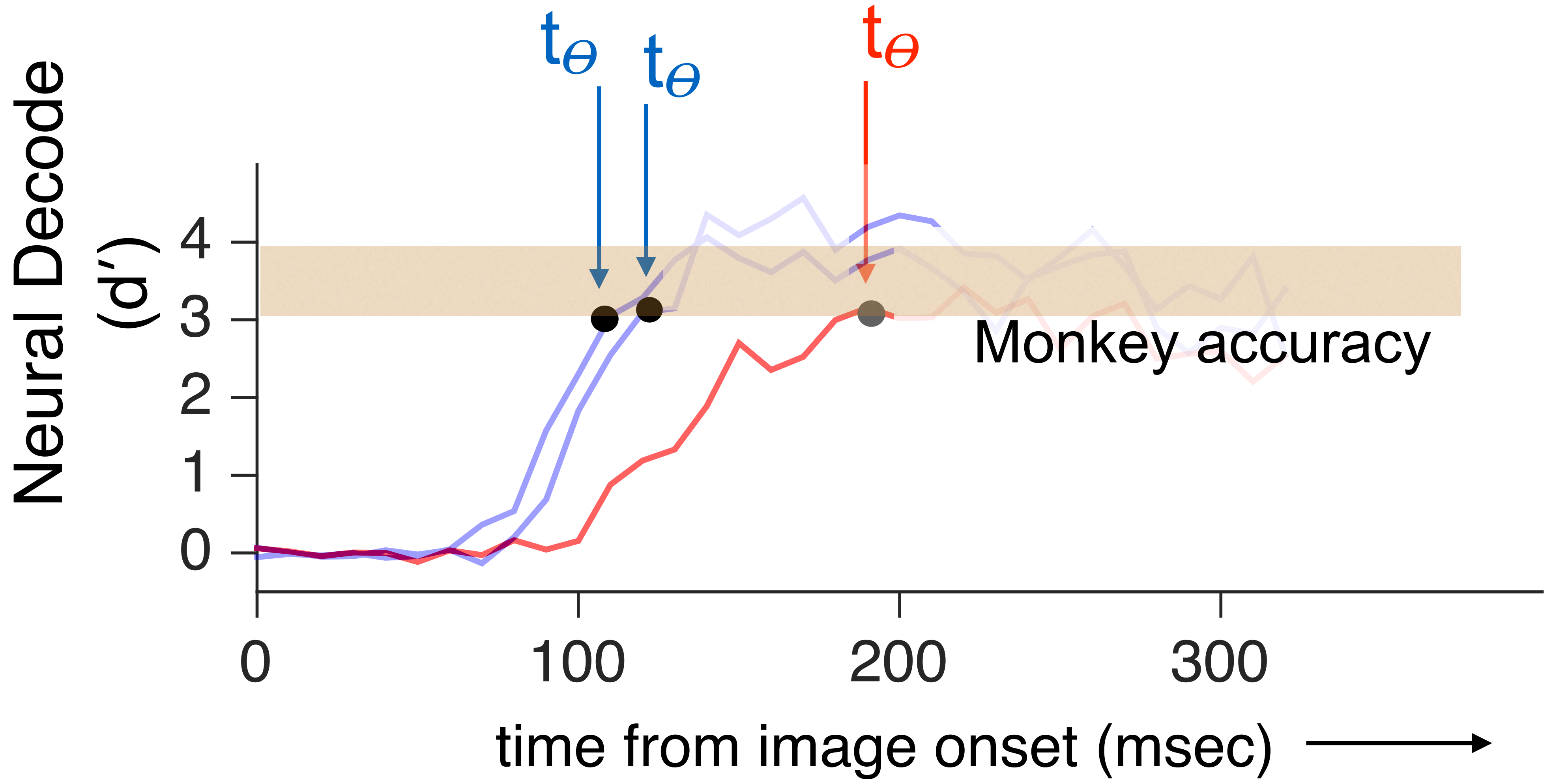




Kar, Kubilius, Issa, Schmidt, DiCarlo. BioRxiv (2018)*

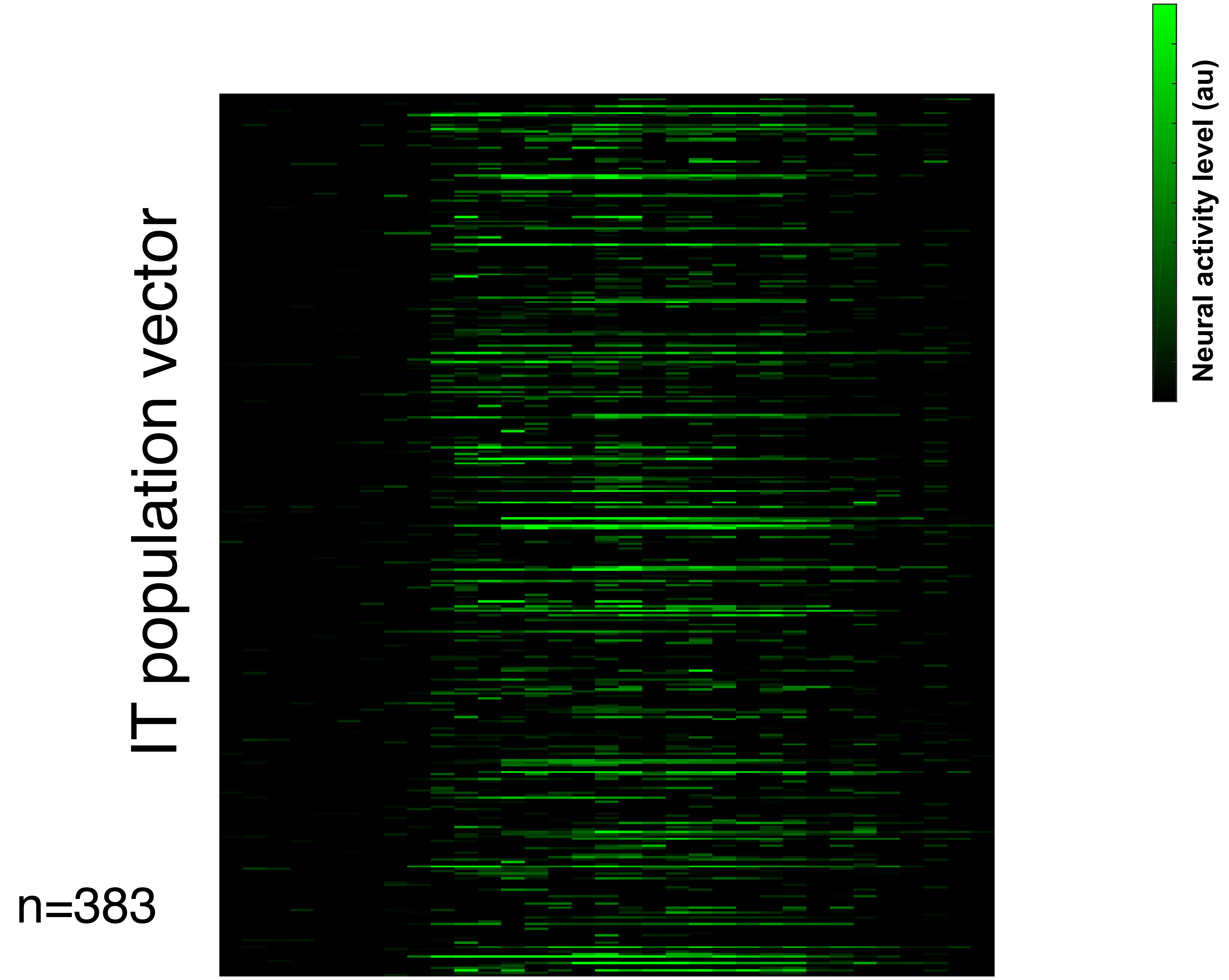
CAR

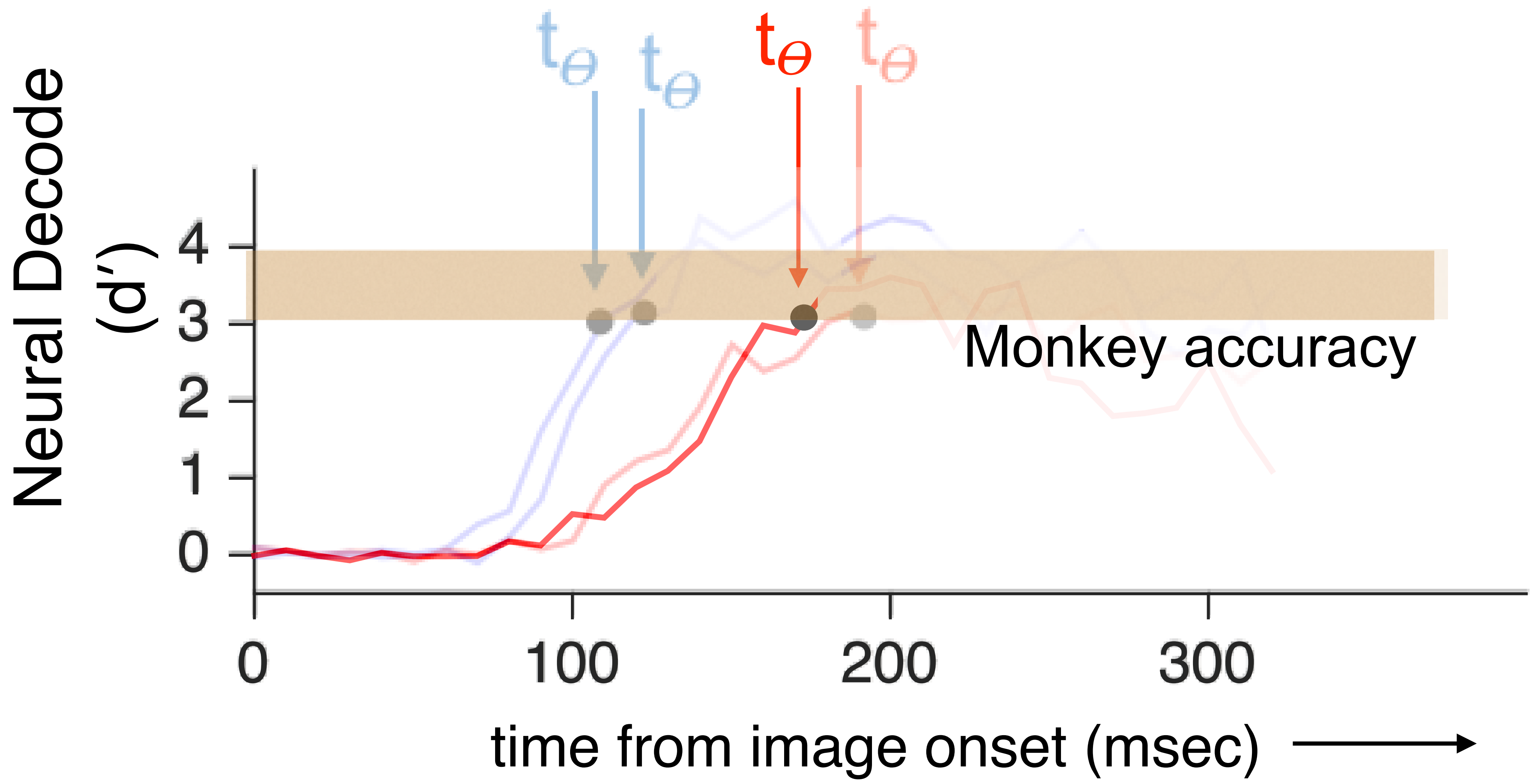




Kar, Kubilius, Issa, Schmidt, DiCarlo. BioRxiv (2018)*

DOG

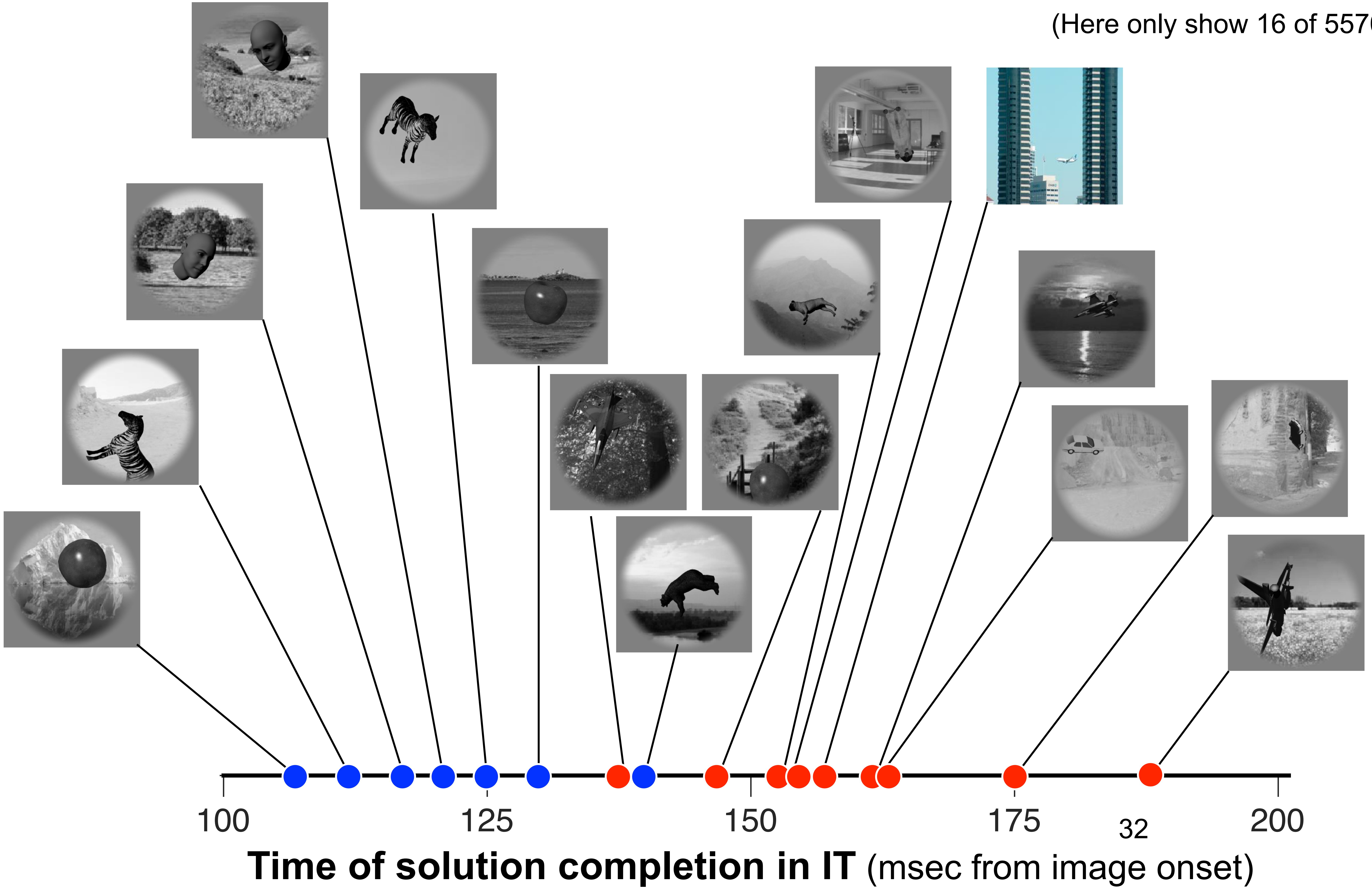




Kar*, Kubilius, Issa, Schmidt, DiCarlo. BioRxiv (2018)

We precisely measured time of brain’s penultimate solution product for thousands of images

(Here only show 16 of 5570 images measured)

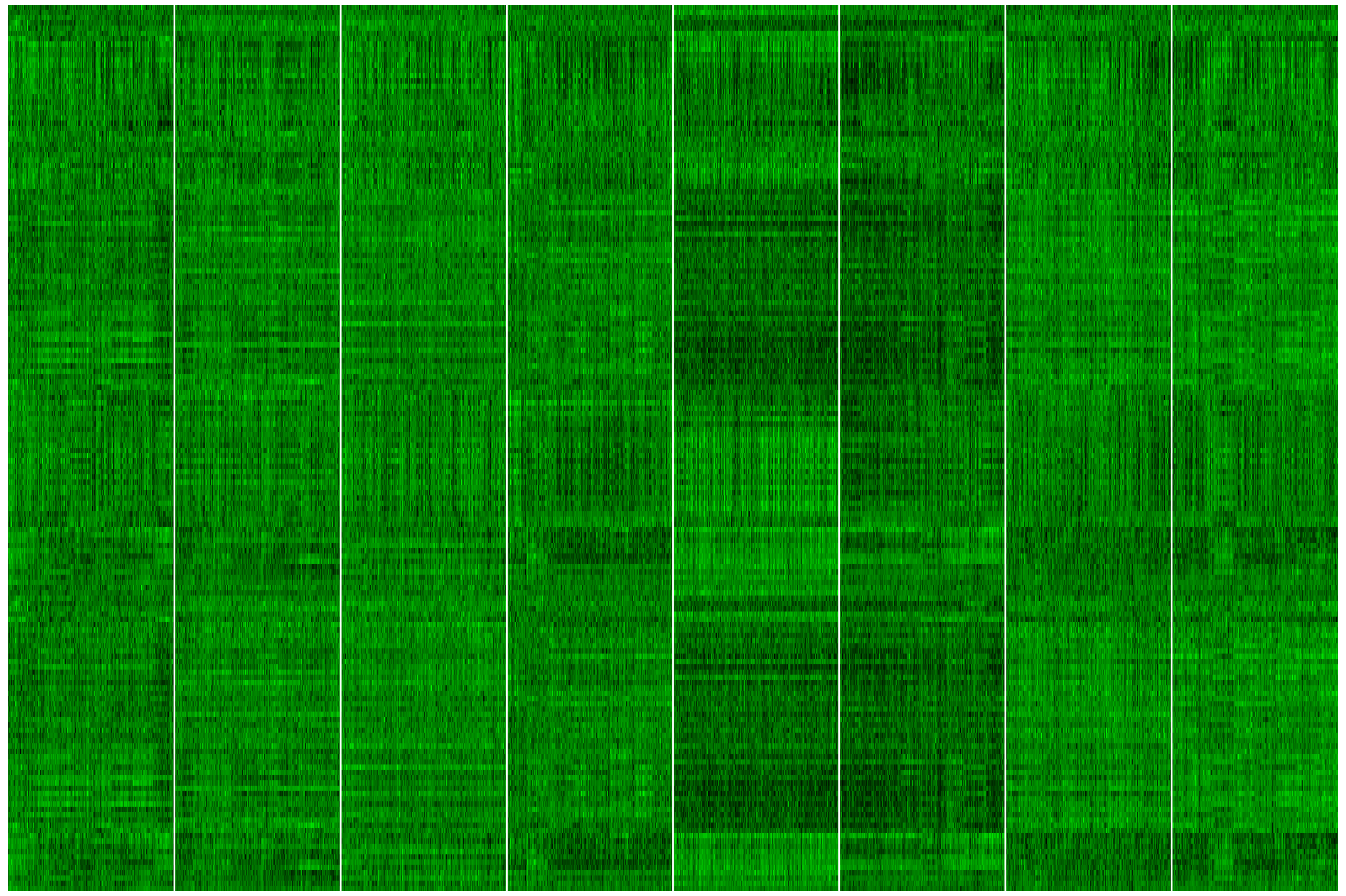


Test against finer grain behavior

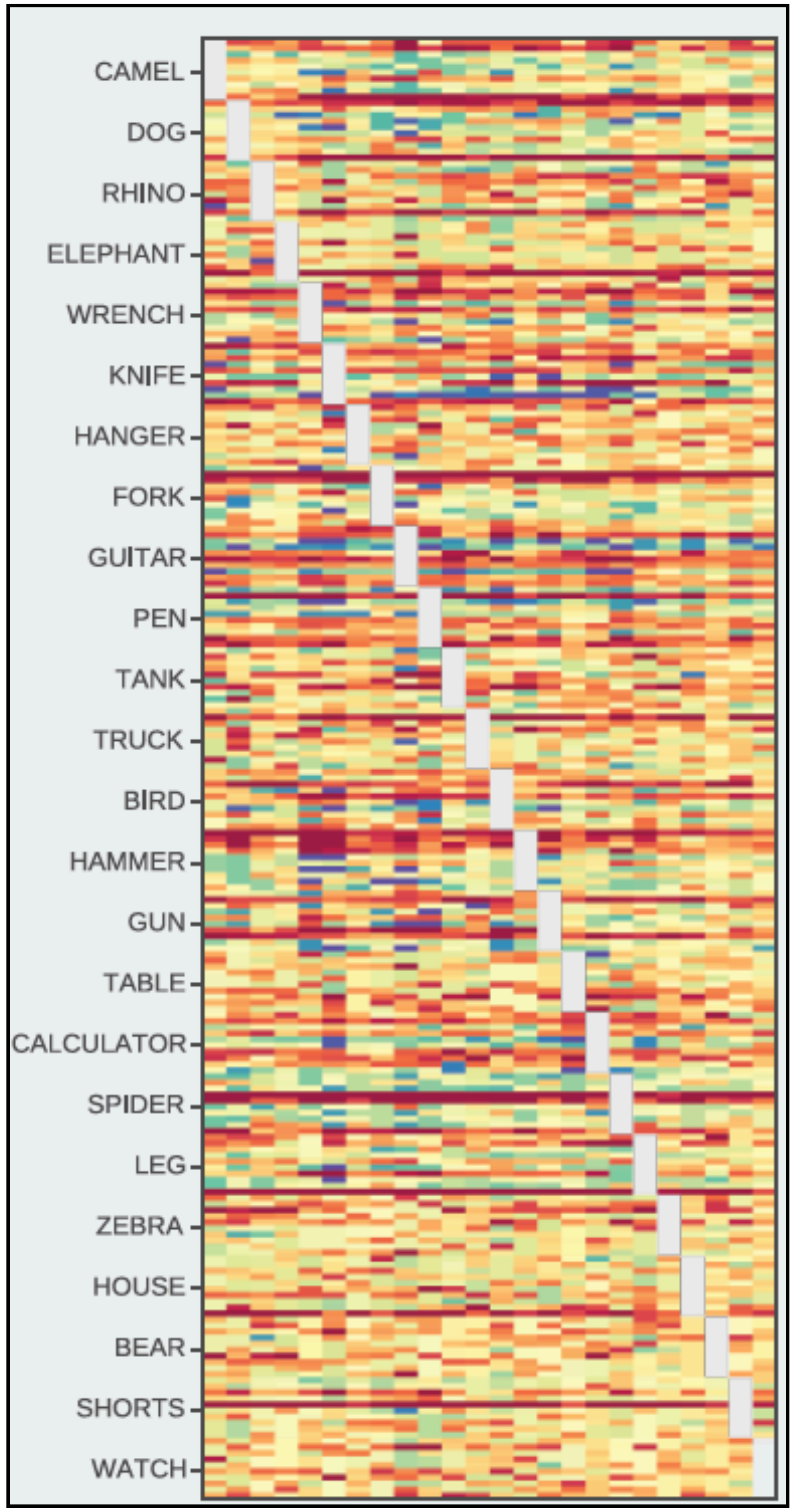
5

Test whether the same population decoding models can explain finer grain behavioral measurements

IT "features"



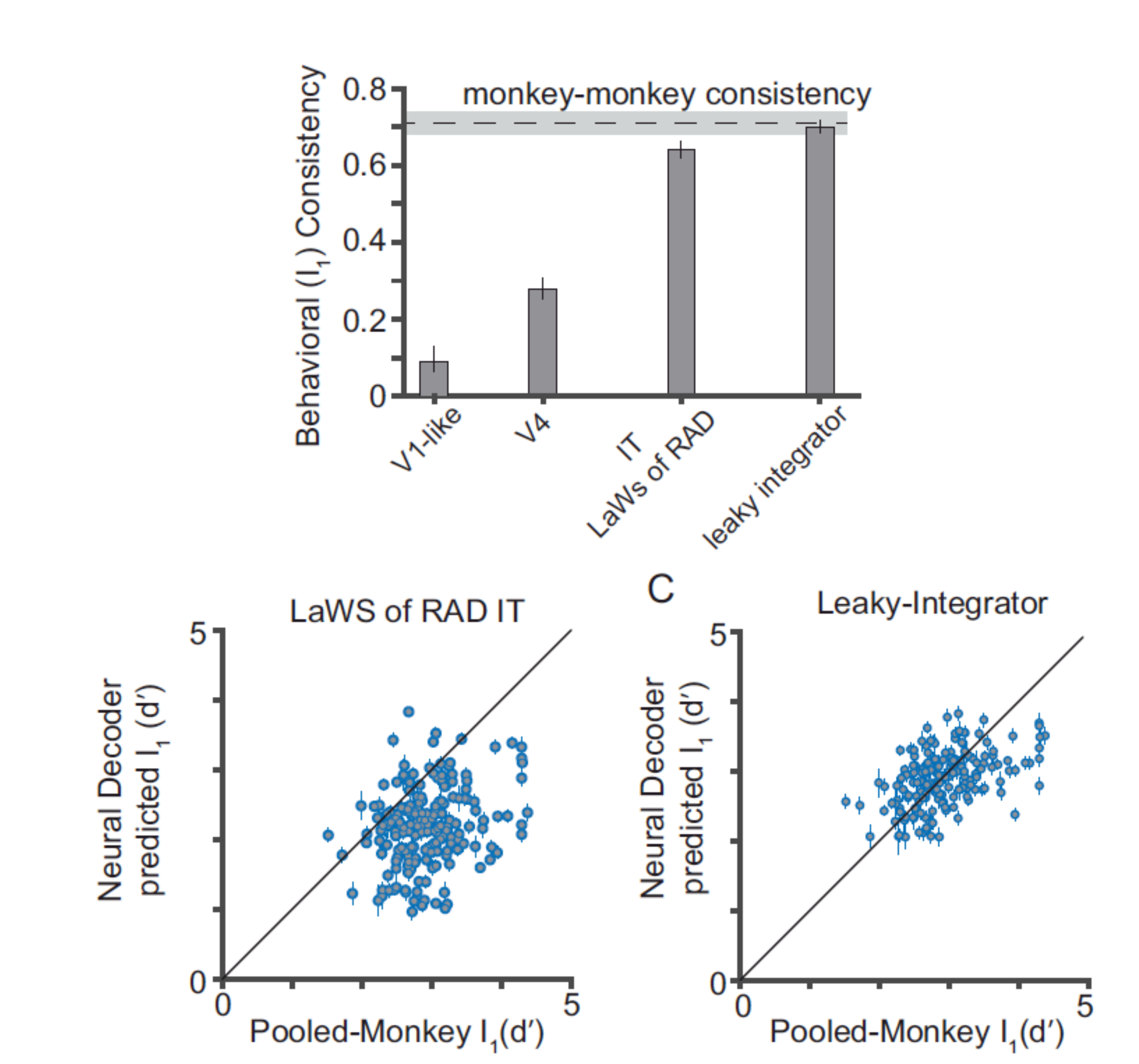
??

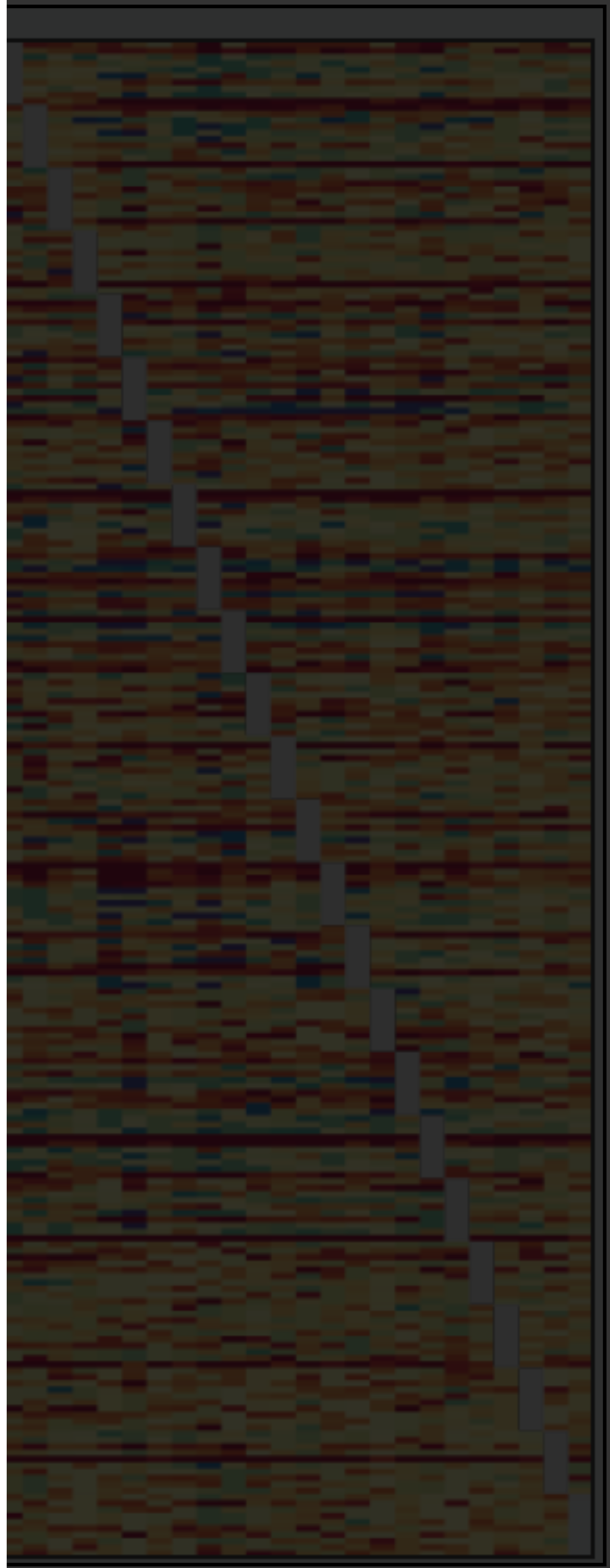


Test against finer grain behavior

5

Test whether the same population decoding models can explain finer grain behavioral measurements





THANKS …

Now you have 19 more days to grill me