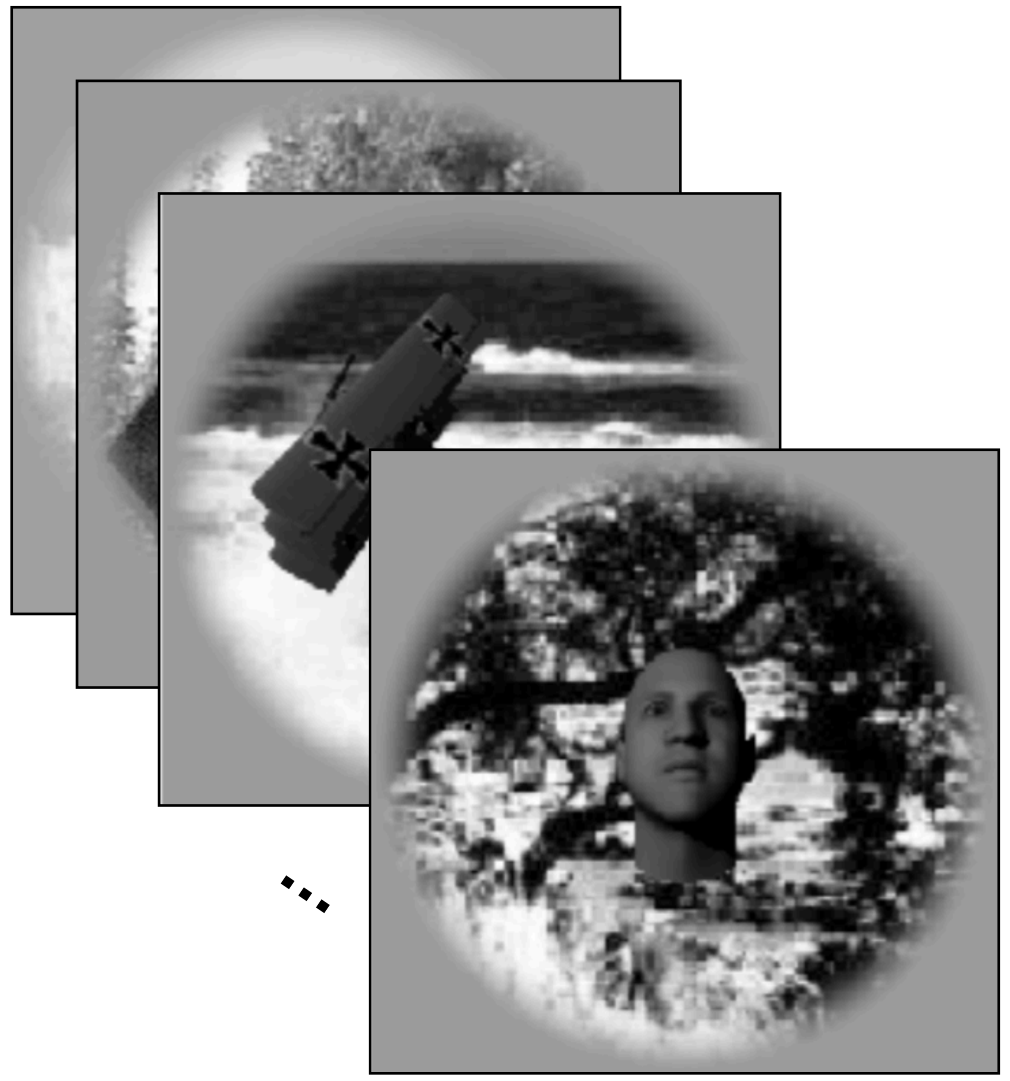
Non-parametric model comparisons

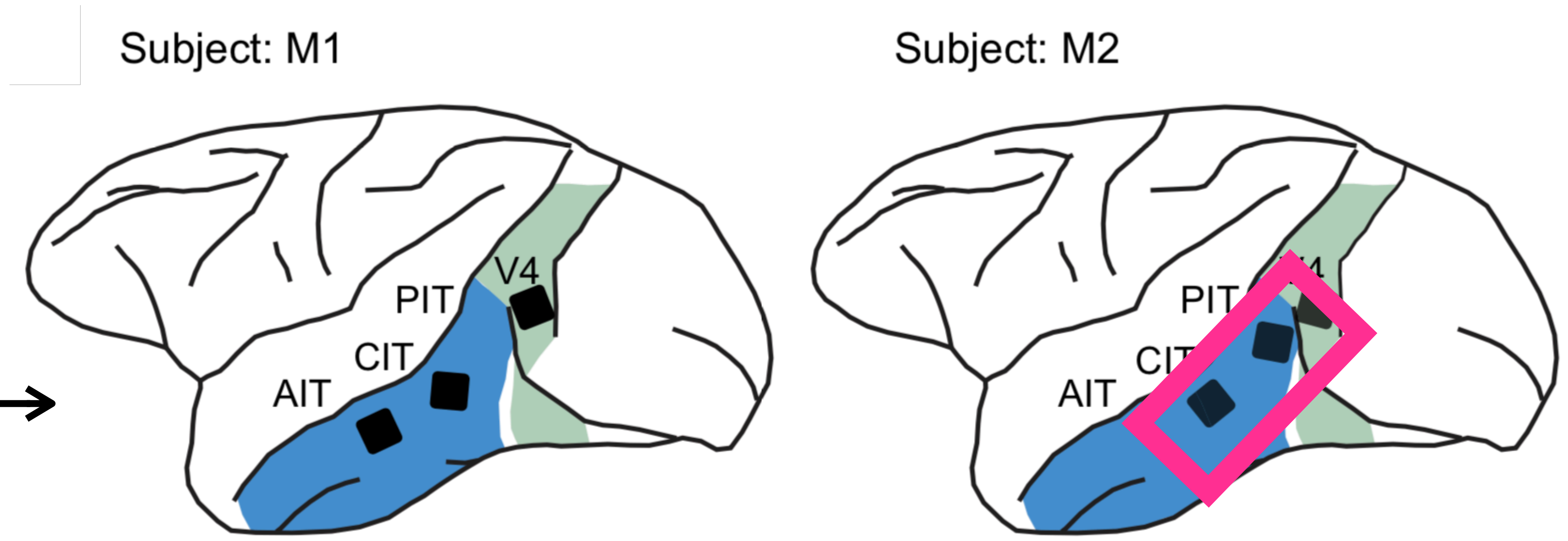
i. Building a linking function between the brain and behavior
ii. Representational Similarity Analysis (RSA)
iii. Linear mappings between responses to shared inputs
iv. A familiar visualization :)

tyler bonnen: bonnen@stanford.edu
—Anthony Wagner & Daniel Yamins

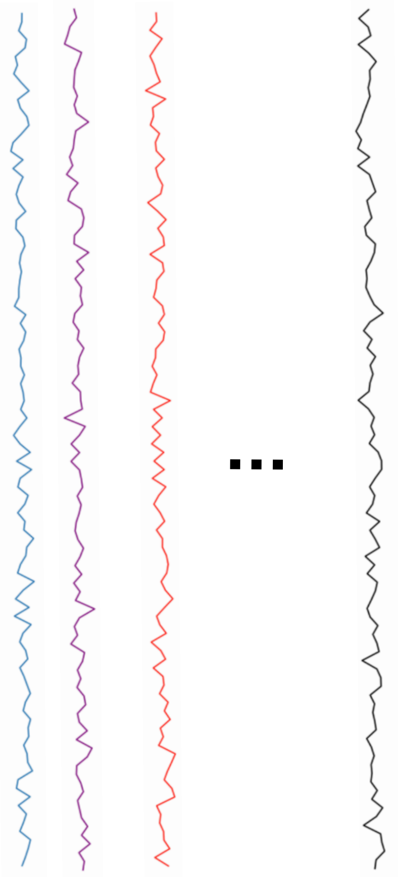
Building a linking function

Majaj et al. 2015





Neural responses



How can relate these measurements?

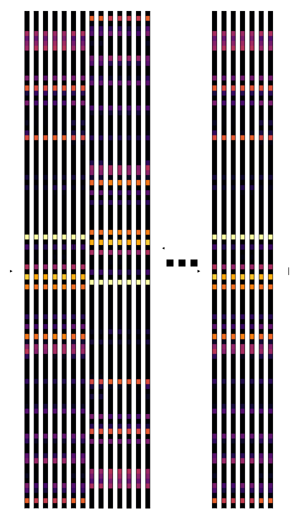
***n* = 104**

Behavioral responses

d’ (basically, accuracy)
Reaction time,

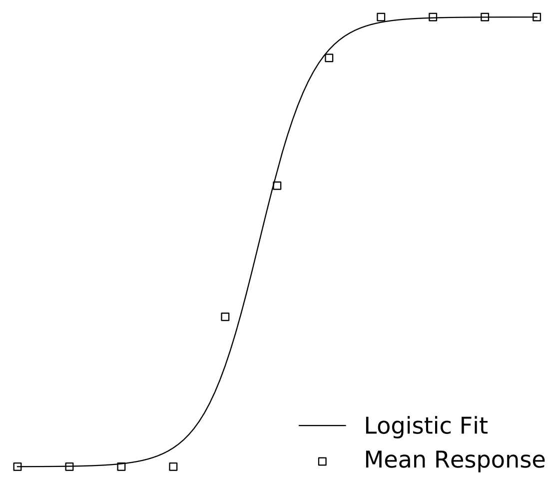
Building a linking function between neural and human behavior

N feature vectors
(i.e. neural responses)



Tran-test split

Build Logistic
Regression model
on training data



Determine performance
on test data

e.g. [0,1,0, 1, 1 … 0]

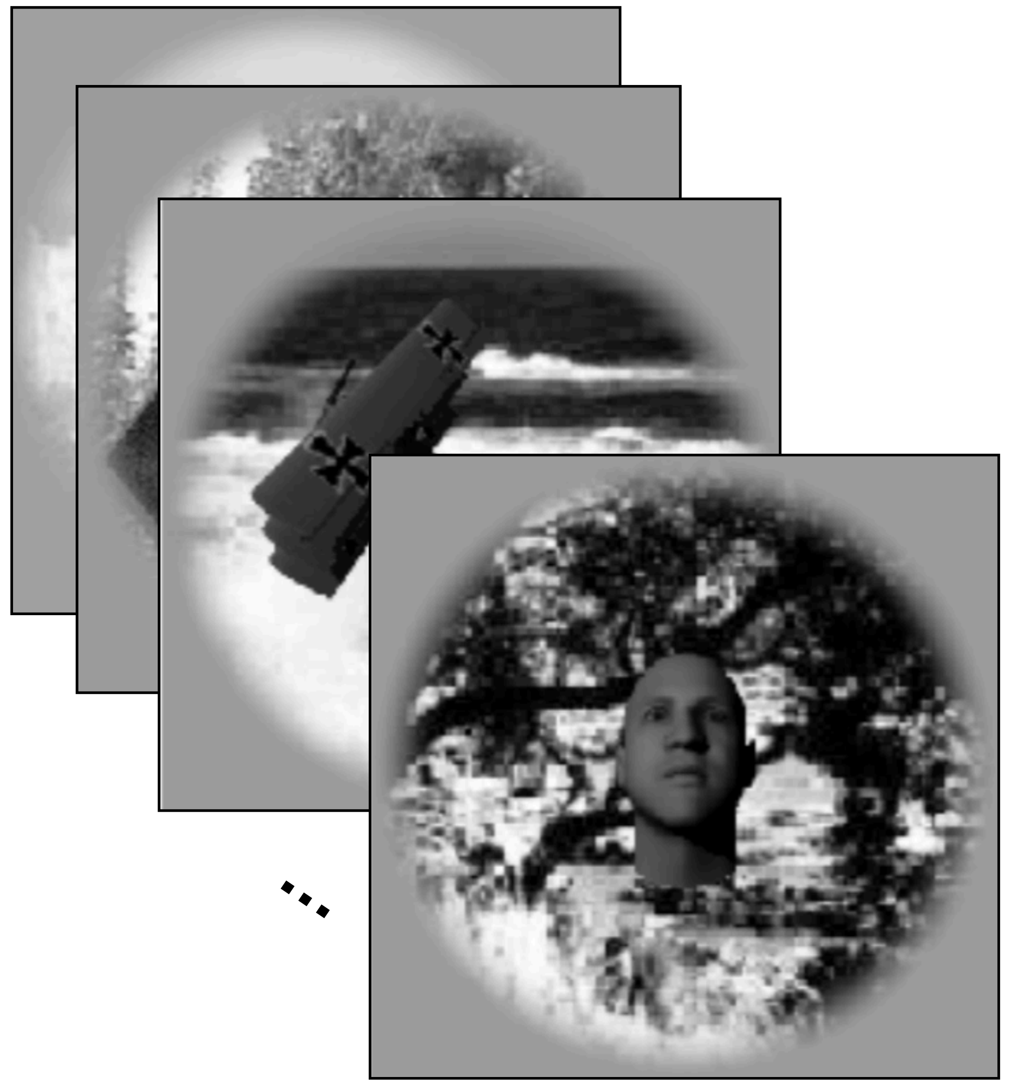
Compute d’ from
Model behavior

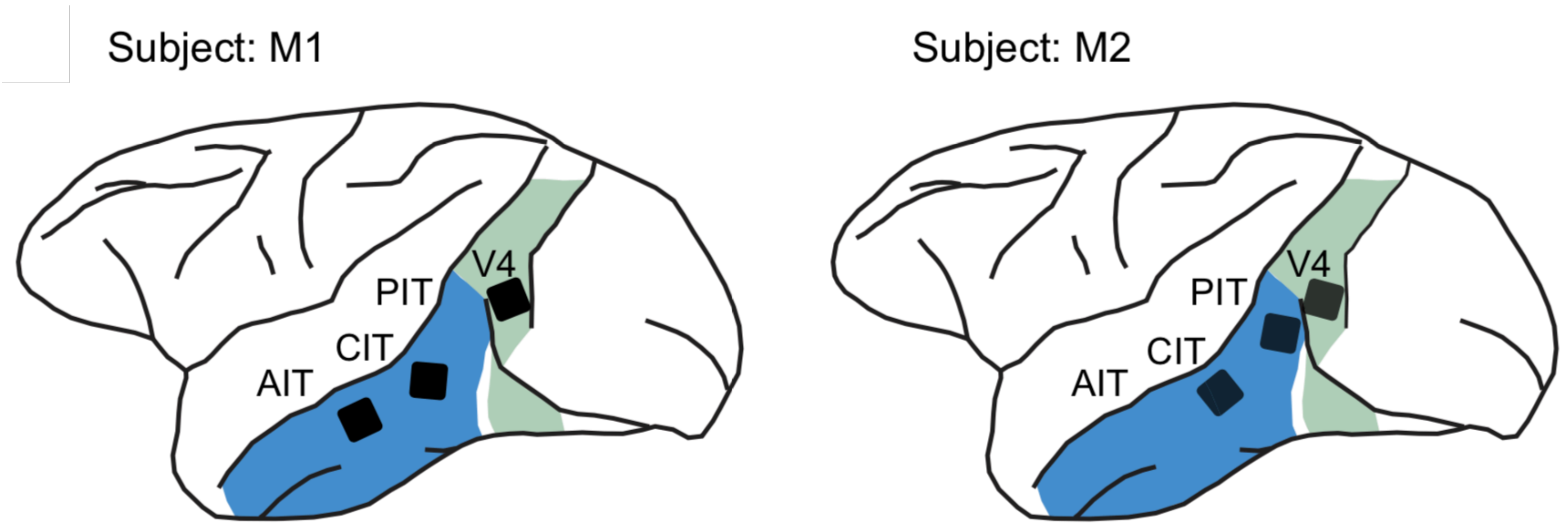
e.g. 3.1

For each object classification task:

1. Segregate feature vectors (i.e. neural responses) training and testing data
2. Build logistic regression model to predict object (e.g. face001 vs. not face001)
3. Determine performance on test data
4. Compute d' (a more sophisticated accuracy measure) from model performance

Majaj et al. 2015





Neural responses

A weighted, linear readout of macaque IT (not V4) approximates human object classification behaviors

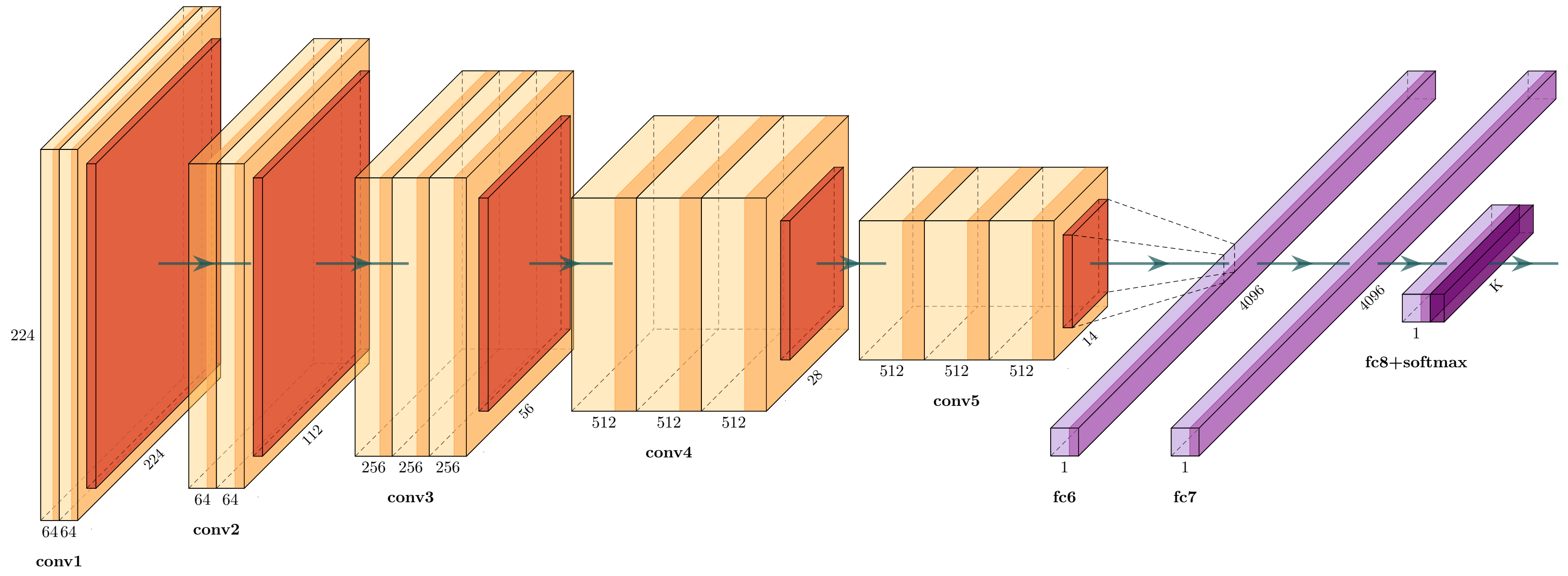
Behavioral responses

0 2 4 6

Human d′on tasks

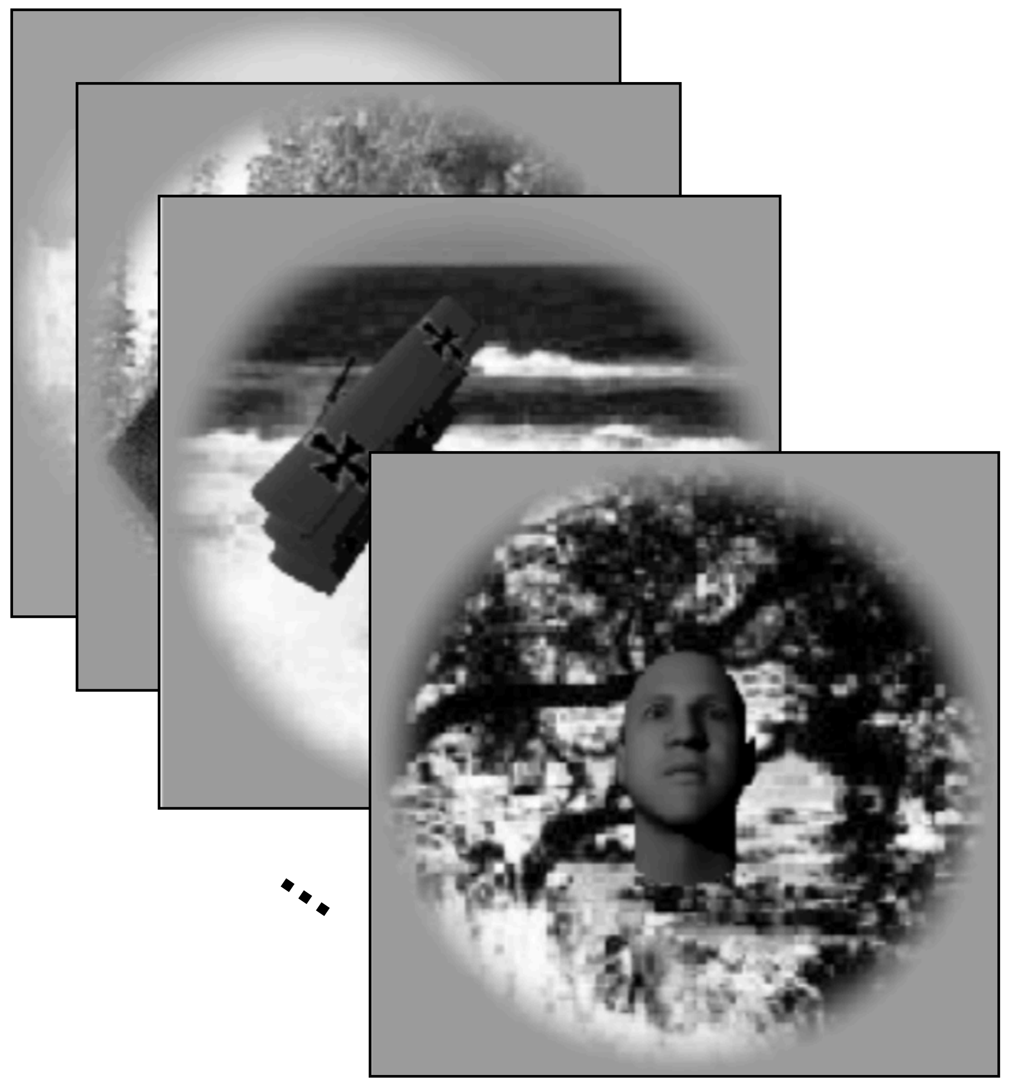
***n* = 104**

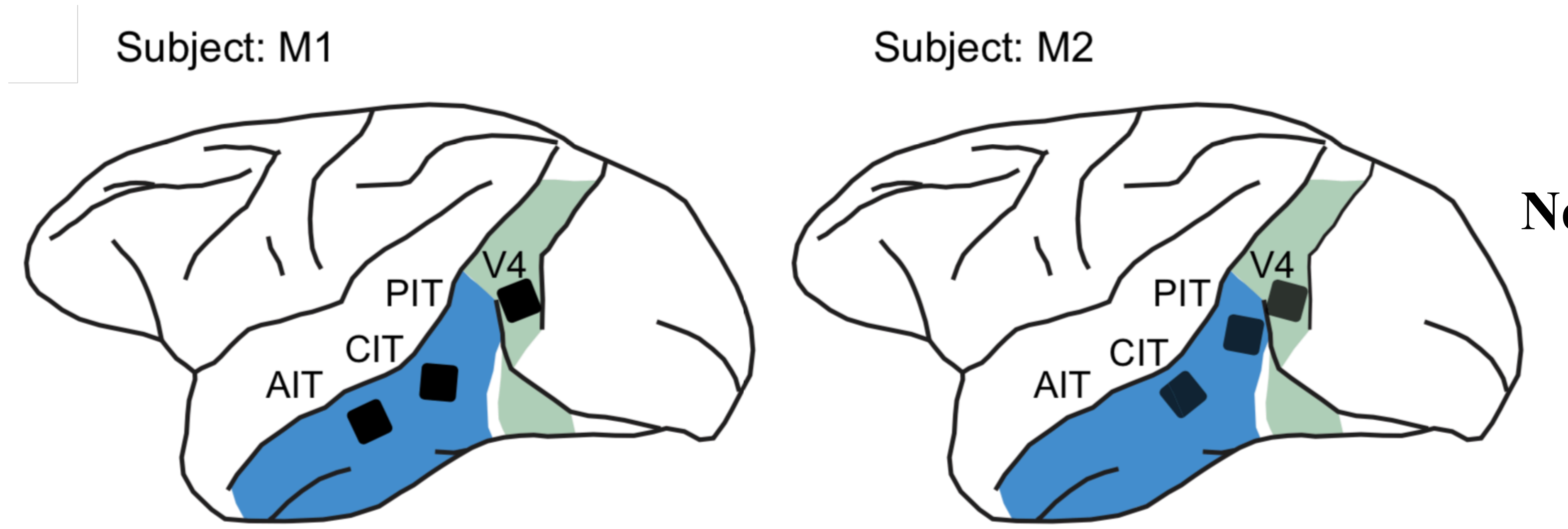
Representational Similarity Analysis (RSA)



Convolutional Neural Networks (CNNs) predict neural responses throughout the VVS directly from experimental stimuli

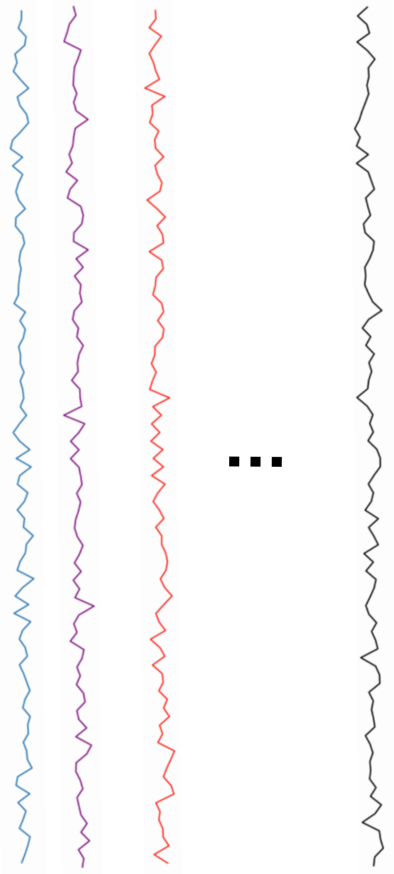
V1: Cadena et al. 2019 | V4: Bashivan et al. 2019 | IT: Yamins et al. 2014 | Behavior: Rajalingham 2018





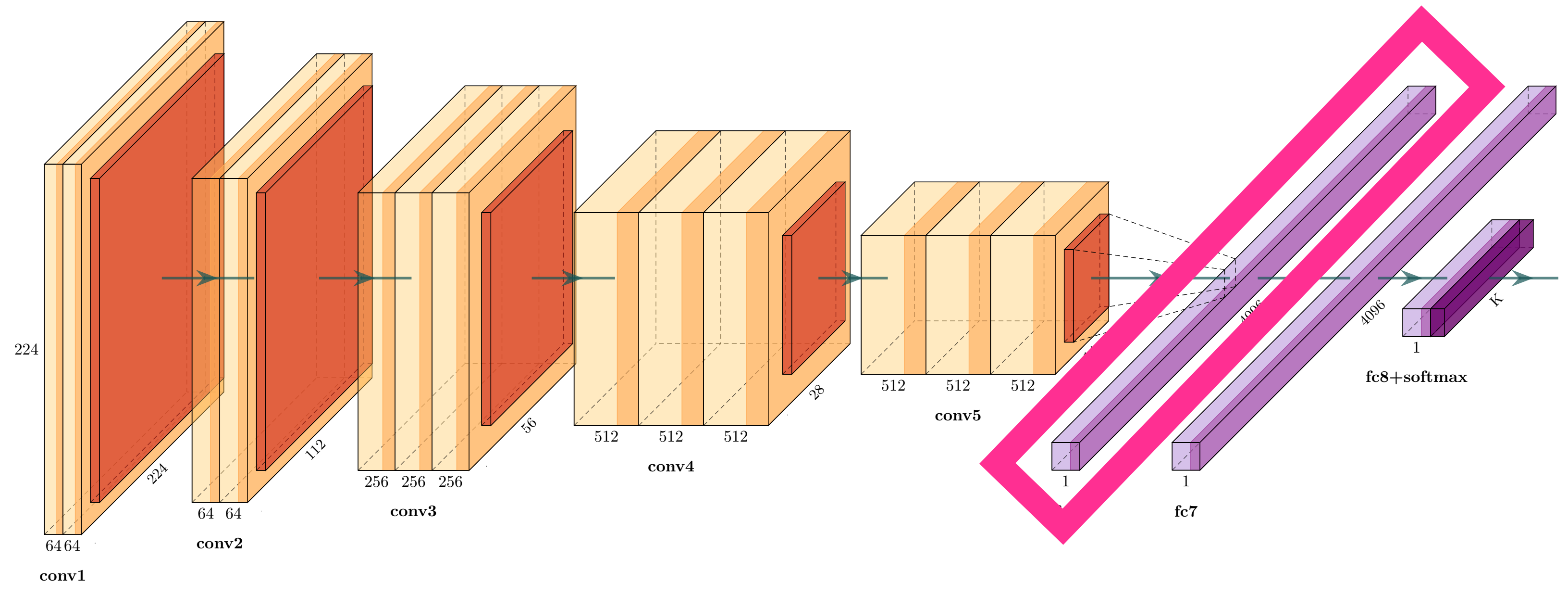
Neural responses

168 (neurons) x 64 (objects)

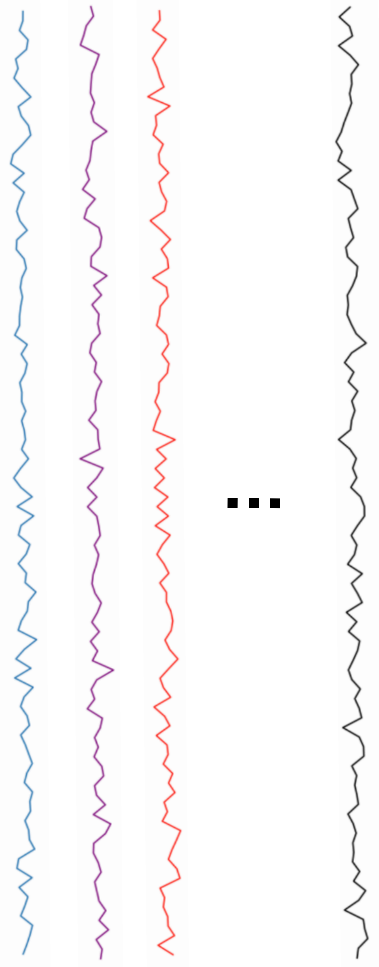


How can relate these measurements?

4096 (“neurons") x 64 (objects)



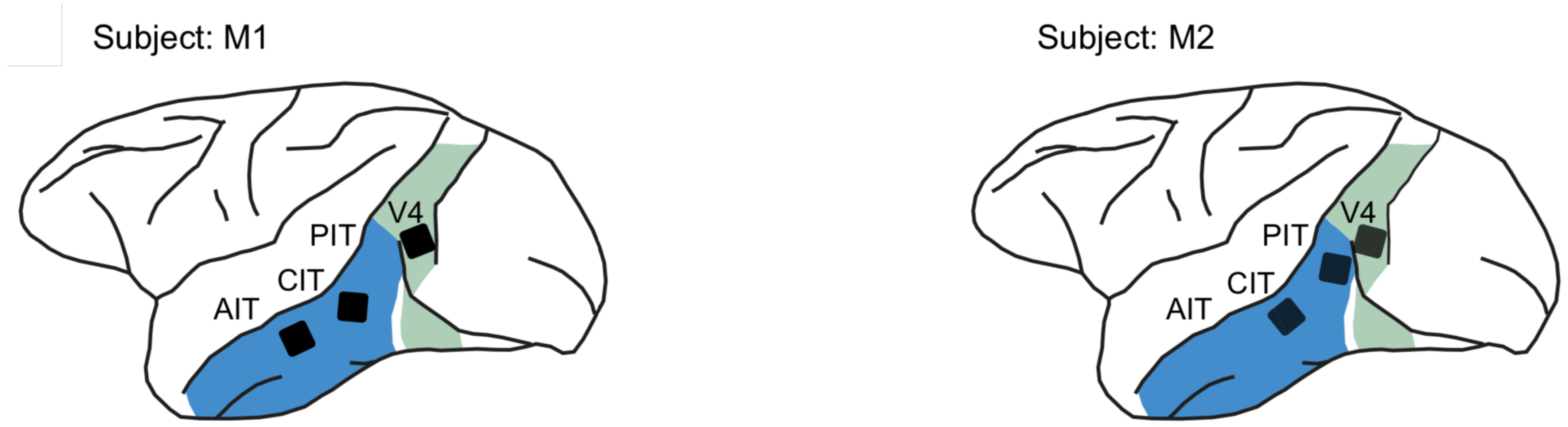
Model responses



In order to compare model and neural responses, here we can

i. preserves task relevant information (e.g. relationship between stimuli)
ii. while being agnostic to units of measurement (e.g. n neurons)

Let's work through an example that's simpler, but still a real problem:
Compare neural responses across monkeys with different number of electrodes

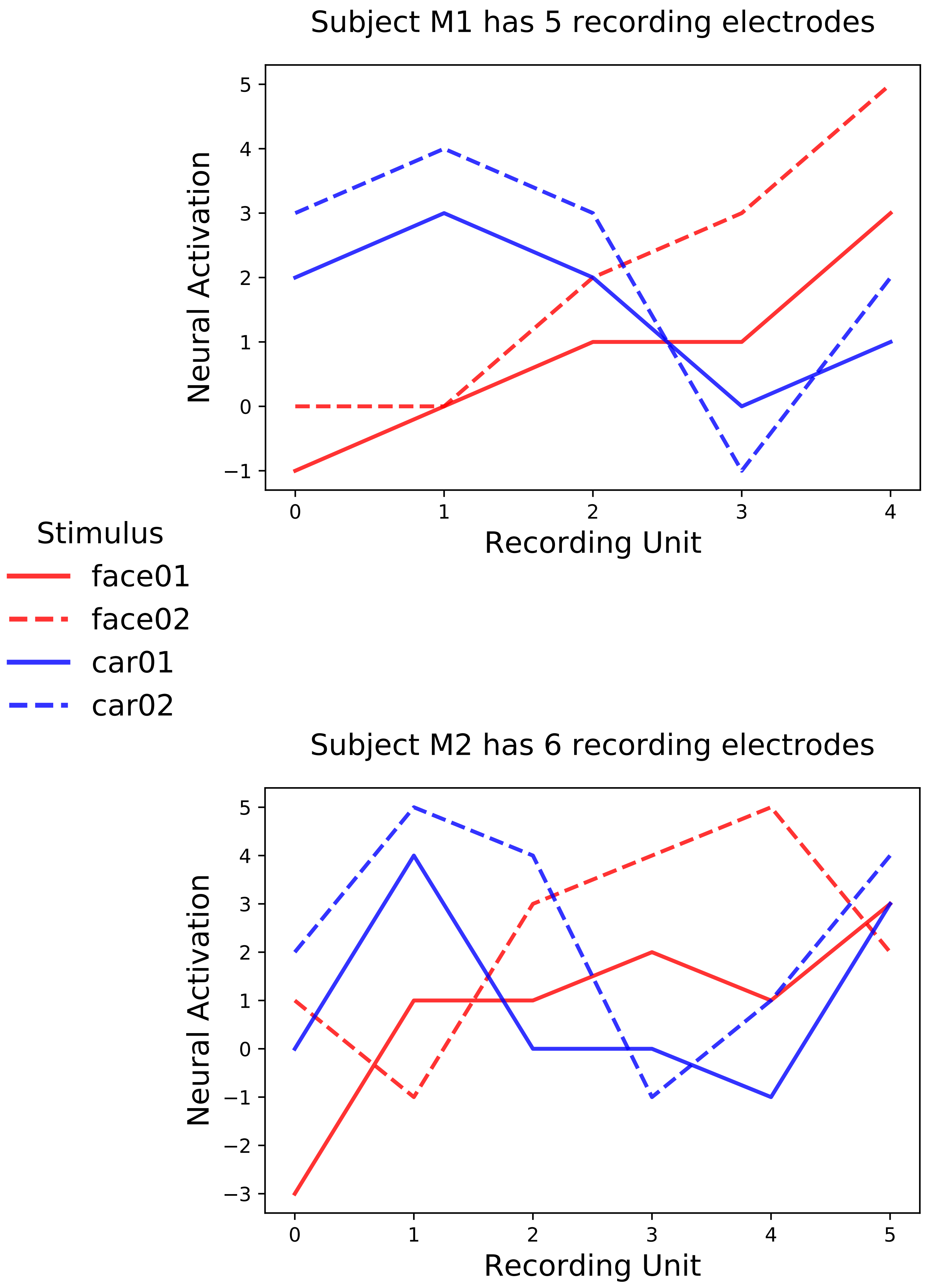


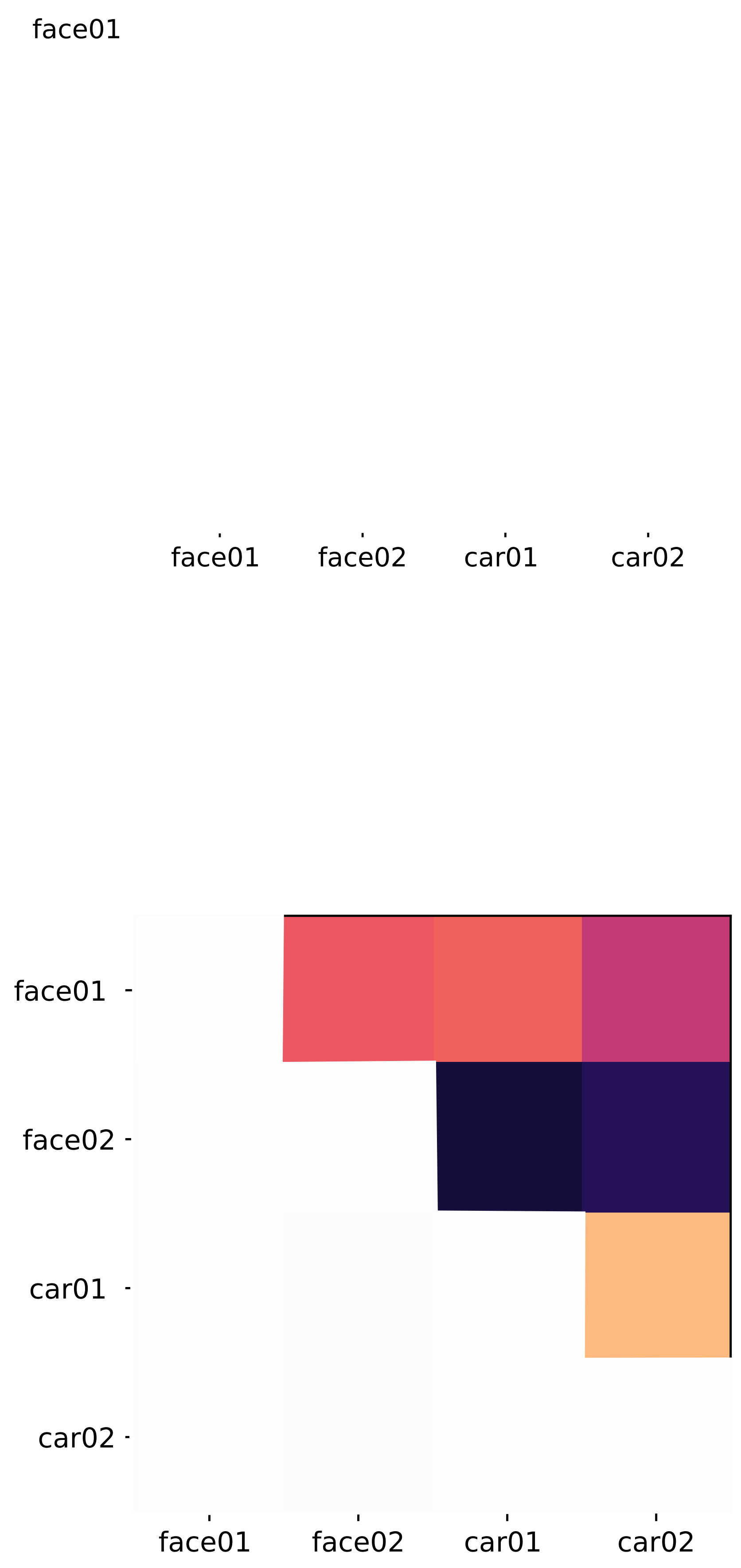
Data in native format has mismatched dimensions but there seems to be shared structure

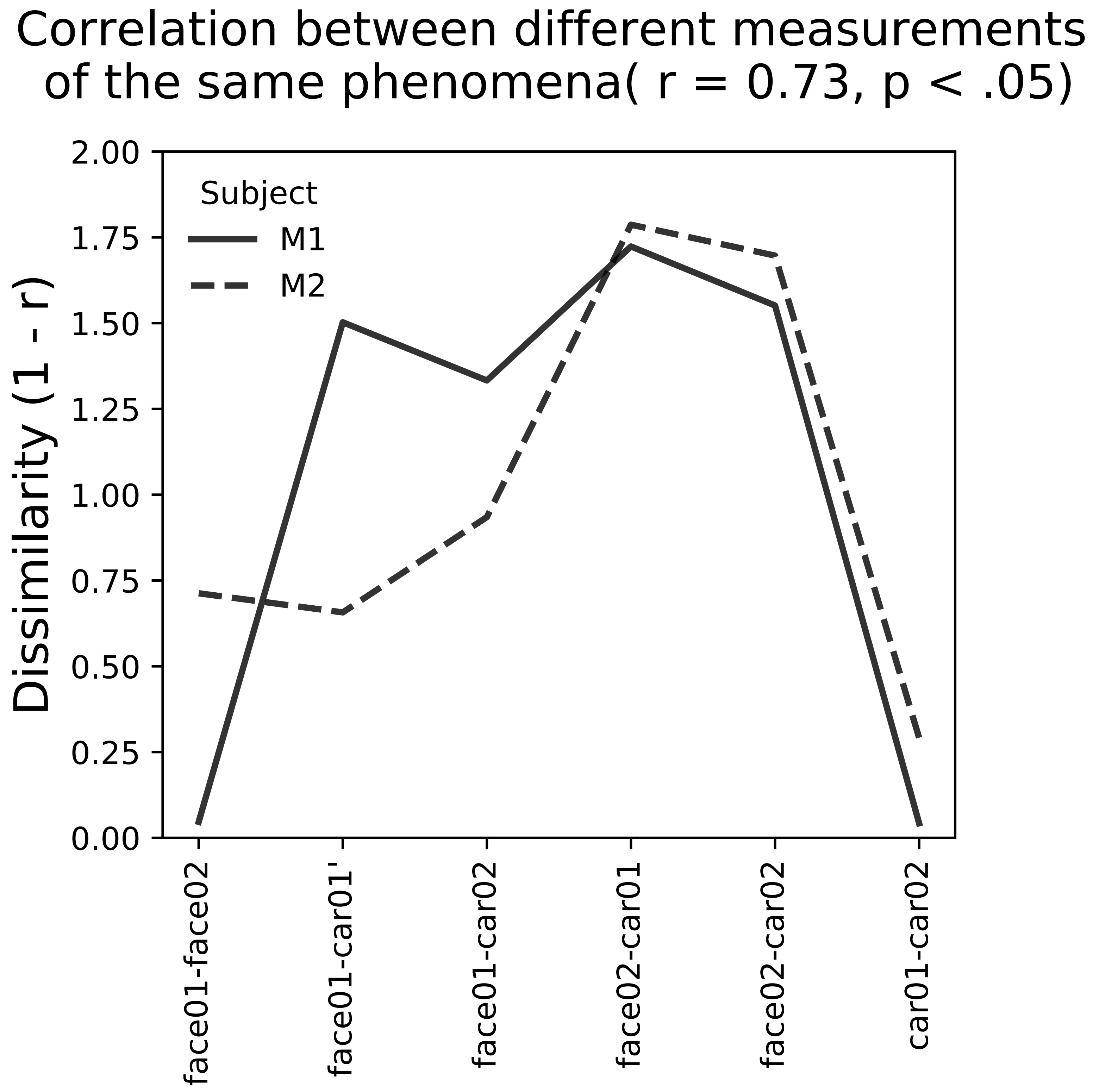
Generate a Representation Dissimilarity Matrix (RDM) —pairwise 1-r

Extract off-diagonal

Compute correlation between off-diagonals from M1 & M2 RDMs





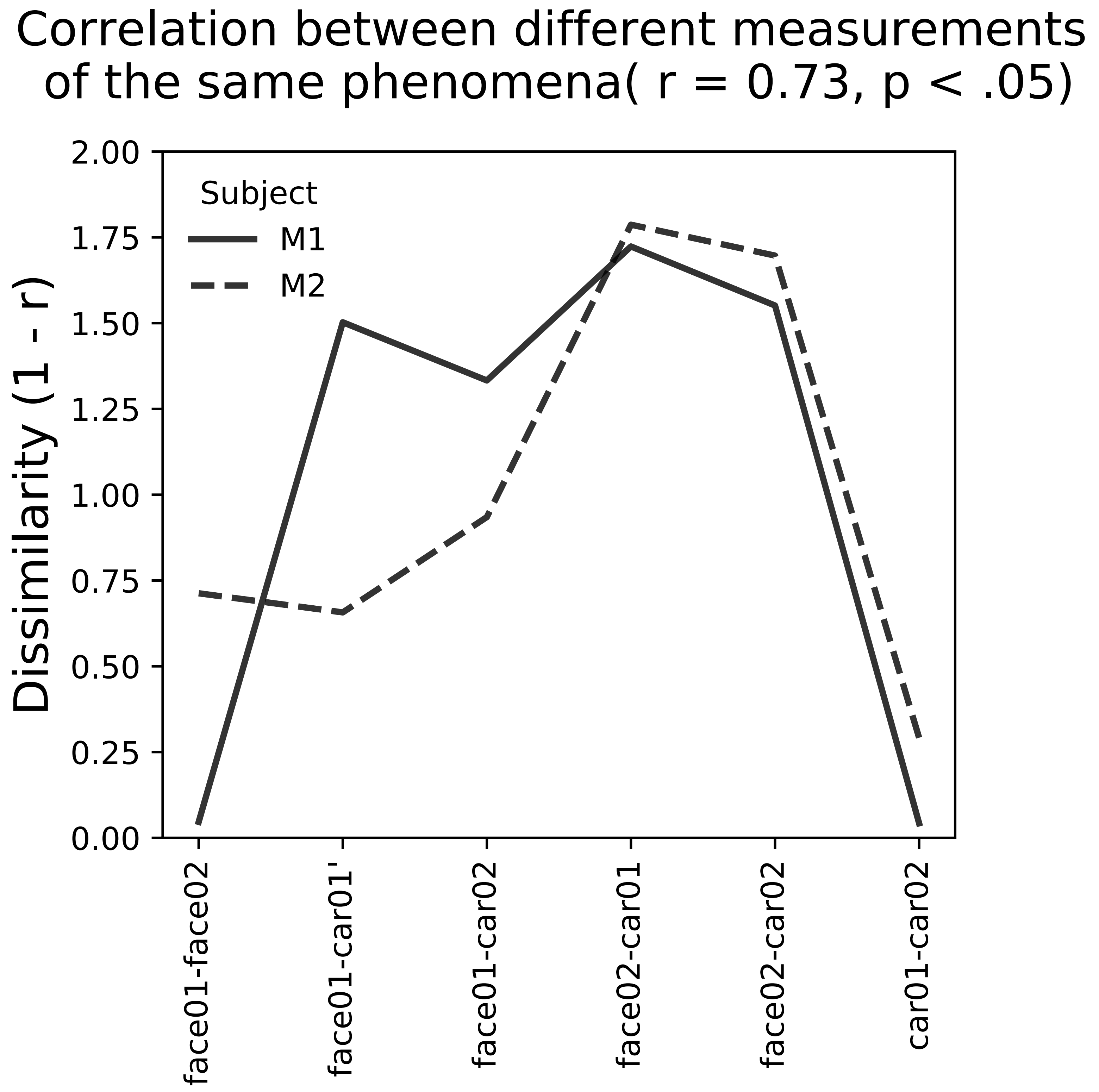


This protocol is called a Representational Similarity Analysis (RSA)

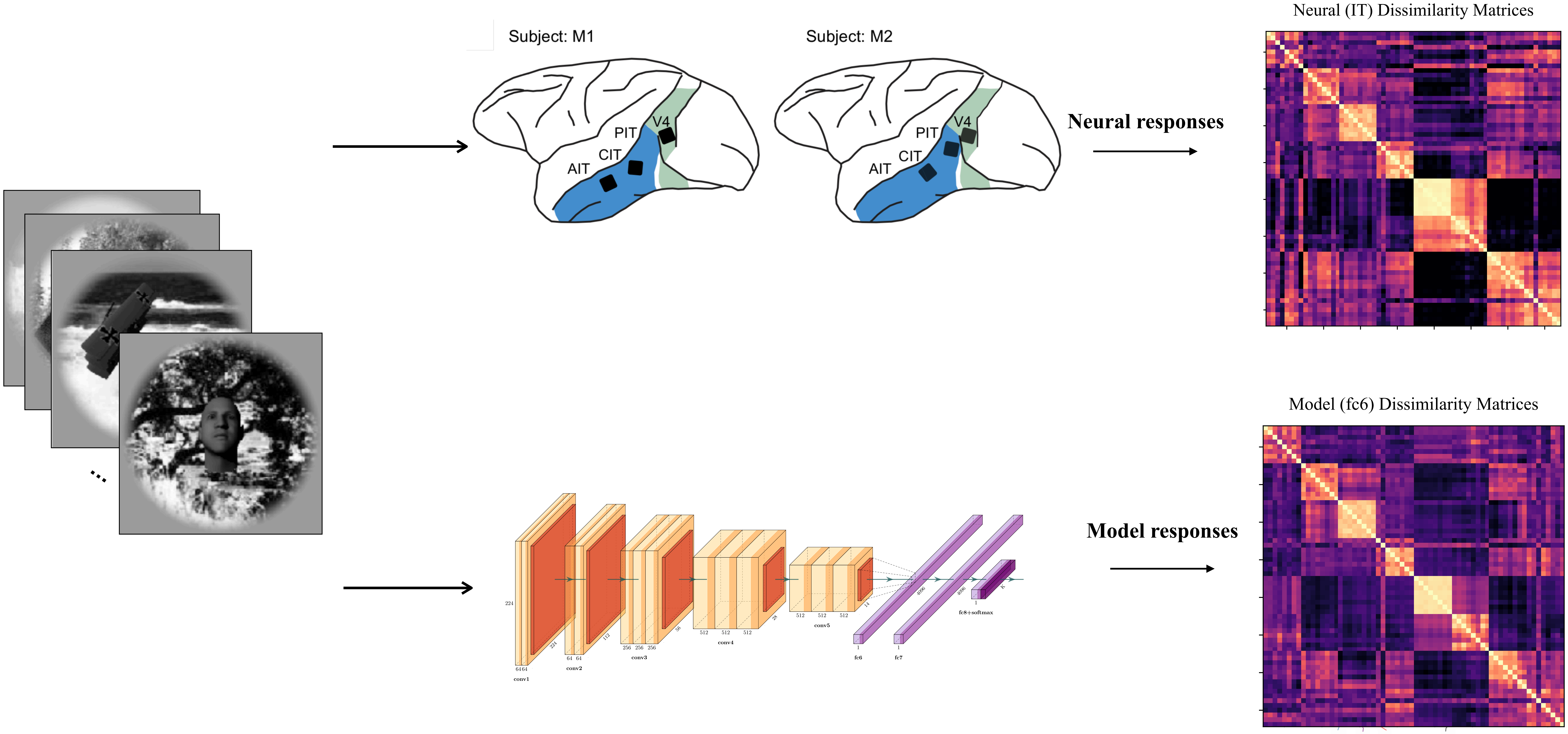
How do we generate p values?

- We generate an empirical null: compute correlation between shuffled RDMs for n iterations
- Compute probability of observed r directly
  - p_value = mean(empirical_null > r_observed)

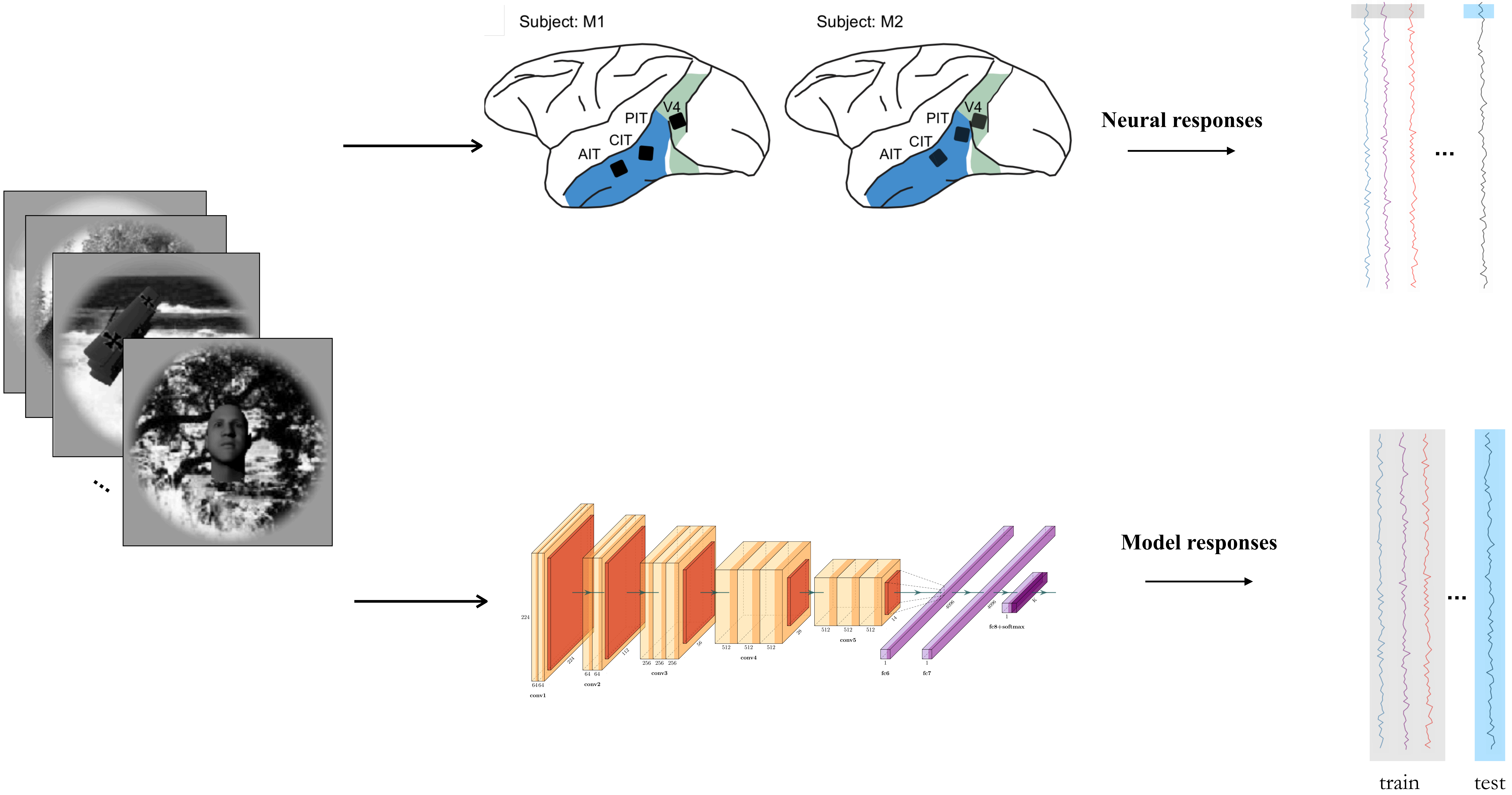
Typically we only report this final (RSA) correlation value

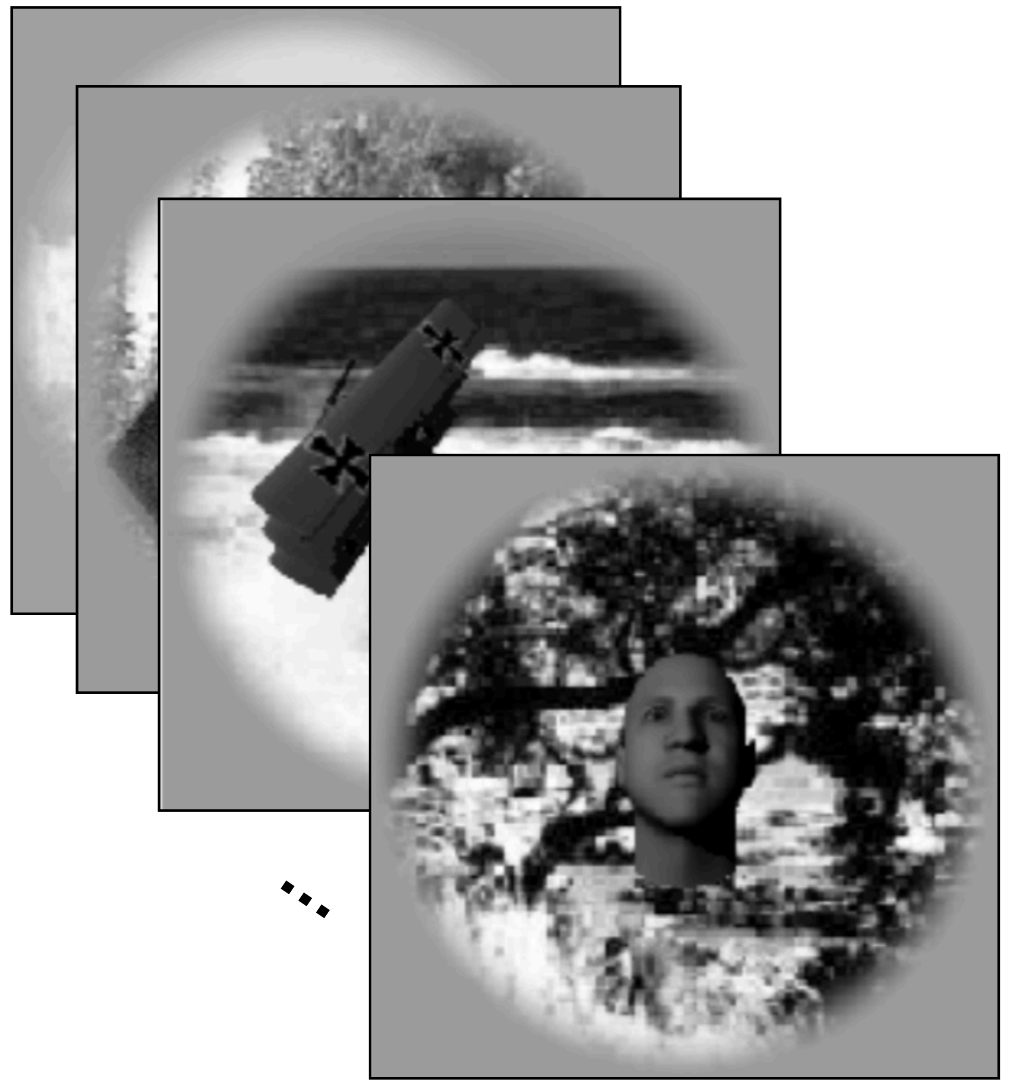


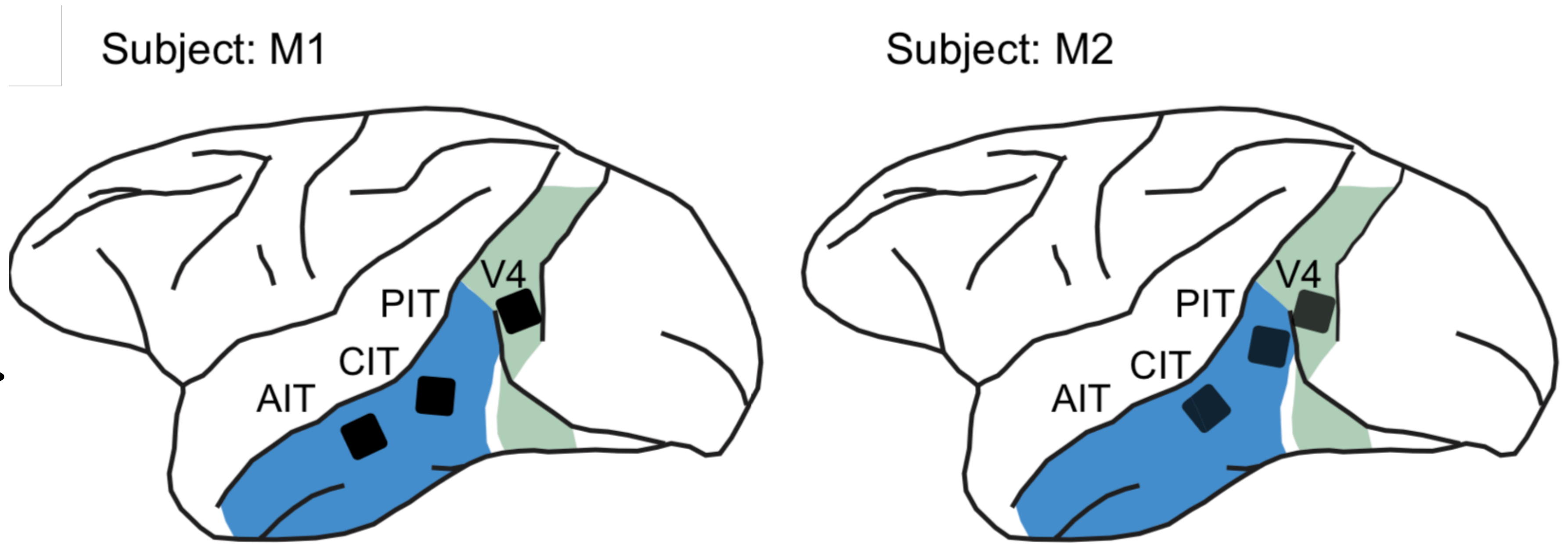
Now lets go back to the model comparison we're interested in

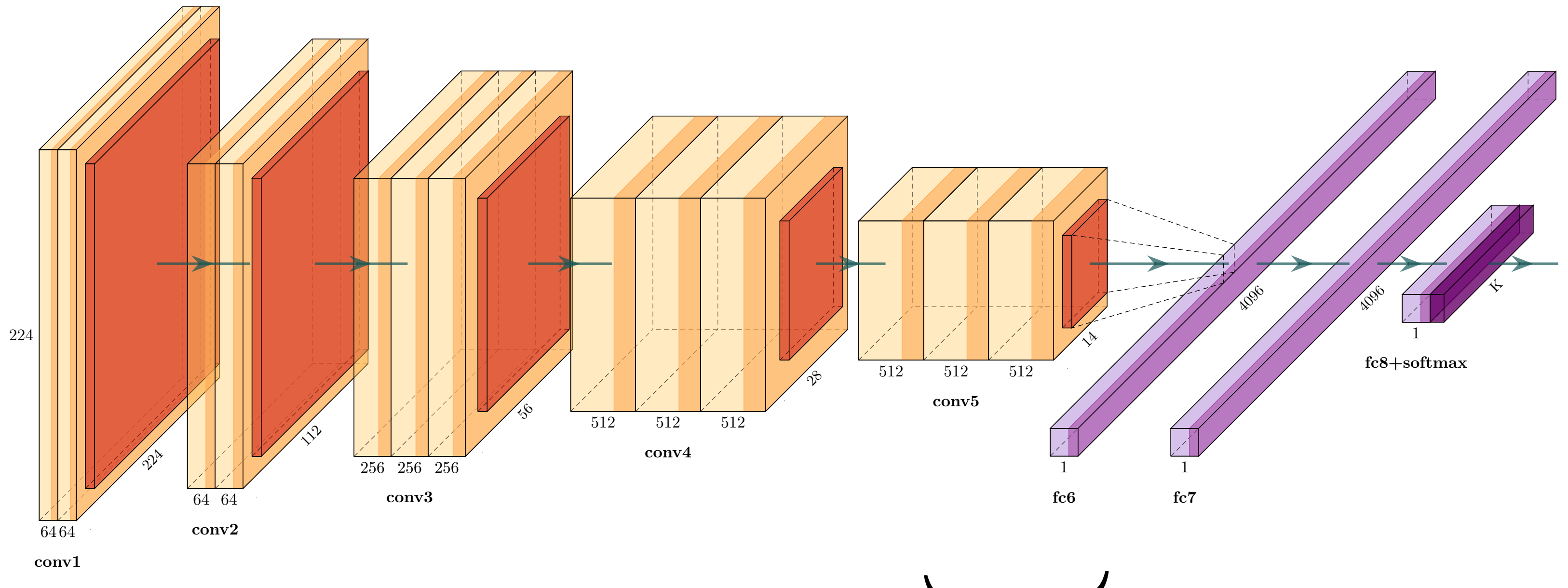

Subject: M1
Subject: M2
PIT
V4
CIT
AIT
Neural responses
Neural (IT) Dissimilarity Matrices
Model responses
Model (fc6) Dissimilarity Matrices
...

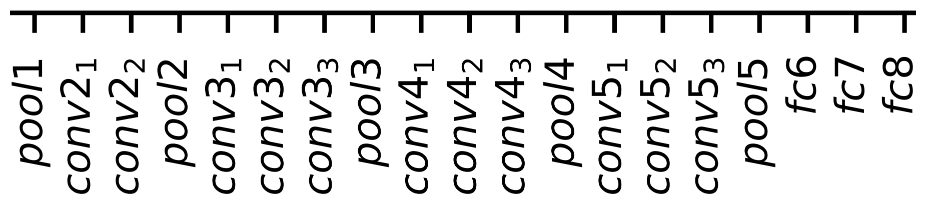
Learning a mapping between responses to shared stimuli


Subject: M1
Subject: M2
V4
PIT
CIT
AIT
Neural responses
Model responses
train
test









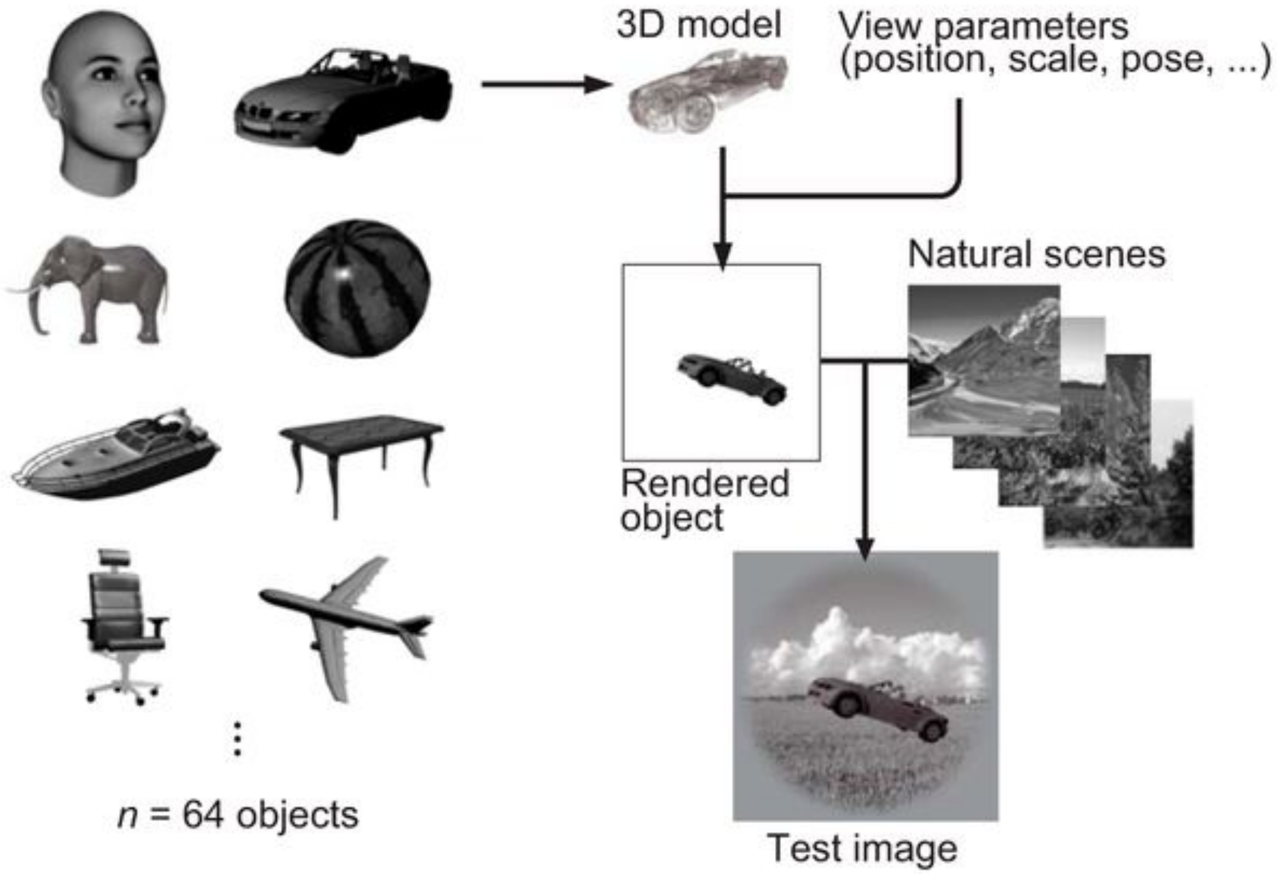
“V4 like”

“IT like”

Using computational models as a common metric space across data types

What if we want to create a whole new behavioral experiment, but still make use of the same images and neural data?

Generating a novel task that leverages previous data



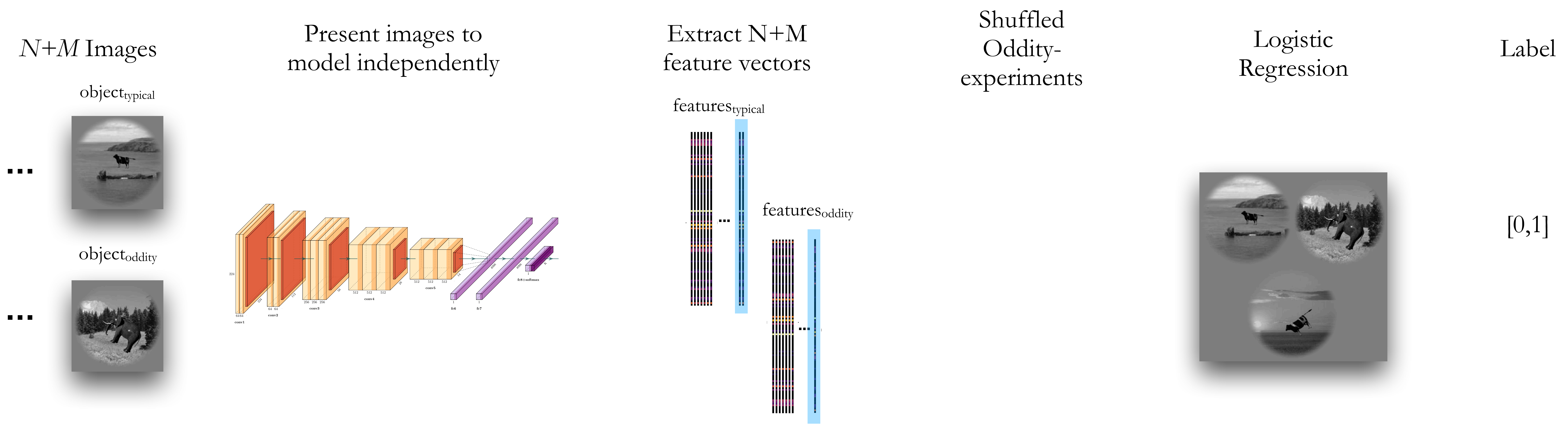
Majaj et al. 2015

Typical Objects

Oddity

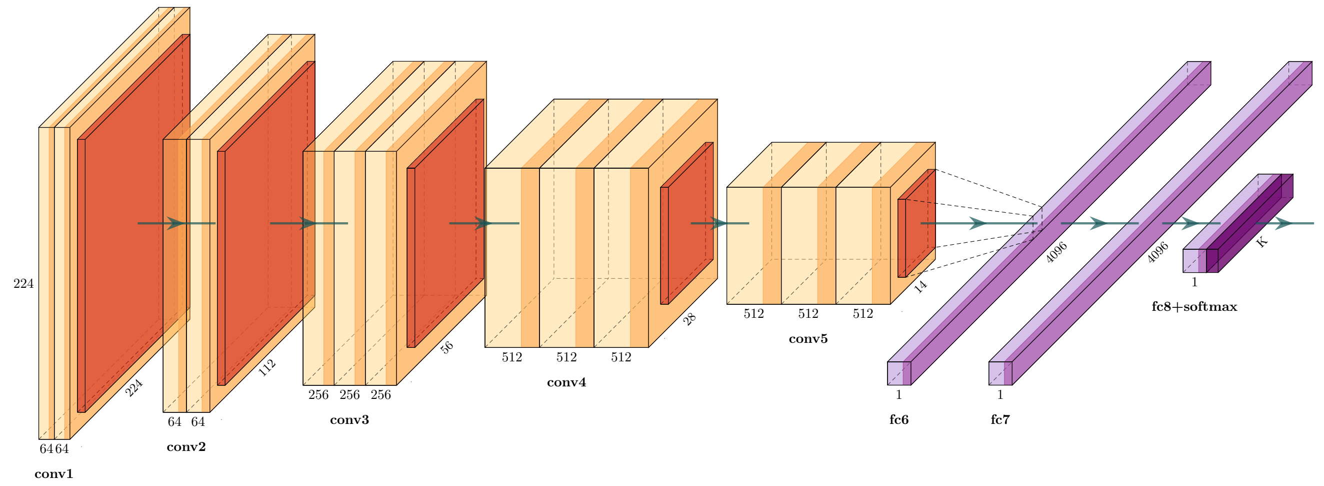
Single Trial

Experimental Protocol

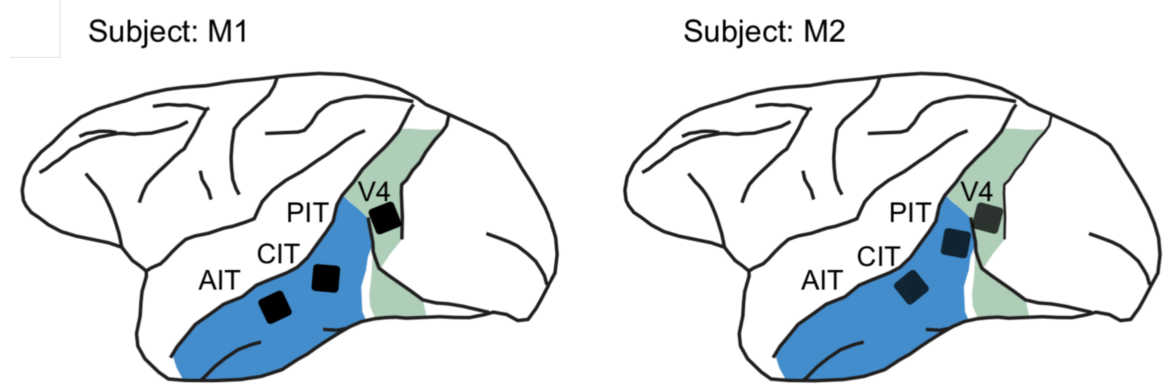


For each trial (e.g. cow_image34, cow_image75, elephan_image34):

1. Pass all N examples of object$_{typical}$ and all M examples of object$_{oddity}$ to model
2. Extract these M+N feature vectors from an "IT like" layer
3. Hold out *this trial* feature vectors (e.g. cow_features34, cow_features75, elephant_features34)
4. Generate training data/oddity experiments: sampling without replacement from typical and oddity features
5. Build logistic regression model from pseudo-experiments
6. Evaluate model on this trial (e.g. cow_image34, cow_image75, elephan_image34)
7. Repeat over n permutations



Model responses



Neural responses

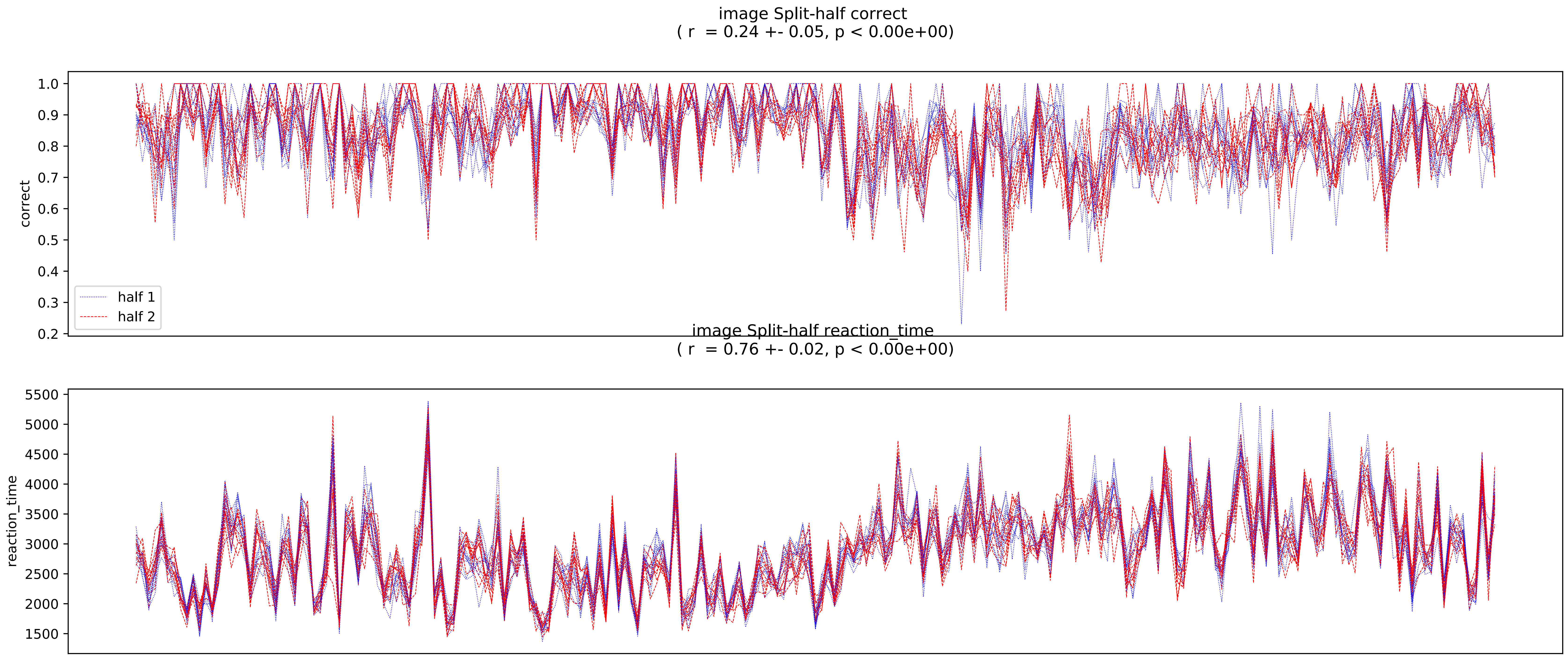
***n* = 297**

Behavioral responses

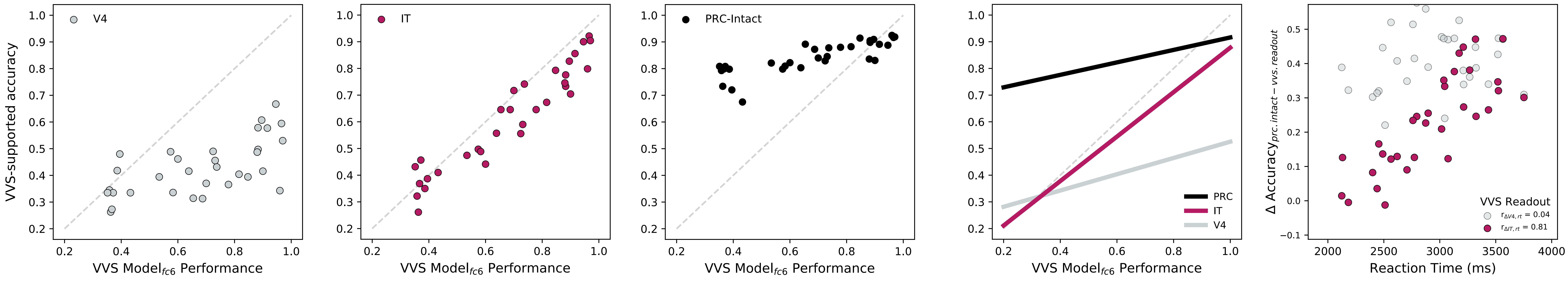
3-way oddity tasks

A quick note on reliability

- What's the most variance we could hope a model could explain?
  - The variance that's reliably evident in different samples of the data
  - Often called the split-half reliability, or referred to as the "noise ceiling"
- When we can, we should look at reliability across multiple resolutions



Relating neural, model, and behavioral data



- But you've already seen this, haven't you :)

Questions?