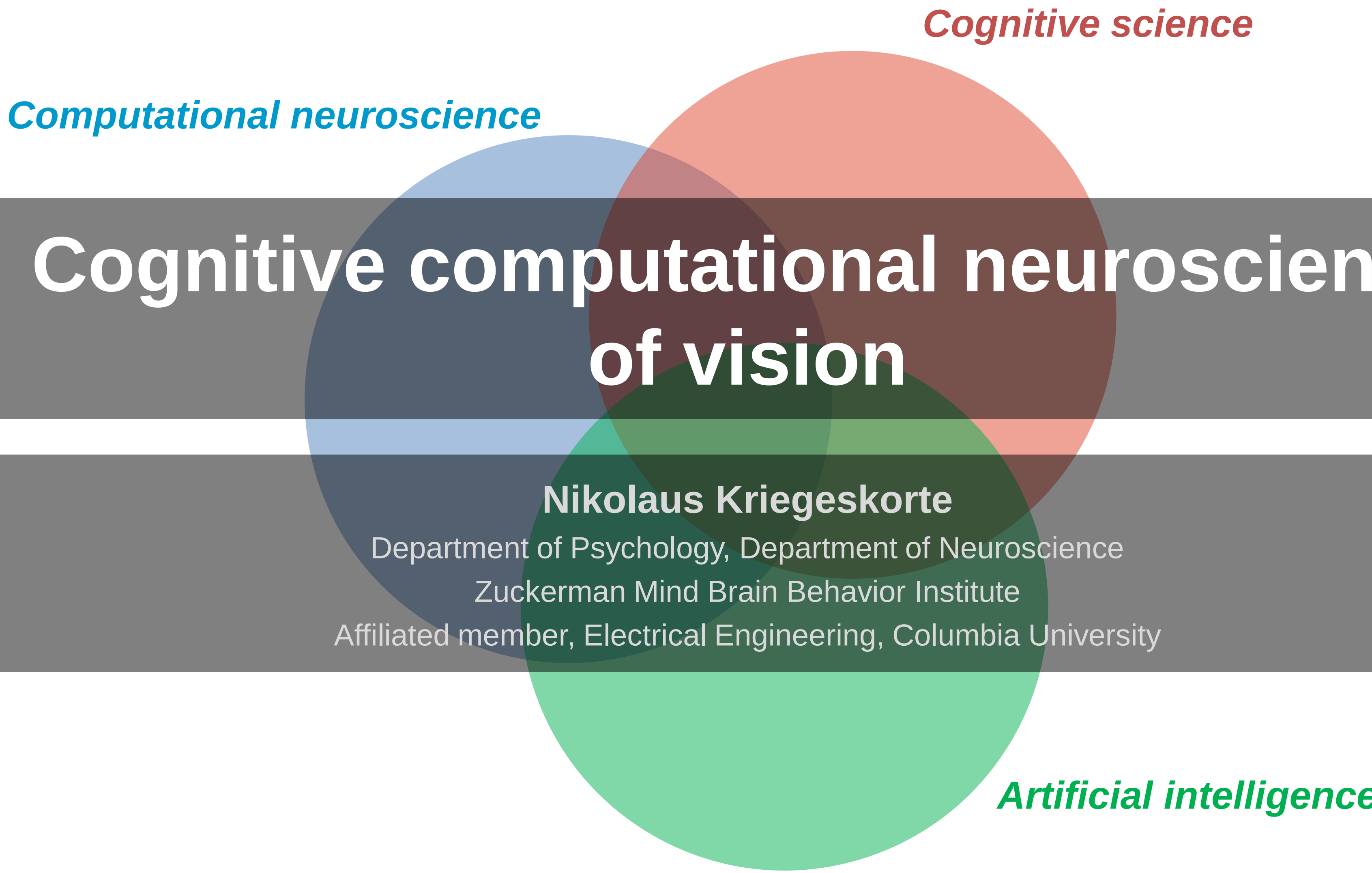
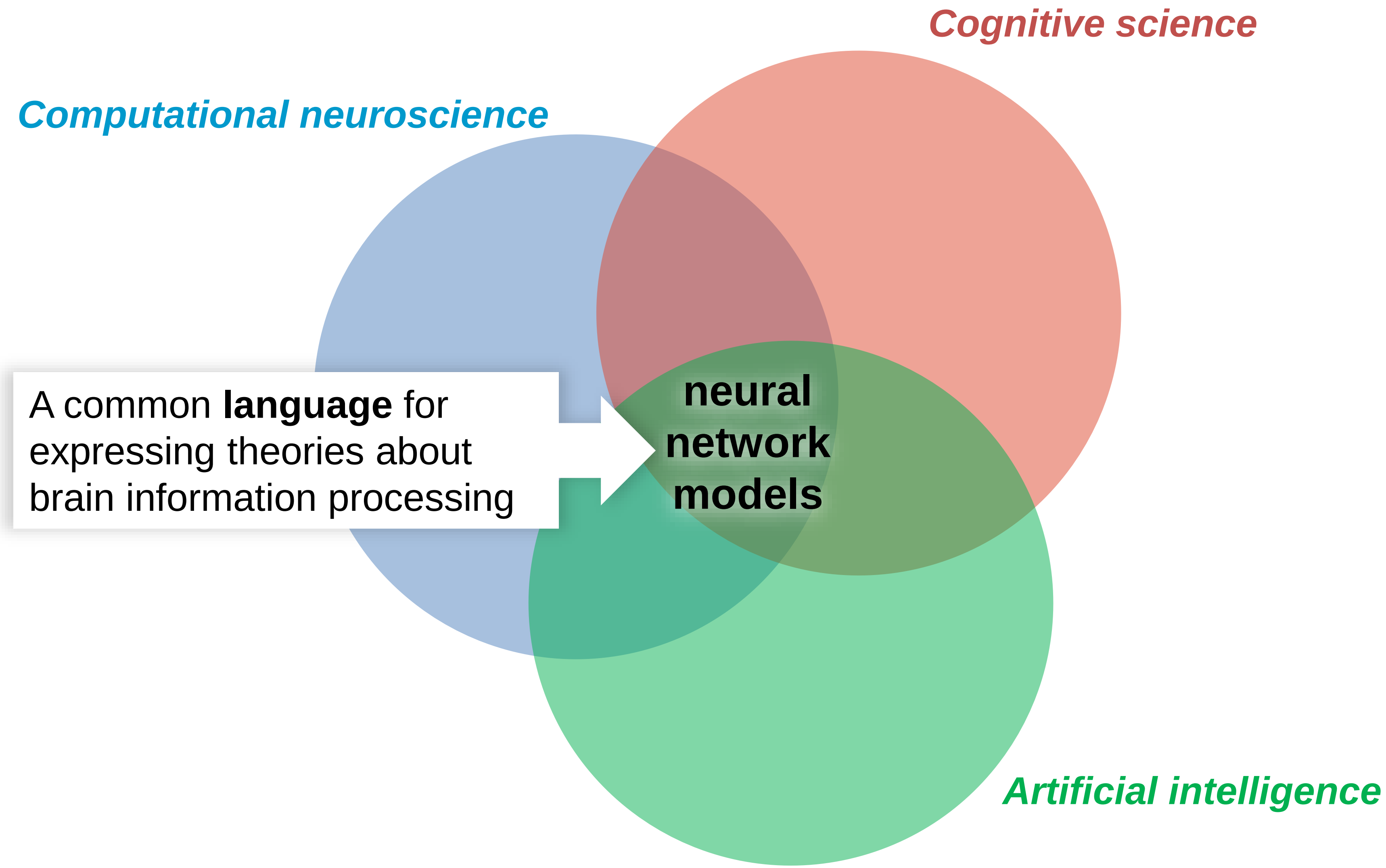


Cognitive computational neuroscience of vision

Nikolaus Kriegeskorte
Department of Psychology, Department of Neuroscience
Zuckerman Mind Brain Behavior Institute
Affiliated member, Electrical Engineering, Columbia University

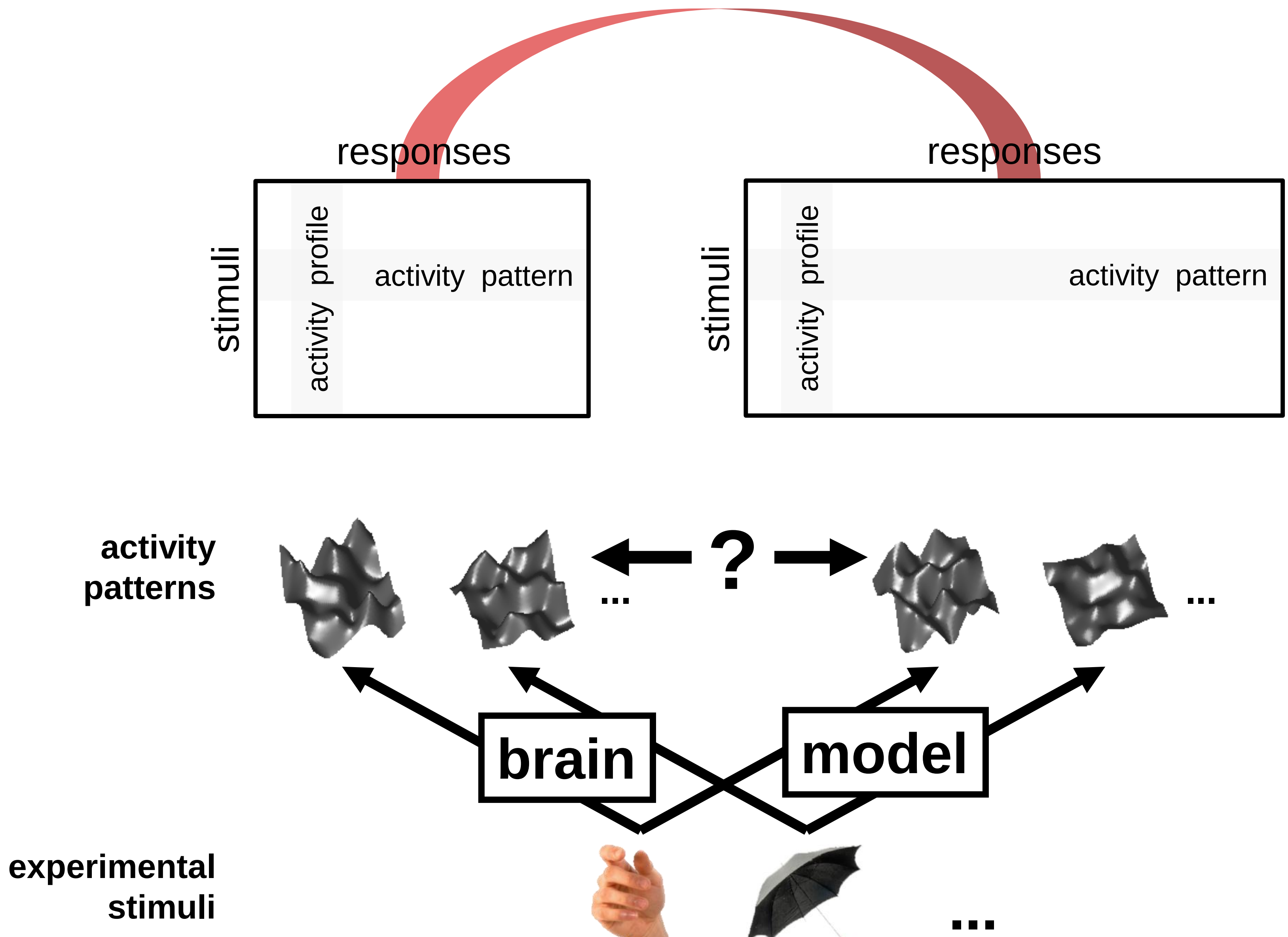
Cognitive computational neuroscience



Kriegeskorte & **Douglas** 2018

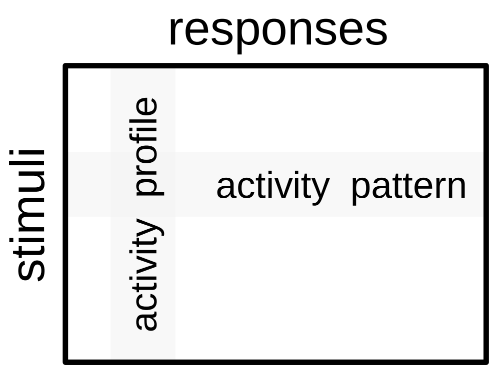
How can we test neural network models with brain-activity data?

Predicting representational spaces



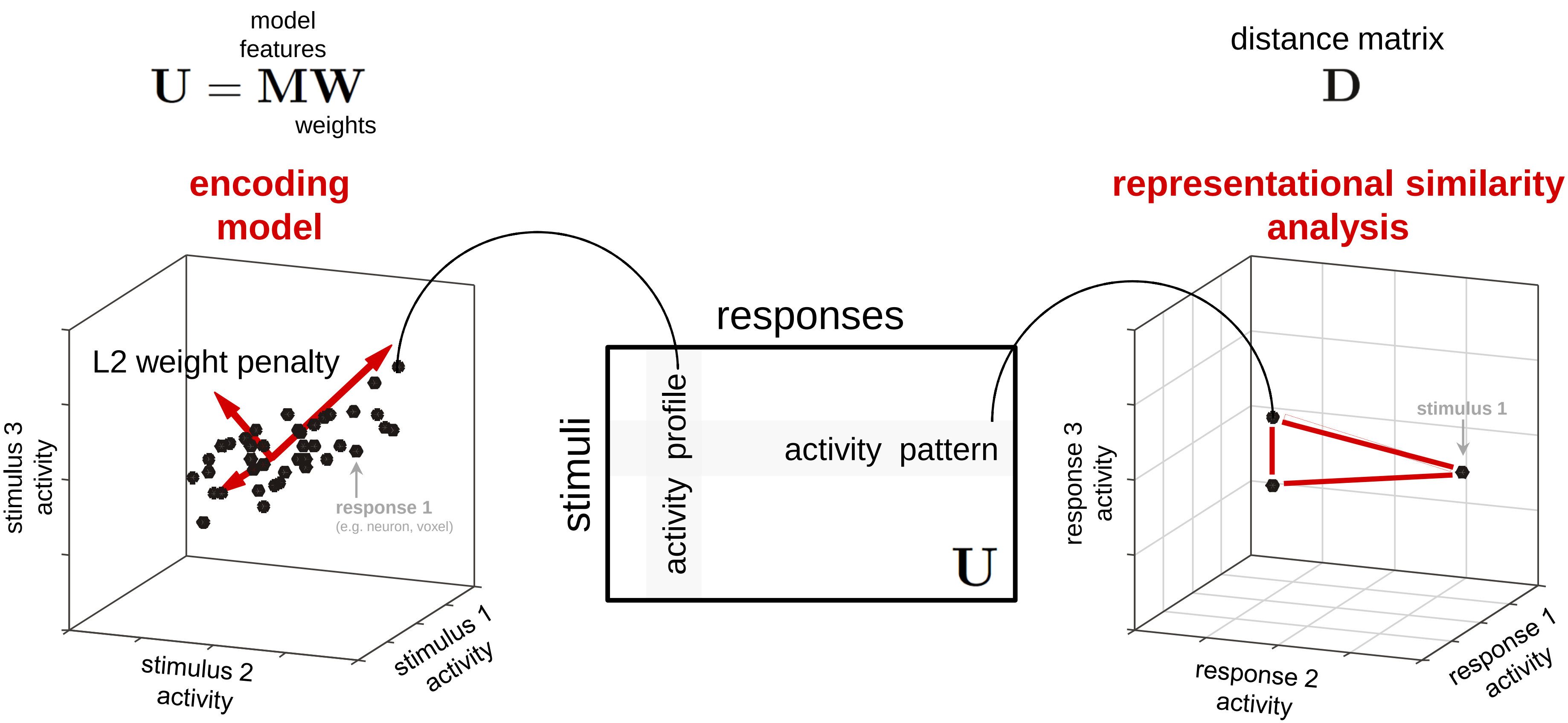
Diedrichsen & **Kriegeskorte** 2017, **Kriegeskorte** & **Diedrichsen** 2019

Predicting representational spaces



Diedrichsen & **Kriegeskorte** 2017, **Kriegeskorte** & **Diedrichsen** 2019

Predicting representational spaces



Diedrichsen & Kriegeskorte 2017, Kriegeskorte & Diedrichsen 2019

Predicting representational spaces

model each response separately

model stimulus-by-stimulus matrix of summary statistics

model features

$$\mathbf{U} = \mathbf{MW}$$

weights

second-moment matrix

$$\mathbf{G}$$

$$diag\left(\mathbf{CGC}^T\right)$$

$$-\frac{1}{2}\mathbf{HDH}$$

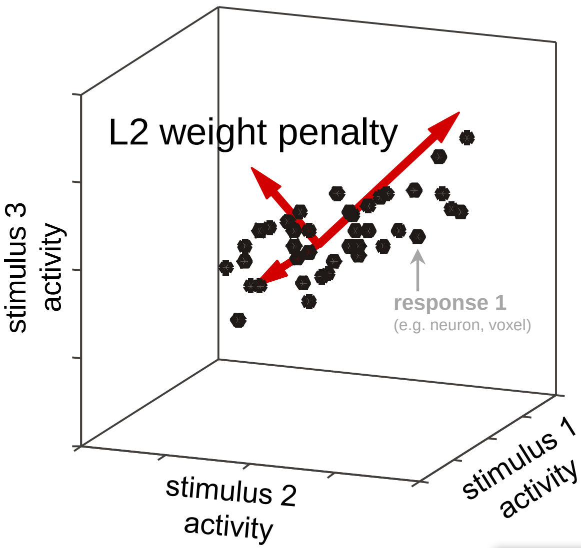
distance matrix

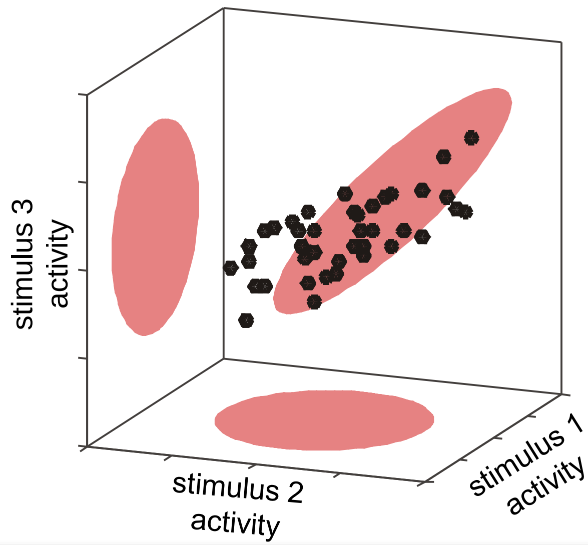
$$\mathbf{D}$$

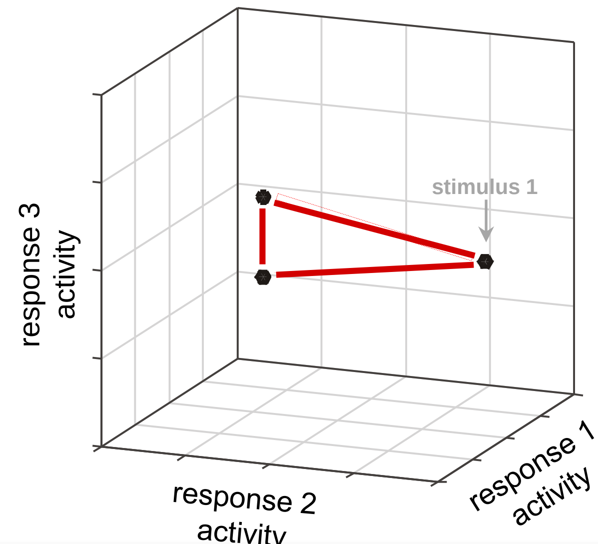
encoding model

pattern component model

representational similarity analysis





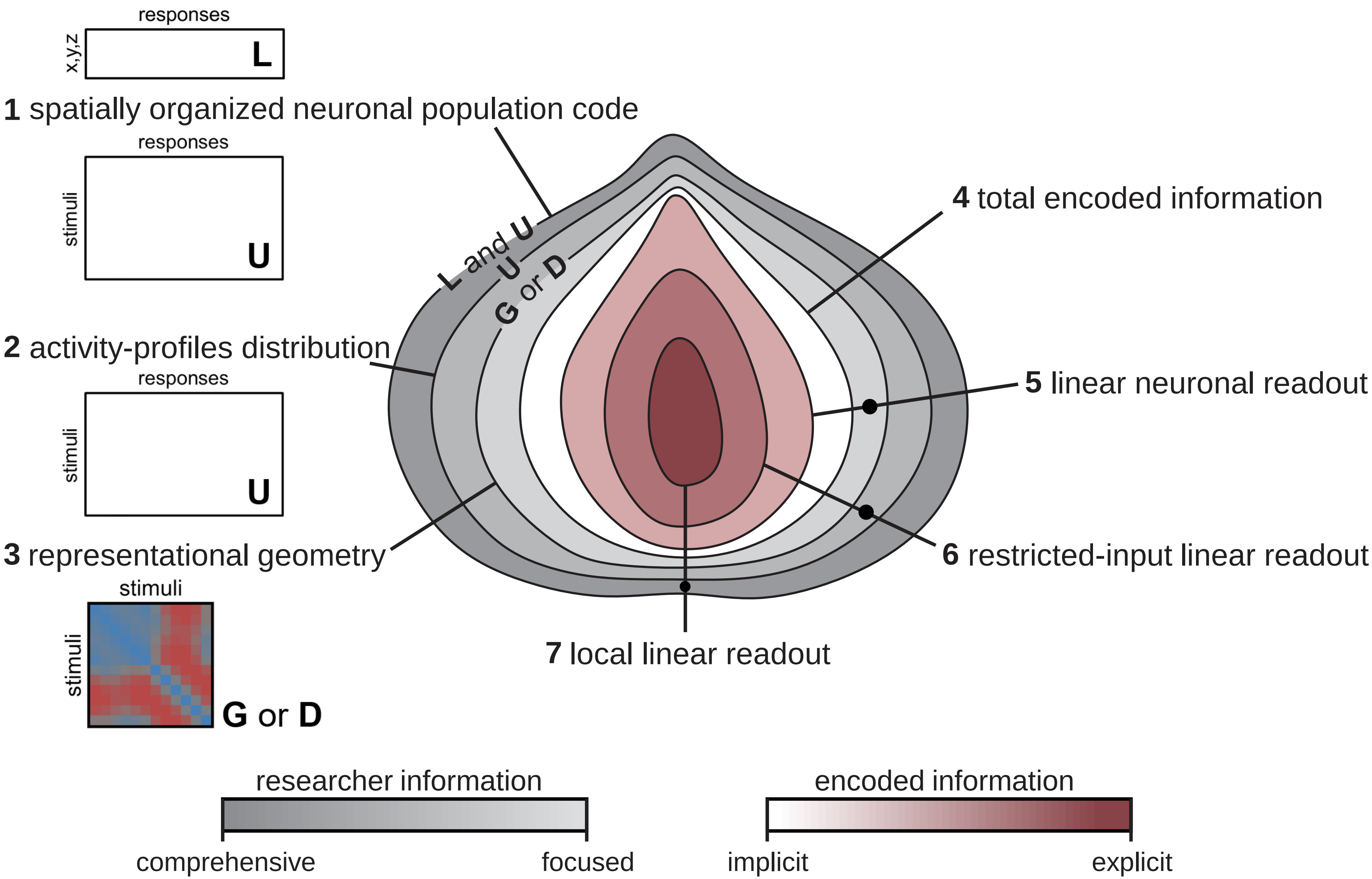


model acti

Core commonality: All three test hypotheses about the second moment of the activity profiles.

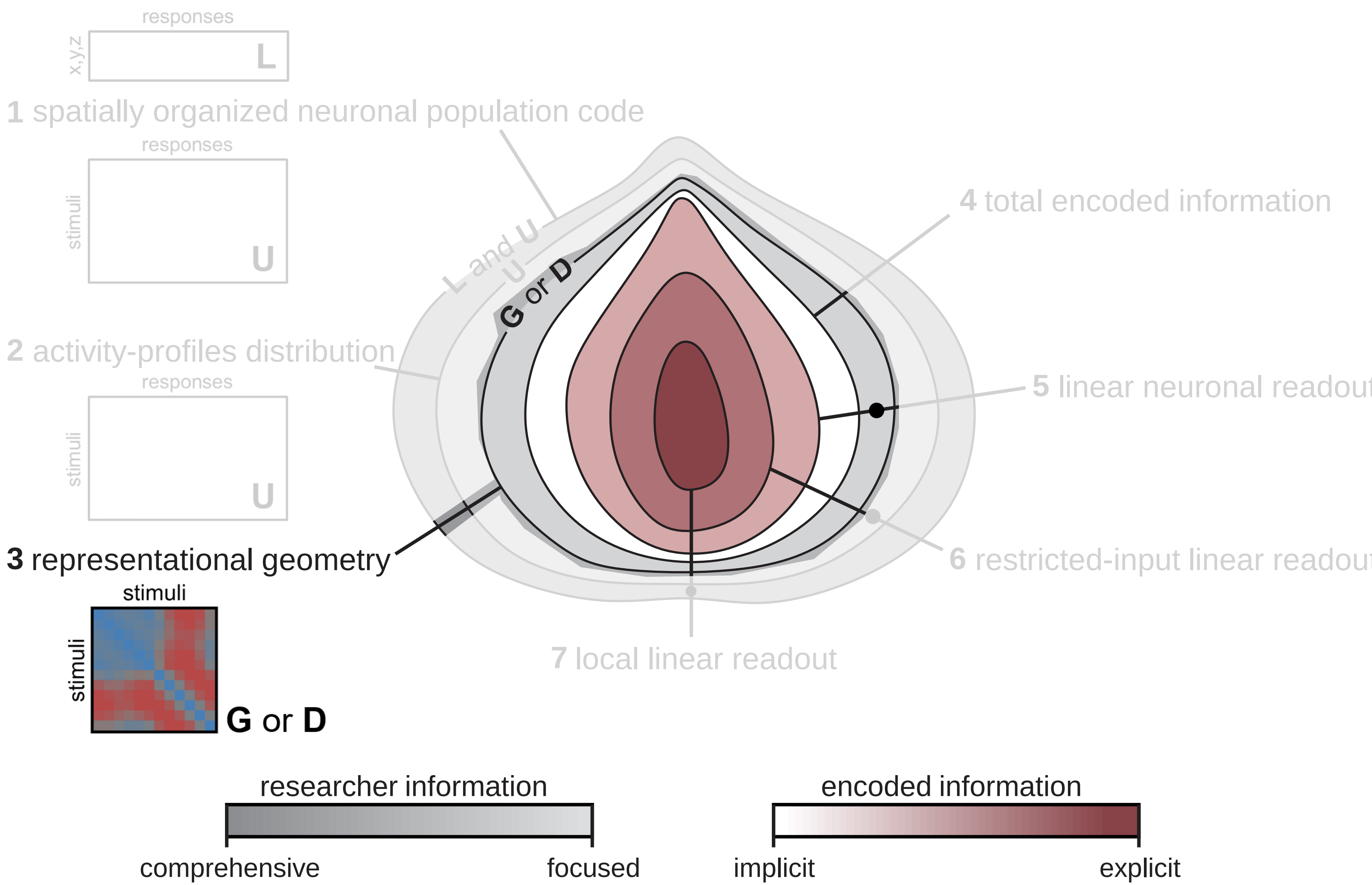
Diedrichsen & Kriegeskorte 2017, Kriegeskorte & Diedrichsen 2019

The onion of brain representations



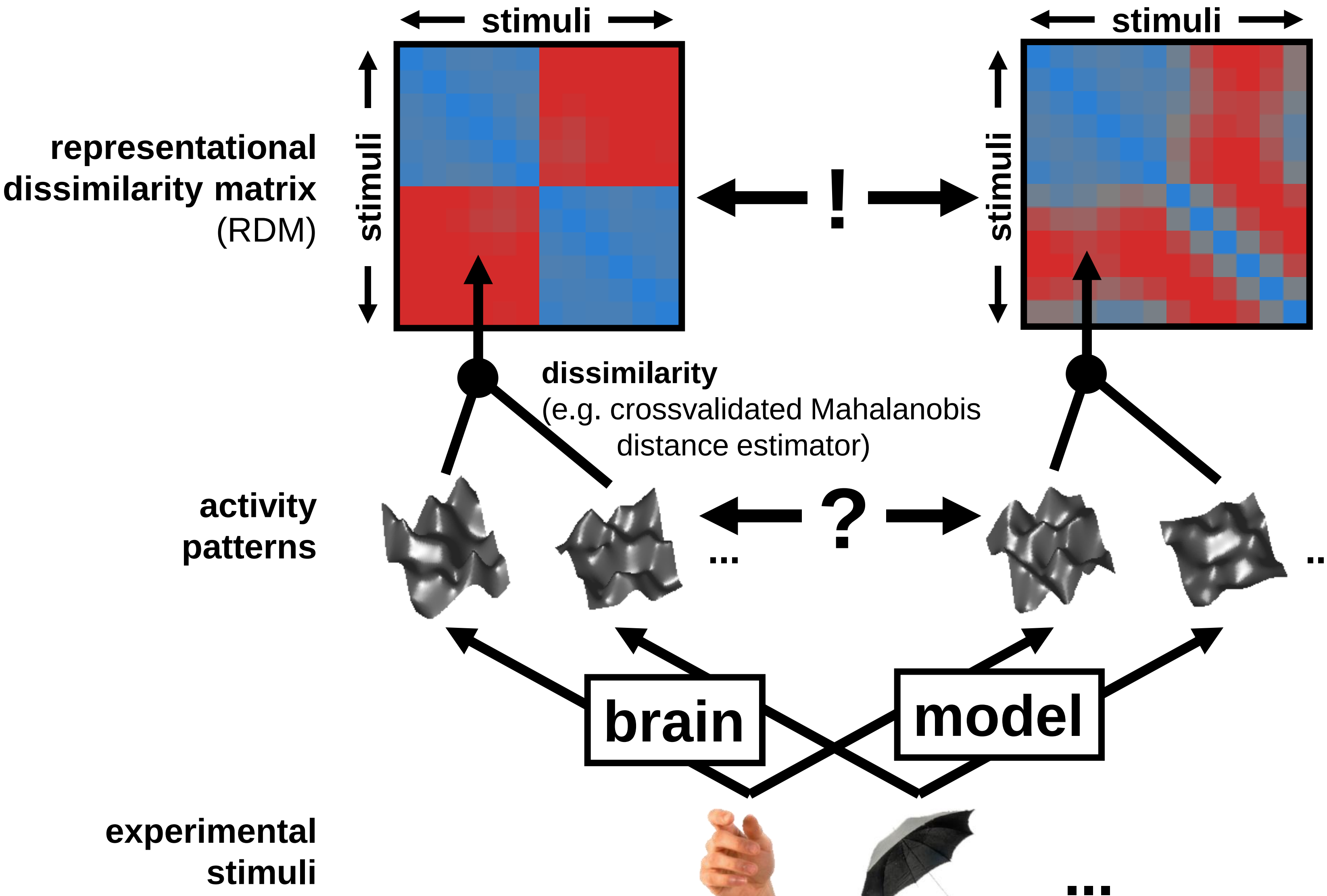
Kriegeskorte & **Diedrichsen** 2019

The onion of brain representations



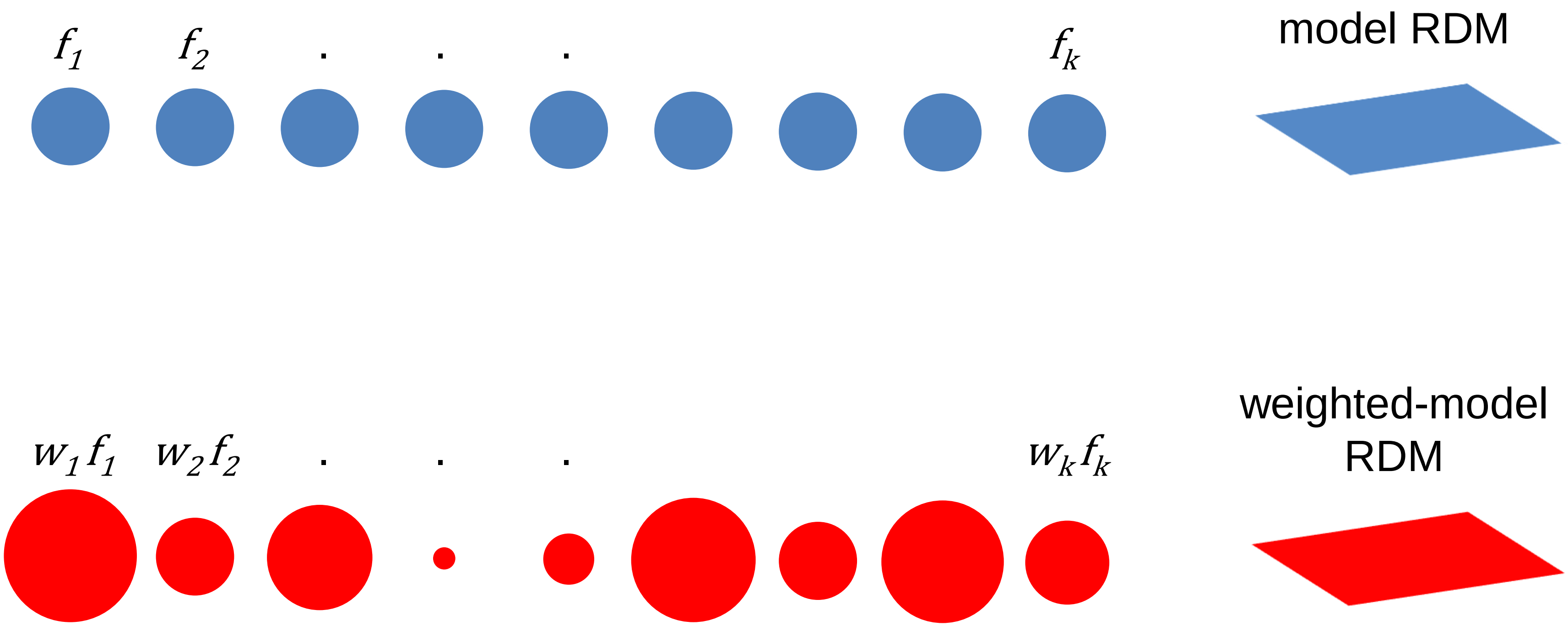
Kriegeskorte & **Diedrichsen** 2019

Representational similarity analysis



Kriegeskorte et al. 2008

Representational feature weighting with non-negative least-squares



Representational feature weighting with non-negative least-squares

predicted distance

weight k model feature k

$$\hat{d}_{i,j}^2 = \sum_{k=1}^{n} [w_k f_k(i) - w_k f_k(j)]^2$$

stimuli i, j

$$= \sum_{k=1}^{n} w_k^2 \cdot [f_k(i) - f_k(j)]^2$$

The squared distance RDM of weighted model features equals a weighted sum of single-feature RDMs.

$w_1{}^2 \cdot$ feature-1 RDM

$+ w_2{}^2 \cdot$ feature-2 RDM

⋮

$+ w_k{}^2 \cdot$ feature-k RDM

= weighted-model RDM

$$\mathbf{w} = \arg\min_{\mathbf{w}\in\mathbf{R}^{+n}} \sum_{i\neq j} [d_{i,j}^2 - \hat{d}_{i,j}^2]^2 = \arg\min_{\mathbf{w}\in\mathbf{R}^{+n}} \sum_{i\neq j} \left[d^2 - \sum_{k=1}^{n} w_k{}^2 \cdot \mathrm{RDM}_k \right]_{i,j}^2$$

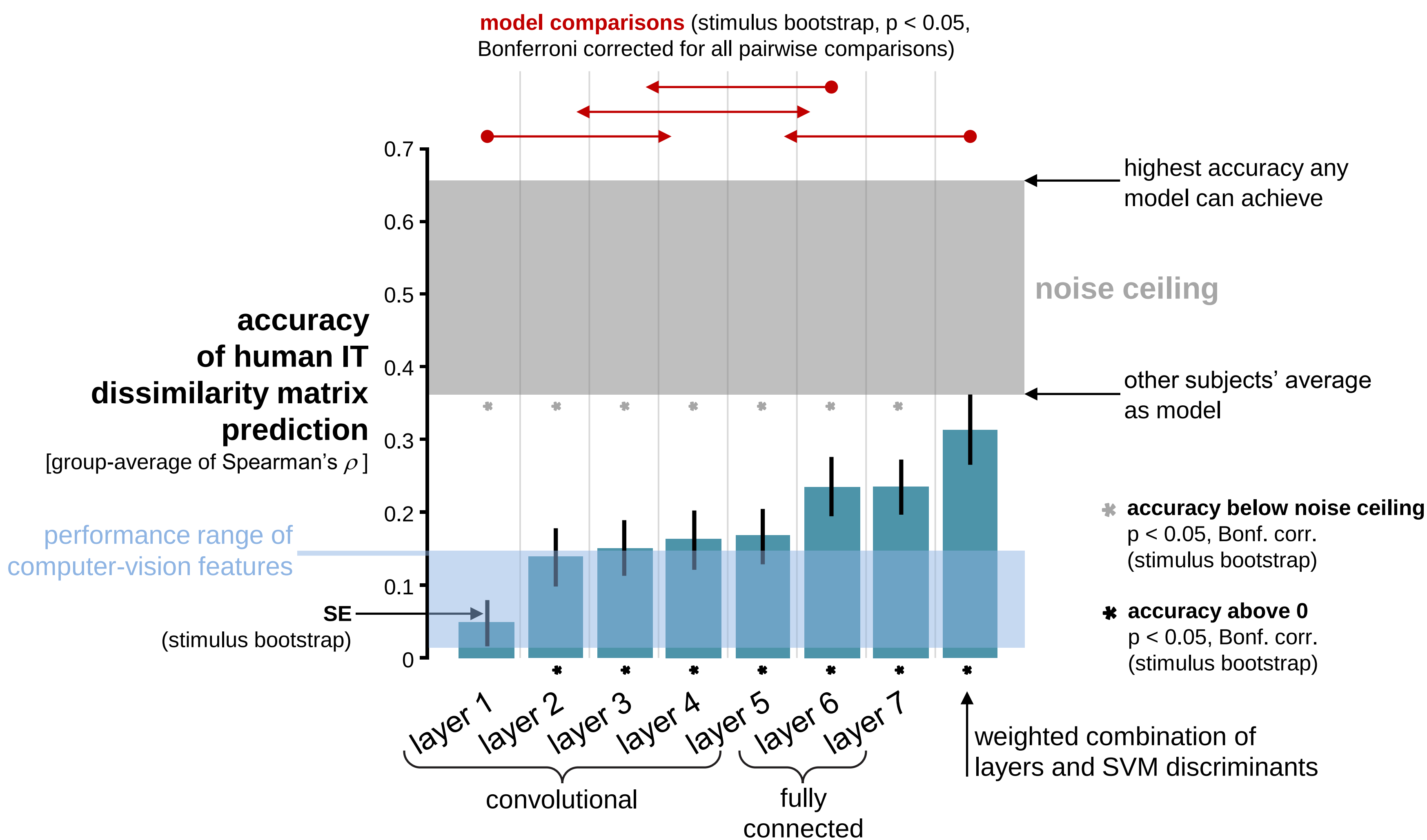
w_k weight given to model feature k

$f_k(i)$ model feature k for stimulus i

$d_{i,j}$ distance between stimuli i,j

w is the weight vector $[w_1\ w_2\ \ldots\ w_k]$ that minimizes the sum of squared errors

Deep convolutional networks predict IT representational geometry

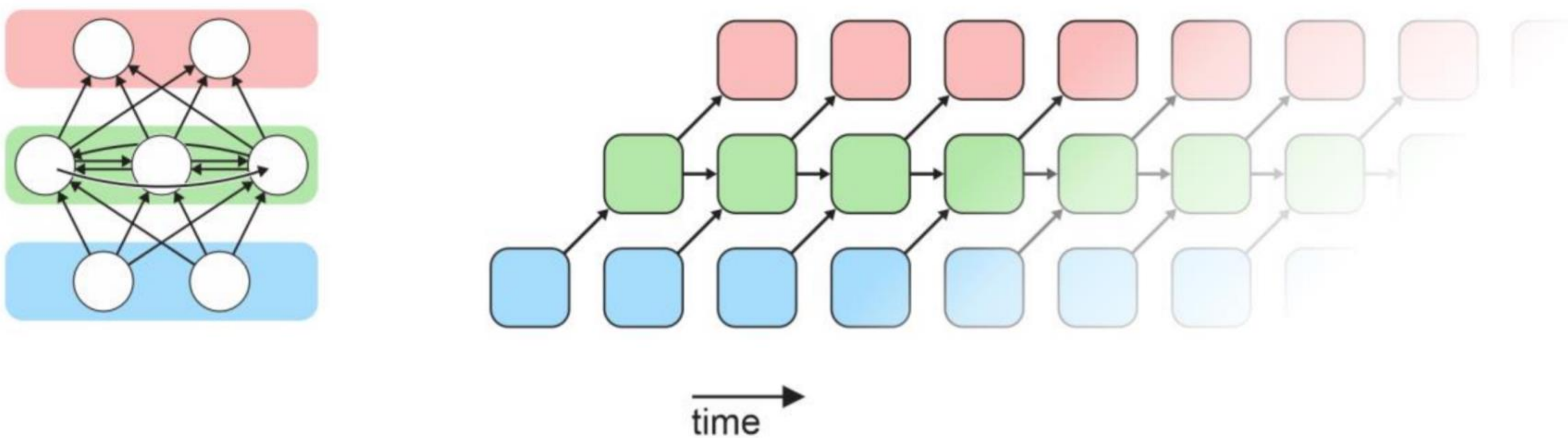


Khaligh-Razavi & **Kriegeskorte** 2014, **Nili** et al. 2014 (RSA Toolbox), **Storrs** et al. (in prep.)

Do *recurrent* neural networks provide better models of vision?

Courtney Spoerer

Recurrent networks can recycle their limited computational resources over time.

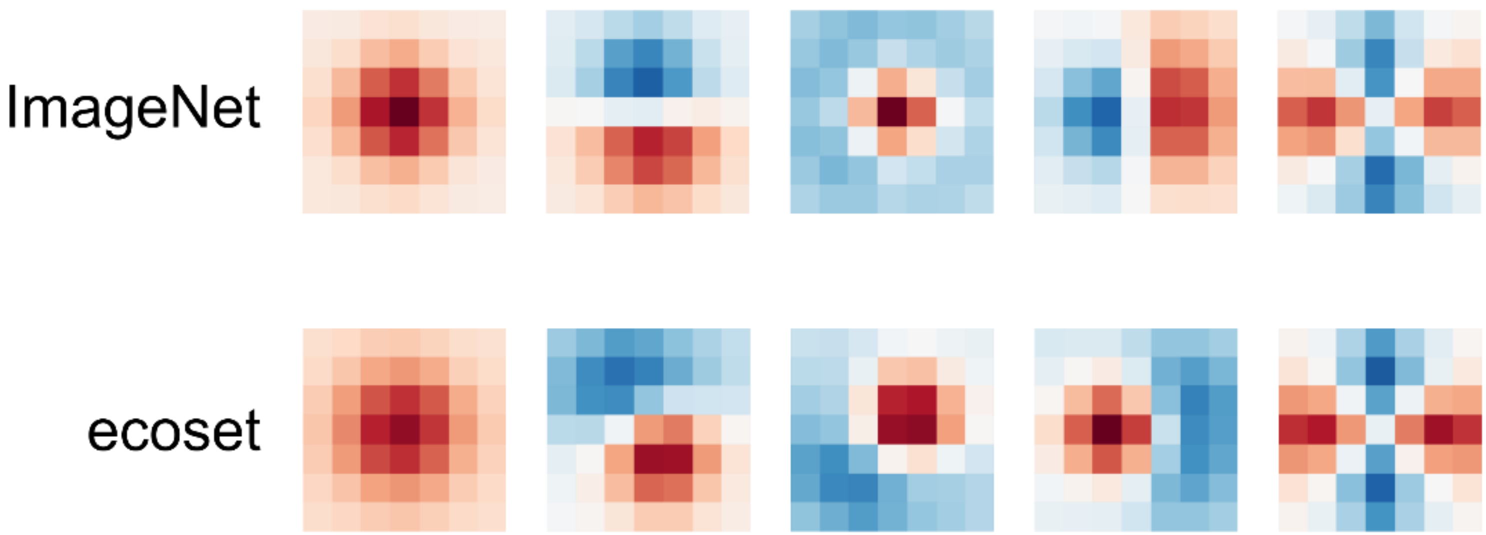


This might boost the performance of a physically finite model or brain.

Kriegeskorte & Golan 2019

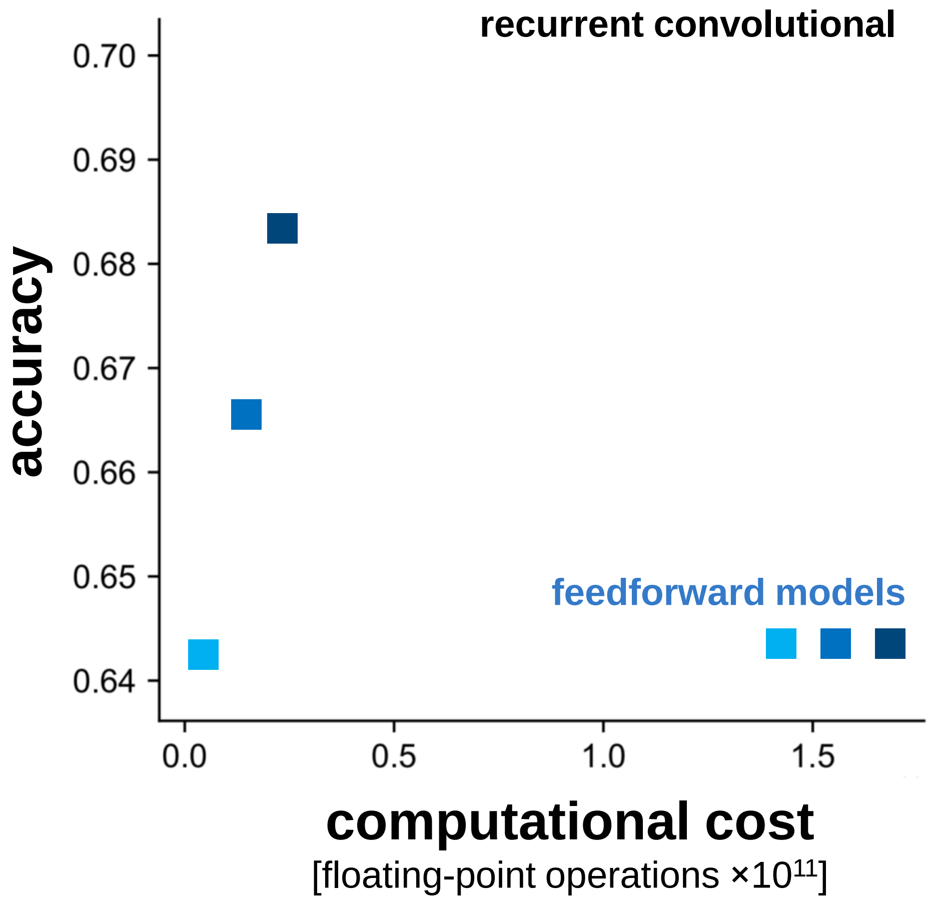
Layer 1 lateral connectivity is consistent with primate V1 connectivity

RCNN, layer 1, lateral connectivity templates
(first 5 principal components)



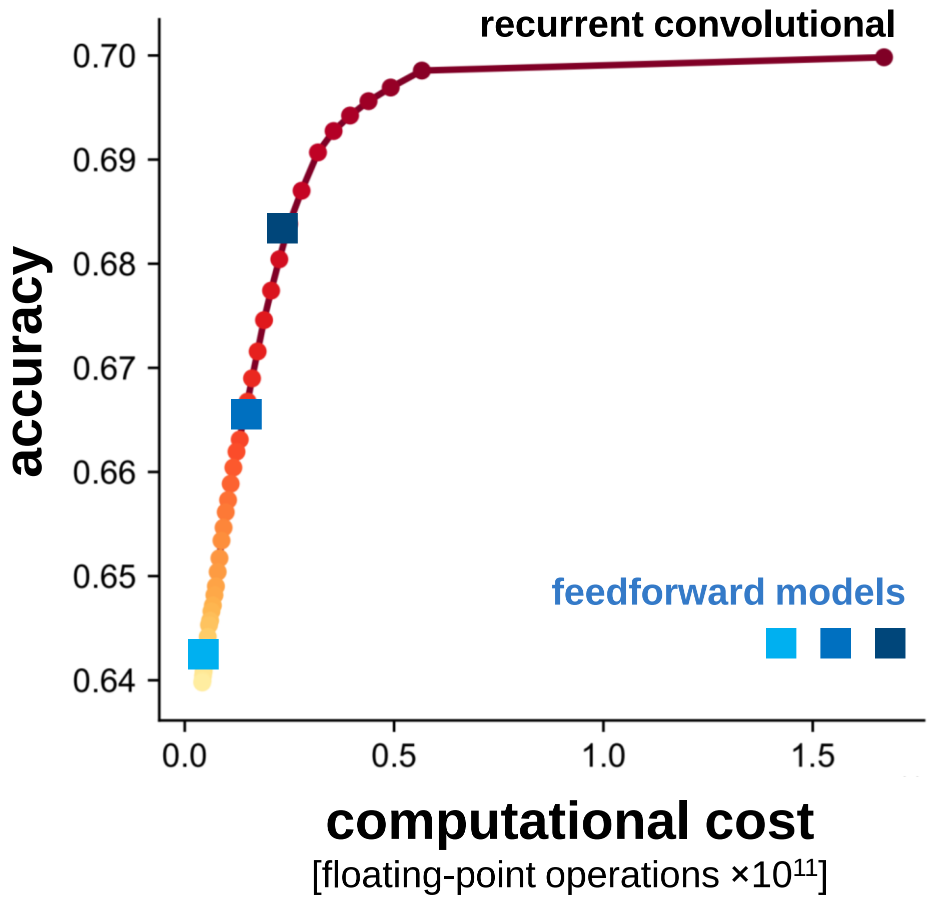
Spoerer et al. pp2019

Recurrent models can trade off speed of computation for accuracy



Spoerer et al. pp2019

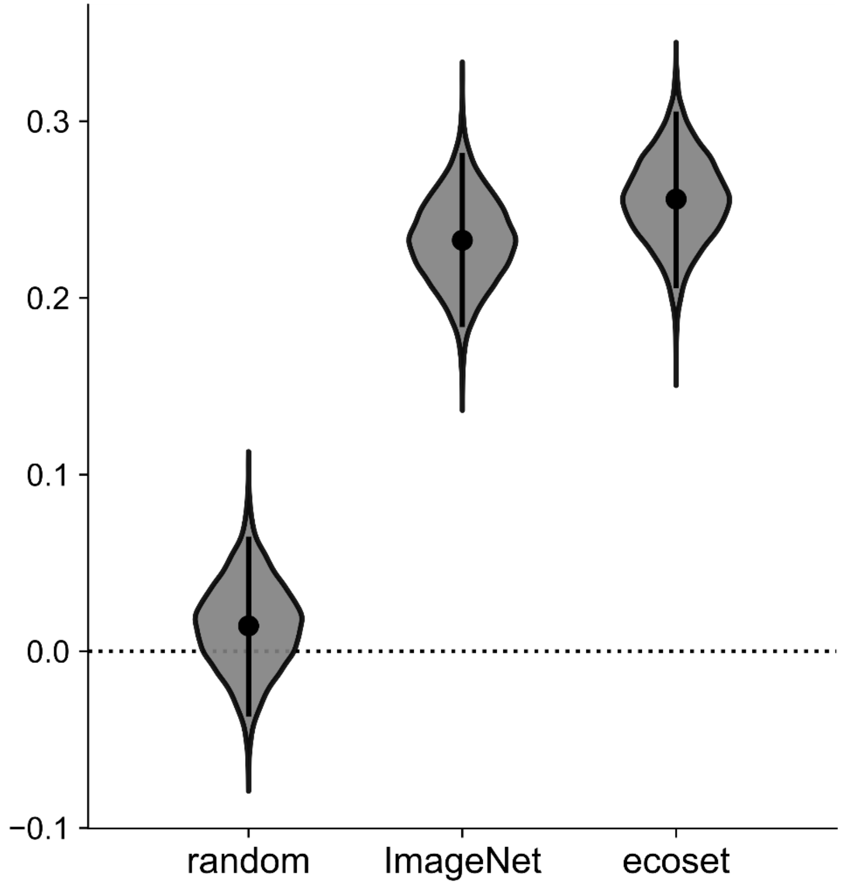
Recurrent models can trade off speed of computation for accuracy

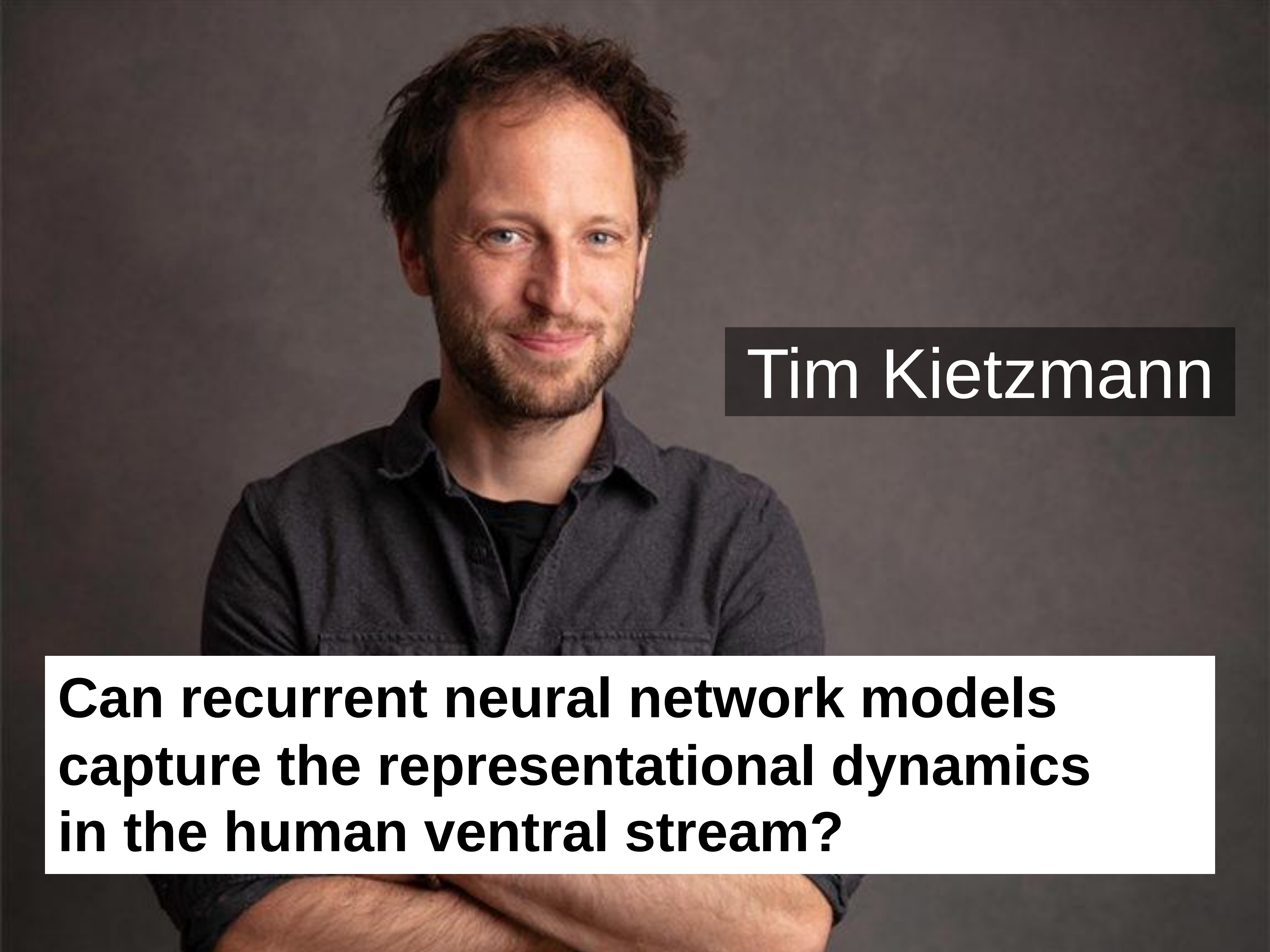


Spoerer et al. pp2019

RCNN reaction times tend to be slower for images humans are uncertain about

correlation between human certainty and RCNN reaction time [Spearman]



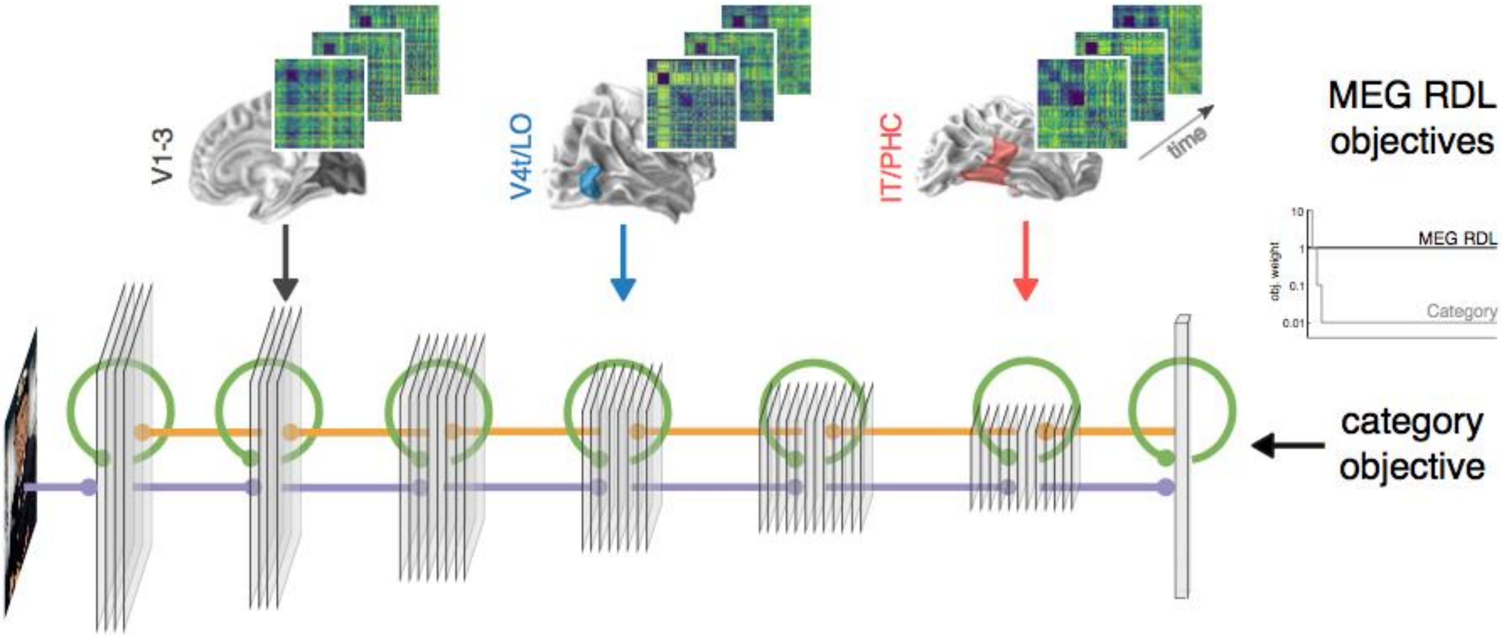


Tim Kietzmann

Can recurrent neural network models capture the representational dynamics in the human ventral stream?

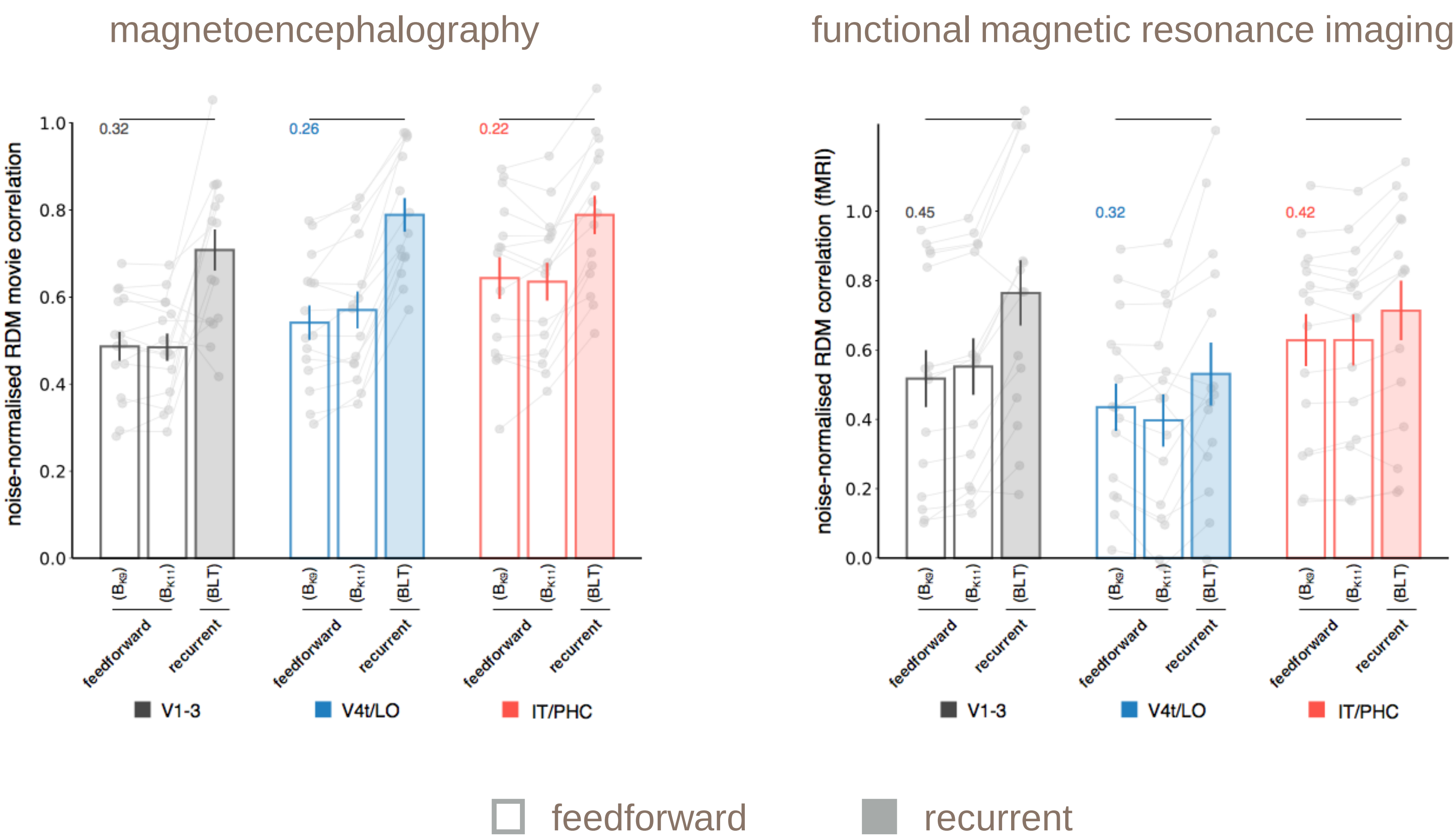
Fitting model representational dynamics
with *deep representational distance learning*

McClure & Kriegeskorte 2016



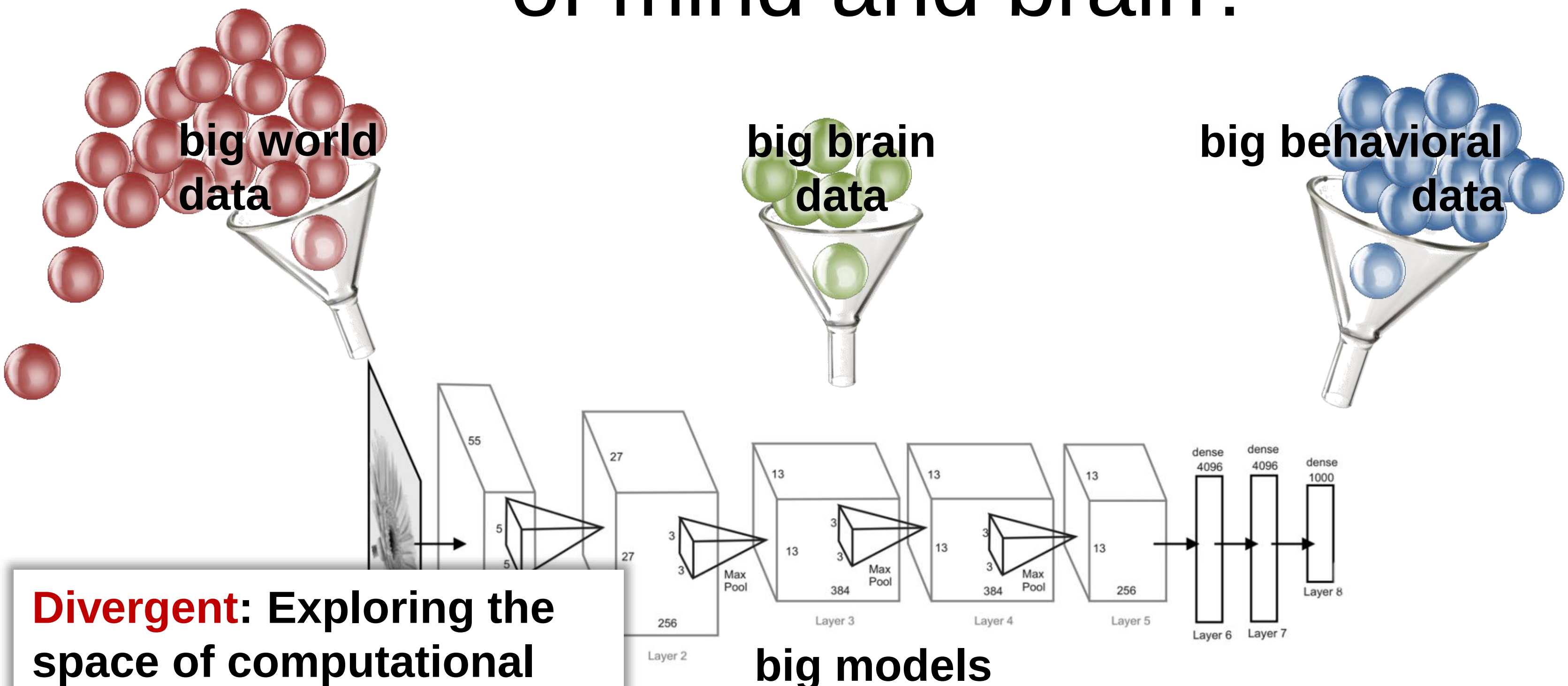
Task: find an image-computable network to model the first 300ms
of representational dynamics of the ventral stream.

Recurrent models better explain representations and their dynamics



Recurrent networks significantly outperform ramping feedforward models in predicting ventral-stream representations (MEG and fMRI).

How can we build neural network models of mind and brain?



Divergent: Exploring the space of computational models with world data

- **Training**
  - different sets of stimuli
  - different tasks
- **Units**
  - stochasticity
  - context-modulation
- **Architecture**
  - skipping connections
  - recurrent connections

Convergent: Constraining models with brain and behavioral data

- **inferential model selection** (model parameters learned for a task)
- **reweighting** of units
- **linear remixing** of units
- **deep learning of model parameters** from brain-activity data