Representational Similarity Analysis

2019/08/17

Choong-Wan Woo

Computational Cognitive Affective Neuroscience Lab (Cocoan lab)
Center for Neuroscience Imaging Research (CNIR), IBS
Dept. of Biomedical Engineering, SKKU

Key references

frontiers in SYSTEMS NEUROSCIENCE
ORIGINAL RESEARCH ARTICLE
published: 24 November 2008
doi: 10.3389/neuro.06.004.2008

Representational similarity analysis – connecting the branches of systems neuroscience

Nikolaus Kriegeskorte[1,]*, Marieke Mur[1,2] and Peter Bandettini[1]

[1] *Section on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA*
[2] *Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands*

- the original RSA paper
- Cited by 1293

OPEN ACCESS Freely available online
PLOS COMPUTATIONAL BIOLOGY

A Toolbox for Representational Similarity Analysis

Hamed Nili[1]*, Cai Wingfield[2], Alexander Walther[1], Li Su[1,3], William Marslen-Wilson[3], Nikolaus Kriegeskorte[1]*

1 MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom, **2** Department of Computer Science, University of Bath, Bath, United Kingdom, **3** Department of Experimental Psychology, University of Cambridge, Cambridge, United Kingdom

Decoding Neural Representational Spaces Using Multivariate Pattern Analysis

James V. Haxby,[1,2] Andrew C. Connolly,[1] and J. Swaroop Guntupalli[1]

[1]Department of Psychological and Brain Sciences, Center for Cognitive Neuroscience, Dartmouth College, Hanover, New Hampshire 03755;
email: james.v.haxby@dartmouth.edu, andrew.c.connolly@dartmouth.edu, swaroopgj@gmail.com

[2]Center for Mind/Brain Sciences (CIMeC), University of Trento, Rovereto, Trentino 38068, Italy

Annu. Rev. Neurosci. 2014. 37:435–56

Representational similarity analysis

Let’s start with the following three questions:

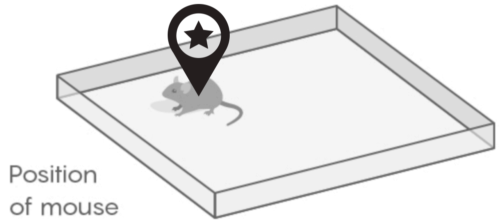
What is…

Representation?

Representational space?

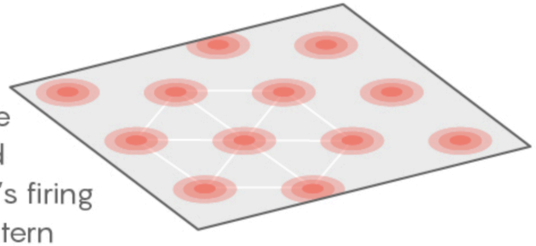
Representational geometry?

Representation?



A spatial location can be represented
in very different forms across different systems

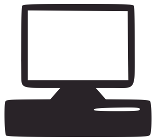
Mouse entorhinal cortex



One
grid
cell's firing
pattern

Maps

36°41'37.6"N 126°39'21.6"E
(longitude and latitude)



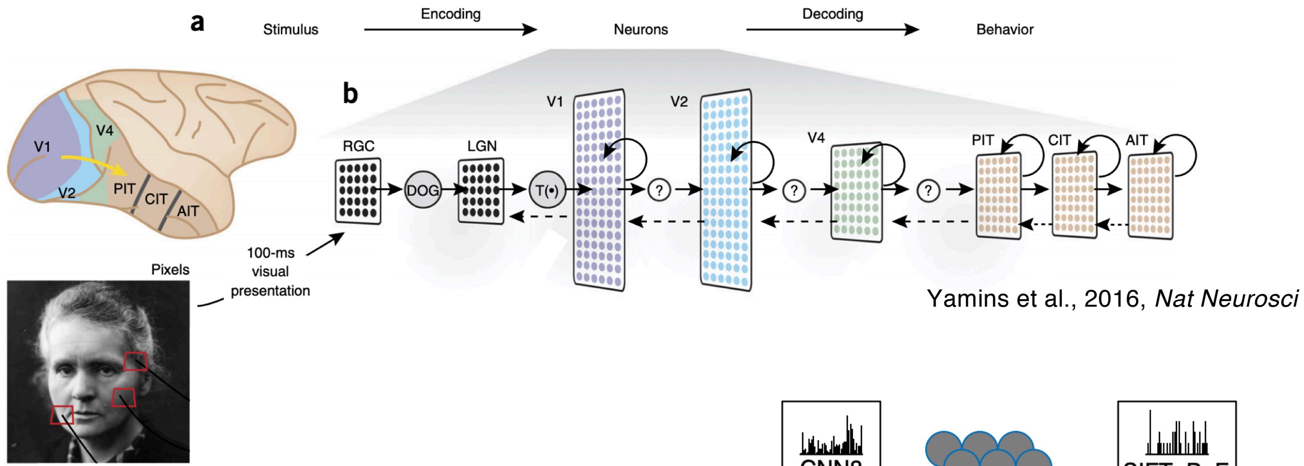
Computer

Binary code

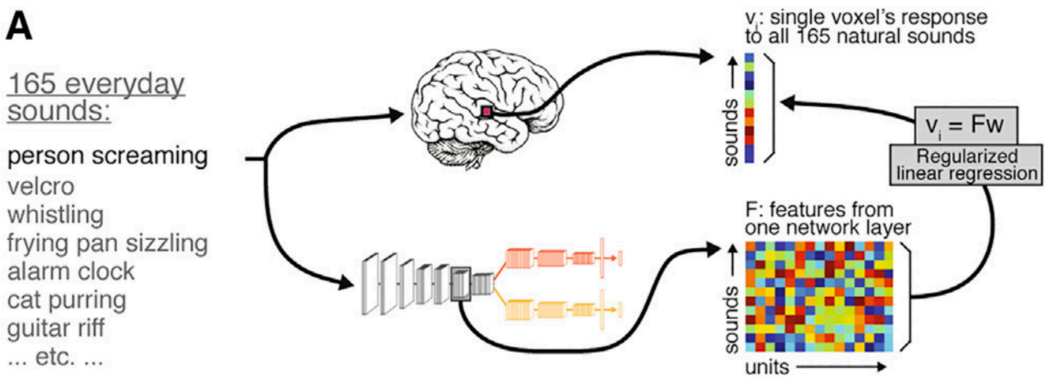
Image sources:
https://www.quantamagazine.org/the-brain-maps-out-ideas-and-memories-like-spaces-20190114/
Behrens et al., 2018, *Neuron*

Representation

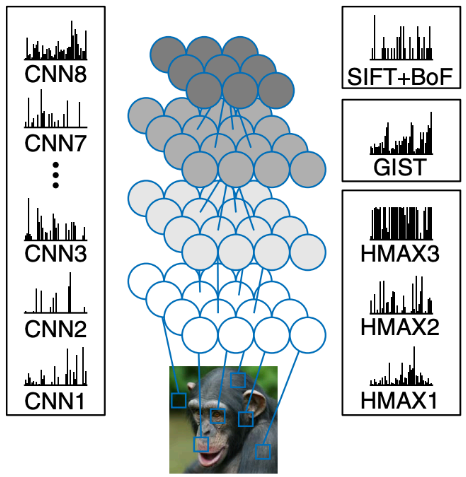
David Marr:
"A representation is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this."



Yamins et al., 2016, *Nat Neurosci*



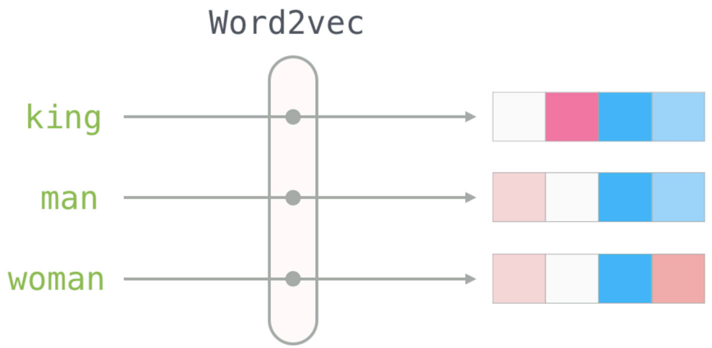
Kell et al., 2018, *Neuron*

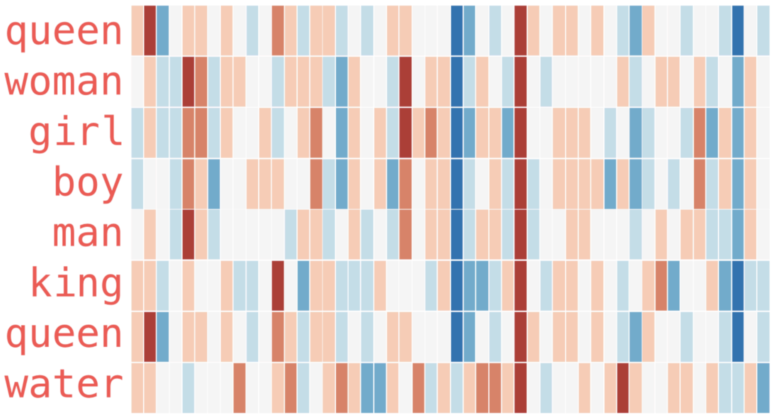


Horikawa et al., 2017, *Nat Comms*

Embedding = Vector representation

E.g.,

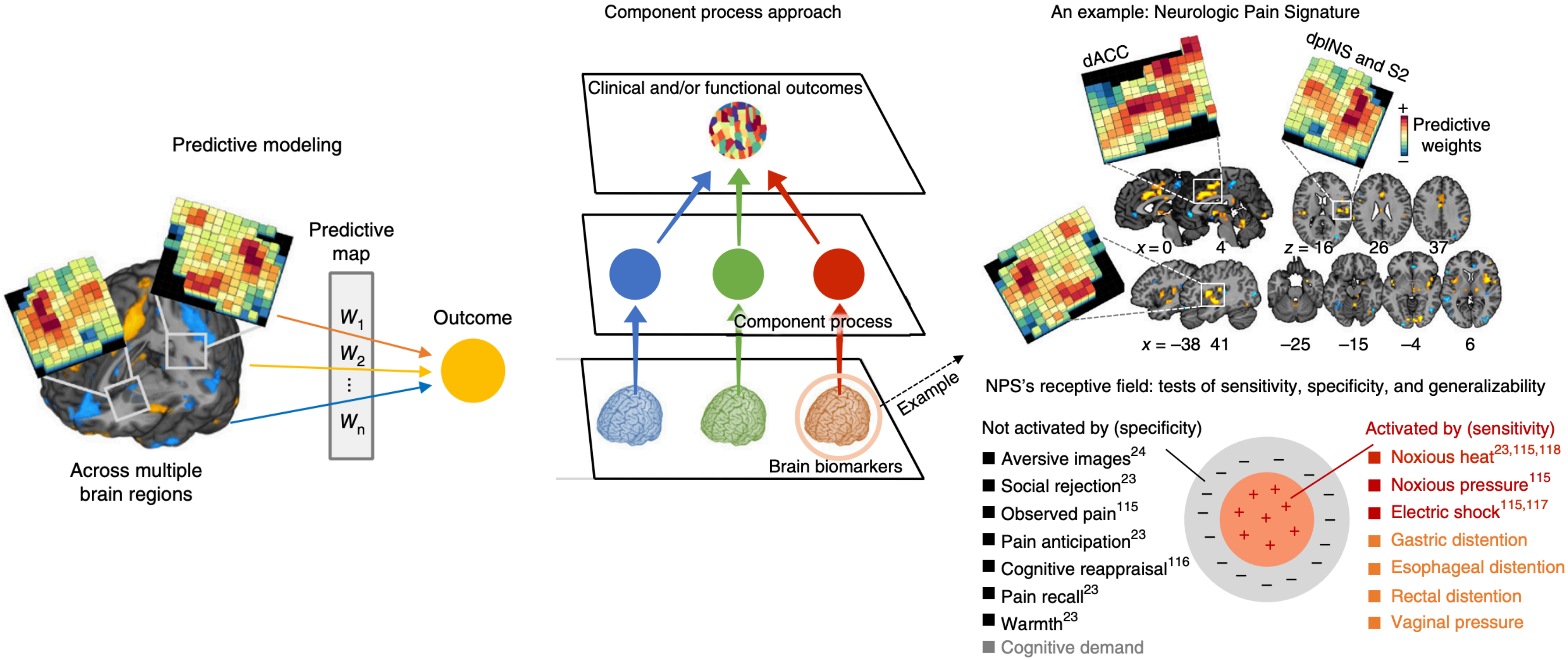




Only dimensionality reduction?
Could involve up-sampling (e.g., eye jitter)

Image sources:
http://jalammar.github.io/illustrated-word2vec/

Building brain biomarkers
= identifying good brain representations



Woo et al., 2017, *Nat Neurosci*; Wager et al., 2013, *NEJM*

Some challenges in studying representations

1. Difficult to identify (or develop) good representations
2. Difficult to use (or compare) representations from different systems (or models)

   e.g., different dimensions, different shapes, etc.

Some challenges in studying representations

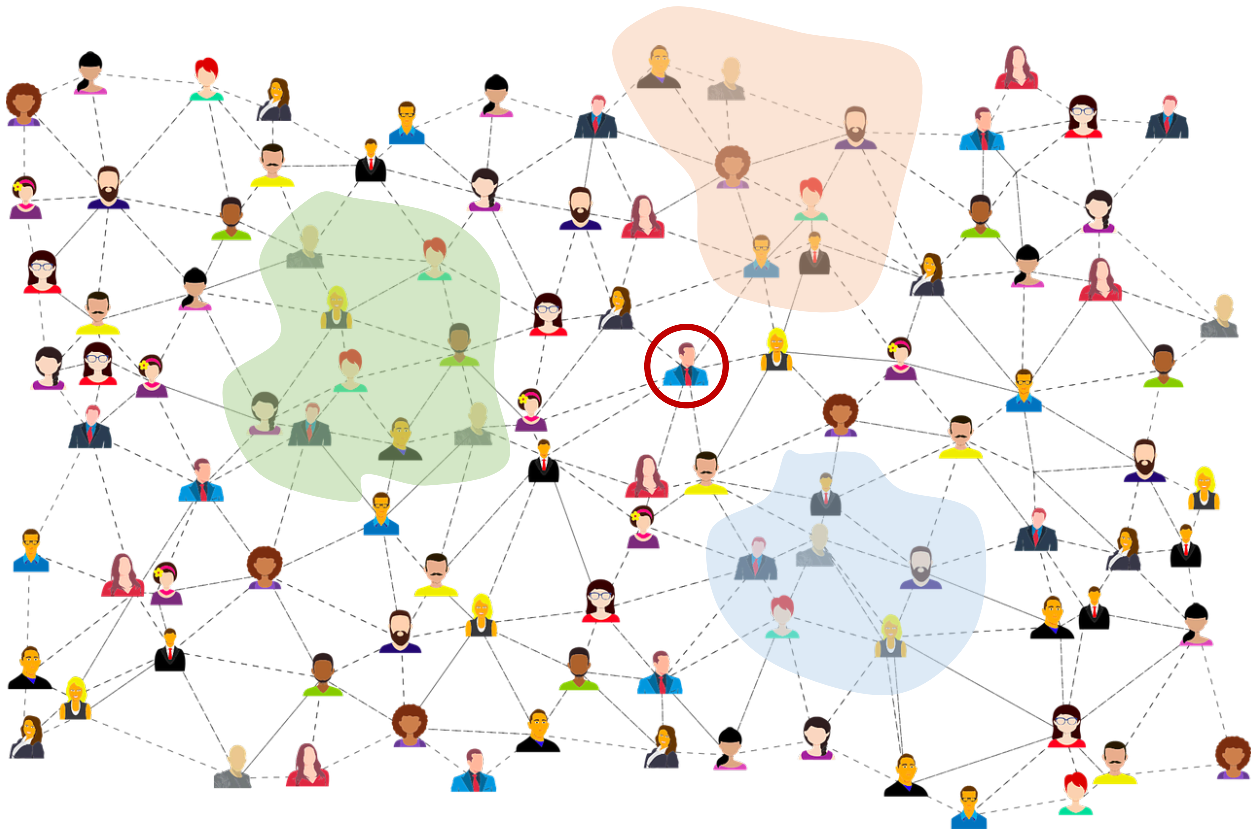
What is your representations of people around you?

What is your (hidden) vector representations of your friends, family members, colleagues, etc.??

Power? Fun? Appearance? Knowledge? Utility? 썸?

It is really difficult to know!

Some challenges in studying representations



An easy way around:
See how you (an internal model) categorize (or link) people!

Compare your categories (or links) with other people → you may find your soulmates!

Representational similarity analysis?

Examining the distances among people from my point of view

And comparing it with your distance for people

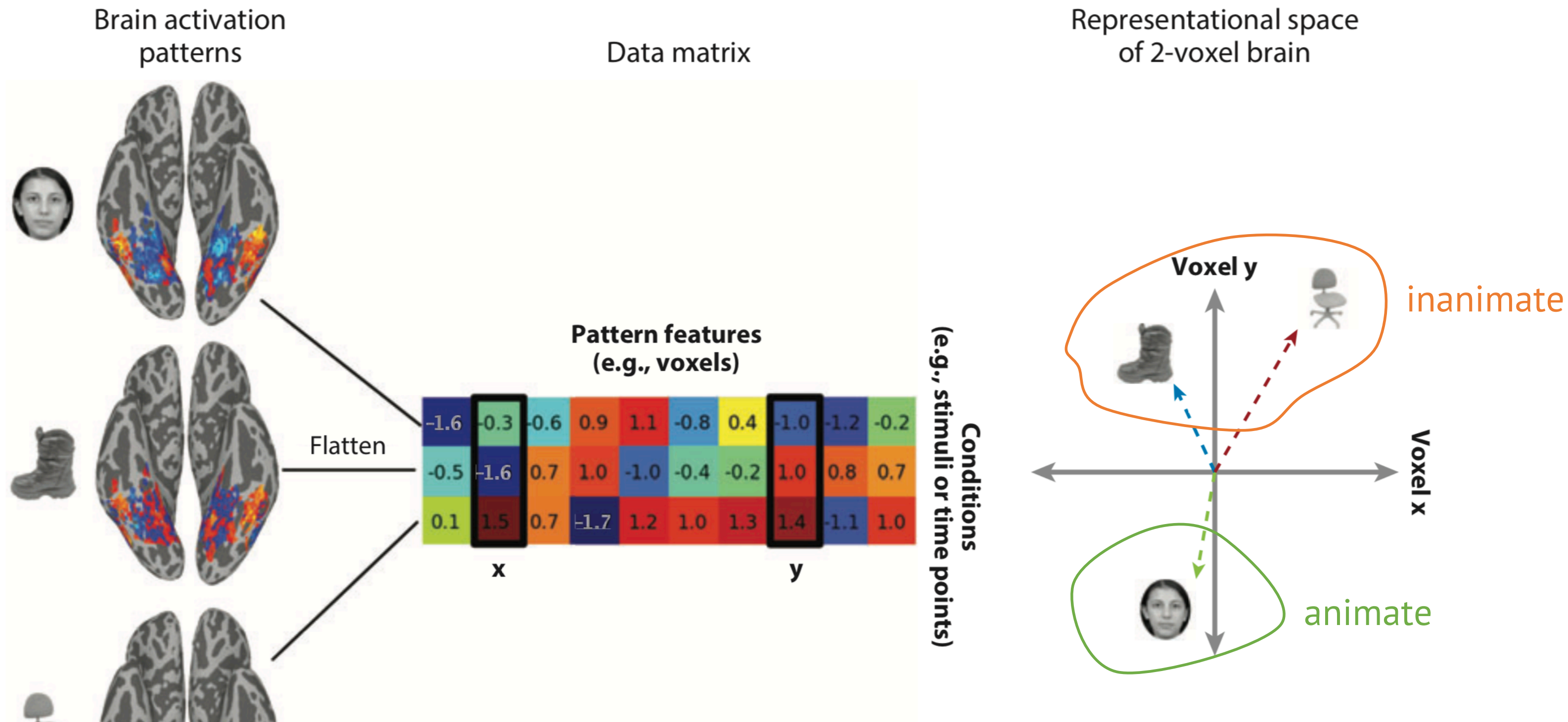
Representational space

REPRESENTATIONAL SPACE

Representational space is a high-dimensional space in which each neural response or stimulus is expressed as a vector with different values for each dimension. In a neural representational space, each pattern feature is a measure of local activity, such as a voxel or a single neuron. In a stimulus representational space, each feature is a stimulus attribute, such as a physical attribute or semantic label.

Haxby et al., 2014, *Annu Rev Neurosci*

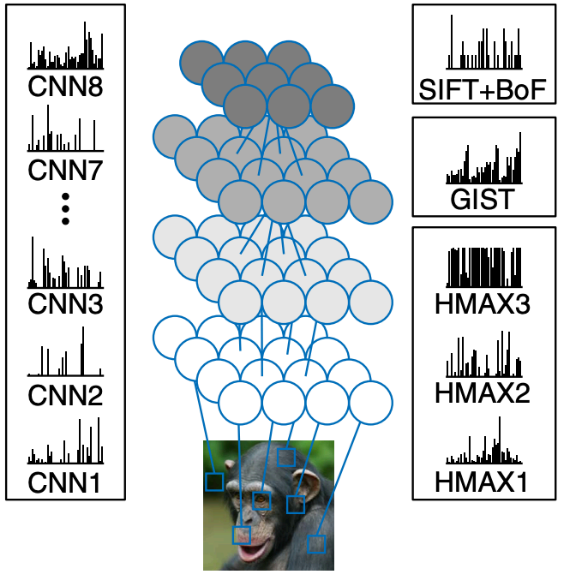
Representational space

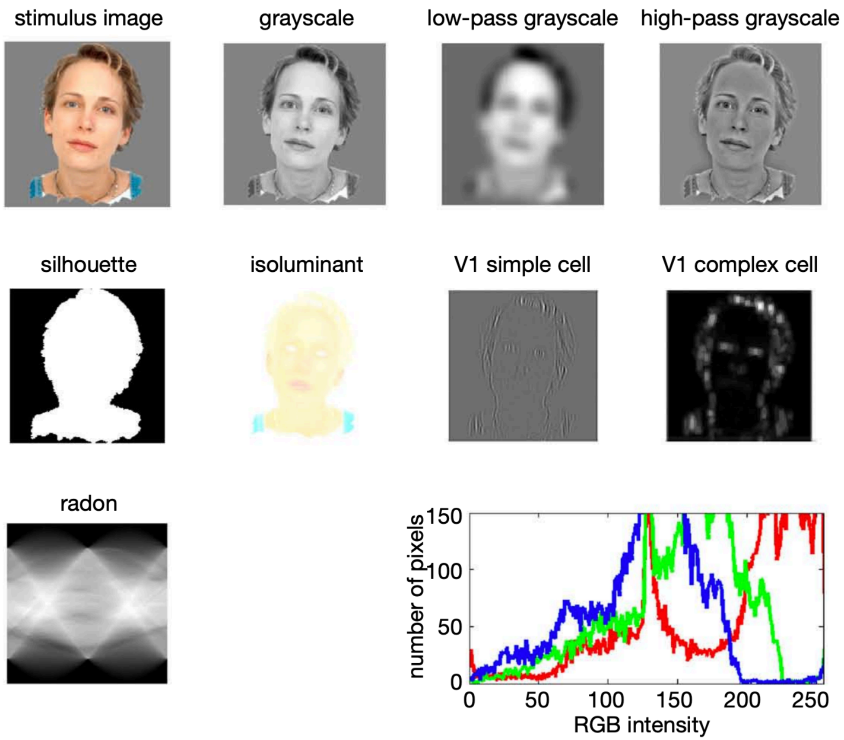


Any types of embedding can be used! (e.g., animate vs. inanimate?)

Haxby et al., 2014, *Annu Rev Neurosci*

Representational space



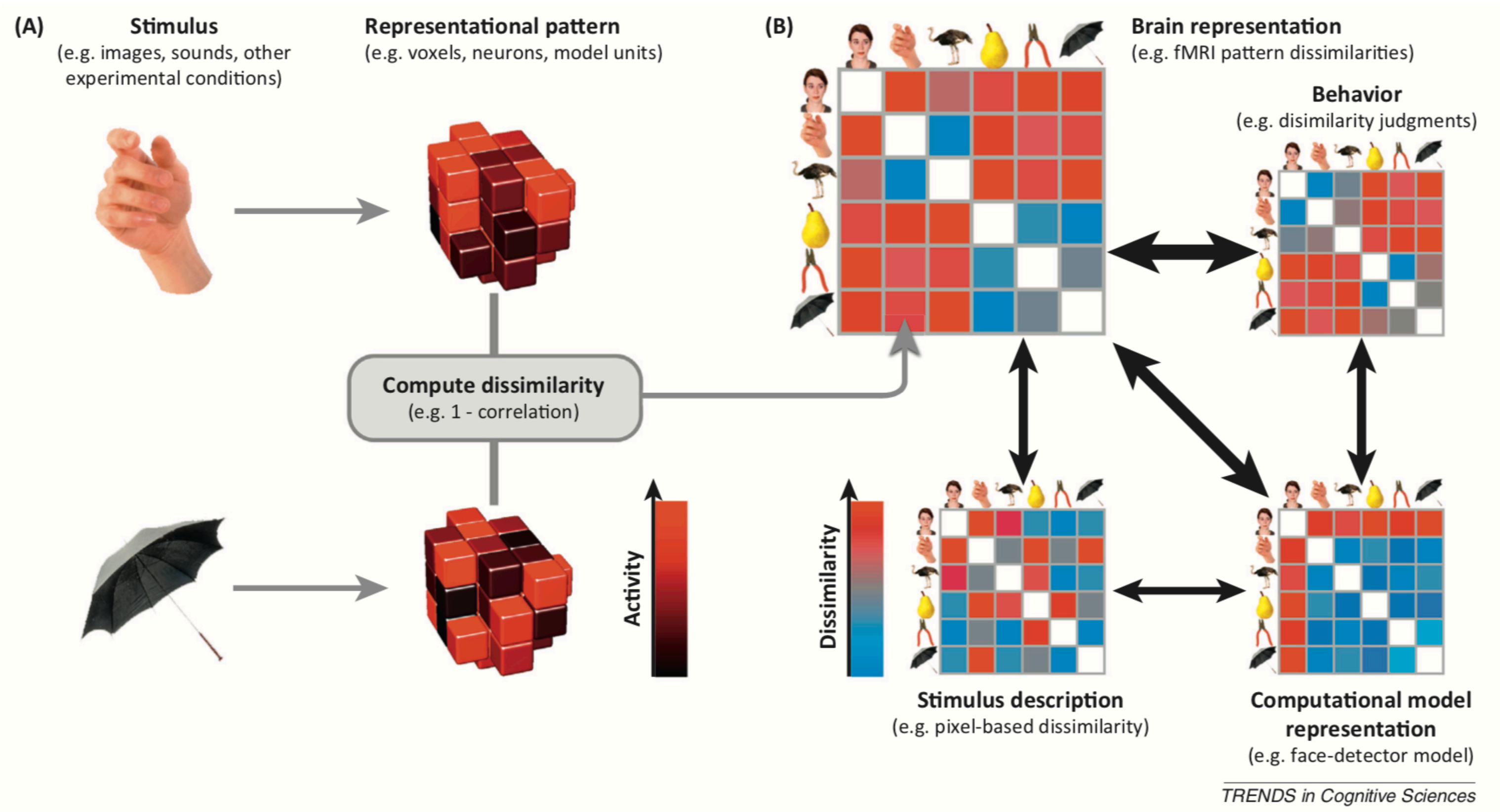


Any embeddings can be used!

Horikawa et al., 2017, *Nat Comms*

Kriegeskorte et al., 2008, *Front Sys Neurosci*

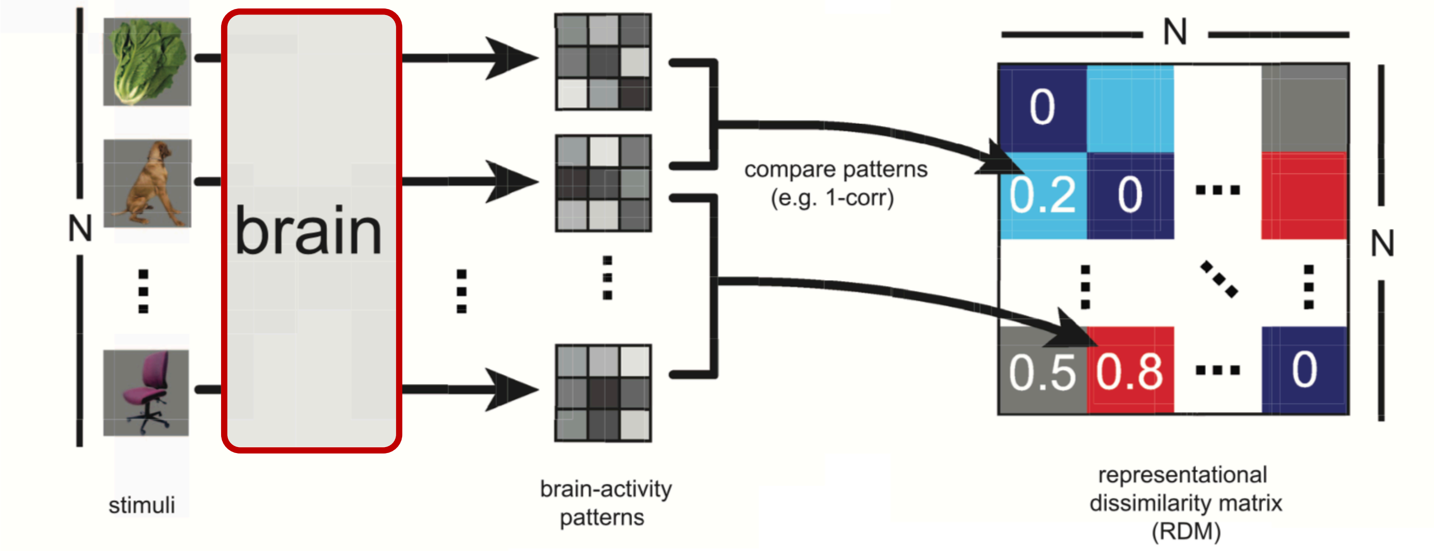
Representational dissimilarity matrix



Kriegeskorte et al., 2013, *TICS*

Representational dissimilarity matrix (in more detail)

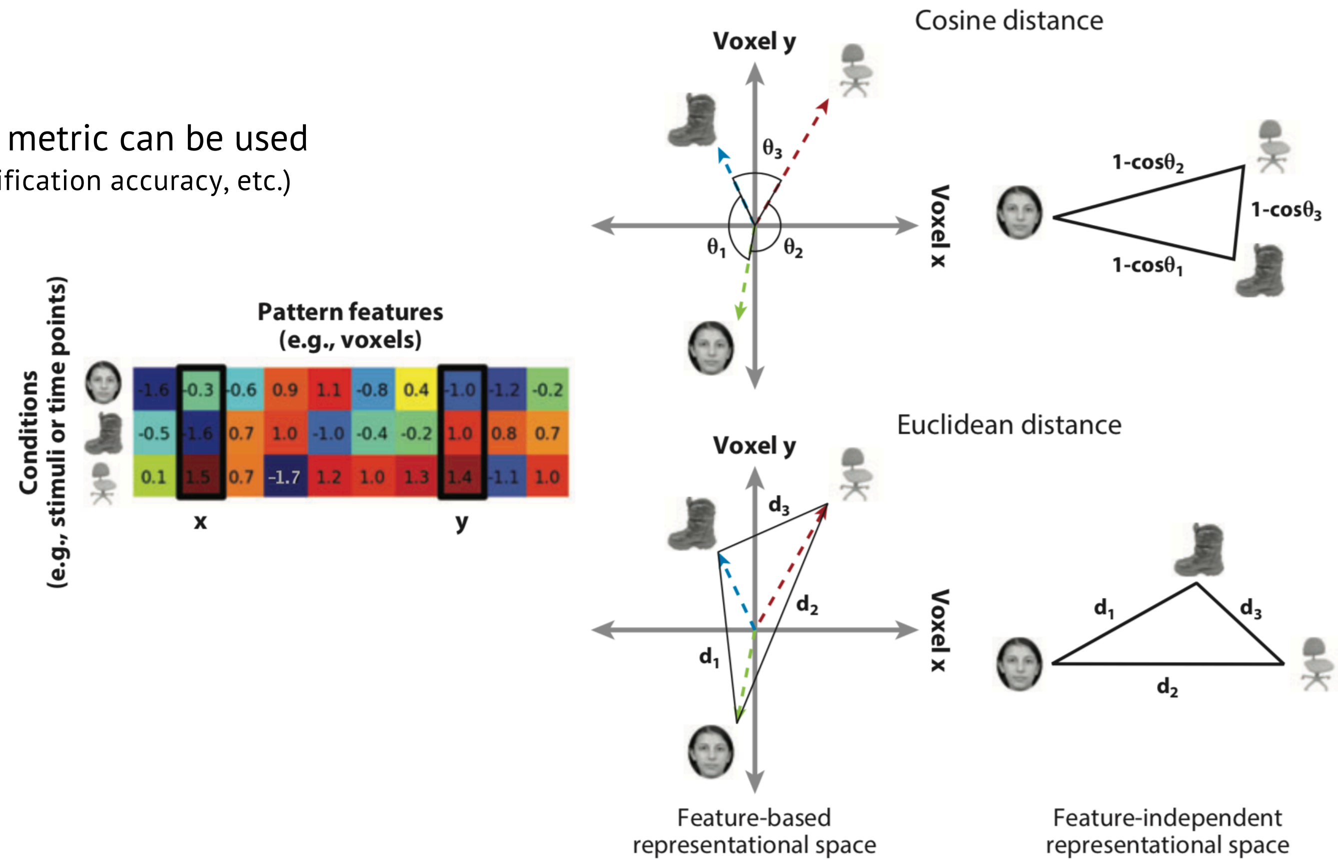
Brain or
Other model representations



Nili et al., 2014, *PLoS Comp Biol*

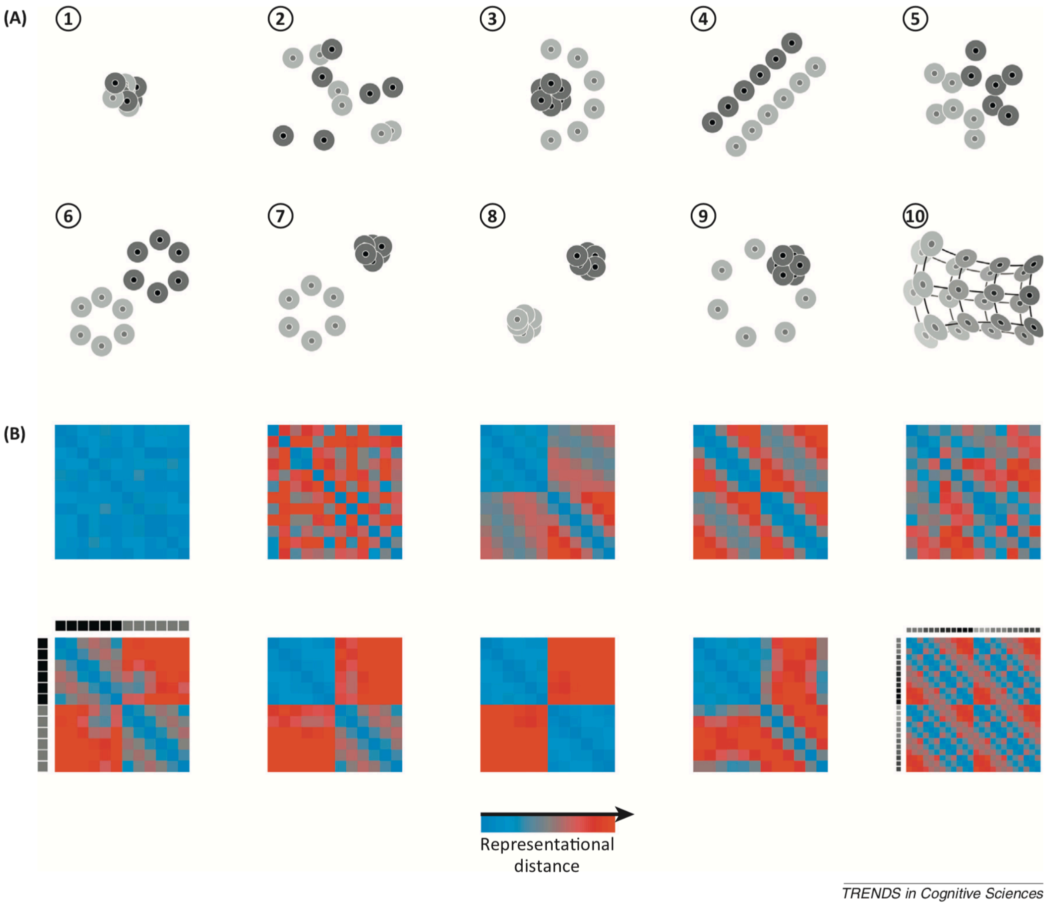
Representational dissimilarity matrix (in more detail)

Other types of distance metric can be used
(e.g., cosine, Euclidean, classification accuracy, etc.)



Nili et al., 2014, *PLoS Comp Biol*

Representational dissimilarity matrix (in more detail)

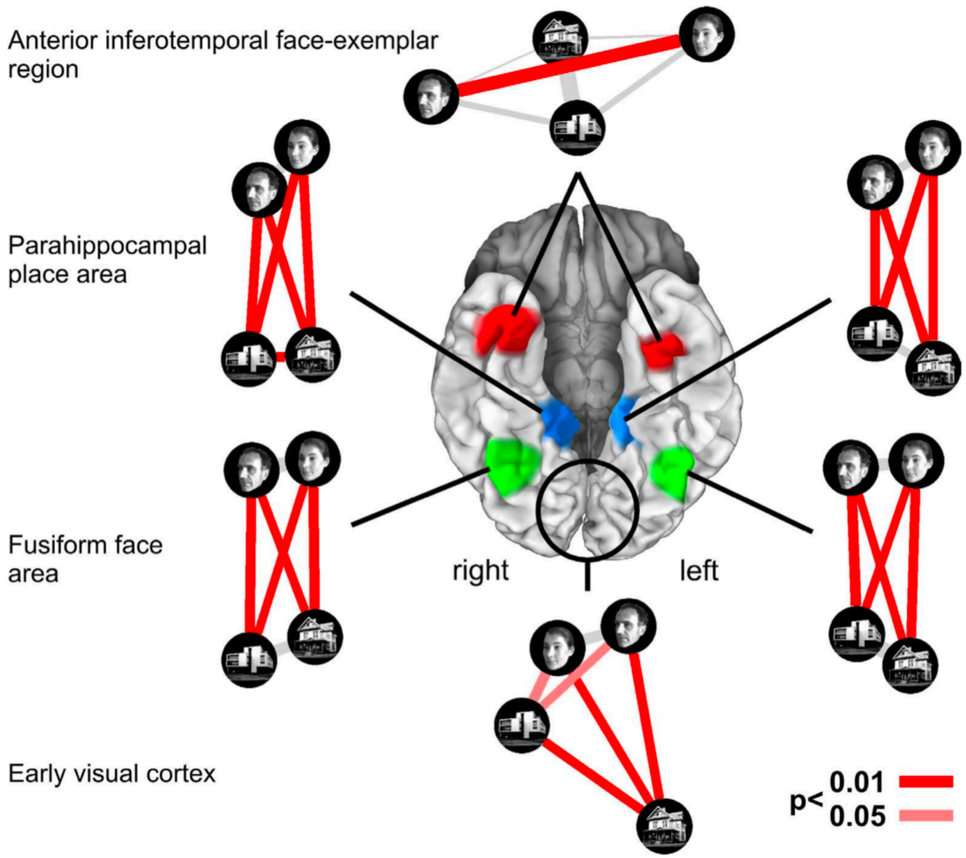


Representational geometry

Representational dissimilarity matrix

Kriegeskorte et al., 2013, *TICS*

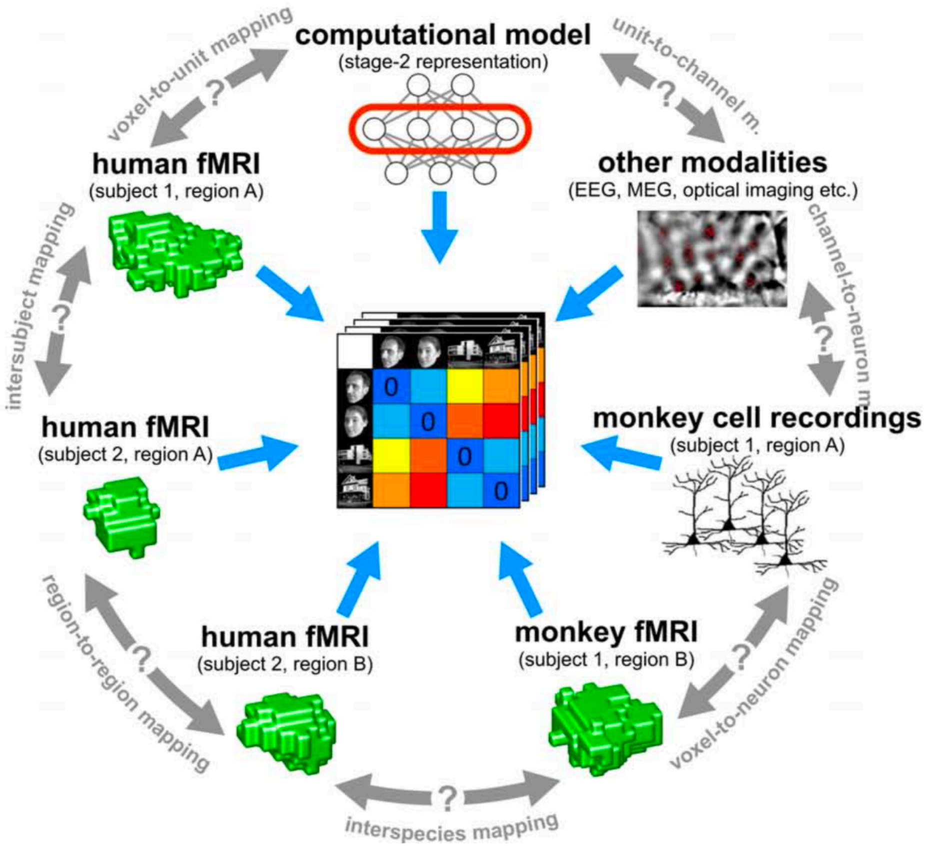
Representational dissimilarity matrix (in more detail)



Different representational geometry
From different representational dissimilarity patterns of different brain regions

Kriegeskorte et al., 2008, *Front Sys Neurosci*

Basic idea: You can compare any models/systems using RDM!



Kriegeskorte et al., 2008, *Front Sys Neurosci*

Three steps of representational similarity analysis

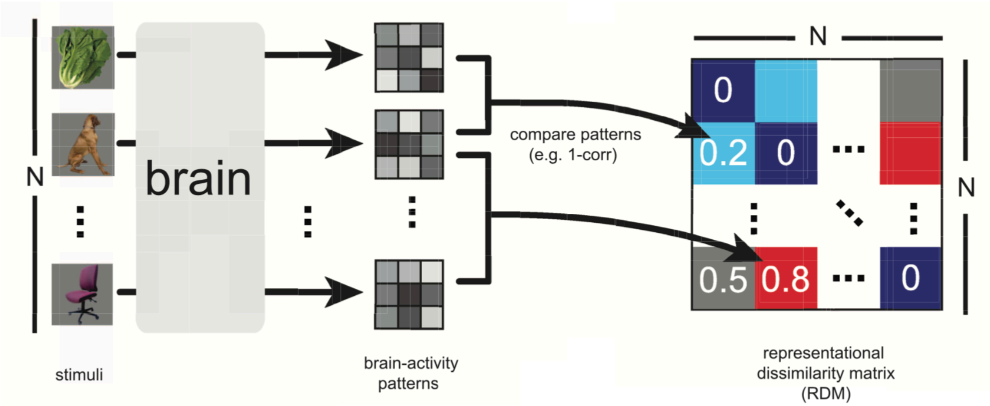
Step 1: Computing and visualizing RDMs

Step 2: Comparing brain and model RDMs

Step 3: Statistical inference

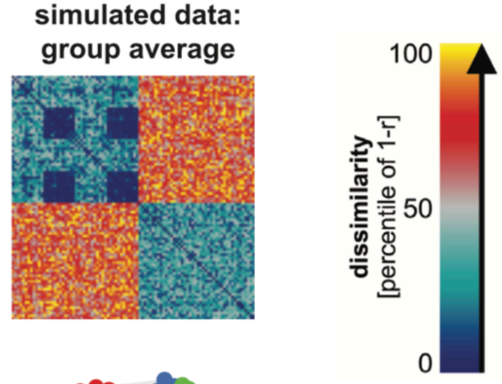
Nili et al., 2014, *PLoS Comp Biol*

Step 1: Computing and visualizing RDMs

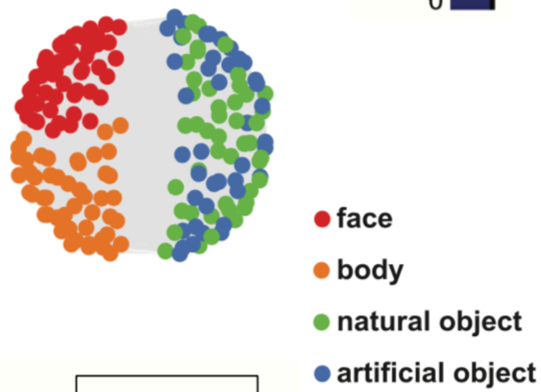


Nili et al., 2014, *PLoS Comp Biol*

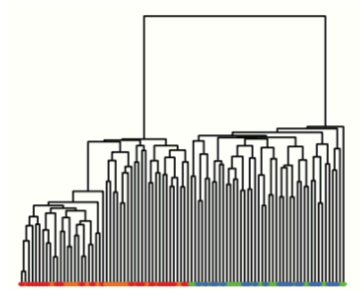
RDM



Dimensionality Reduction method for visualization (MDS, t-SNE, PCA, UMAP, etc.)



Clustering

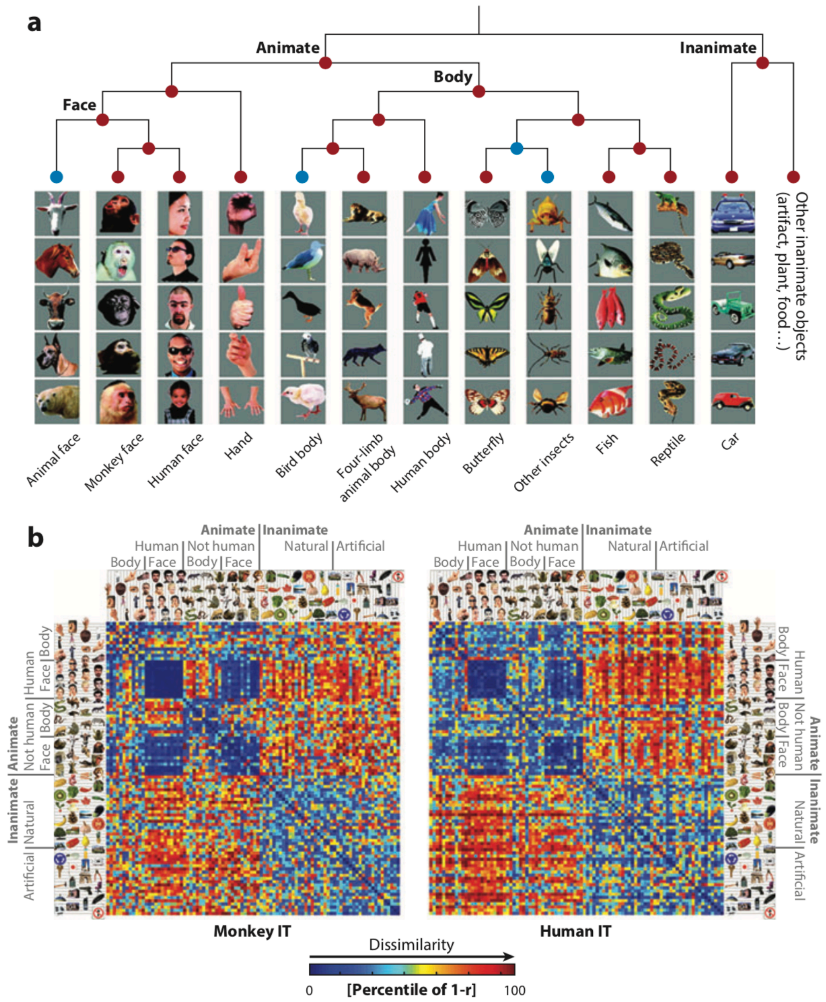


Step 1: Computing and visualizing RDMs

Real study examples:

Dendrogram derived from
multiple single-unit recording
in macaque inferior temporal (IT) cortex
(Kiani et al., 2007, *J. Neurophysiol.*)

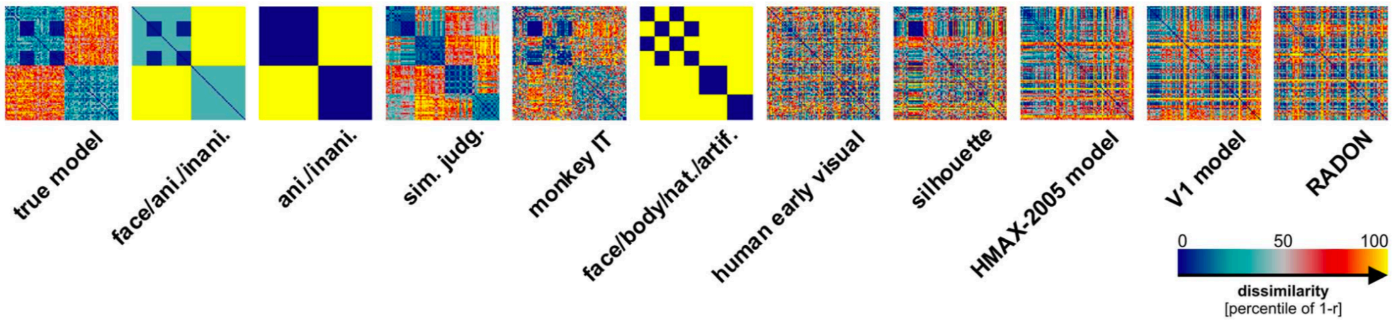
Cross-modal, cross-species comparisons of RDM
for a common set of stimuli
(Kriegeskorte et al., 2008, *Neuron*)



Nili et al., 2014, *PLoS Comp Biol*

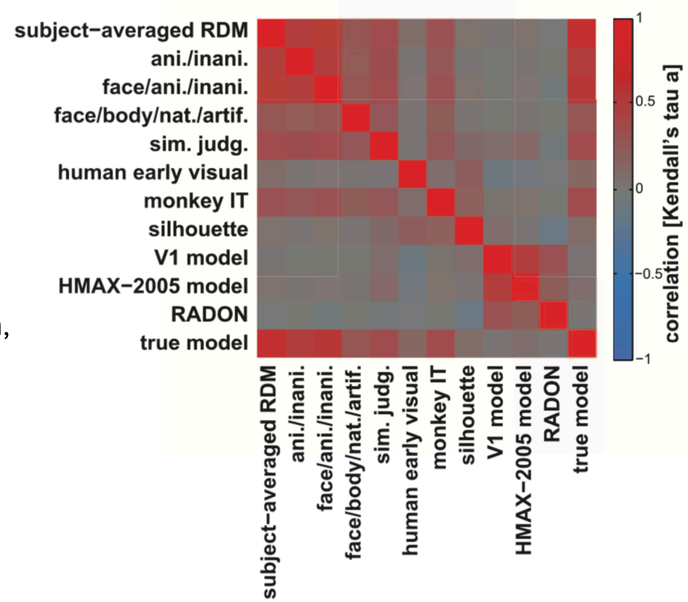
Step 2: Comparing brain and model RDMs

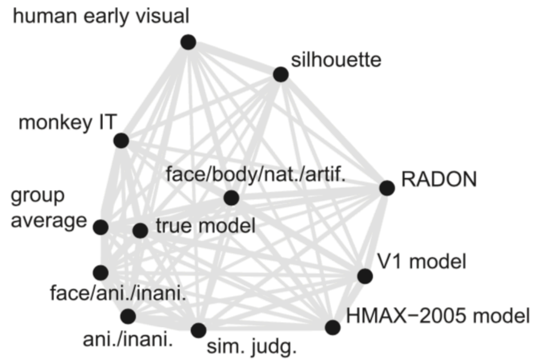
Many model RDMs



Comparing the RDMs: using any metric that quantifies the extent to which two matrices are "in agreement"

e.g., Spearman's rank correlation, Kendall's tau a, etc.





Again, dimensionality reduction for visualization (MDS, PCA, t-SNE, force-directed graph drawing, etc.)

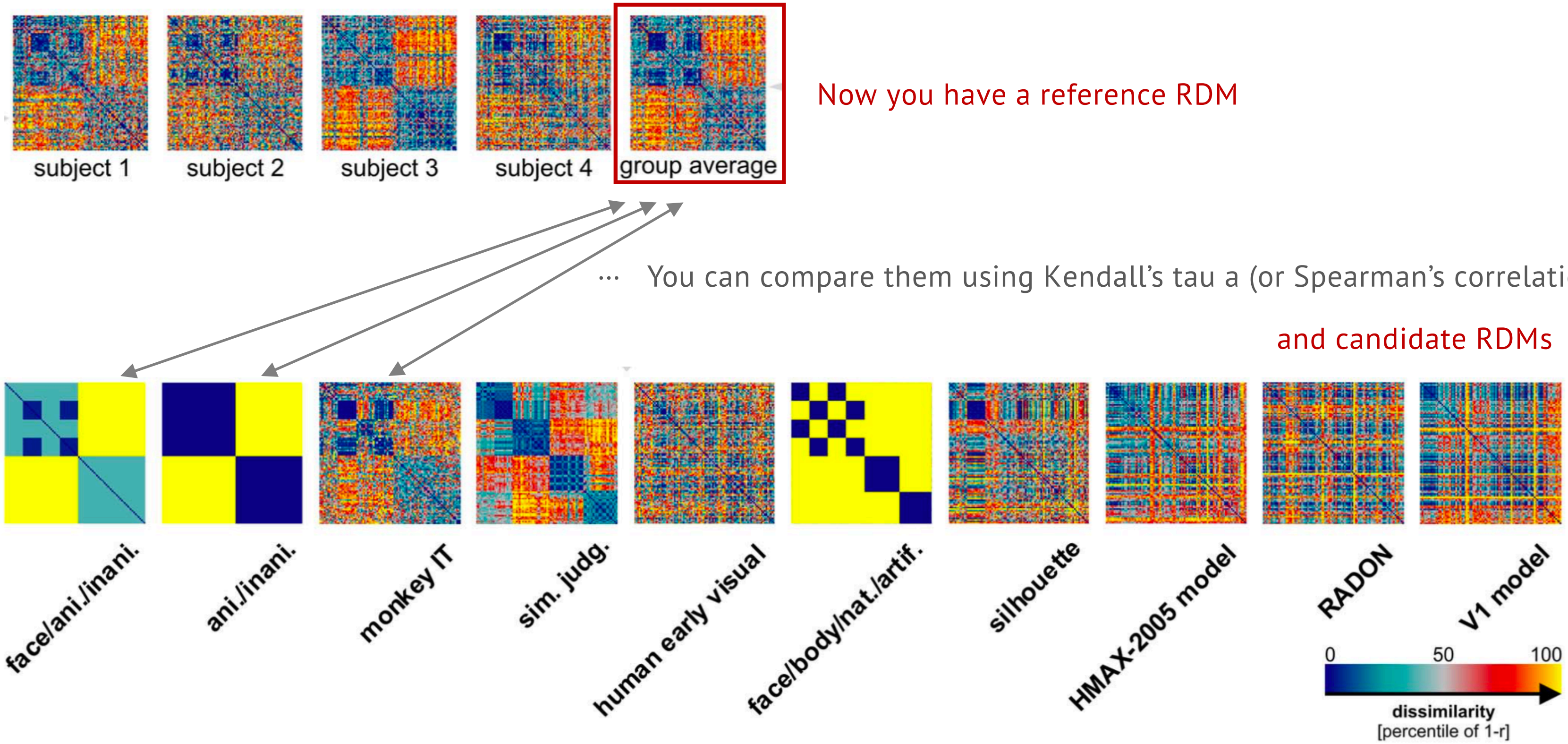
Nili et al., 2014, *PLoS Comp Biol*

Step 3: Statistical inference

Now you have a reference RDM

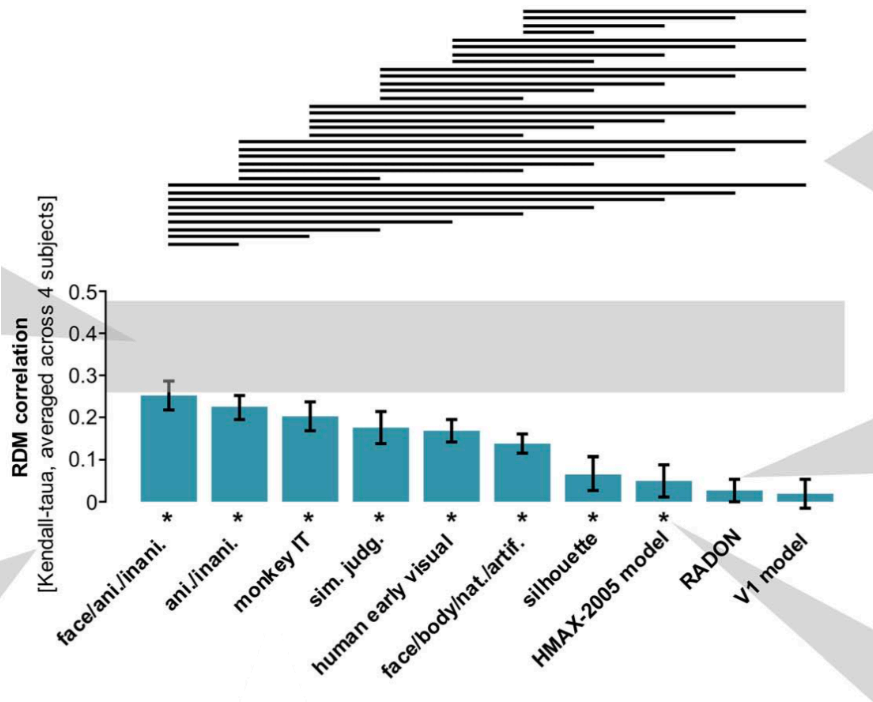
… You can compare them using Kendall's tau a (or Spearman's correlation)

and candidate RDMs



Nili et al., 2014, *PLoS Comp Biol*

Step 3: Statistical inference



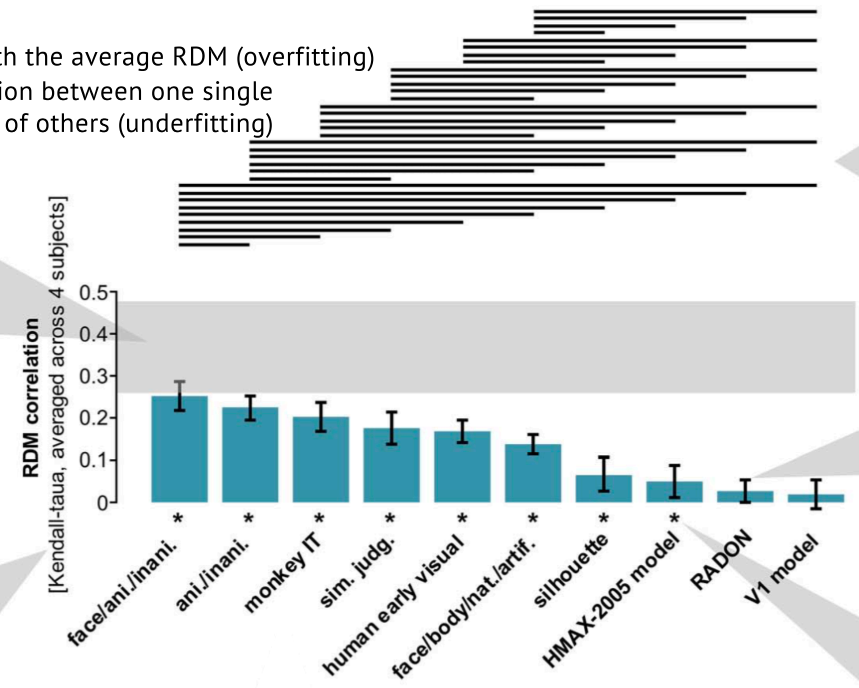
Performance is measured, as in Fig. 4, by Kendall's tau a between the reference RDM and the candidate RDMs.

Nili et al., 2014, *PLoS Comp Biol*

Step 3: Statistical inference

Noise ceiling:

- **Upper bound:** average correlation with the average RDM (overfitting)
- **Lower bound:** leave-one-out correlation between one single subject's RDM with the average RDM of others (underfitting)



The **noise ceiling** indicating the expected performance of the true model is much wider here than in the simulated data of Fig. 4. This reflects the fact that only 4 subjects entered this analysis, and the representation of human IT is thus much less precisely defined.

Performance is measured, as in Fig. 4, by Kendall's tau a between the reference RDM and the candidate RDMs.

Pairwise comparisons are based on bootstrap resampling of the stimulus set. This procedure simulates the variability of the estimates across random samples of stimuli (from an imaginary population of stimuli). Multiple testing is accounted for by controlling the expected FDR at 0.05.

Error bars indicate the standard error of the mean based on the bootstrap resampling of the stimulus set.

Each candidate RDM's relatedness to the reference RDM is now tested using a stimulus-label randomization test, which does not require multiple subjects. This time stars instead of p values were chosen to indicate significance. Multiple testing was accounted by controlling the expected FDR at 0.05.

Other analysis options (from Nili et al.):

1. **Representational connectivity:** Comparing brain RDMs among different regions
2. **Searchlight RSA:** Each region serves as a model
3. Use **classification performance** as a distance metric

Nili et al., 2014, *PLoS Comp Biol*

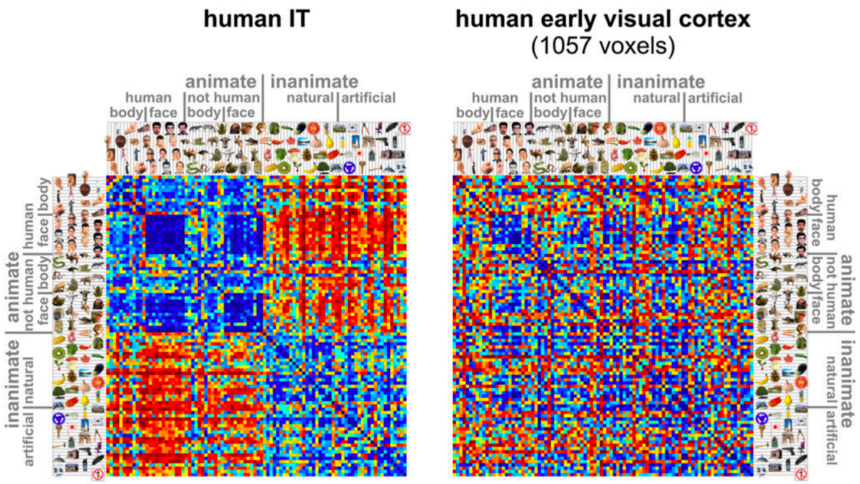
Other RSA options

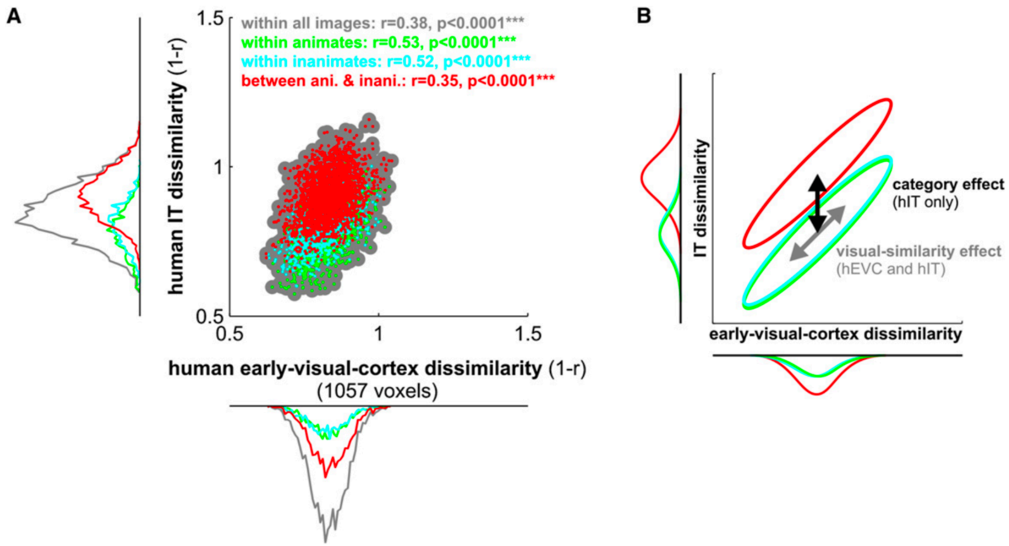
Other analysis options

1. Representational connectivity: Comparing brain RDMs among different regions
2. Use classification performance as a distance metric
3. Searchlight RSA: Each region serves as a model
4. Conducting RSA in the GLM context

Nili et al., 2014, *PLoS Comp Biol*

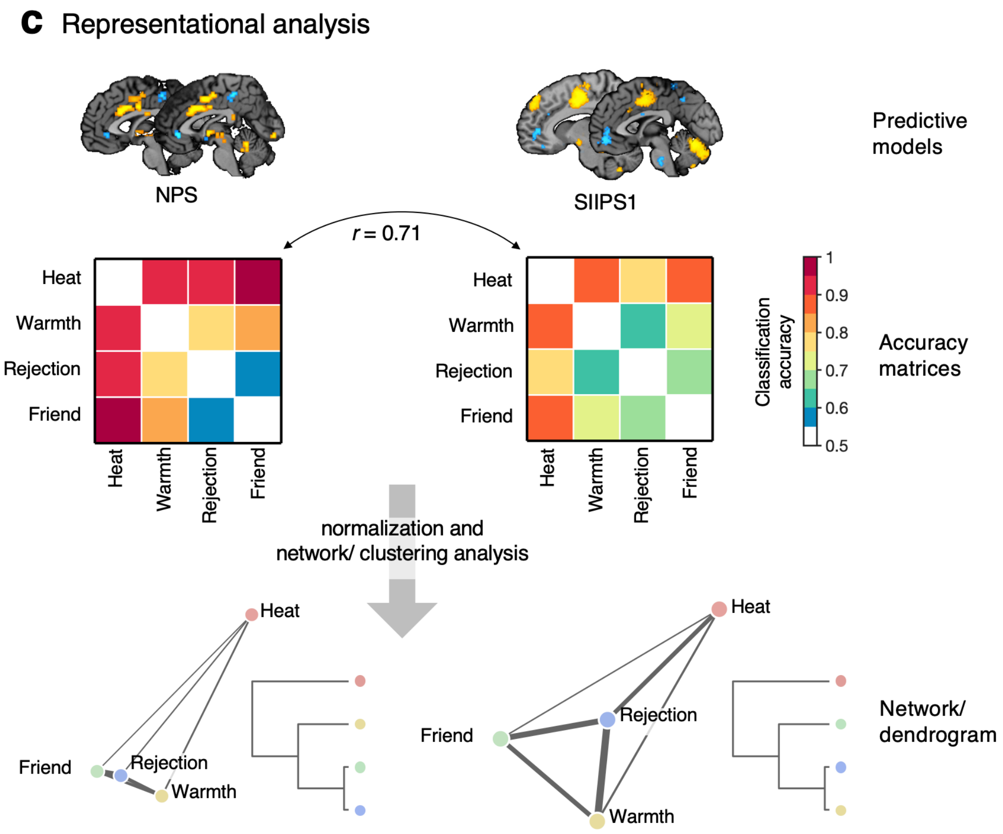
Example: 1. Representational connectivity





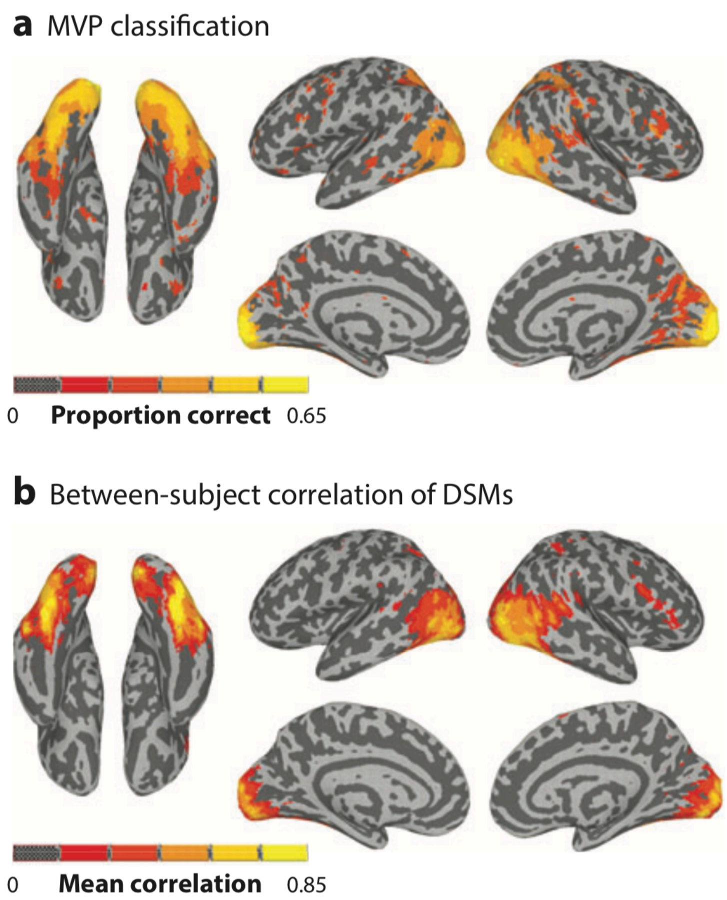
Kriegeskorte et al., 2008, *Neuron*

Example: 2. Accuracy as distance



Kohoutová et al., in revision

Example: 3. Searchlight RSA

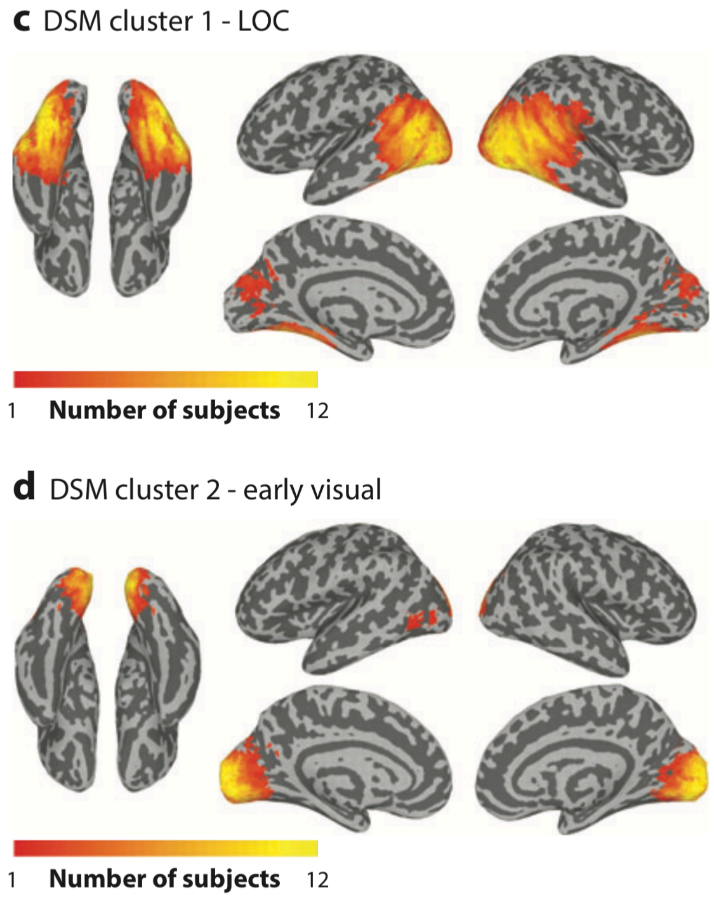


In addition to the MVPA, we can calculate the RDM (in this figure, DSM) for each person, and see between-subject consistency of the RDMs

Haxby et al., 2014, *Annu Rev Neurosci*

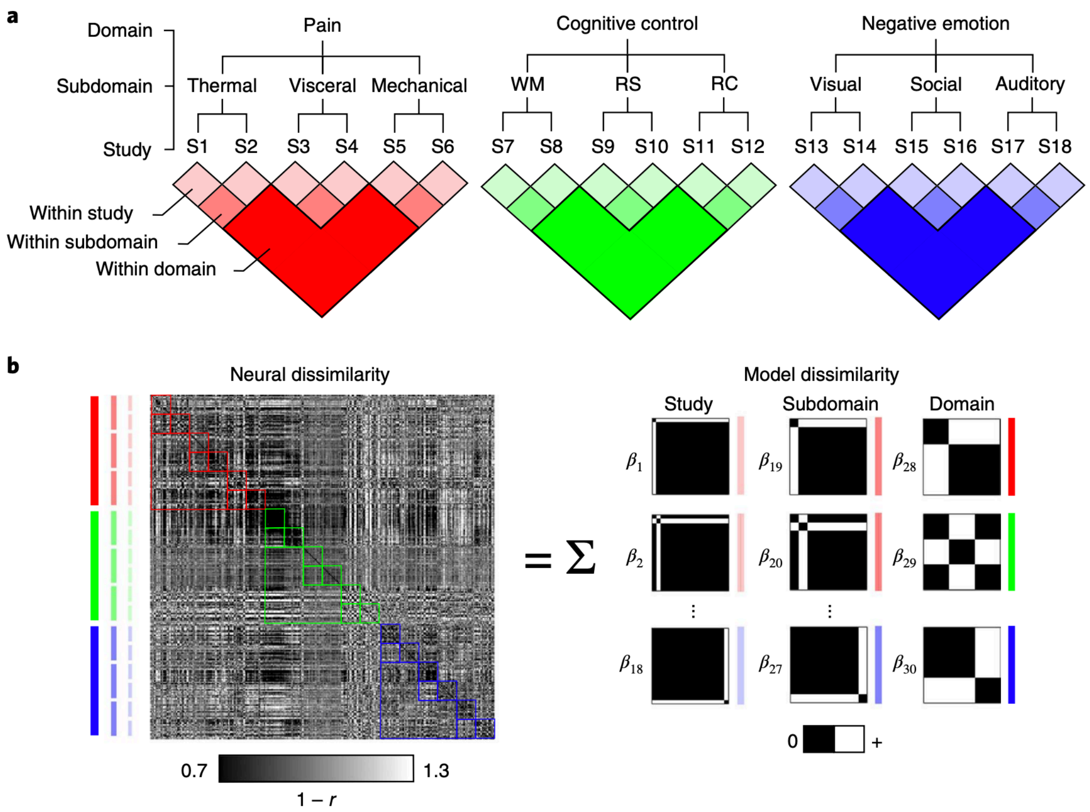
Example: 3. Searchlight RSA

We can also cluster the regions based on the RDM patterns (for each individual or for group)



Haxby et al., 2014, *Annu Rev Neurosci*

Example: 4. RSA in the GLM context

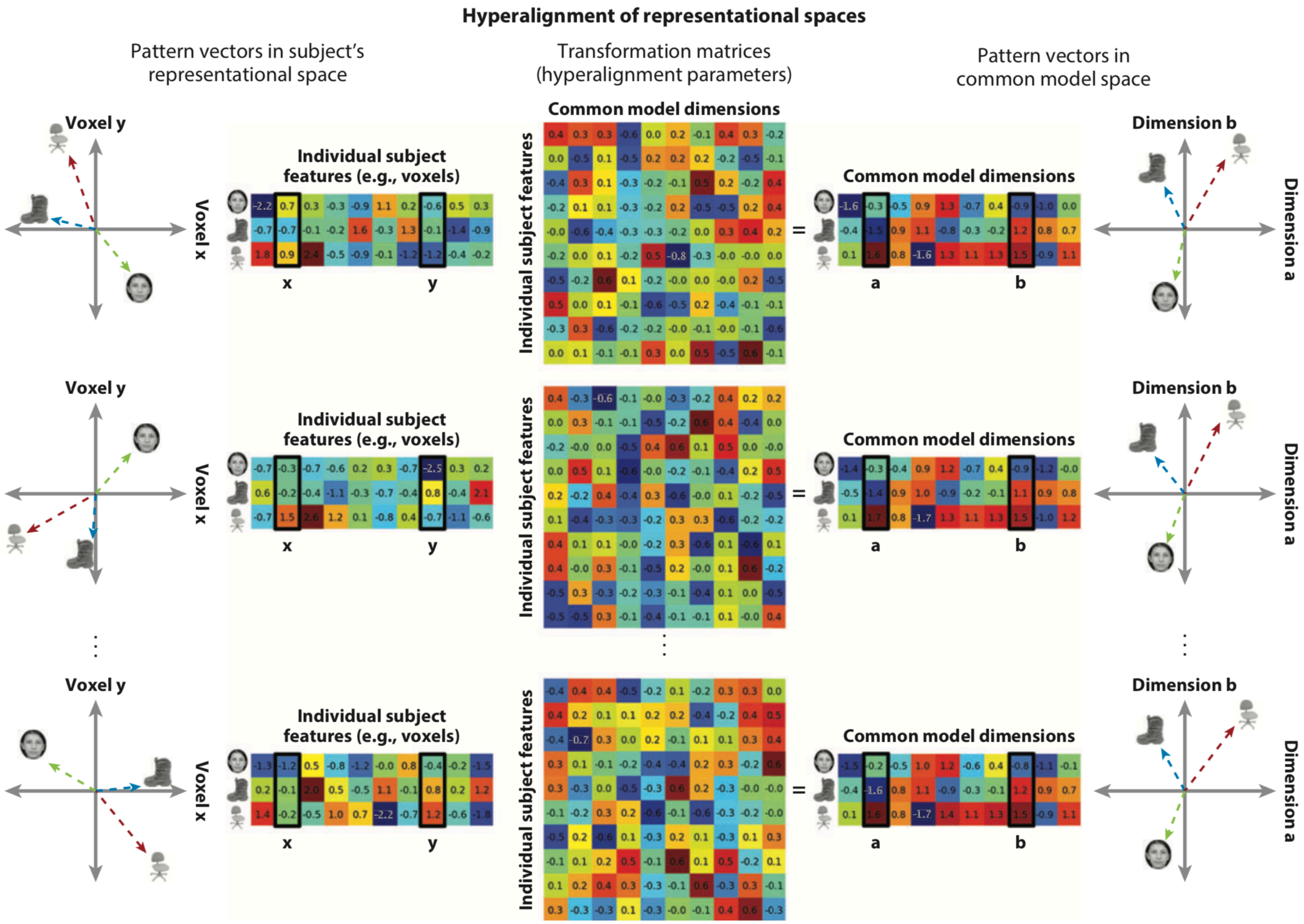


Kragel et al., 2018, *Nat Neurosci*

Hyper-alignment

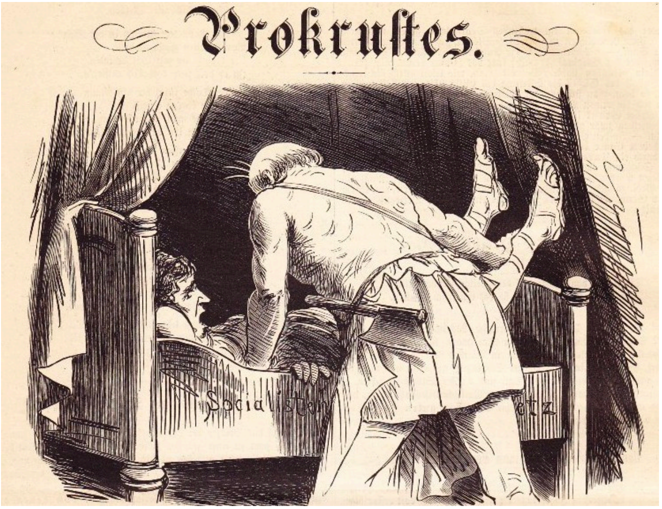
Hyper-alignment:
aligning individual neural representational spaces into a common model space

Iterative **procrustean transformation** to the common representational space



Haxby et al., 2014, *Annu Rev Neurosci*

Procrustes



From Greek mythology,

Procrustes (or “the stretcher”) was a bandit from Attica

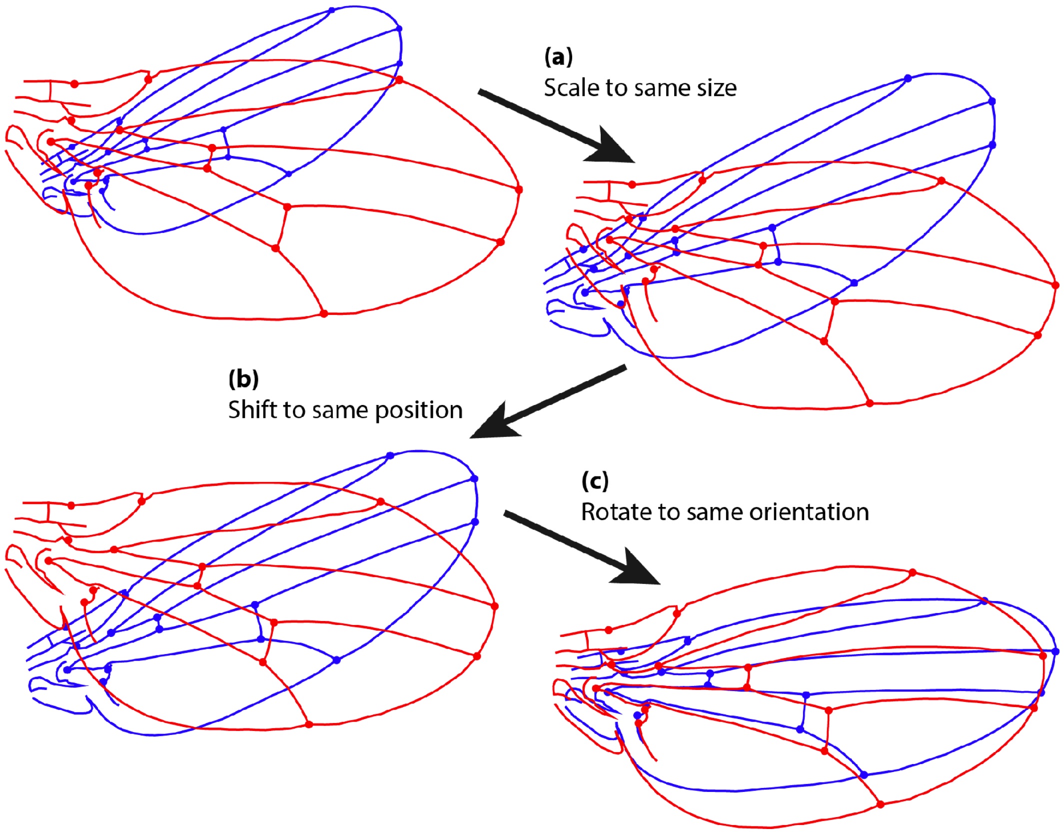
He attached people by stretching them or cutting off their legs, so as to force them to fit the size of an iron bed

The word "Procrustean" is thus used to describe situations where different lengths or sizes or properties are fitted to an arbitrary standard.

(from Wikipedia)
https://en.wikipedia.org/wiki/Procrustes

Haxby et al., 2011, *Neuron*

Procrustes analysis

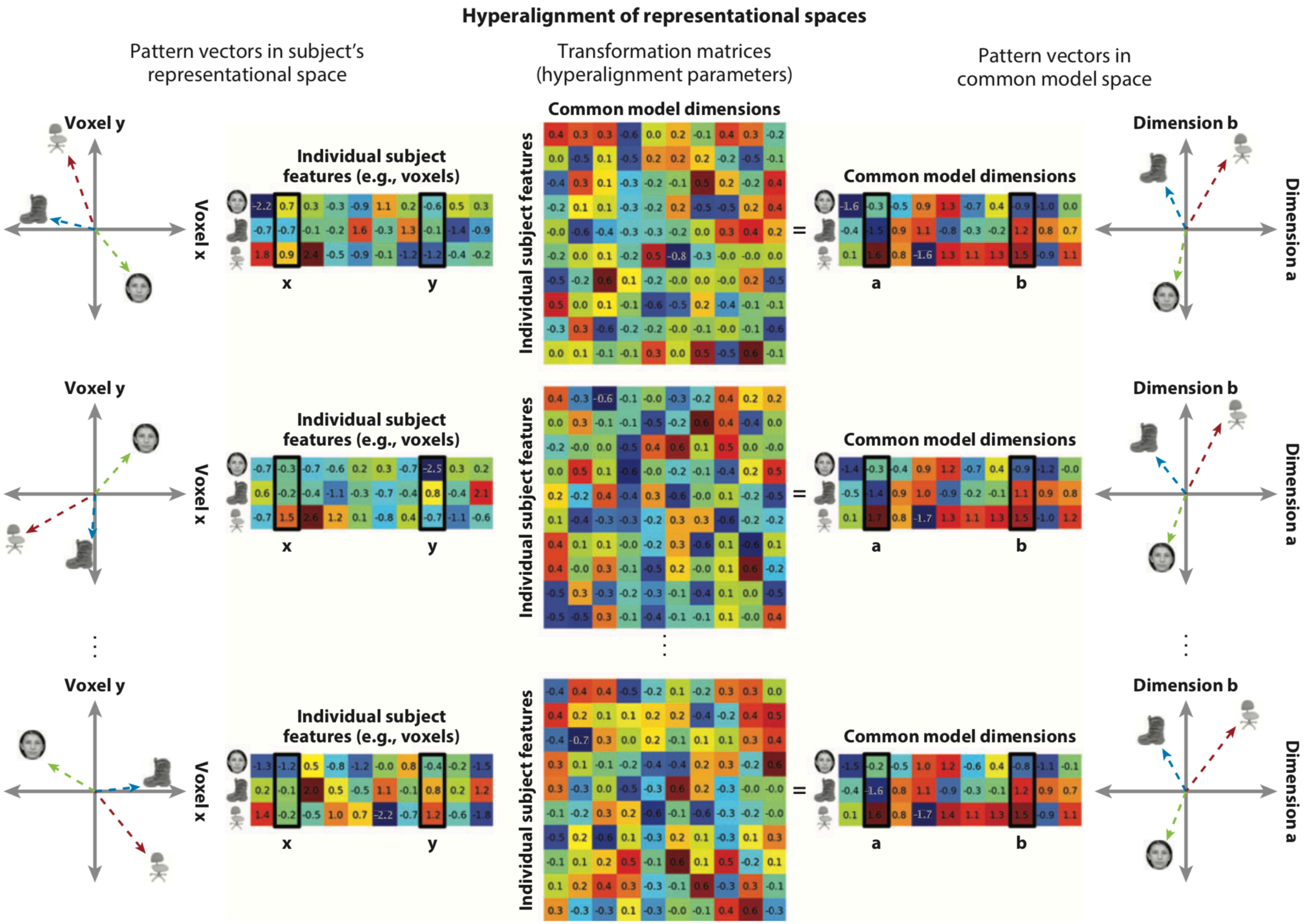


Haxby et al., 2011, *Neuron*

Hyper-alignment

Hyper-alignment:
aligning individual neural representational spaces into a common model space

Iterative **procrustean transformation** to the common representational space



Haxby et al., 2014, *Annu Rev Neurosci*

Thank you for your attention

Hope our representations about “RSA” to be well-aligned

Any questions?

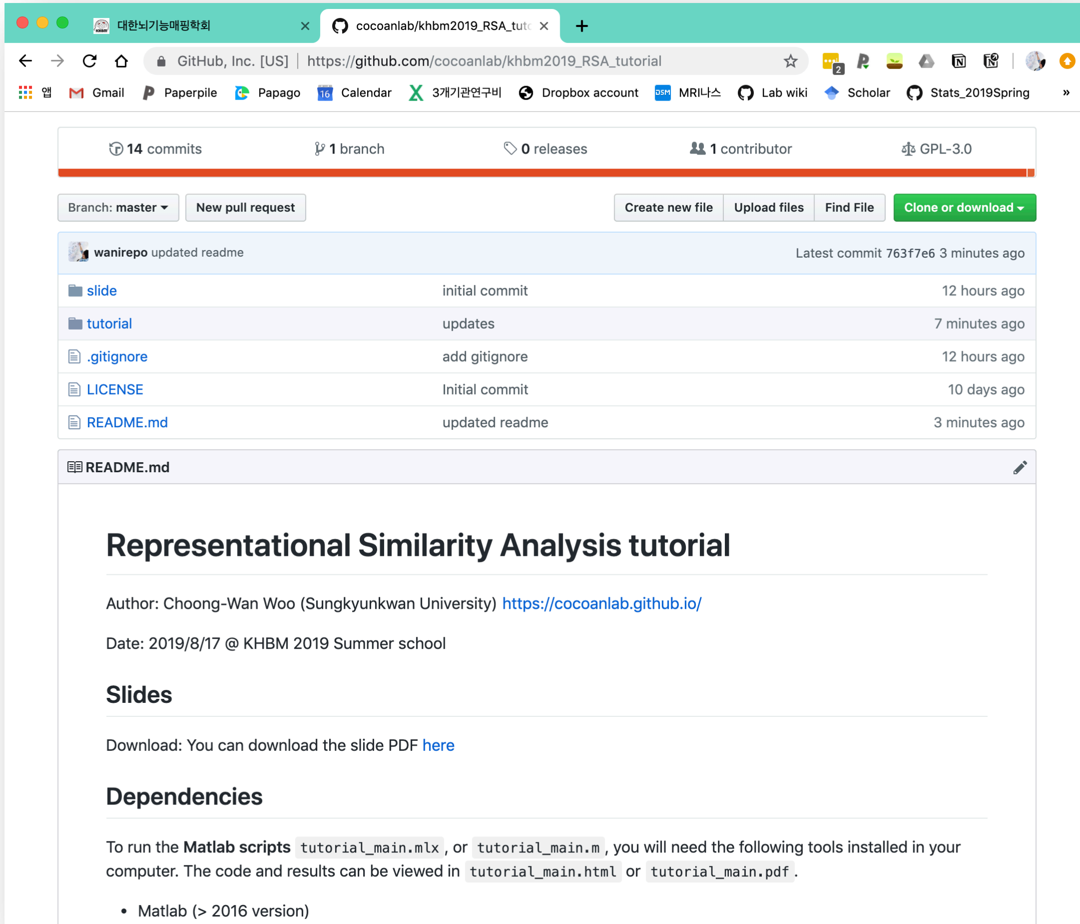
In the tutorial session:

https://github.com/cocoanlab/khbm2019_RSA_tutorial

Step 1: Computing and visualizing RDMs

Step 2: Comparing brain and model RDMs

Step 3: Statistical inference



Data and research question:

nature COMMUNICATIONS

ARTICLE

Received 8 May 2014 | Accepted 25 Sep 2014 | Published 17 Nov 2014 DOI: 10.1038/ncomms6380

Separate neural representations for physical pain and social rejection

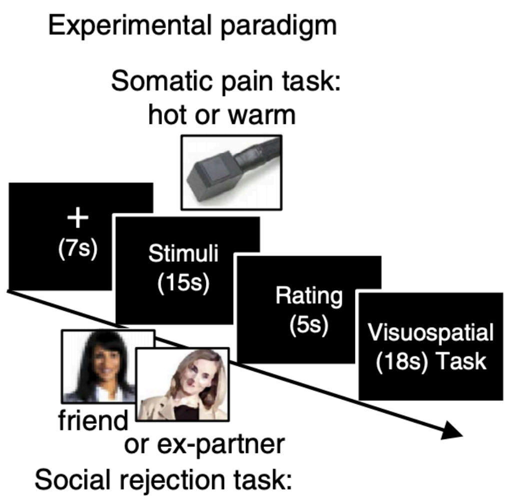
Choong-Wan Woo[1,2], Leonie Koban[1,2], Ethan Kross[3], Martin A. Lindquist[4], Marie T. Banich[1,2], Luka Ruzic[1,2], Jessica R. Andrews-Hanna[2] & Tor D. Wager[1,2]

Current theories suggest that physical pain and social rejection share common neural mechanisms, largely by virtue of overlapping functional magnetic resonance imaging (fMRI) activity. Here we challenge this notion by identifying distinct multivariate fMRI patterns unique to pain and rejection. Sixty participants experience painful heat and warmth and view photos of ex-partners and friends on separate trials. FMRI pattern classifiers discriminate pain and rejection from their respective control conditions in out-of-sample individuals with 92% and 80% accuracy. The rejection classifier performs at chance on pain, and vice versa. Pain- and rejection-related representations are uncorrelated within regions thought to encode pain affect (for example, dorsal anterior cingulate) and show distinct functional connectivity with other regions in a separate resting-state data set ($N = 91$). These findings demonstrate that separate representations underlie pain and rejection despite common fMRI activity at the gross anatomical level. Rather than co-opting pain circuitry, rejection involves distinct affective representations in humans.

Original question: Can we identify specific patterns of fMRI activity for physical pain and social pain, respectively?

- We conducted an fMRI experiment (N = 59) using somatic and social pain tasks.
- All 59 individuals (31 females, M_{age} = 20.8, SD_{age} = 3.0) recently experienced an unwanted break-up with their romantic partners and felt intensely rejected.

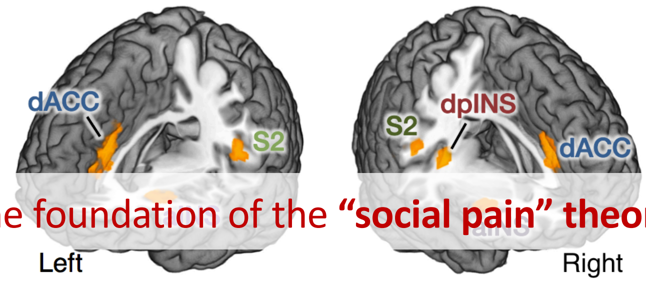
Tasks:



Data and research question:

Traditional mapping results

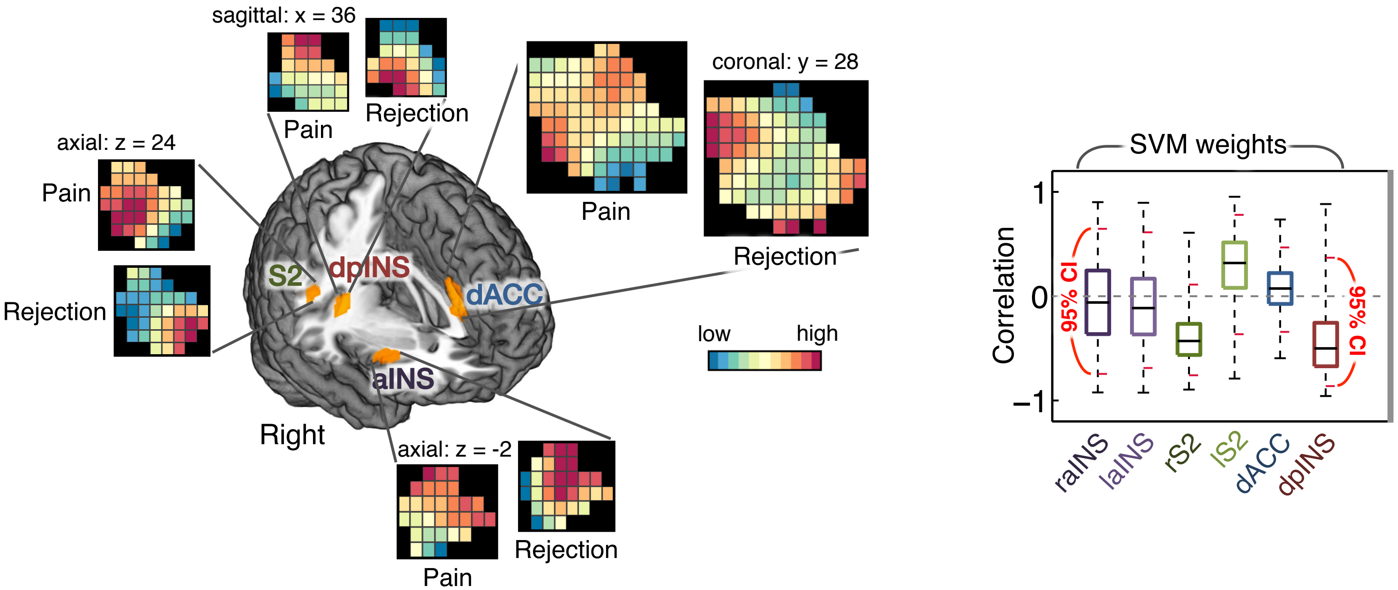
Univariate overlap
between [Heat-pain vs. Warmth] and [Ex-partner vs. Friend]



The foundation of the **"social pain" theory**

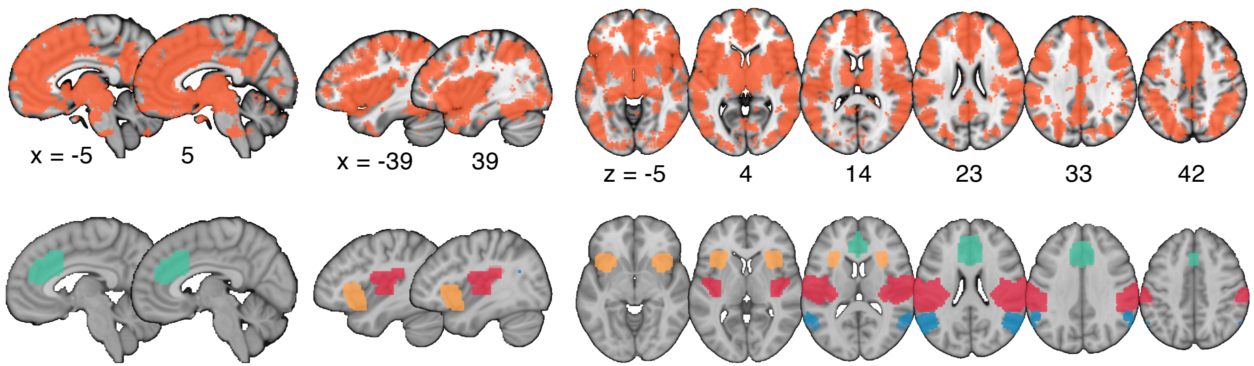
Then, we used the predictive mapping approach (linear support vector machines, SVMs) to obtain specific multivariate fMRI pattern for pain and rejection on the same dataset.

Data and research question:



Data and research question:

1. Computing and visualizing RDMs for each participant, for each region

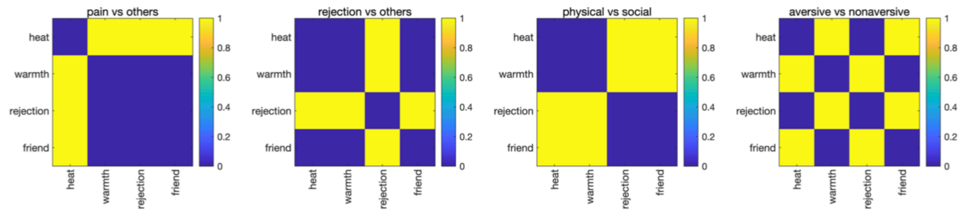


Whole-brain mask
based on neurosynth
(Woo et al., 2014)

aMCC
aINS
S2/dpINS
TPJ

2. Computing four model RDMs with these ROI RDMs

1) Heat vs. others
2) Rejection vs. others
3) Physical vs. social
4) Aversive vs. non-aversive



Question: Which one of these is the best-supported model based on the representational similarity patterns from data?