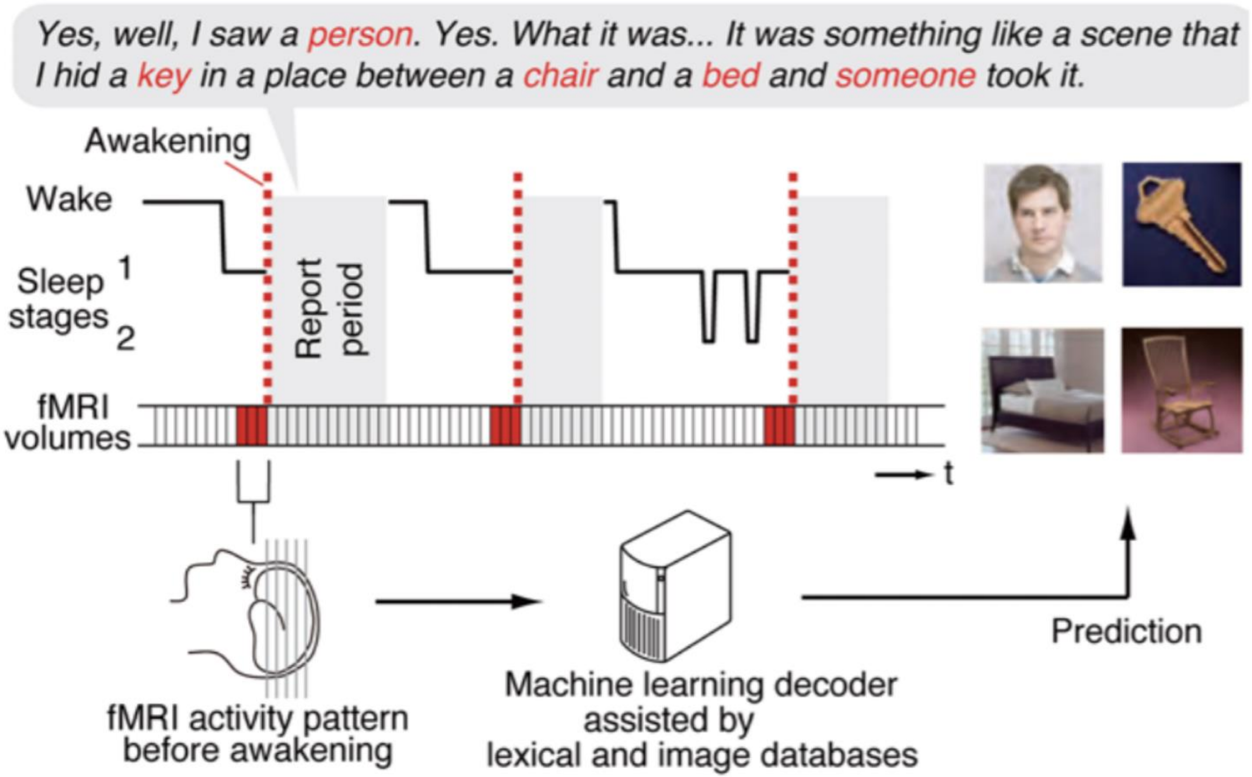
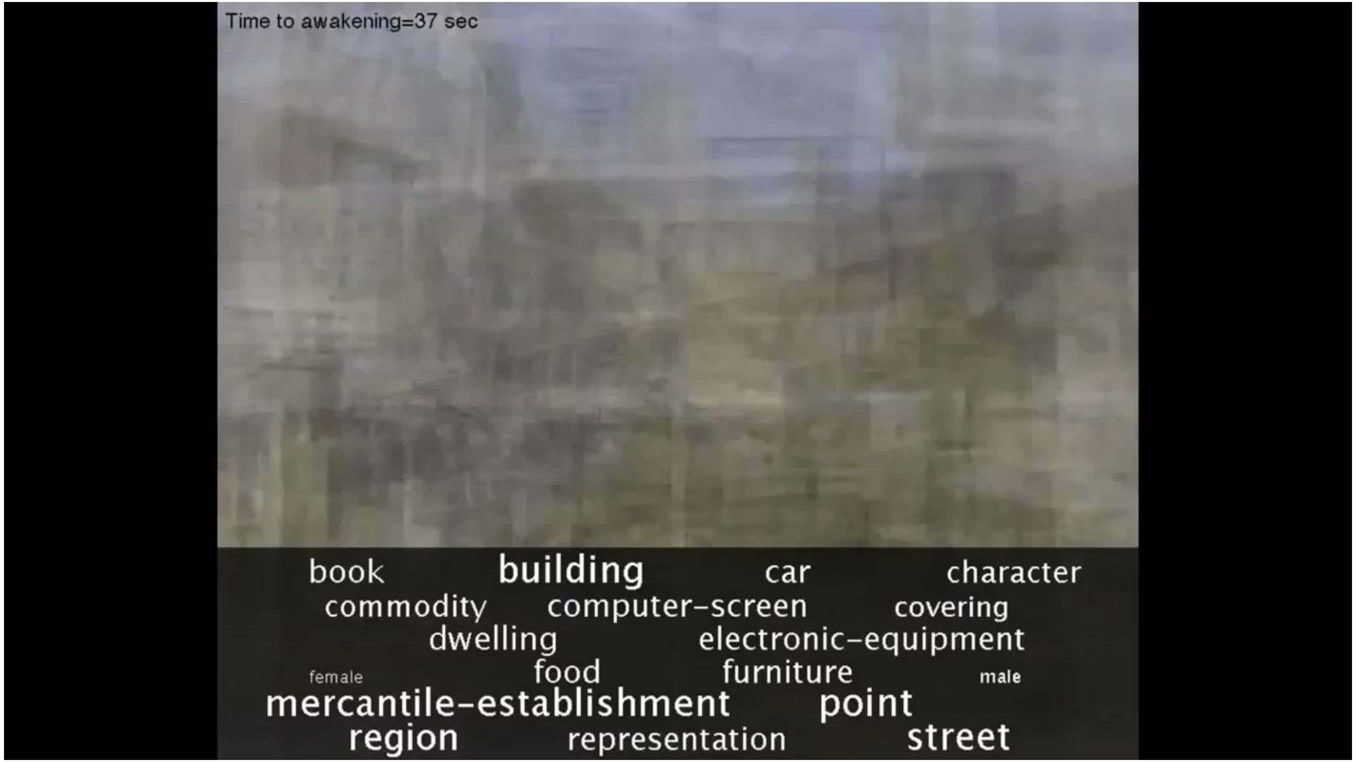
Multivariate pattern analysis

Multivariate pattern analysis: An early example study



Kamitani & Tong 2013

Multivariate pattern analysis: An early example study



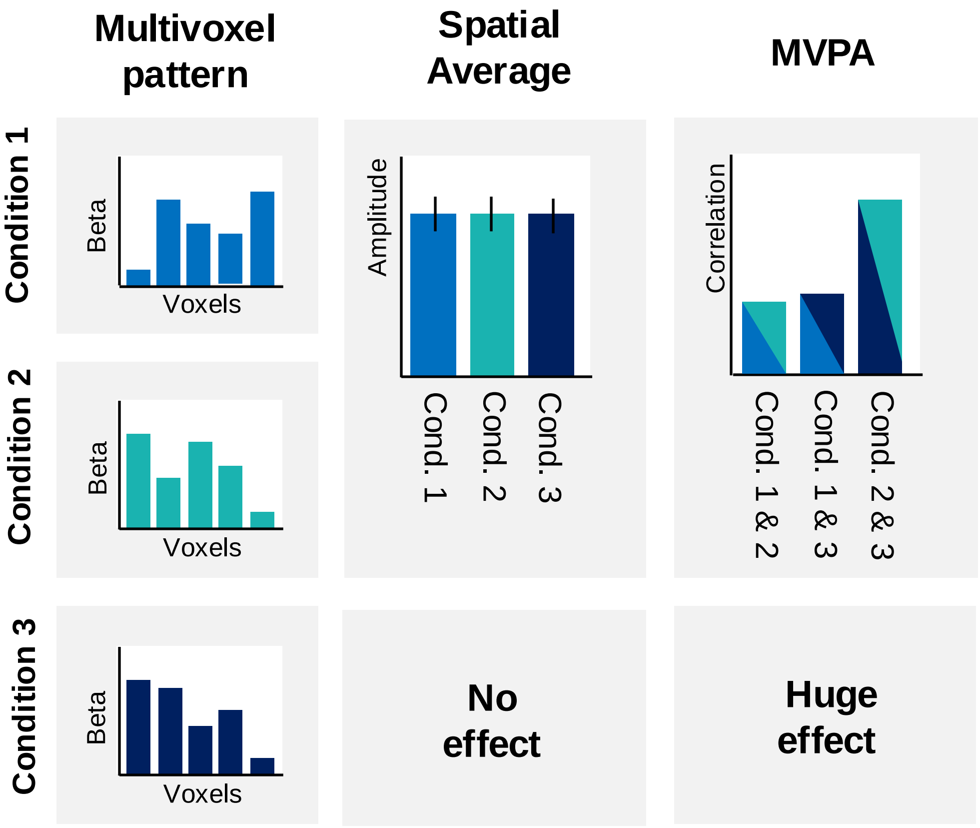
Reconstructing visual imagery during sleep – How is this possible?

Kamitani & Tong 2013

Multivariate pattern analysis (MVPA)

Set of computational techniques to analyze patterns of brain activity by combining information across multiple voxels

What is MVPA and why would we use it?



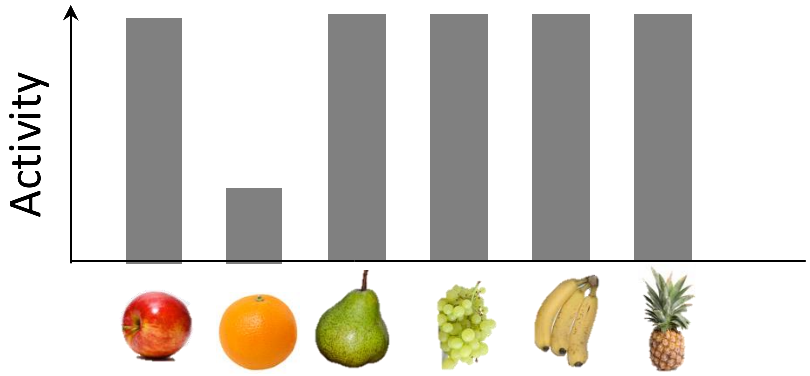
MVPA utilizes the distribution of activity across voxels, rather than just activity level of individual voxels

Univariate
Analyzing each voxel separately (e.g, GLM)

Multivariate
Combine information of multiple voxels (e.g., Decoding)

What is MVPA and why would we use it?

Thought experiment



Typical univariate result interpretation

- Region response to all fruit but oranges
- Region more active = more involved in task

Typical multivariate result interpretation

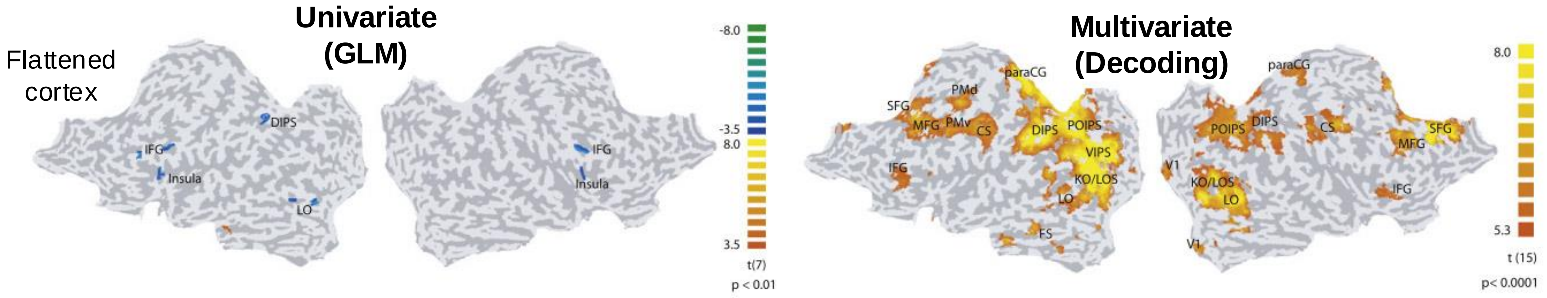
- Region carries information about oranges
- Pattern more distinct = more involved

MVPA is more than a technique, it is a mindset

- It asks about the presence of information
- Useful even if it is not clear what that information is (e.g., for brain-computer interfaces)
- Prediction does not equal explanation

What is MVPA and why would we use it?

1) **Increased sensitivity** compared to classical GLM approaches



Example: BOLD signals reflecting behavioral choice during categorization

Li et al (2009)

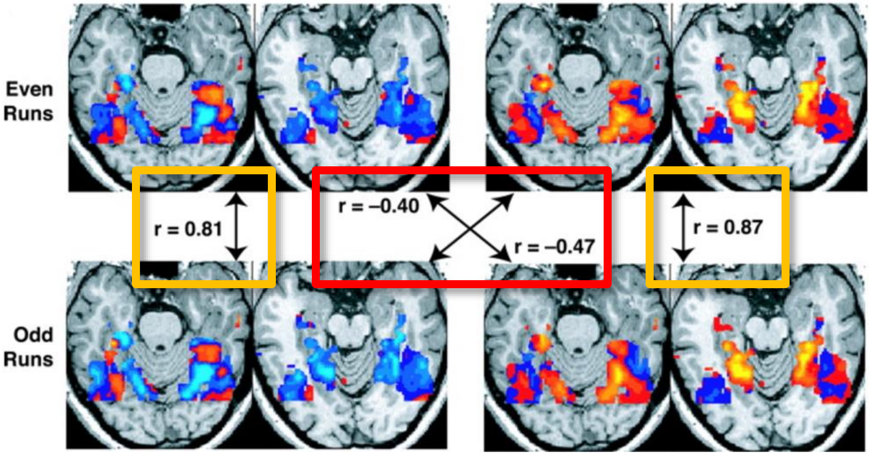
2) **Abstracting away from measurement** (e.g., representational similarity) allows comparisons between species, data modalities & computational models

Will become clearer in the lecture!

MVPA – First demonstration

Visual stimuli from different categories

Distributed category-selective brain activity



Correlation matrix

Within-category and across-category similarity

Even runs \ Odd runs	Faces	Houses
Faces	r=0.81	r=-0.47
Houses	r=-0.40	r=0.87

Critical logic

If category matters, then:

Within-category similarity

>

Across-category similarity

Haxby et al. 2001

There are many MVPA techniques. This lecture will focus on:

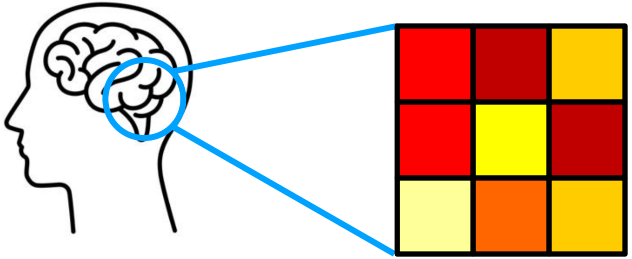
- **Decoding**
- **Representational similarity analysis**

Decoding

Decoding – What is it?

Encoding (e.g., GLM)
Predicting brain activity using stimulus- or behavioral features

Stimulus or Behavior | **Participant** | **fMRI data**



Reconstructing stimulus- or behavioral features from brain activity
Decoding

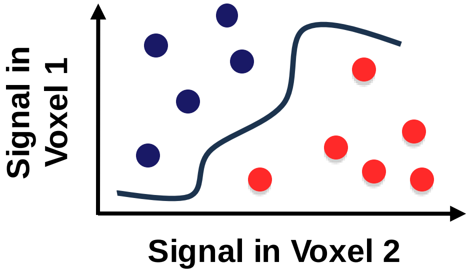
Decoding – What is it?

Two basic types of decoders

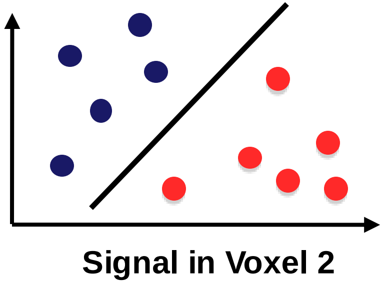
- **Continuous outcomes** (e.g., location in VR) are predicted by **Regression Models**
- **Categorical outcomes** (e.g., cat vs. dog) are predicted by **Pattern Classifiers**

Two basic types of Pattern Classifiers

Non-linear classifiers

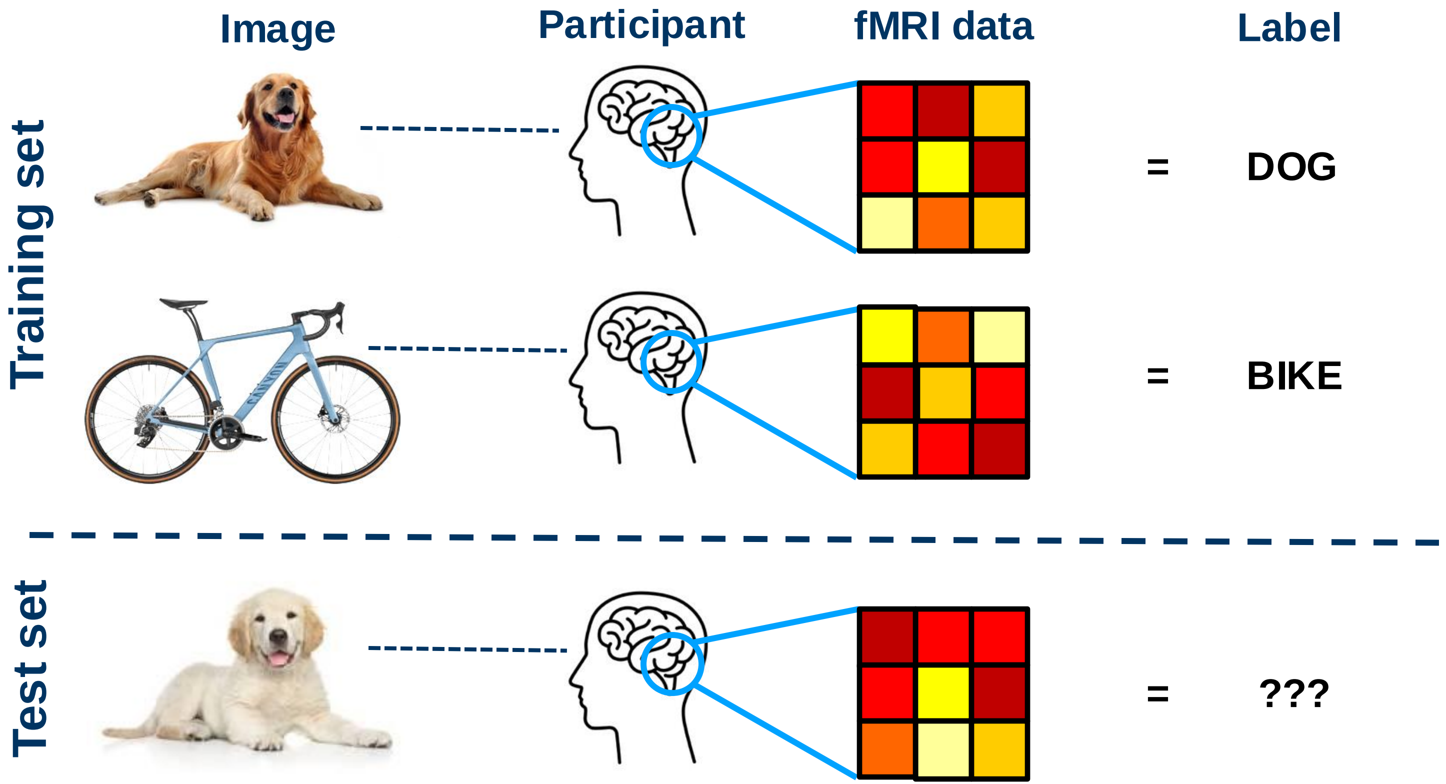


Linear classifiers



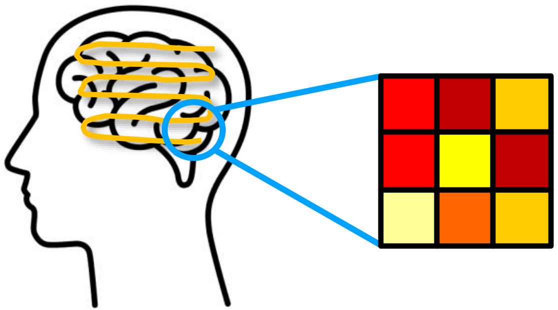
Most common decoders fMRI

Decoding – Pattern classification



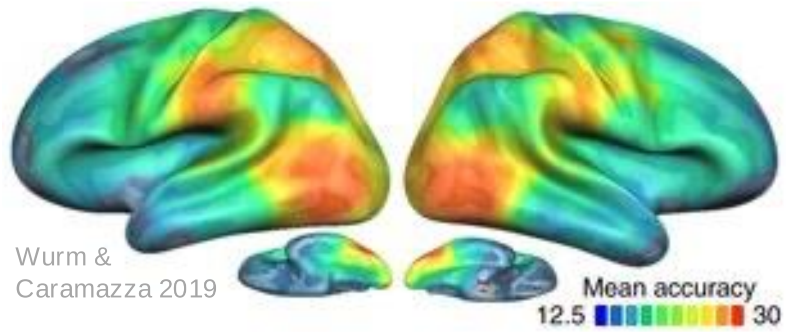
Decoding – Searchlight-based approaches

Analysis logic



- Center a sphere on each voxel and extract multivoxel pattern
- Train & test the decoder for sphere
- Assign decoding performance to center voxel

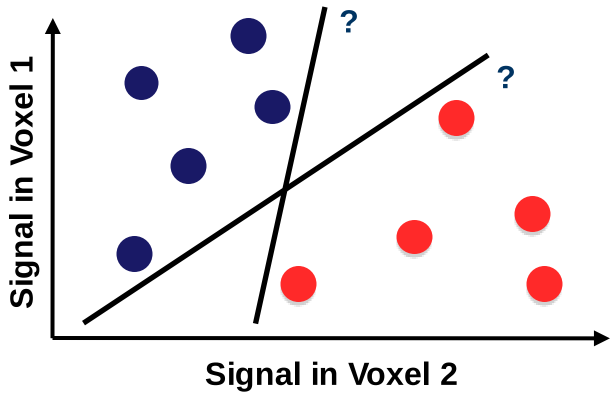
Decoding from all across the cortex

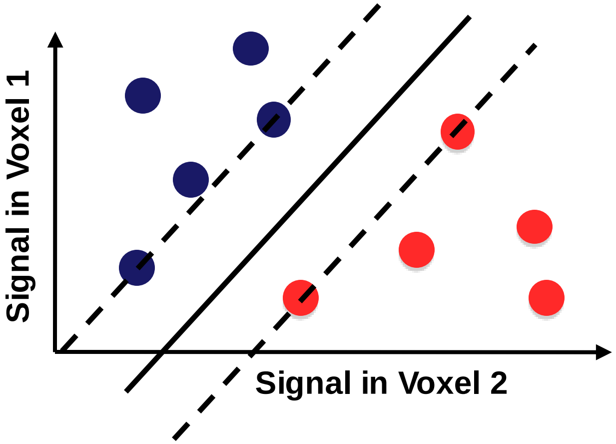


Decoding – Support Vector Machine (SVM)

Problem
How do we decide on a **decision boundary** that separates the classes?

SVM solution
1) Find **support vectors** (points on dashed lines).
2) The **decision boundary maximizes margin** between them (solid line)



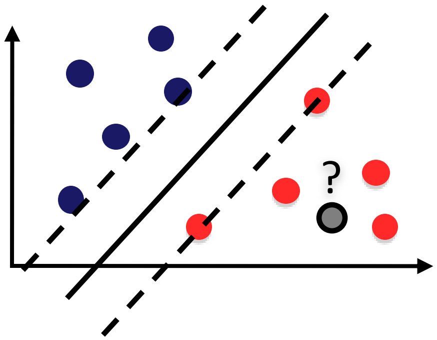


In fMRI, SVMs are trained on **more than two voxels**. Rather than 1D lines, they find **Hyperplanes** (i.e., high dimensional decision boundaries)

Common because SVMs are robust and versatile in high-dimensional settings

Decoding – Common classifiers in fMRI *

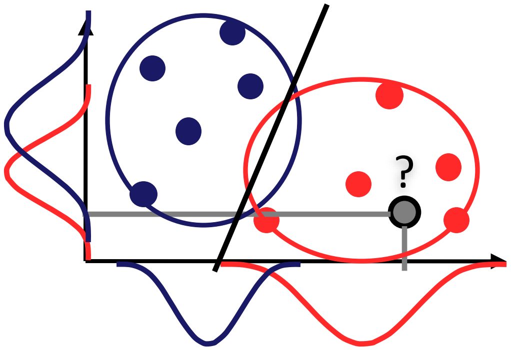
Support Vector Machine



Maximizes margin
(distance between closest
points of different classes)

Robust & versatile classifier
for high-dimensional data,
computationally expensive

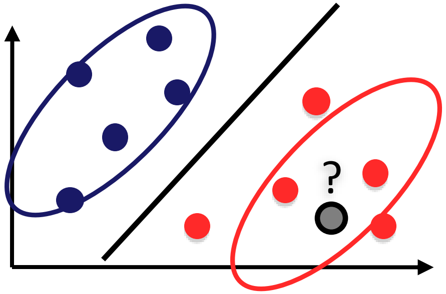
Gaussian Naïve Bayes



Computes Bayesian
probability of belonging
to a specific class

High accuracy, Fast to
compute, assumes normality
and independence of voxels

Linear Discriminant Analysis



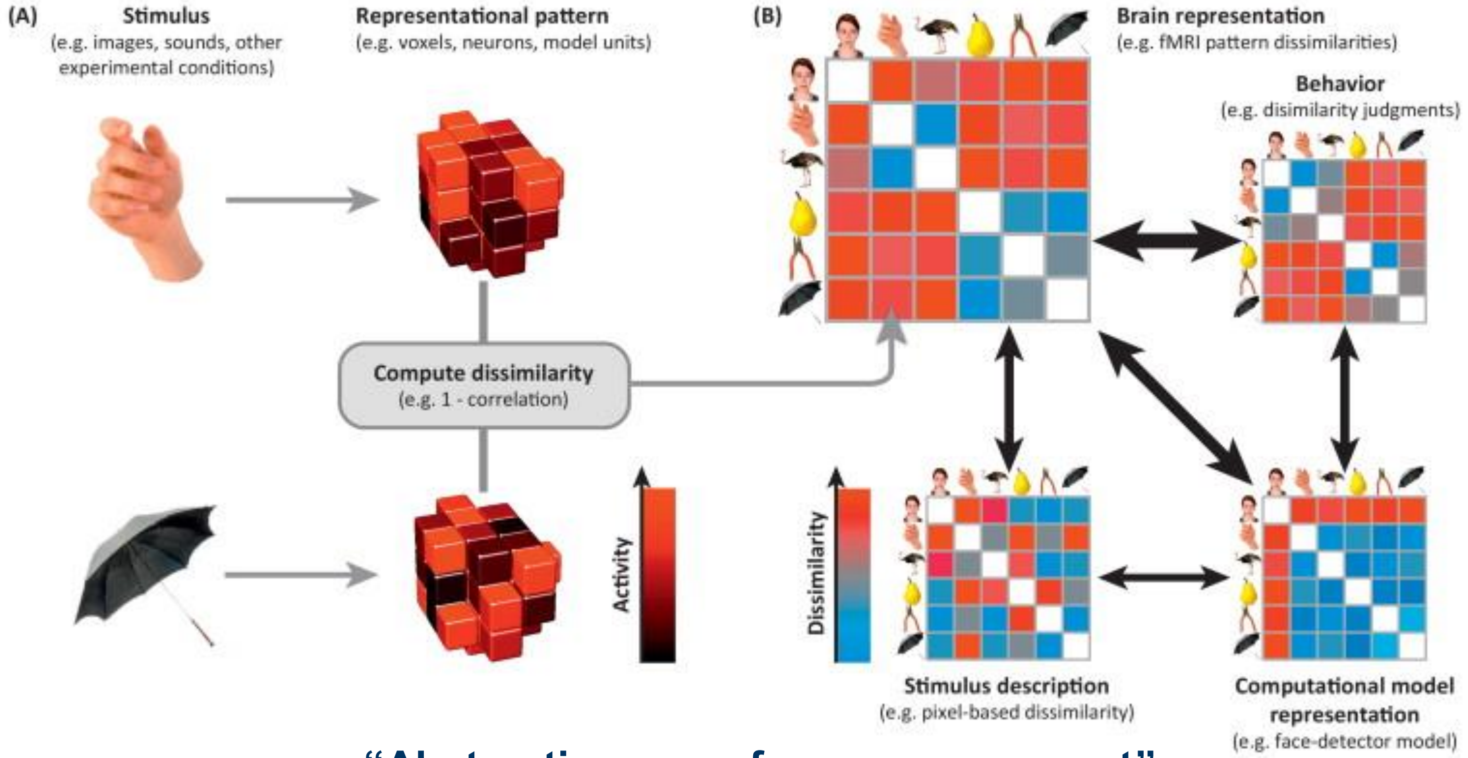
Maximizes ratio of
between-class variance
to within-class variance

High accuracy even in
small samples, assumes
normality and equal
covariance across classes

*Many more: Logistic regression, deep learning models…

Representational similarity analysis (RSA)

Representational similarity analysis (RSA)

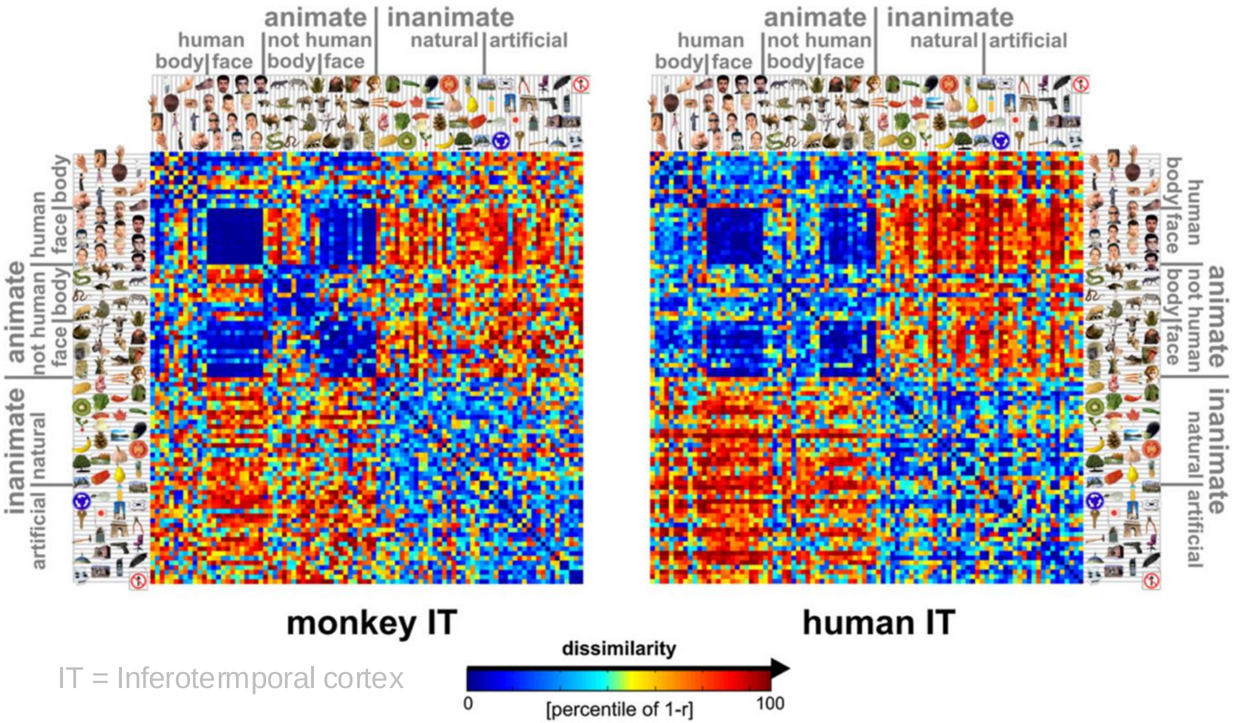


“Abstraction away from measurement”

Representational dissimilarity matrices (RDMs) are compared, NOT the BOLD signal

Kriegeskorte & Kievit 2013

Representational similarity analysis (RSA)



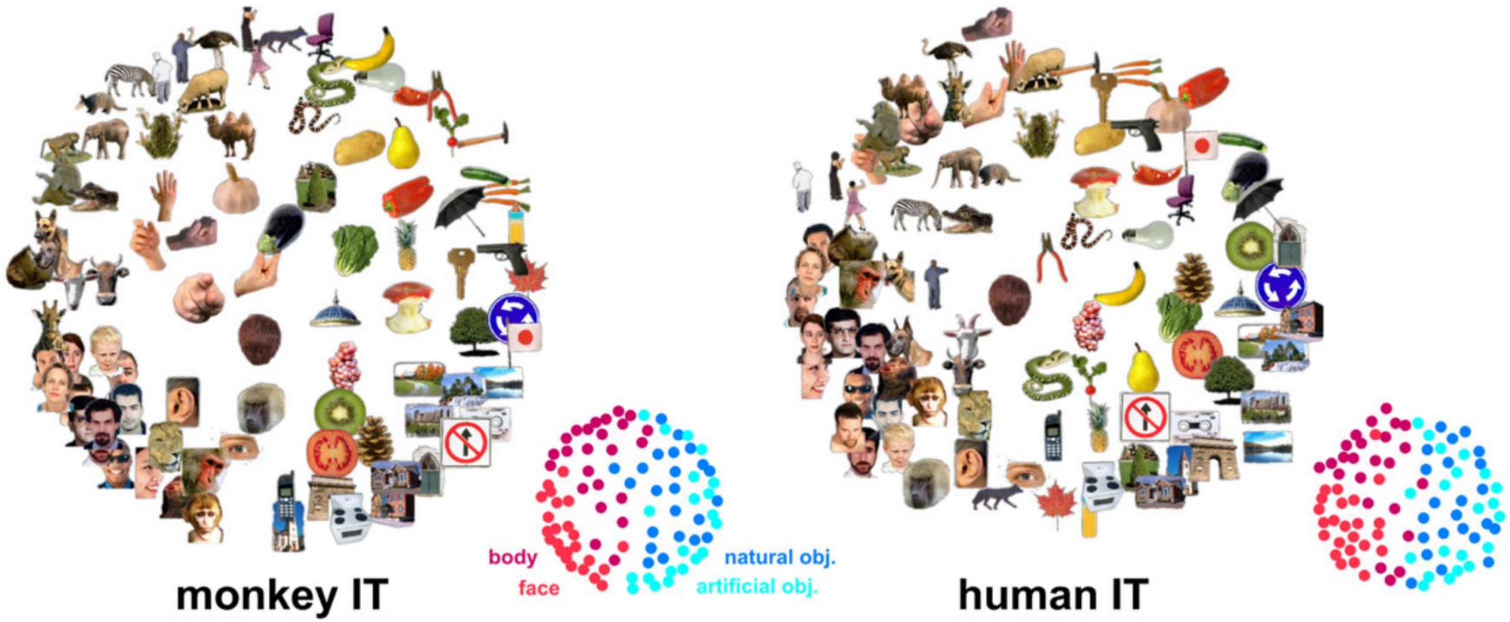
IT = Inferotermporal cortex

Comparisons across species

Different BOLD responses in monkeys and humans, but RDMs are similar

Kriegeskorte et al. 2008

Representational similarity analysis (RSA)



Multidimensional scaling

Embedding stimuli in a 2D space based on their representational dissimilarity.
Intuitive visualization of representational dissimilarities (or “**Representational spaces**”)

Kriegeskorte et al. 2008

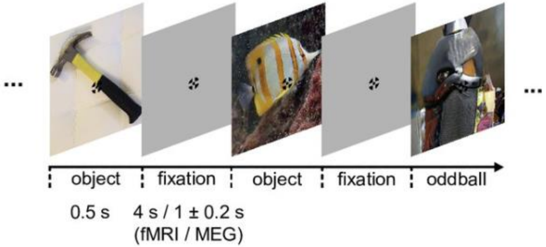
Representational similarity analysis (RSA)

Similarity judgement task outside MRI

"Which is the odd one out?"

Viewing task in the MRI

"Is this image an object"?

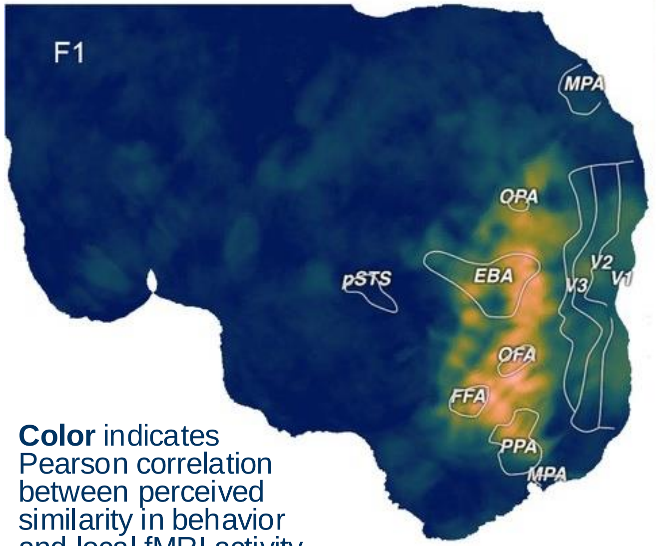


Object similarity

Behavioral similarity matrix
(based on people's responses)

fMRI similarity matrix
(estimated for each local patch of cortex)

Cortical surface of a single participant



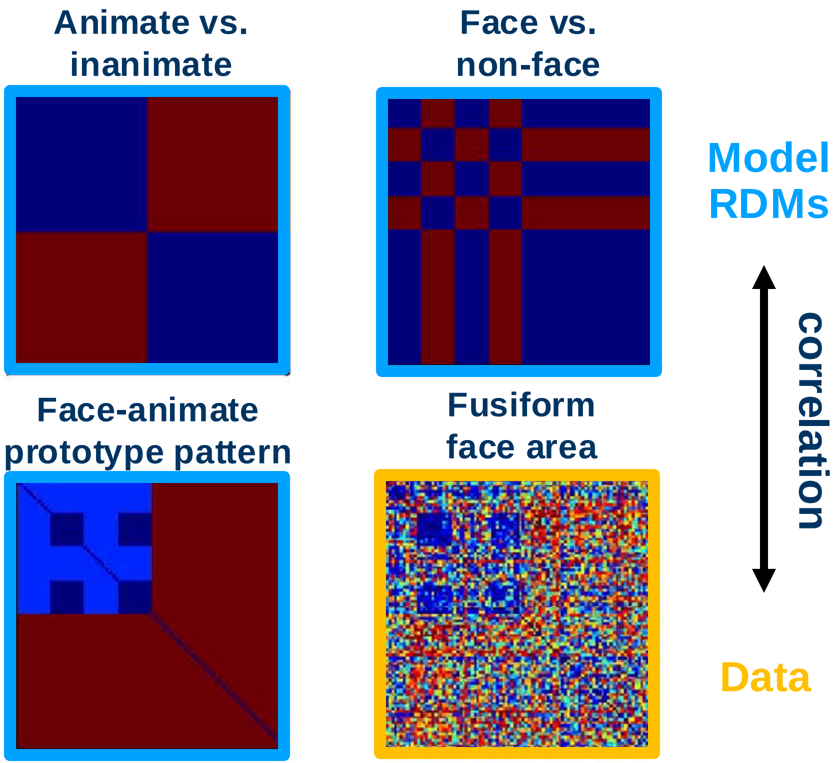
Color indicates Pearson correlation between perceived similarity in behavior and local fMRI activity

Comparing behavior to fMRI data

"Representational spaces" derived from behavioral & neural measurements

Hebart, Contier, Teichmann et al. 2023

Representational similarity analysis (RSA)



Simple hypothesis testing

Representational dissimilarity matrices can be constructed by experimenters (Model RDMs), reflecting hypotheses, and then compared to data RDMs

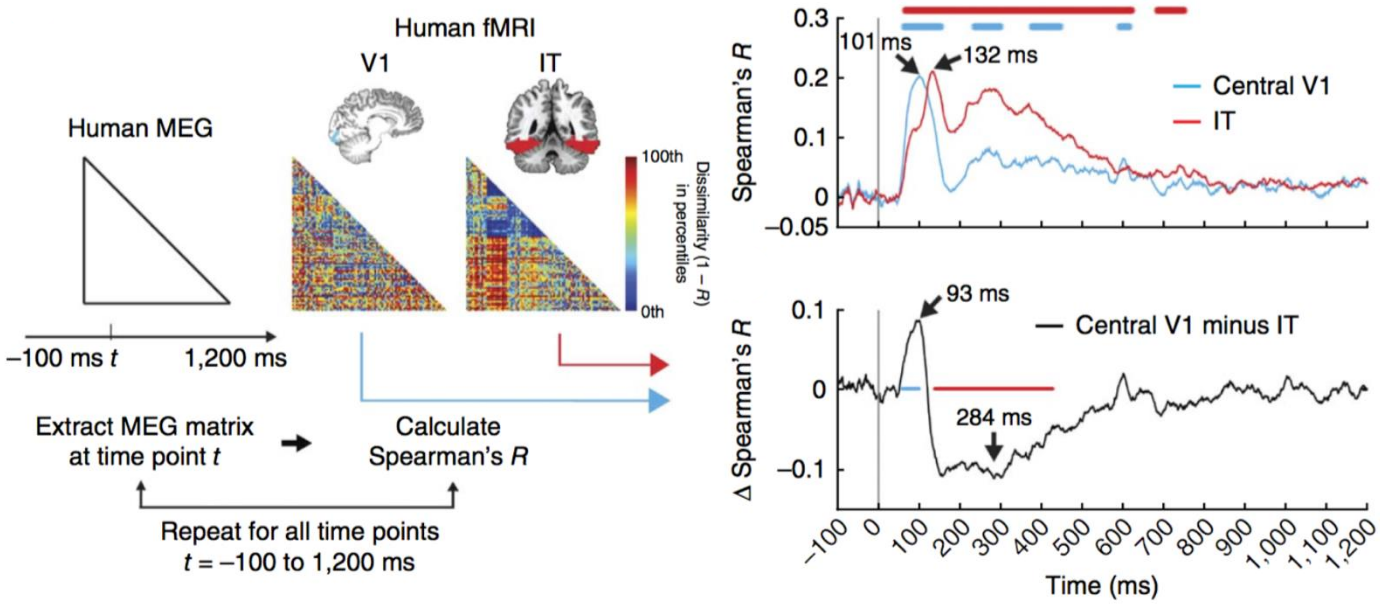
More complex models possible!

(e.g., comparing convolutional neural networks layers to fMRI data)

See e.g., Khaligh-Razavi & Kriegeskorte 2014

Kriegeskorte et al. 2008

Representational similarity analysis (RSA)

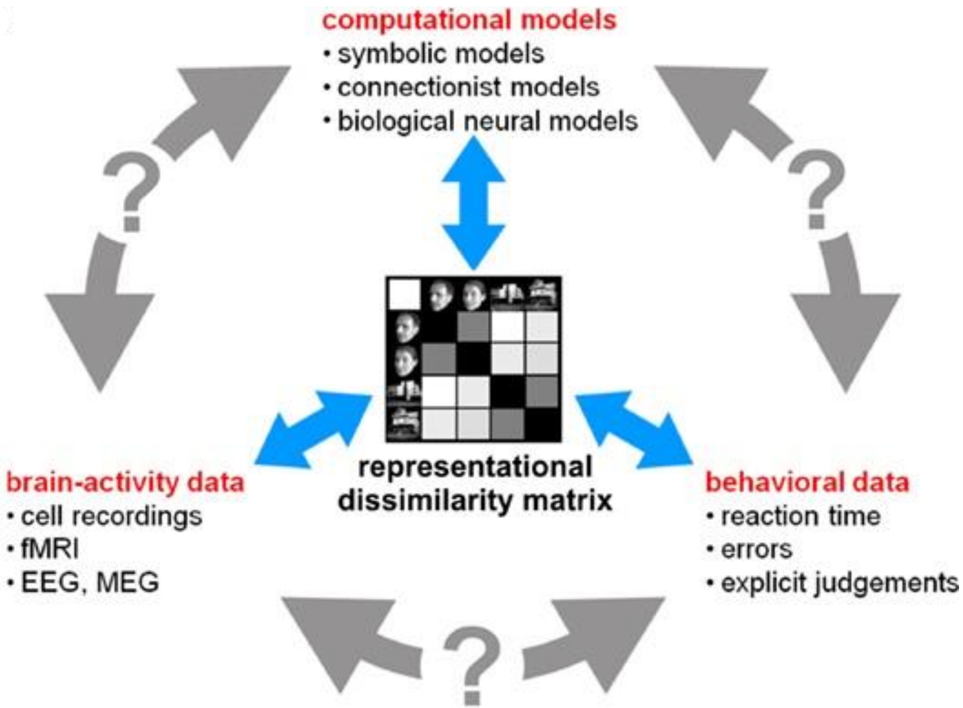


Cichy et al. 2014

Comparing different recording techniques

RSA enables combining different techniques, such as those with high spatial (fMRI) and temporal (e.g., MEG) resolution (i.e., for review see Cichy & Oliva 2020)

Representational Dissimilarity Matrices (RDMs)
are incredibly versatile hubs for relating different representations



Distance metric matters
(Pearson, Spearman, Euclidean distance…)

Data normalization matters
(Z-scoring, multivariate noise normalization…)

Dataset size matters
(More is better)

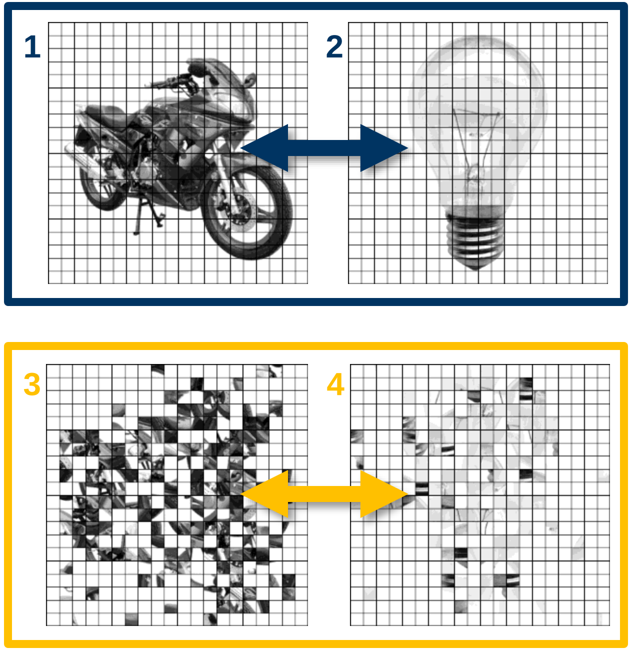
Always check diagonal of RDMs!
(The diagonal reflects SNR)

Cross-validate your results!
(Distance estimates can be corrupted by noise)

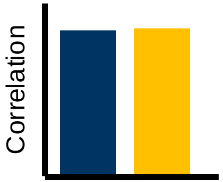
Conclusion & general remarks

Conclusion & general remarks

What is a pattern?



Would you say the pattern has changed?



Same correlation!
Image 1 is as similar to image 2,
as image 3 is to image 4*

* Assuming both images were scrambled in the same way

When we talk about **patterns,** most often, we mean the **distribution** of voxel intensities, not their spatial organization (True for RSA, SVM,...)

Certain techniques are sensitive to **spatial organization** (e.g., convolutional neural networks)

Different **distance metrics** are sensitive to different patterns (e.g., correlation vs. Euclidean distance)

- **Different techniques & metrics are sensitive to different patterns**
- **Reverse inference problem remains (e.g., What exactly drives the decoding?)**

Rare footage of a discussion on “brain decoding” and Representational spaces at the school of Athens

Plato

“Representational spaces are so cool! It’s like we are visualizing the content of the mind… the things up there in the clouds!”

Aristotle

“Errrm, maybe we should keep our feet on the ground down here. We are still measuring MRI signals and blood flow and stuff, not mental representations“

In other words: We can never fully “abstract away” from the measurement process!

Key terms to remember

- Univariate vs. Multivariate approaches
- Multivariate pattern analysis
- Information
- Encoding vs. Decoding
- Classification
- Linear vs. non-linear classifiers
- Searchlight-based analysis
- Support vector machines (SVM)
- Representational similarity analysis
- Representational dissimilarity matrix
- Multidimensional scaling
- Patterns (what are they?)
- Distance metric

Image: amberusa.com

Happy brain decoding!