Multivariate Pattern Analysis

Martin N. Hebart
Laboratory of Brain and Cognition
NIMH

Neural Decoding of Visual Imagery During Sleep

T. Horikawa,[1,2] M. Tamaki,[1*] Y. Miyawaki,[3,1]† Y. Kamitani[1,2]‡

www.sciencemag.org SCIENCE VOL 340 3 MAY 2013

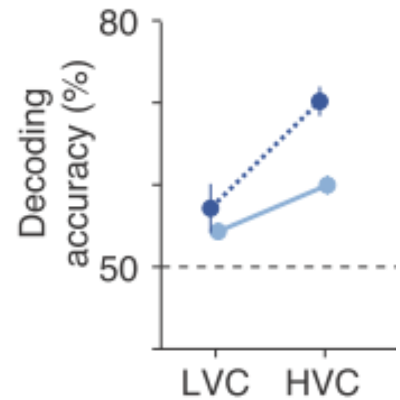
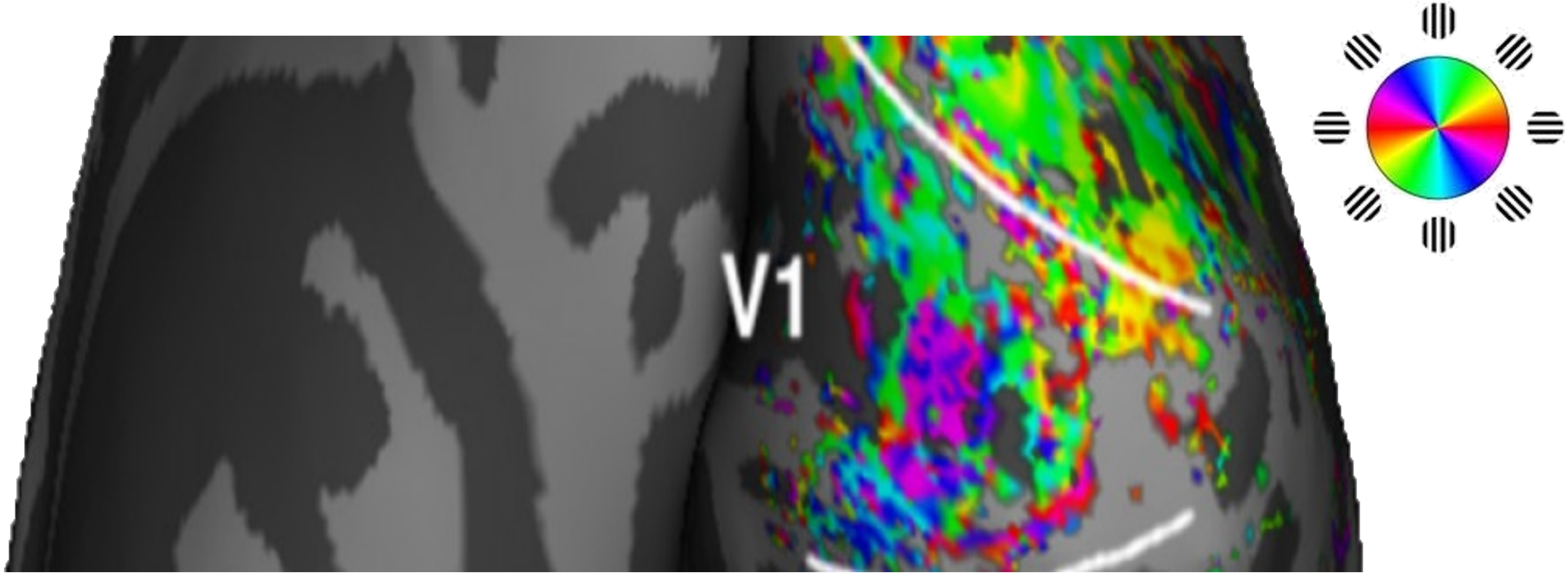


Image courtesy: David Castillo Dominici, http://www.freedigitalphotos.net

What is Multivariate Pattern Analysis?

Combined use of multiple variables measuring the brain (e.g. BOLD signal in multiple voxels) to predict or characterize states of the brain



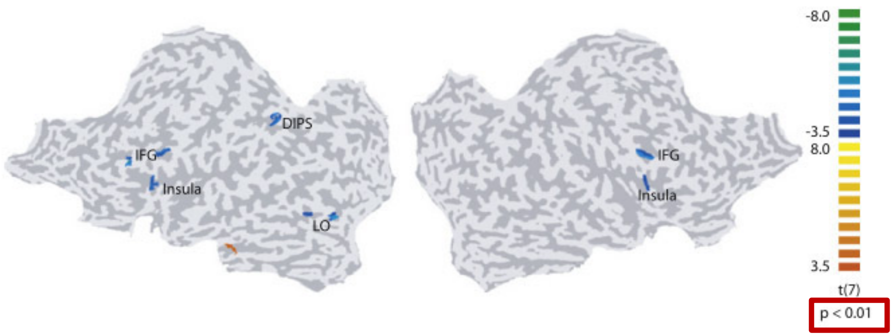
Swisher et al. (2010) – J Neurosci

Why Multivariate Pattern Analysis?

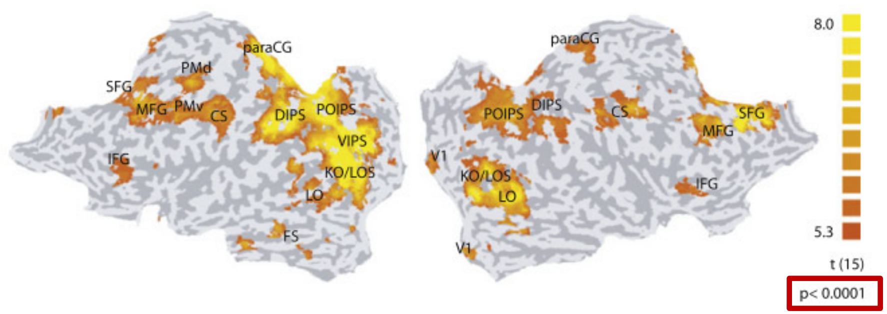
1. Often **higher sensitivity** compared to classical univariate analysis

Example: Representation of perceptual choices

classical univariate analysis



multivariate decoding

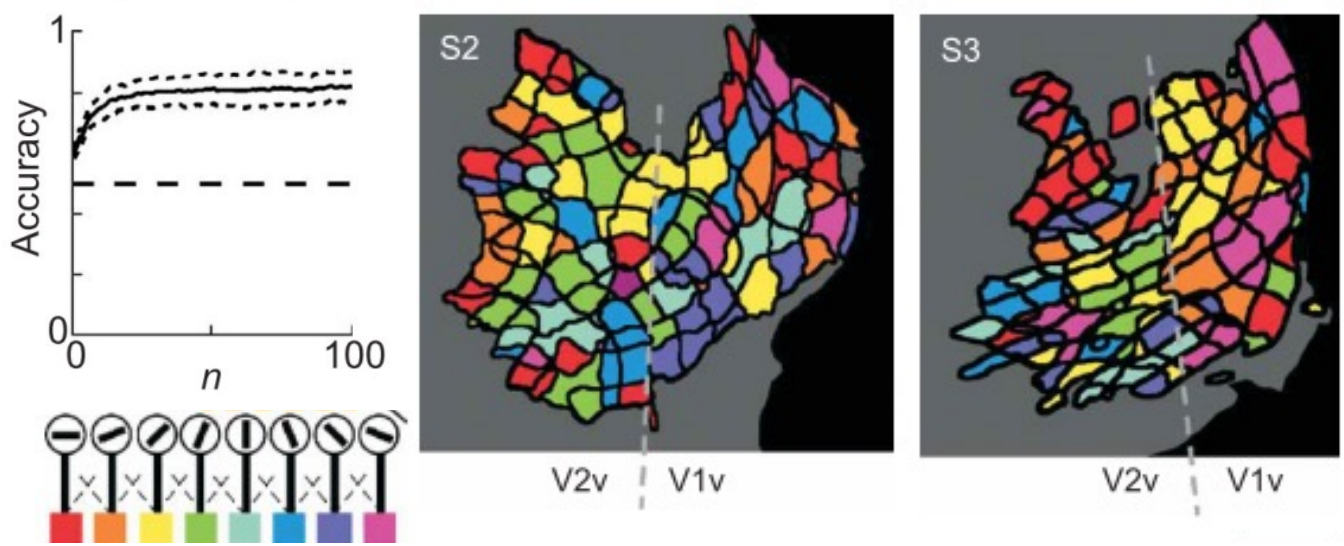


Li et al (2009) - Neuron

Why Multivariate Pattern Analysis?

2. **Higher specificity** allows studying representational content rather than general activation

Example: Representation of orientations in visual cortex



Kamitani & Tong (2005) – Nat Neurosci, Haynes & Rees (2005) – Nat Neurosci

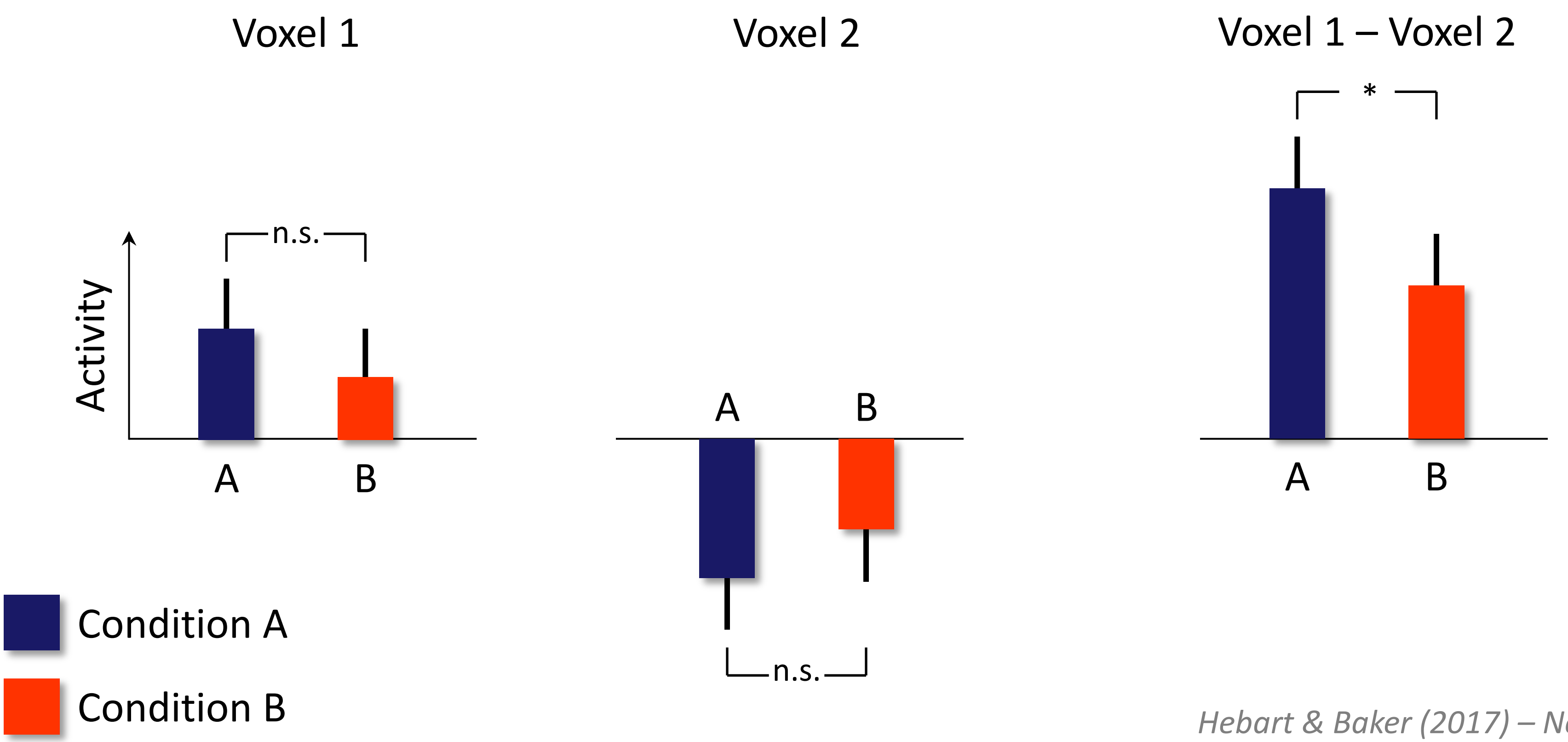
How does MVPA work?

three principles

How Does MVPA Work?

1. Weak information can be combined across voxels

→ Multivariate analysis can enhance signal

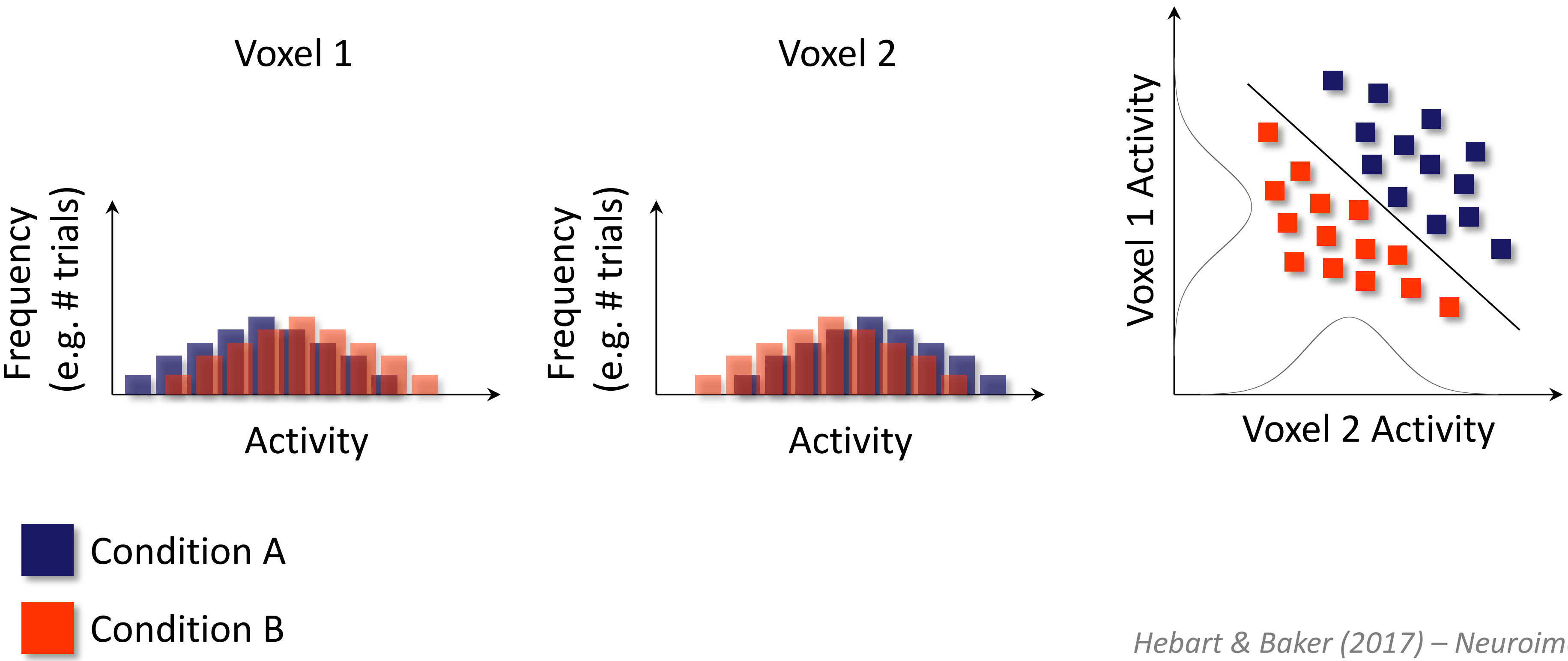


Hebart & Baker (2017) – Neuroimage

How Does MVPA Work?

2. Covariation of voxel information can be used

→ Multivariate analysis can suppress noise

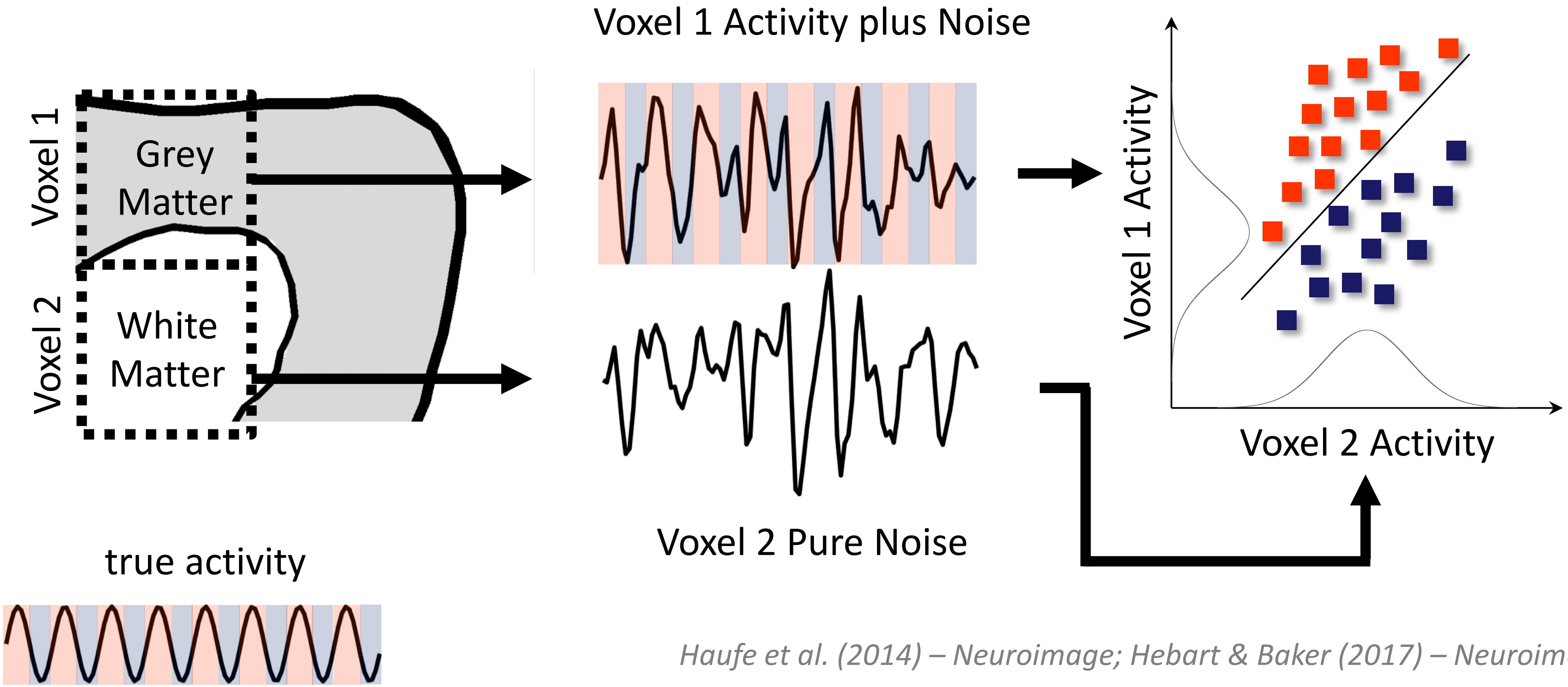


Hebart & Baker (2017) – Neuroimage

How Does How Does MVPA Work? Pattern Analysis Work?

2. Covariation of voxel information can be used

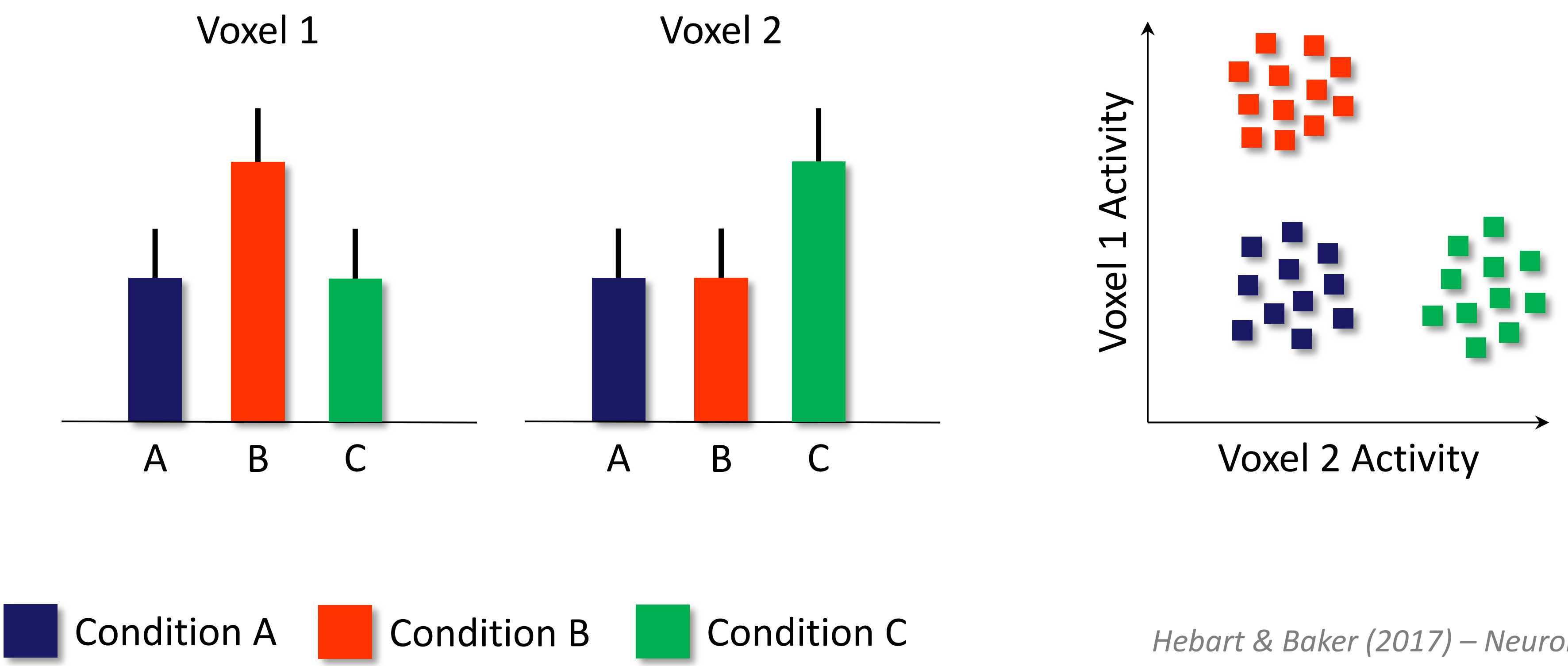
→ Multivariate analysis can suppress noise



Haufe et al. (2014) – Neuroimage; Hebart & Baker (2017) – Neuroimage

How Does MVPA Work?

3. Information becomes accessible that is encoded only in distributed activity patterns

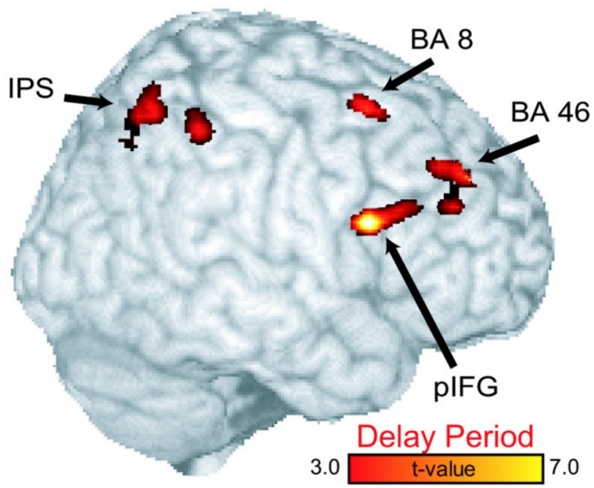


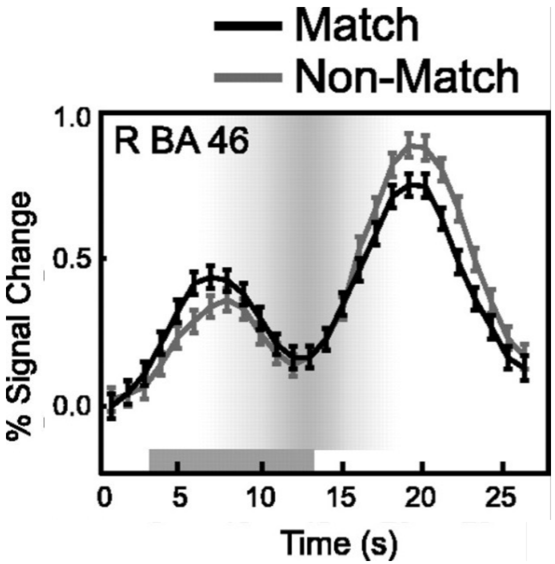
Hebart & Baker (2017) – Neuroimage

ACTIVITY VS INFORMATION

Activity vs. Information

Activity: Tells us about general involvement in cognitive function (e.g. working memory vs. no working memory)

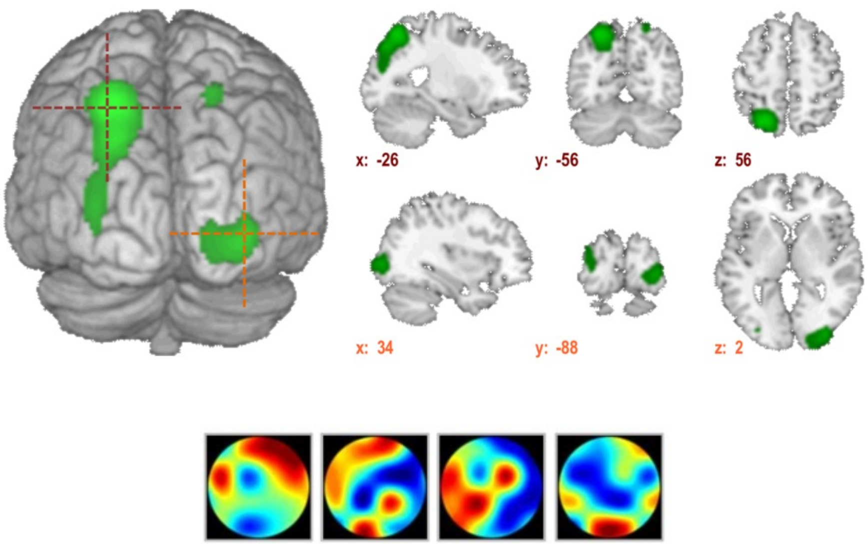


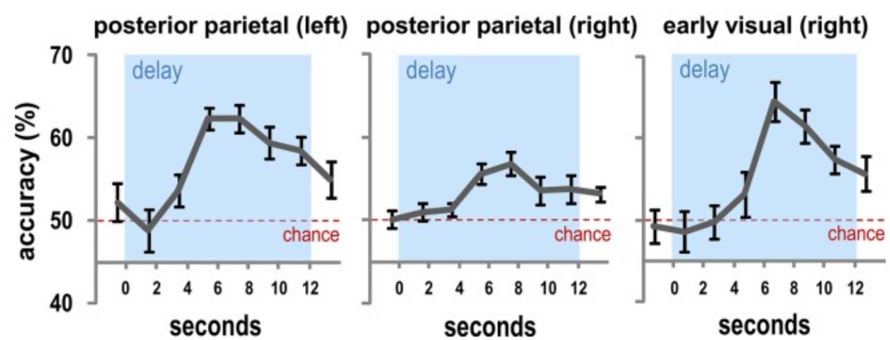


Curtis et al. (2004) – J Neurosci

Activity vs. Information

Information: Tells us about representational content (e.g. memory trace of A vs. memory trace of B)





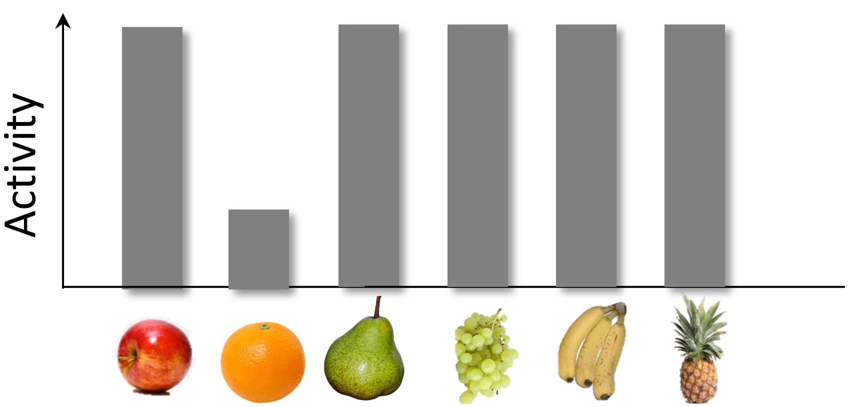
Christophel, Hebart, & Haynes (2012) – J Neurosci, see also Christophel et al (2017) – TiCS

Different Methodological Philosophies

Classical approach: More active = more involved

MVPA: More distinct = more involved

Thought experiment:



Classical approach: Brain region responds to all fruit but oranges

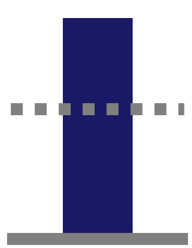
MVPA: Brain region carries information about oranges
(when contrasted with fruit)

Hebart & Baker (2017) – Neuroimage

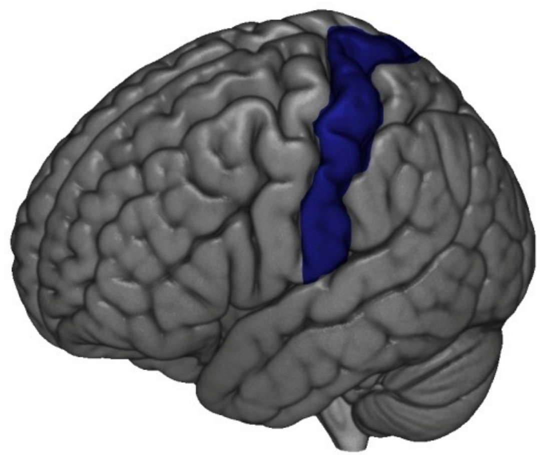
Levels of MVPA analysis

Wholebrain

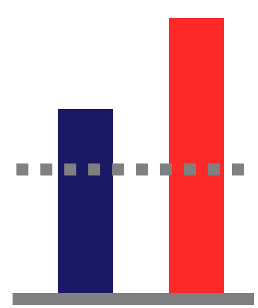
One value per brain



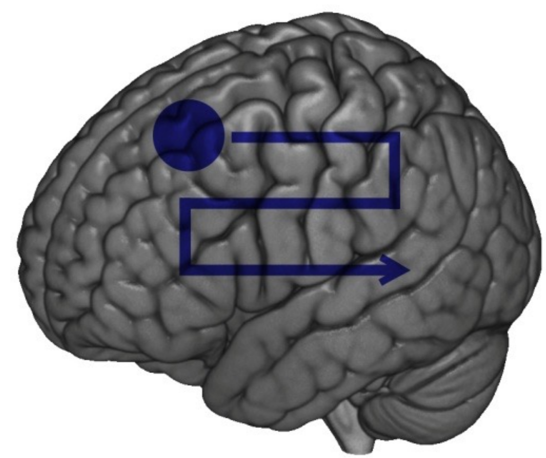
Region of Interest



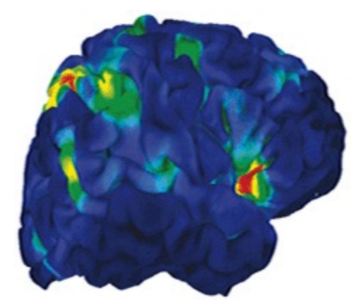
One value per ROI



Searchlight

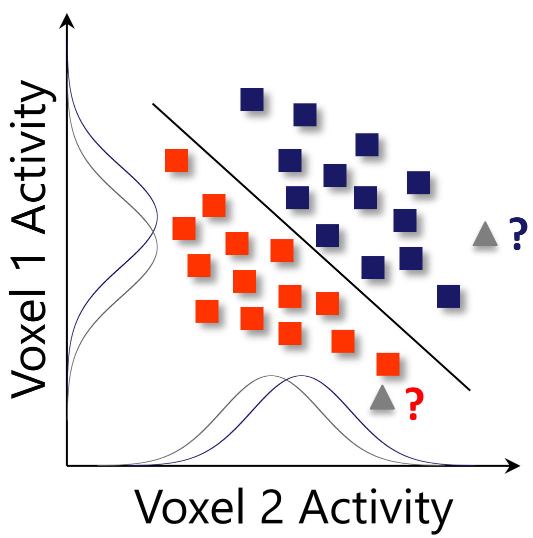


A value per searchlight,
i.e. a map of values

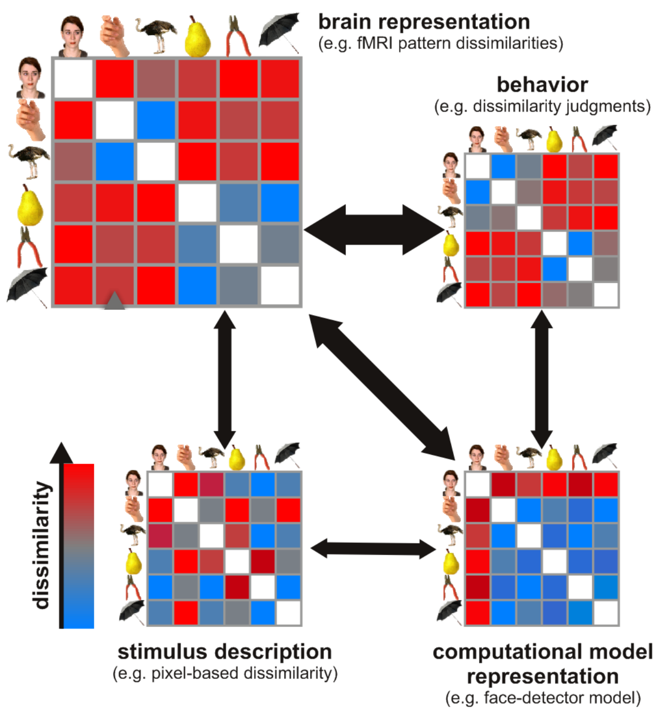


Two main MVPA approaches

Multivariate decoding

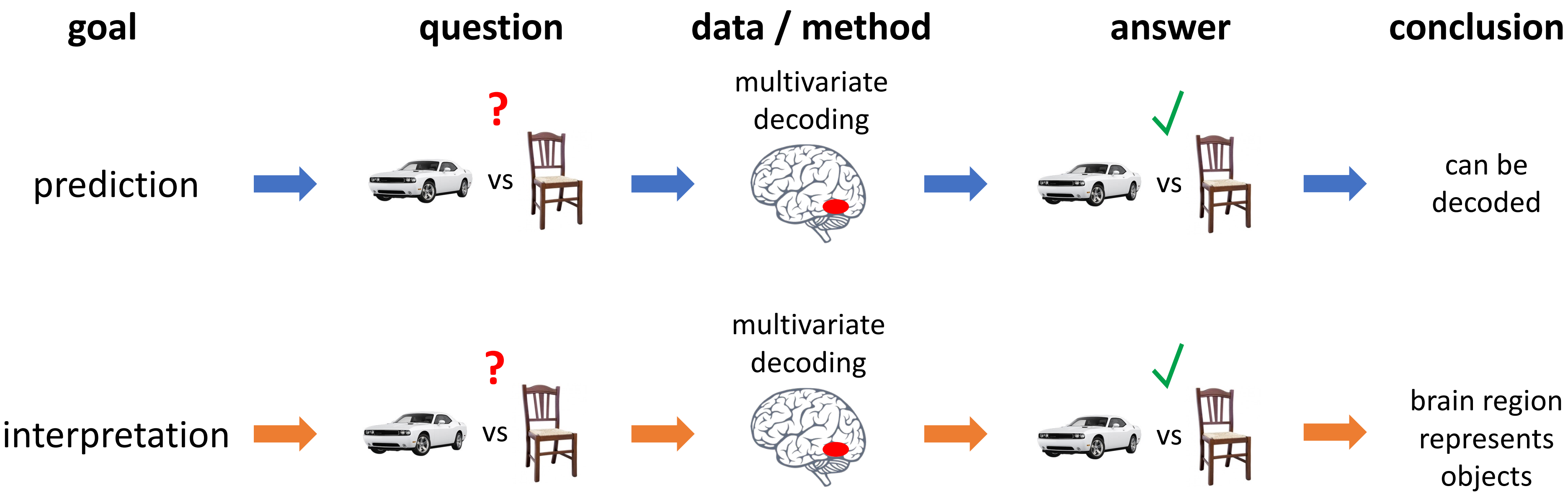


Representational Similarity Analysis



DECODING FOR PREDICTION
VS
DECODING FOR INTERPRETATION

Goals of Decoding: Prediction vs. Interpretation



Hebart & Baker (2017) – Neuroimage

Goals of Decoding: Prediction

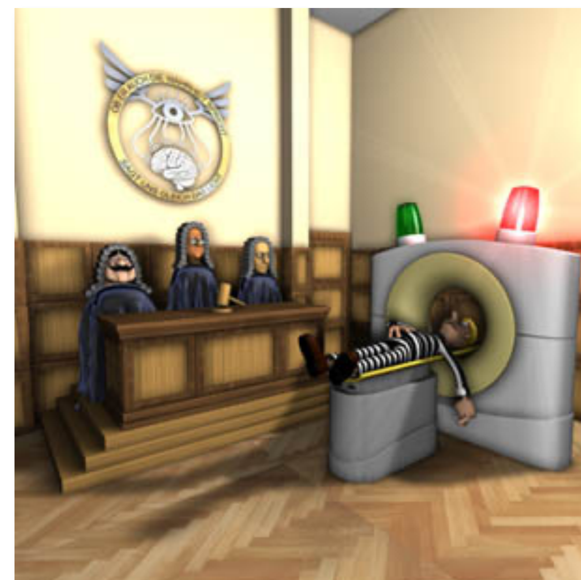
Prediction: Goal is to maximize future correct predictions

→ Any information is useful as long as it increases accuracy

Medical Diagnosis

Brain-Computer-Interface

Lie Detection



http://www.heise.de/tp/artikel/26/26759/26759_2.jpg, http://thinktechuk.files.wordpress.com/2011/10/97877.jpg, http://www.poly.edu/sites/polyproto.poly.edu/files/pressrelease/MRI_Alzheimers_Research.jpg

Goals of Decoding: Prediction

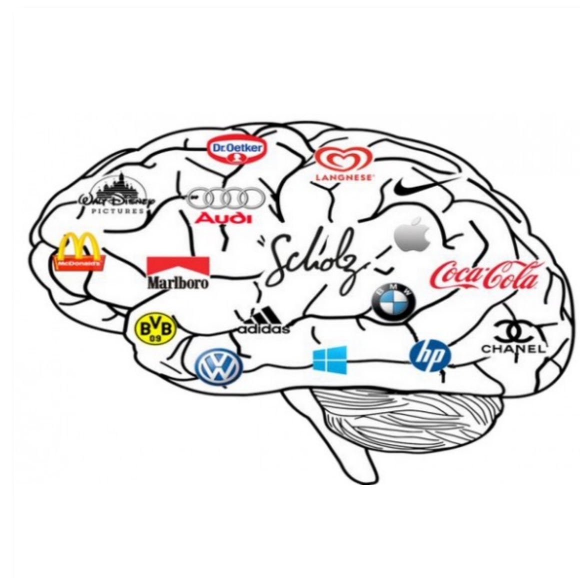
Prediction: Goal is to maximize future correct predictions

→ Any information is useful as long as it increases accuracy

Medical Diagnosis

Brain-Computer-Interface

Neuromarketing

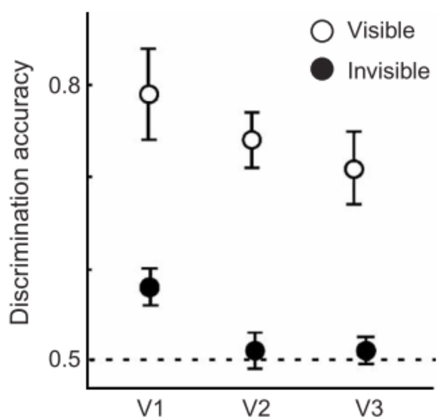


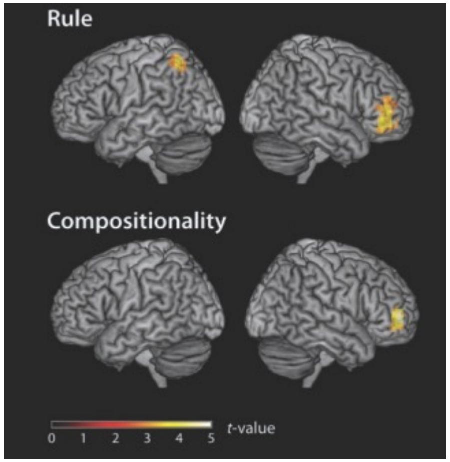
Goals of Decoding: Interpretation

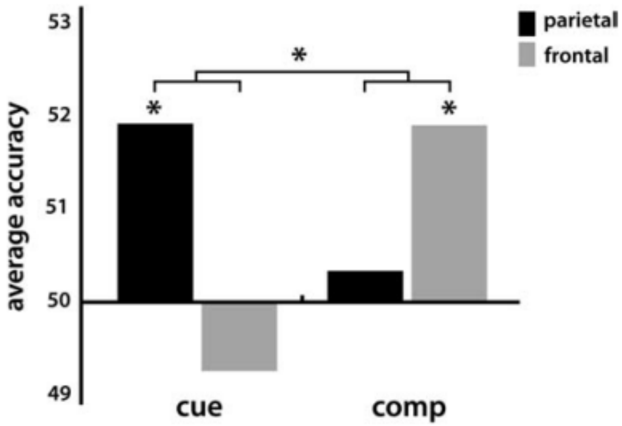
Interpretation: Is there information about XYZ?

→ Sufficient to show above chance accuracy (statistically!)

→ Not all information sources ok, need to rule out confounds



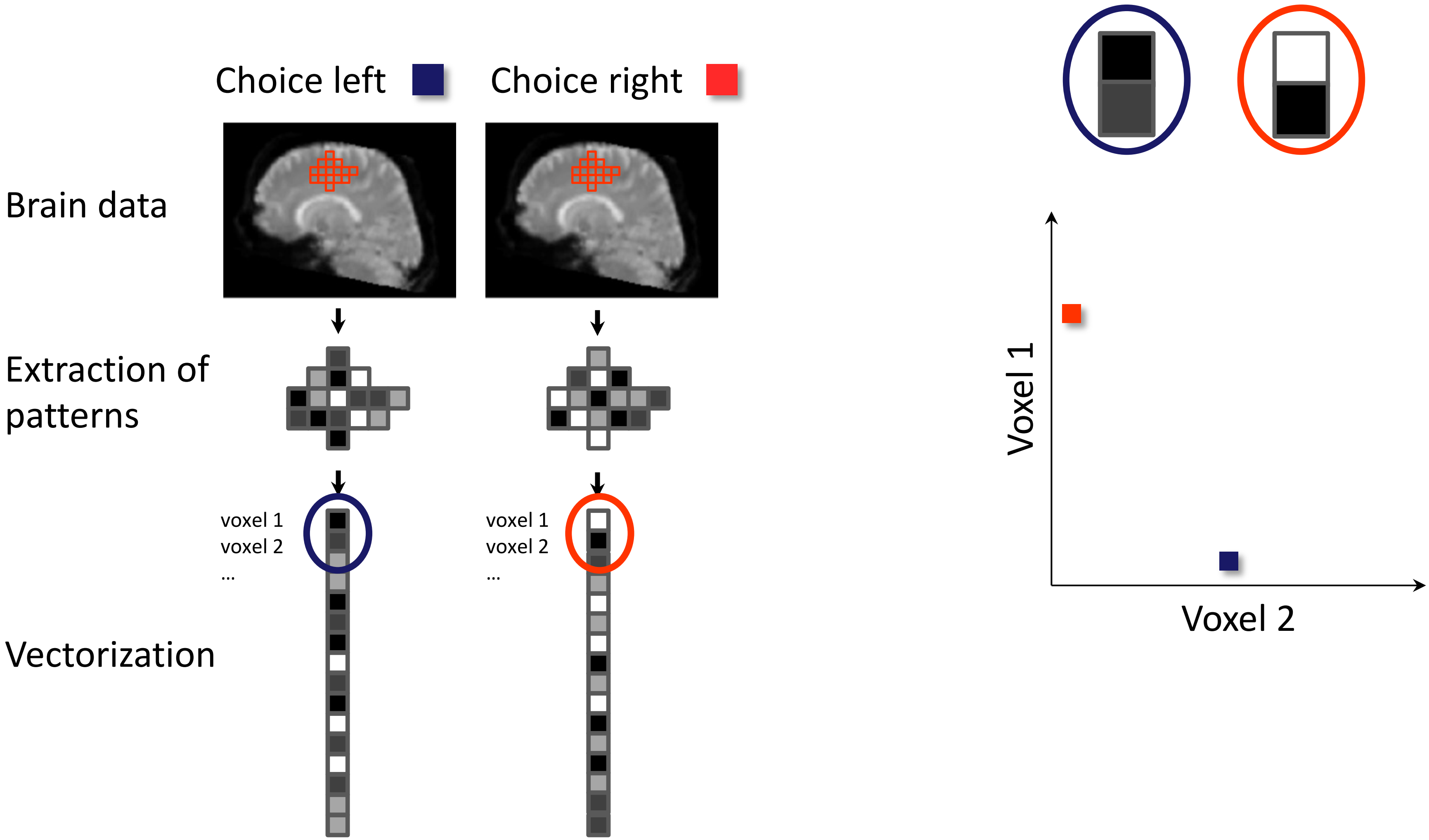




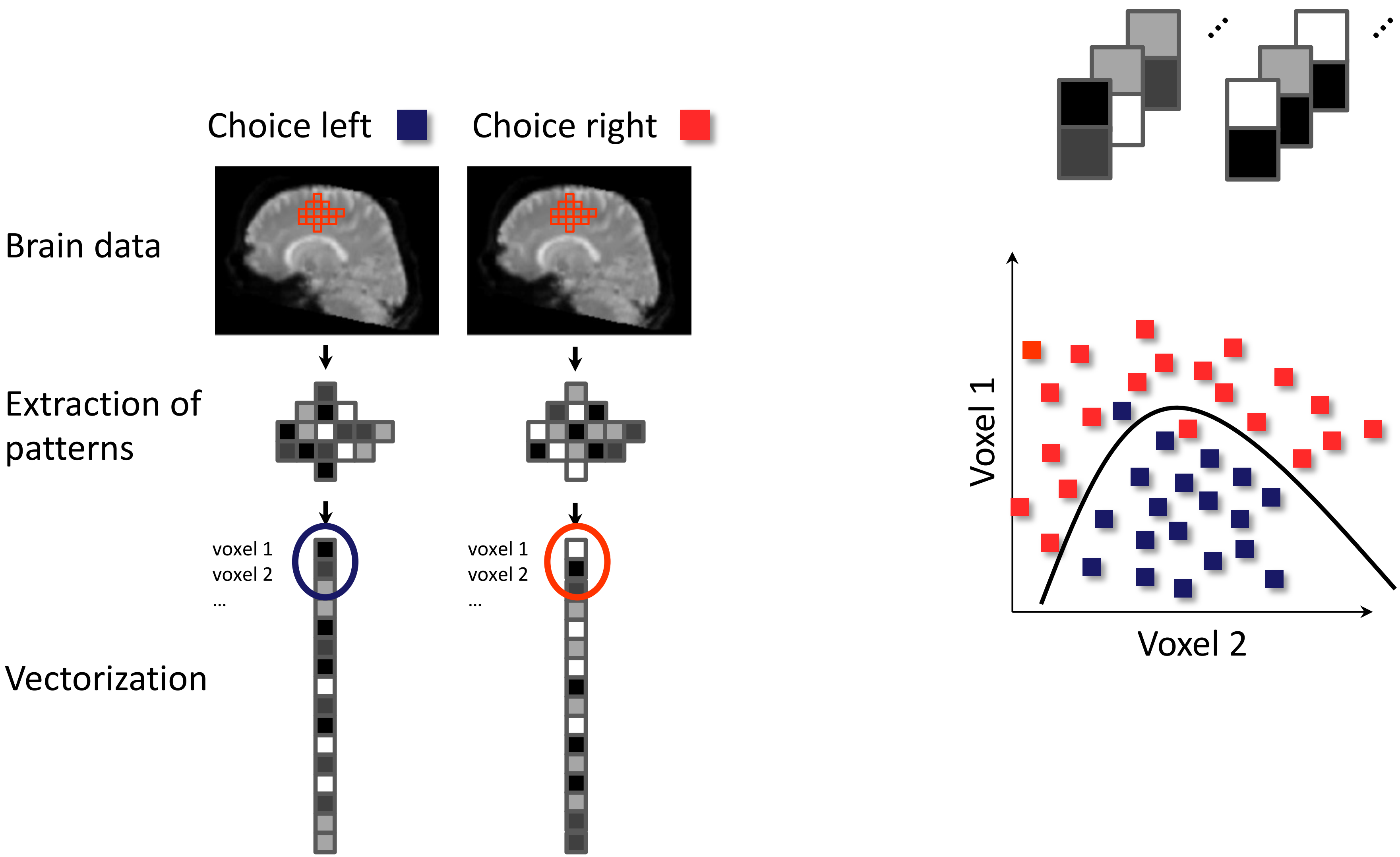
Haynes & Rees (2005) – Nat Neurosci; Reverberi, Görgen, & Haynes (2012) – Cereb Cortex

HOW DOES MULTIVARIATE DECODING WORK?

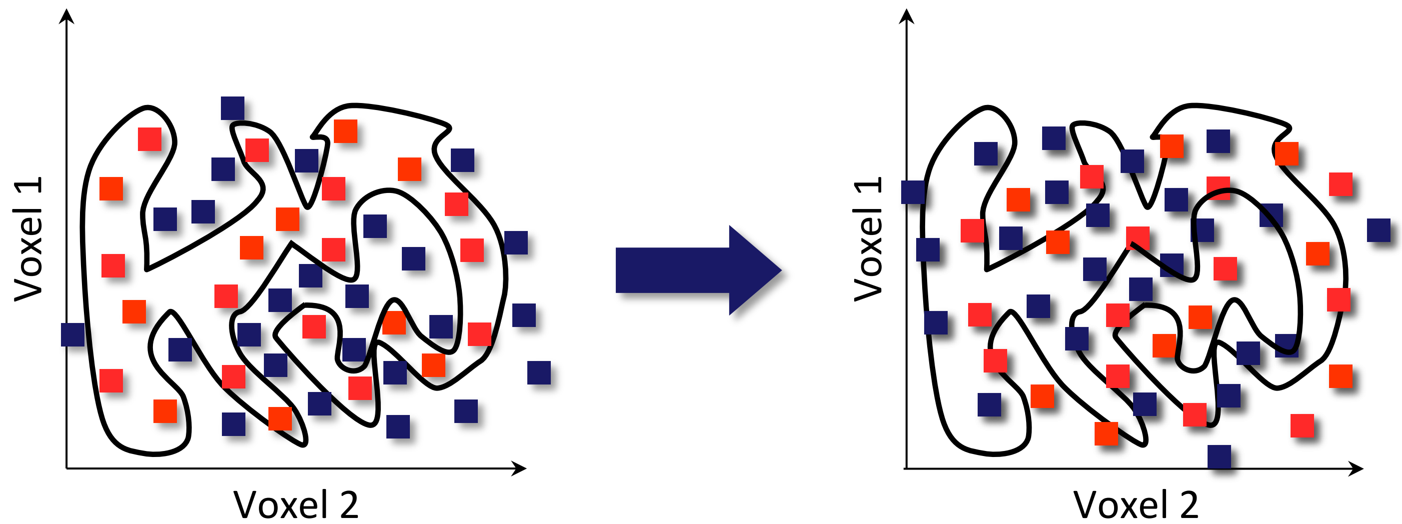
Classification Overview: Example



Classification Overview: Example



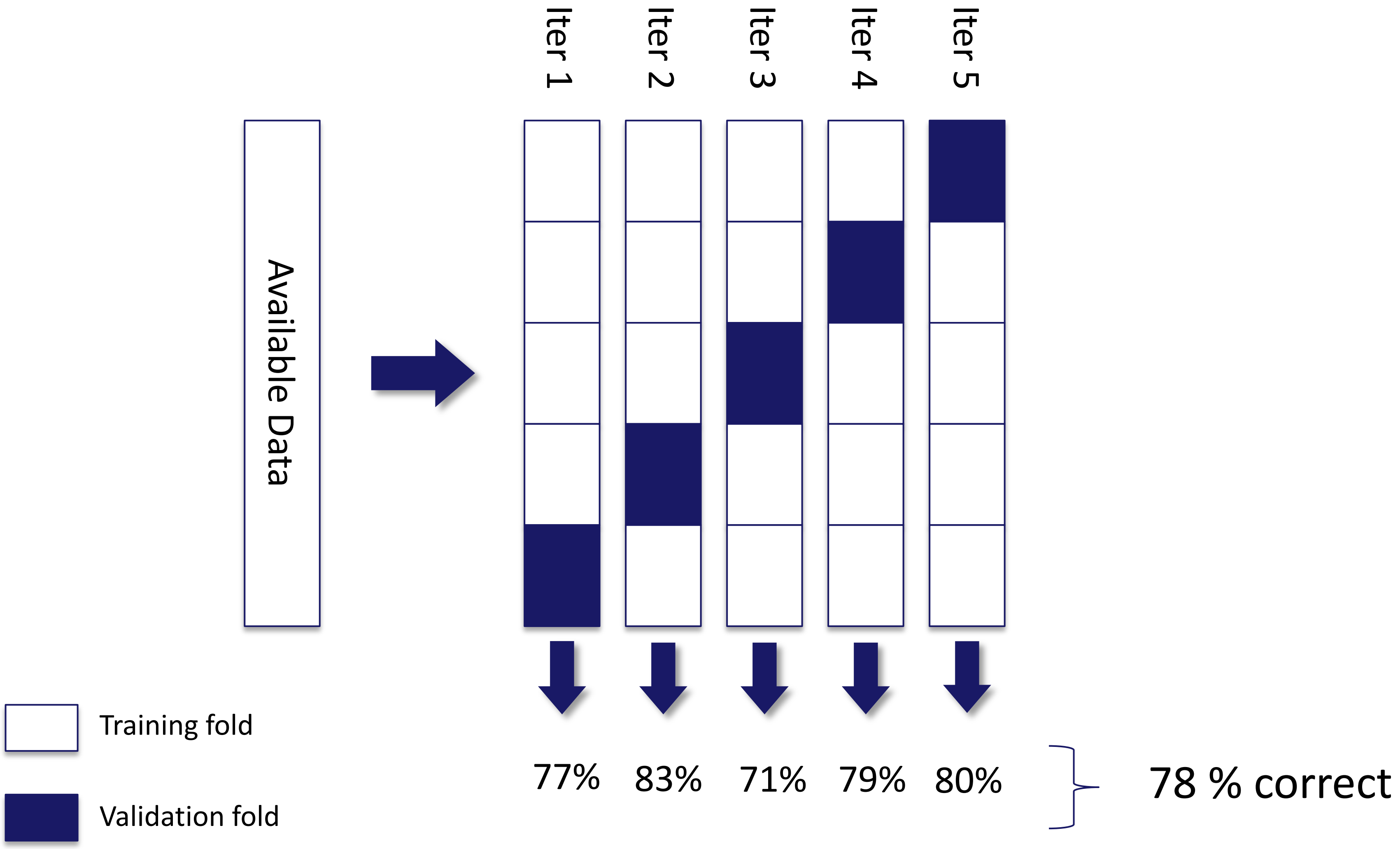
Why Train and Test a Classifier?



- Goal of classification: Finding a model that generalizes beyond noise in the data

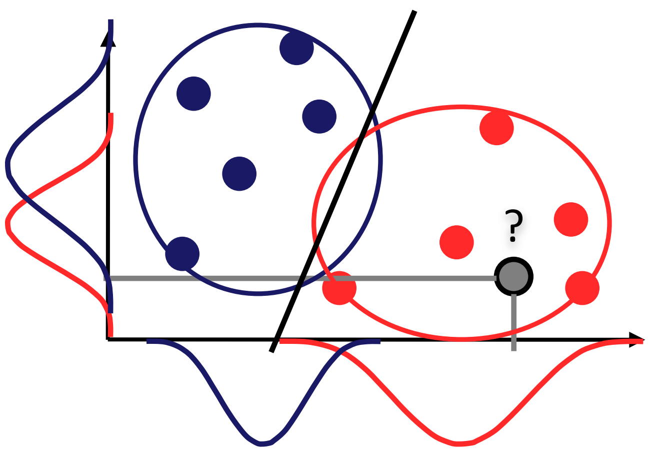
- Way of testing generalization: Test classifier on new data = out-of-sample generalization

Cross-validation



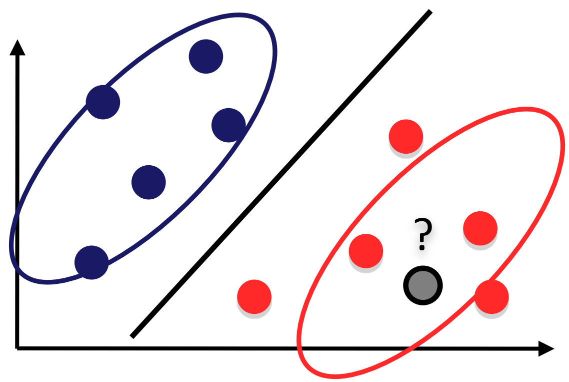
Typical linear classifiers

Gaussian Naïve Bayes



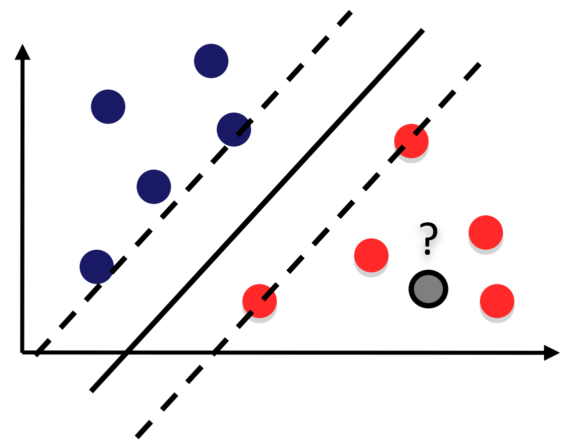
Ignores covariance between voxels

Linear Discriminant Analysis



Considers covariance between voxels

Support Vector Machine

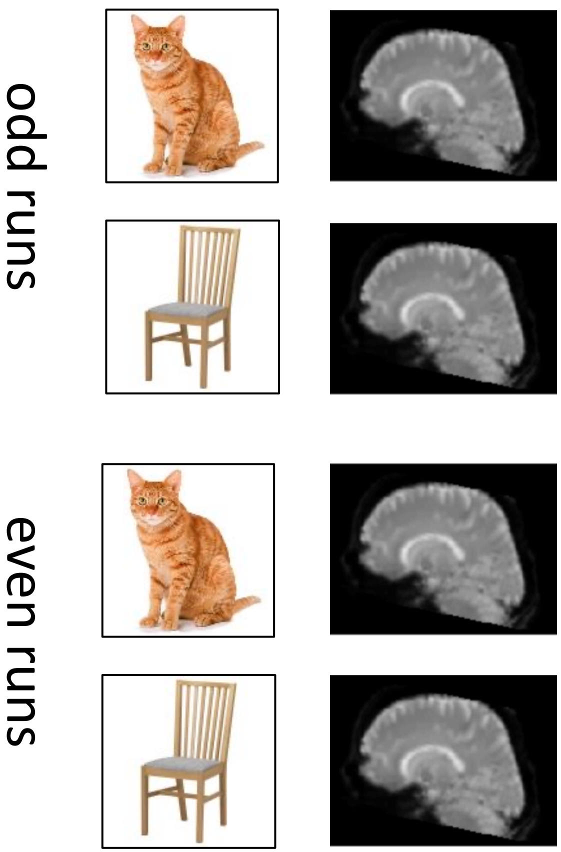


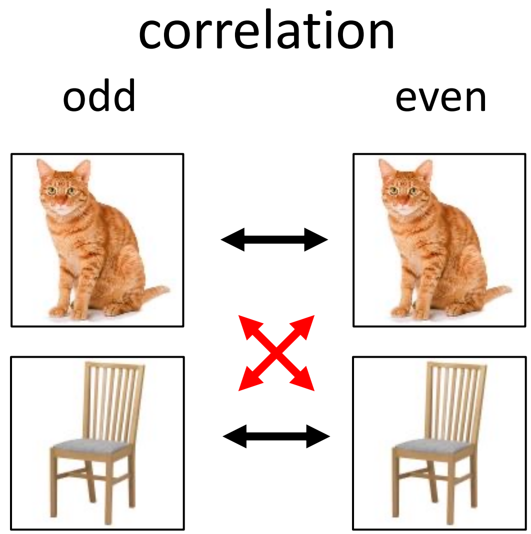
Maximizes margin (distance between closest points of different classes)

- Linear classifiers are the most commonly used classifiers in MVPA
- All share the same formula $y = \Sigma w_i x_i$ but differ in how they find parameters w

Correlation-based classifier

Very simple classifier: find maximal pattern correlation

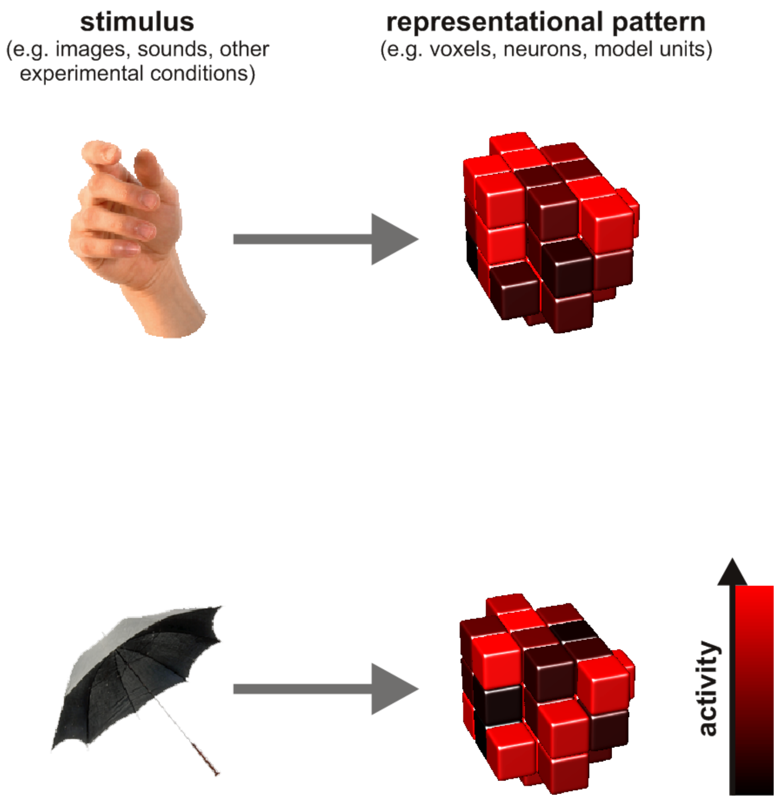




$r_{withincategory} > r_{betweencategory}$?

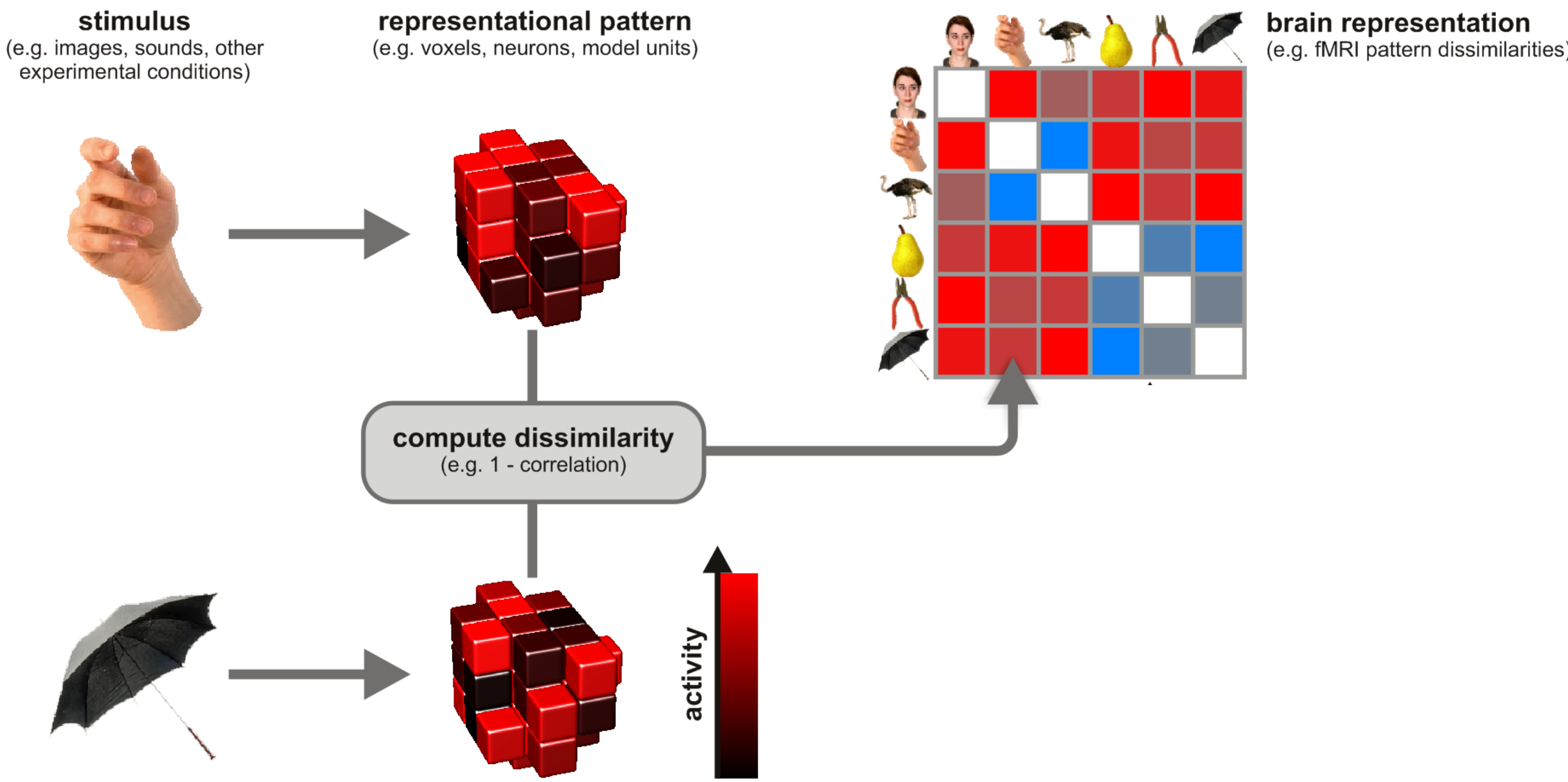
REPRESENTATIONAL SIMILARITY ANALYSIS

Representational similarity analysis



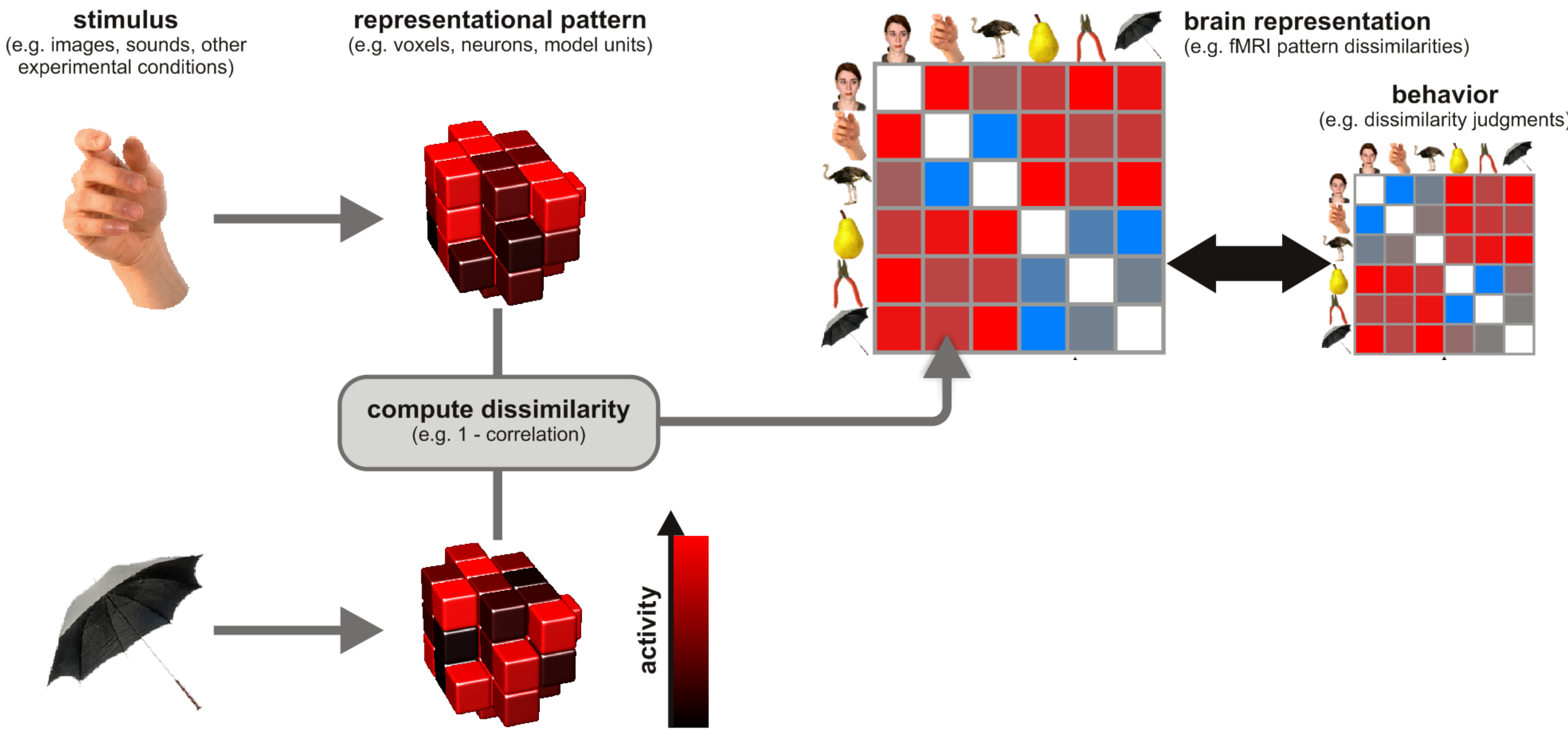
Kriegeskorte & Kievit (2013) – TICS

Representational similarity analysis



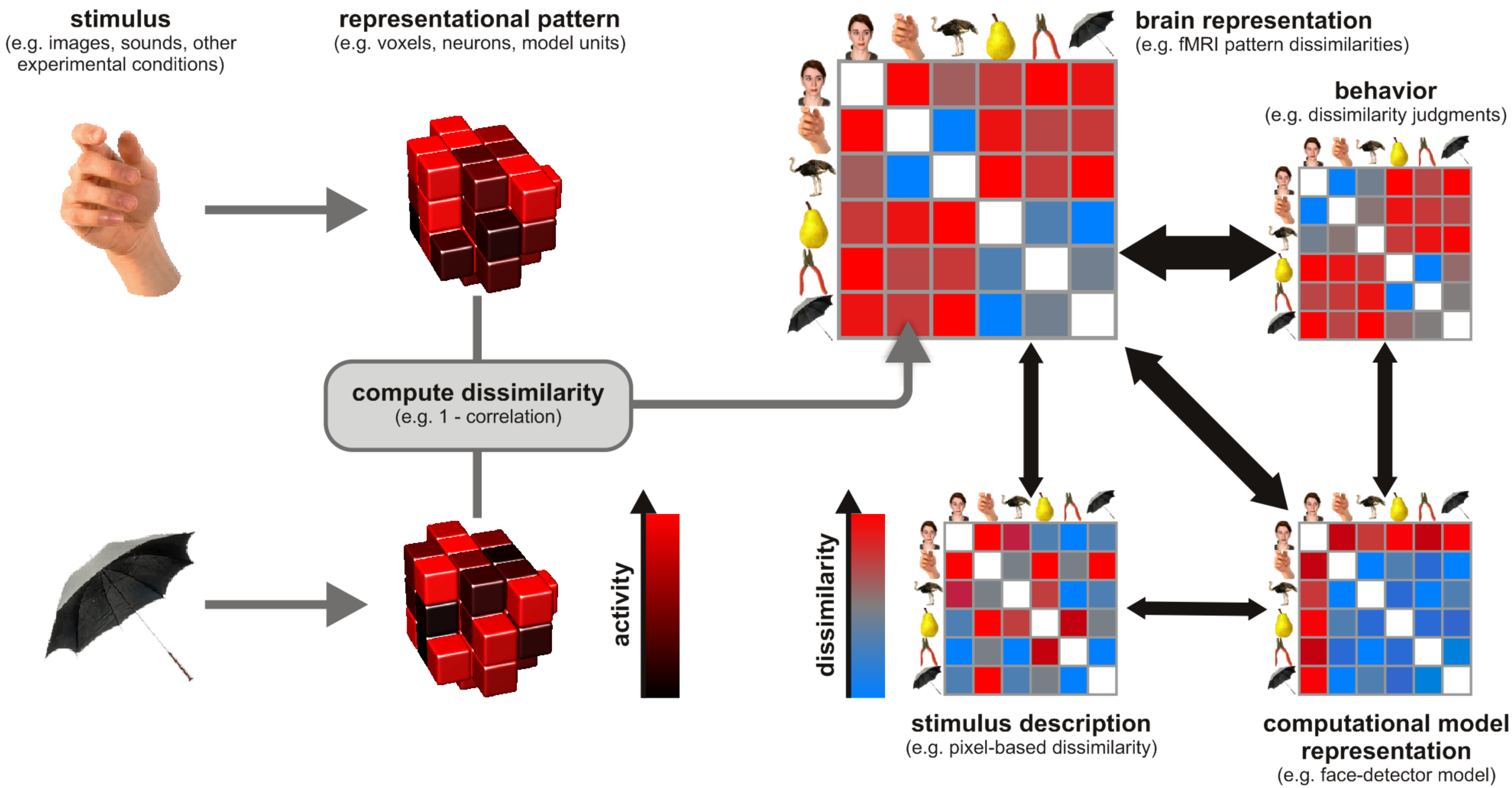
Kriegeskorte & Kievit (2013) – TICS

Representational similarity analysis



Kriegeskorte & Kievit (2013) – TICS

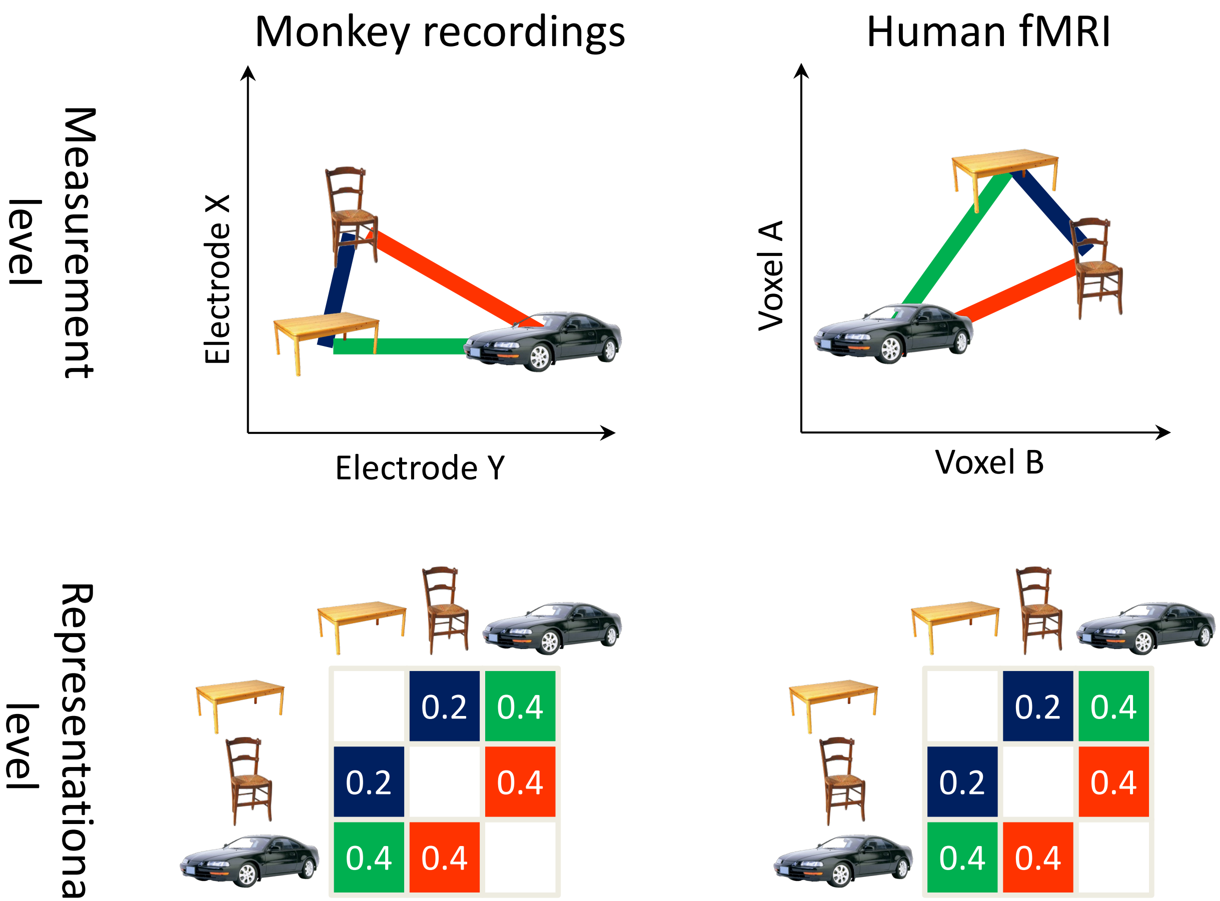
Representational similarity analysis



Kriegeskorte & Kievit (2013) – TICS

Representational similarity analysis

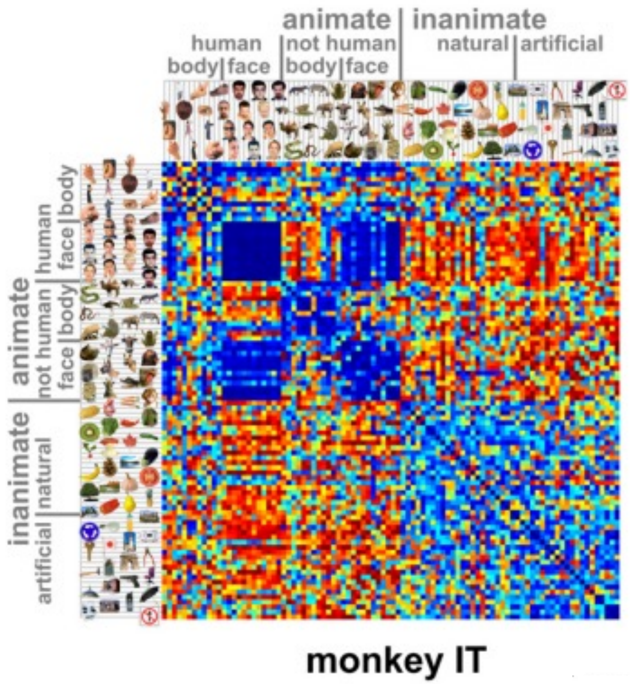
Idea of a representational geometry



Kriegeskorte & Kievit (2013) – TICS

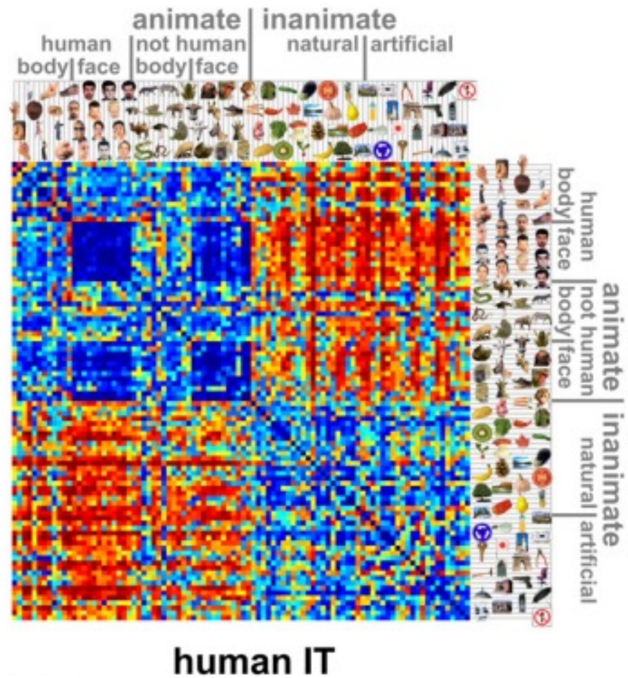
Representational Similarity Analysis

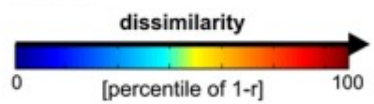
Monkey Dissimilarity Matrix



Comparison

Human Dissimilarity Matrix





Kriegeskorte et al (2008) – Neuron

THREE CONCRETE CHALLENGES

What is the best information measure?

Multivariate decoding

- Area-under-the-curve (AUC) less prone to bias shift than decoding accuracy but no formal comparison exists
- Continuous-valued methods (e.g. signed decision value) perhaps more powerful than accuracy or AUC?
- Alternative encoding-based approach (cross-validated Mahalanobis distance) may unite best of both worlds?

Representational similarity analysis

- Currently unclear whether Pearson correlation or Euclidean distance is better
- Good model of noise covariance may improve fit

Walther et al (2016) – Neuroimage; Guggenmos et al (2018) – Neuroimage

How to deal with multiple experimental variables / confounds?

Problem:

- Multivariate decoding and RSA and optimized for test of single variables at a time
- Interactions in factorial designs (e.g. 2 x 2), unique effects of variables and confounds are difficult to deal with

Possible solutions:

- RSA: Multiple regression approach (but additional assumptions)
- Decoding: Confound regression during cross-validation

Alternative possible solution:

- Run multivariate GLM

Allefeld et al (2014) – Neuroimage; Snoek, Miletic, Scholte (2018) – bioRxiv

What is the correct statistical test for MVPA analyses?

Current standard:

- At MVPA-level: permutation test or randomization test (because assumptions of classical tests don't apply)
- At second-level: t-test, ANOVA, sign-permutation test

Problems:

- MVPA-level: Assumption of exchangeability often violated in permutation tests (specifically for trial-wise analyses)
- Second-level: Random effects test against chance at second-level not valid

Possible solutions:

- MVPA-level: Development of new permutation test for dependent data
- Second-level: Prevalence inference (proportion of significant subjects larger than chance?) or random effects test against baseline condition / region

THREE LESS CONCRETE CHALLENGES

1. How many representational dimensions can we capture with neuroimaging?

2. Is the information we extract from brain patterns used by the brain or just an epiphenomenon?

Williams et al (2007) – Nature Neurosci; Ritchie et al (2016) – Brit J Philos Sci

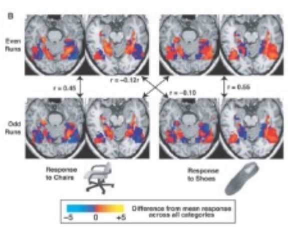
3. Can we understand the brain if we mix the activation-based philosophy with the information-based philosophy?

Hebart & Baker (2017) – Neuroimage

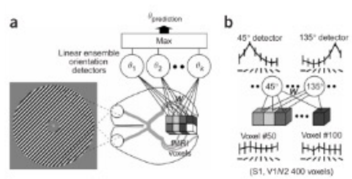
Milestones of MVPA

Edelman et al (1998)

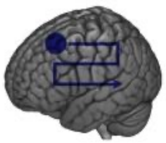
Haxby et al (2001)



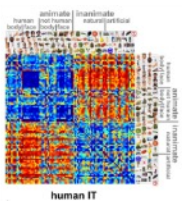
Kamitani & Tong (2005)
Haynes & Rees (2005)



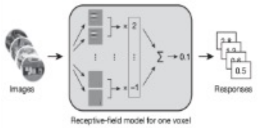
Kriegeskorte et al (2006)



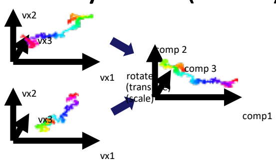
Kriegeskorte et al (2008)



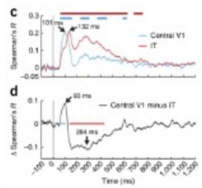
Kay & Gallant (2008)



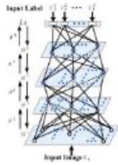
Haxby et al (2011)



Cichy et al (2014)



e.g. Di Carlo



1998 2001 2005 2006 2008 2011 2014 2015/2016

"first" MVPA study

first multivariate decoding study

popularization of multivariate decoding

searchlight approach

representational similarity analysis

model-based encoding methods

hyperalignment

combination of fMRI and MEG using RSA

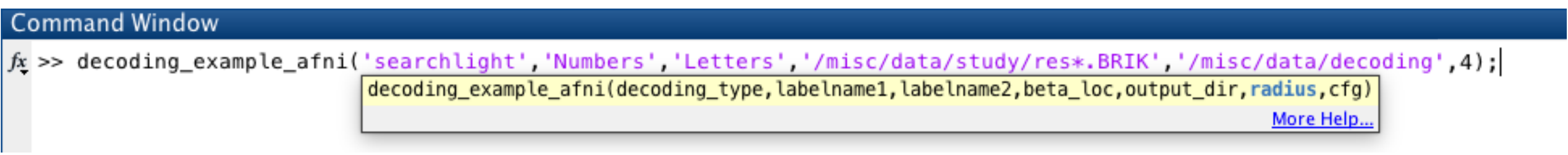
Demonstration of homologies in feed-forward architecture btw artificial neural networks and visual cortex

Summary

- MVPA is often more sensitive than classical univariate approaches and can reveal representational content
- MVPA works by combining information across voxels and using their covariation
- Investigating brain activity vs. informational content are two different approaches
- We use cross-validation to reduce avoid bias from overfitting

The Decoding Toolbox

- Fast and easy to use MVPA software package in Matlab (for Python, we recommend PyMVPA, Scikit-Learn, and MNE Python)
- Provides searchlight, ROI and wholebrain analyses
- Comes with a wide range of options, classifiers and similarity analysis
- Runs with SPM and AFNI



https://sites.google.com/site/tdtdecodingtoolbox/

Hebart MN*, Görgen K*, Haynes JD (2015). The Decoding Toolbox (TDT): A versatile software package for multivariate analyses of functional imaging data. Front. Neuroinform. 8:88.

Thank you for your attention

Suggested Readings

Beginners

Tong & Pratte (2012) – Decoding patterns of human brain activity
Haxby et al (2014) – Decoding neural representational spaces
Haynes (2015) – A primer on pattern-based approaches

More Advanced

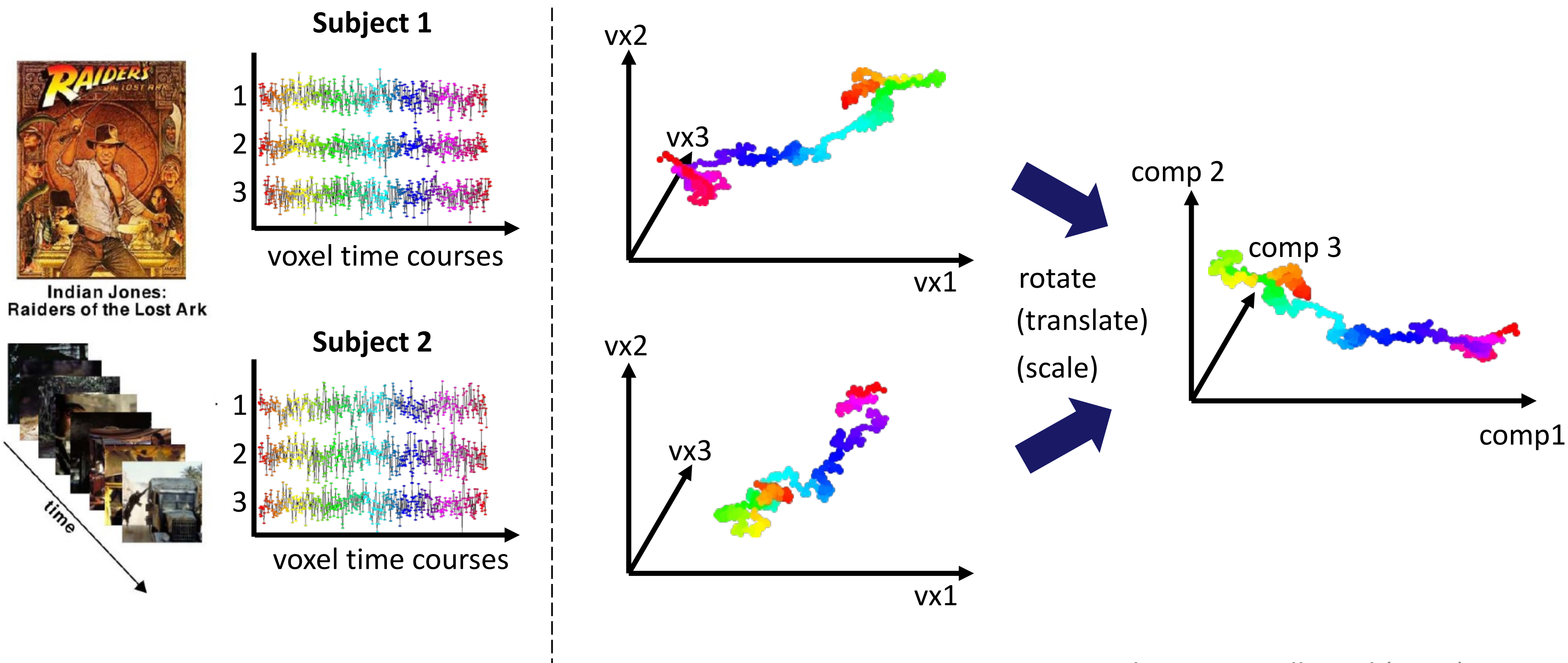
Pereira et al (2009) – Machine learning classifiers and fMRI: a tutorial
Hebart & Baker (2017) – Deconstructing multivariate decoding

Representational models

Kriegeskorte & Kievit (2013) – Representational geometry
Diedrichsen & Kriegeskorte (2017) – Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis

Hyperalignment

- Brings subjects functionally in common space
- Allows predicting one brain from another



Haxby, Guntupalli et al (2011) – Neuron