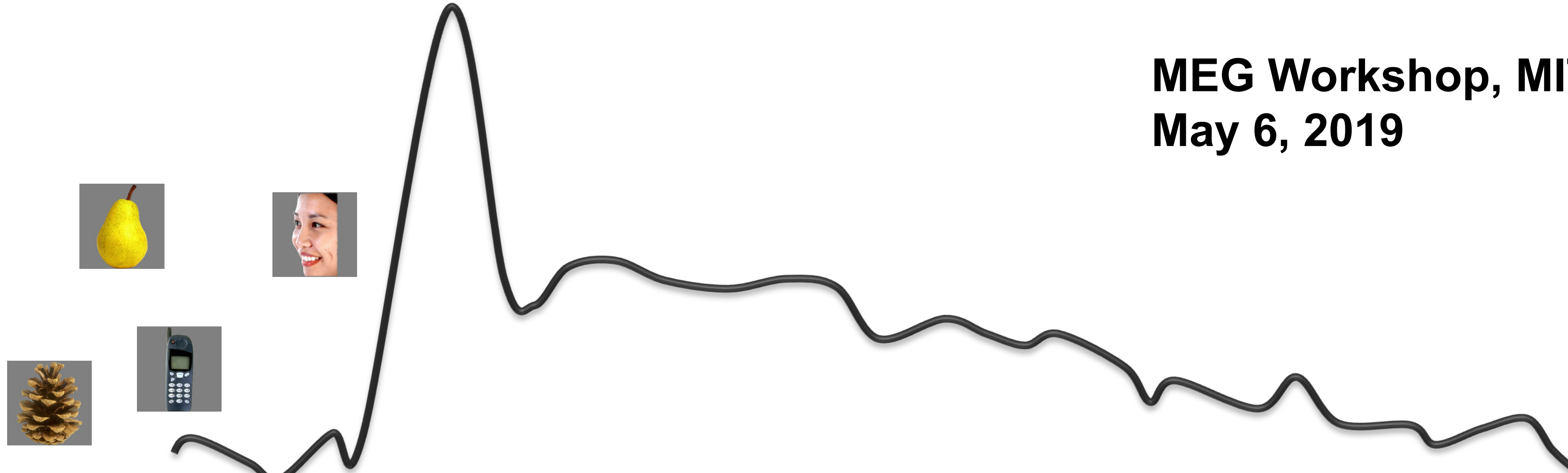
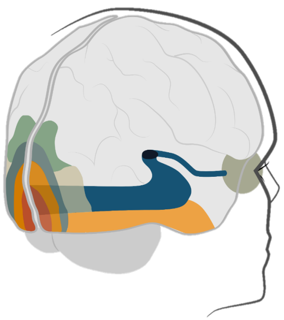


MEG Workshop, MIT
May 6, 2019



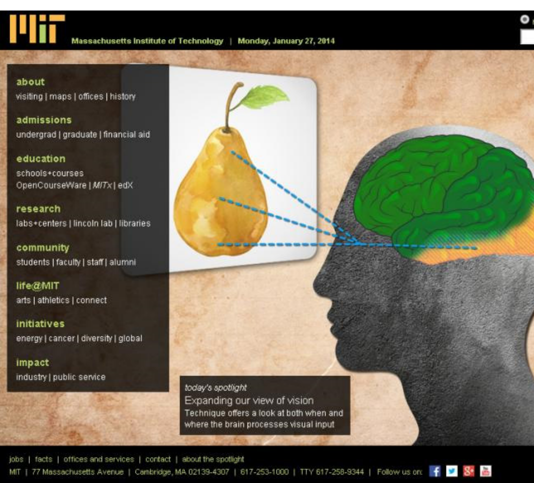
Decoding cognitive function with MEG: Recent advances, challenges, and future prospects

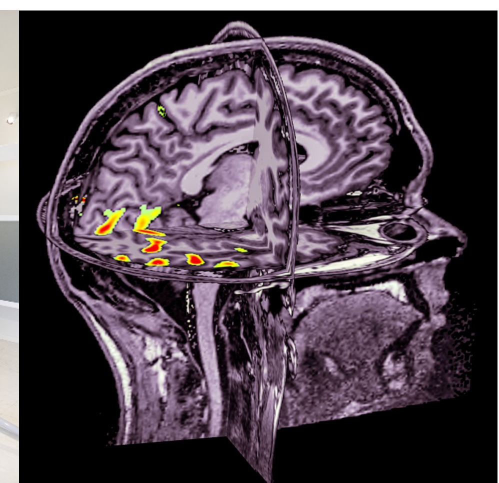
Dimitrios Pantazis

http://pantazislab.org

Principal Research Scientist

McGovern Institute for Brain Research at MIT

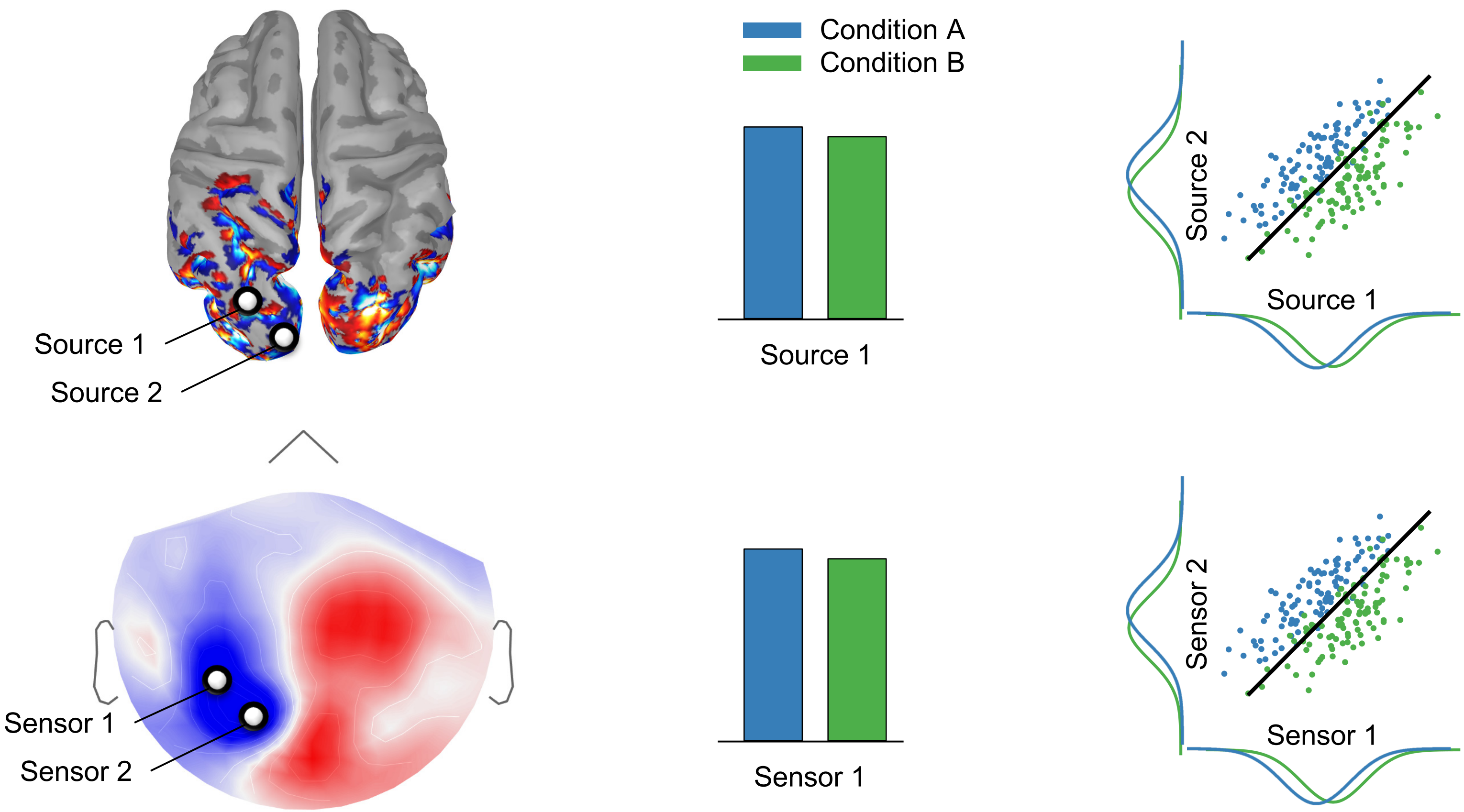




Decoding patterns in neuroimaging

Univariate methods treat each variable (e.g. sensor or source voxel) as an independent piece of data

Multivariate methods extract the information contained in distributed patterns of activity across multiple variables (*multivariate pattern analysis, multivariate classification, multivariate decoding*)



Decoding patterns in neuroimaging

Multivariate methods have been used in fMRI to decode:

- visual features
- visual objects and scenes
- top-down attentional processes
- imagery and working memory
- episodic memory
- phonological representations and language processing
- decisions

Multivariate methods are increasingly popular in MEG.

Prediction: brain-computer interfaces; disease progression; neuroimaging-based lie detectors

Interpretation: study brain function

Information encoded in MEG signals

Recent studies using multivariate pattern classification in MEG/EEG:

Simple visual features, e.g. position and orientation of contrast edges (Carlson et al., 2011; Isik et al., 2013; Ramkumar et al., 2013; Cichy et al., 2015; Wardle et al., 2016; Pantazis et al., 2017; Groen et al., 2017)

Complex visual patterns, e.g. representation of objects and scenes (Isik et al., 2014; Cauchoix et al, 2013; Carlson et al., 2013; Cichy et al., 2014; Clarke et al., 2014; Barragan-Jason et al, 2015; Kaneshiro et al., 2015; Cichy et al., 2016; Nemrodov et al., 2016; Groen et al., 2017; Contini et al., 2017; Cichy et al., 2017; Dima et al., 2018; Grootswagers et al., 2018; Kozunov et al., 2018; Hebart et al, 2018; Khaligh-Razavi et al., 2018; Mohsenzadeh et al., 2018)

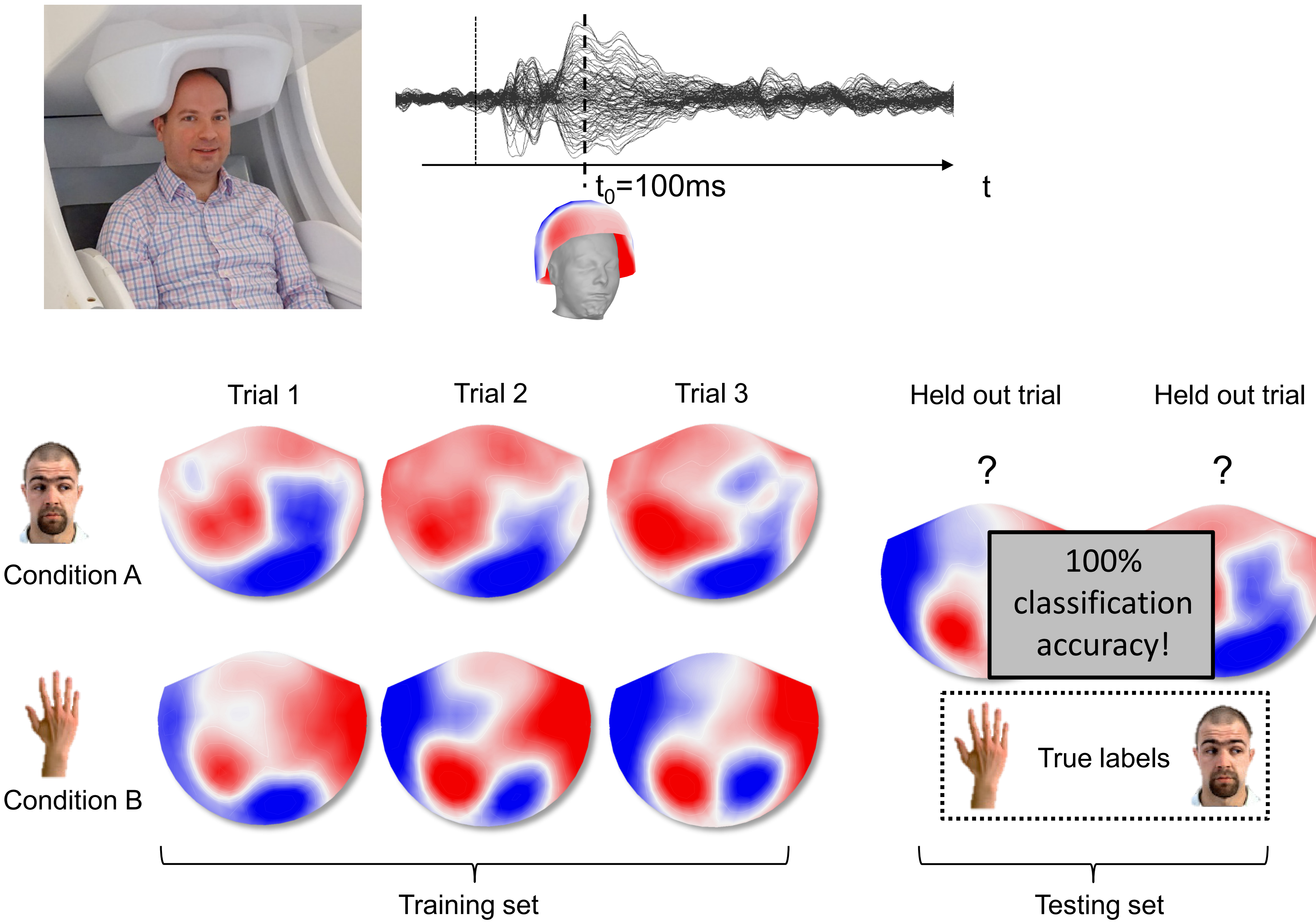
Auditory representations (King et al., 2013 & 2014; Teng et al., 2017)

Temporal maintenance of information and working memory (Carlson et al., 2011; Isik et al., 2014; Cichy et al., 2014; King et al., 2014 & 2016; Pantazis et al., 2017; Spaak et al., 2017)

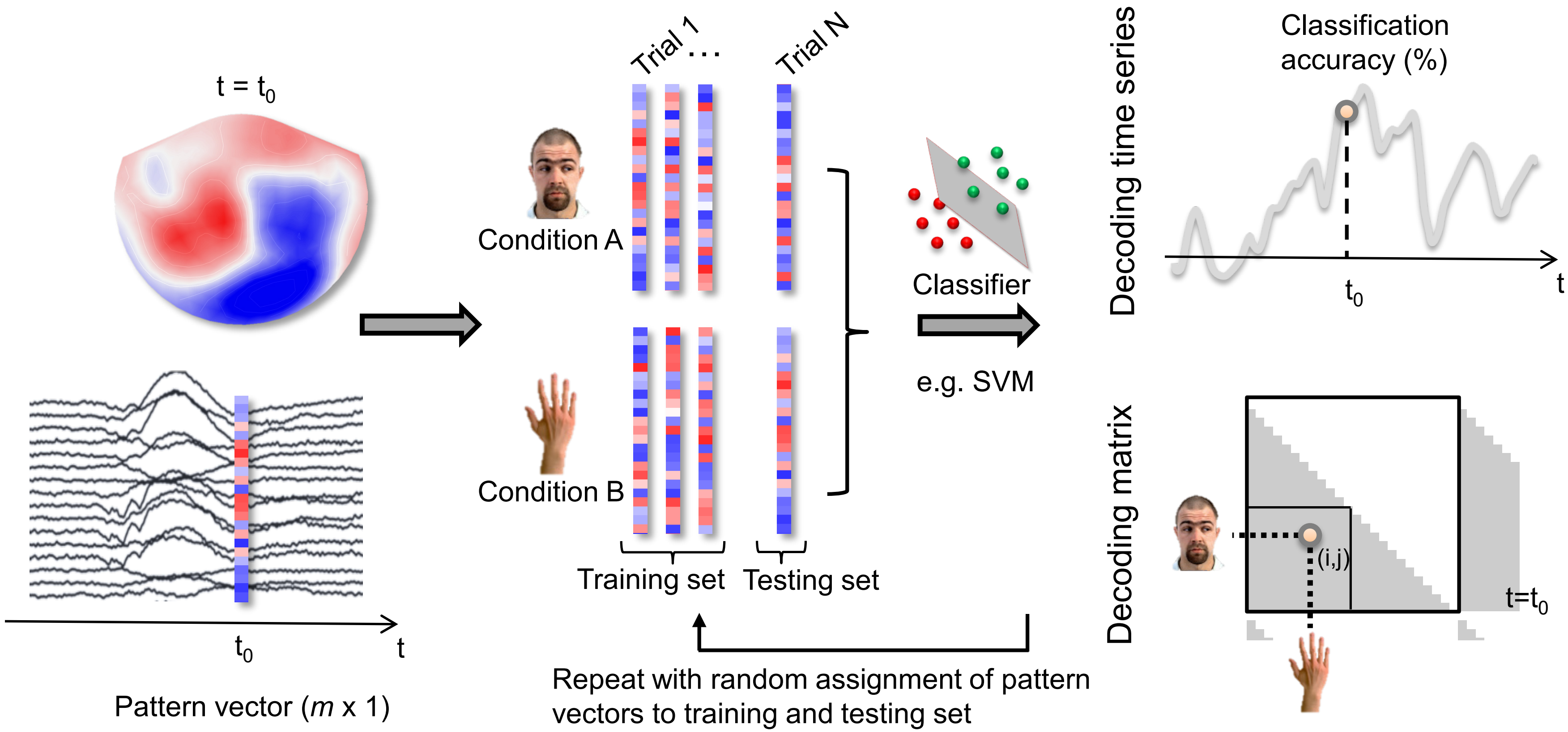
Visual motion (Bekhti et al., 2017), mental arithmetic and numerical symbols (Pinheiro-Chagas et al., 2018; Teichmann et al., 2018).

Methods (Haufe et al., 2014; Kaplan et al., 2015; Cichy and Pantazis, 2017; Hebart et al., 2017; Guggenmos et al, 2018; Vidaurre et al. 2018)

Conceptual framework of MEG decoding

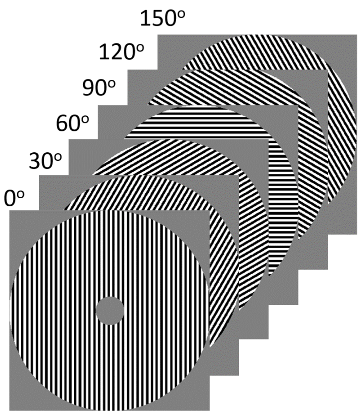


Time-resolved MEG decoding

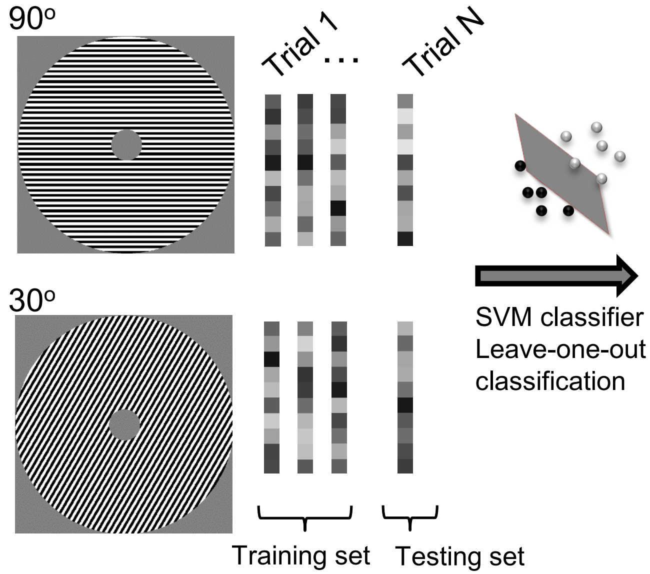


Example: Decoding the orientation of contrast edges

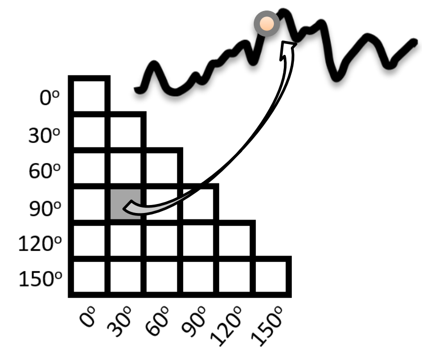
Stimulus set



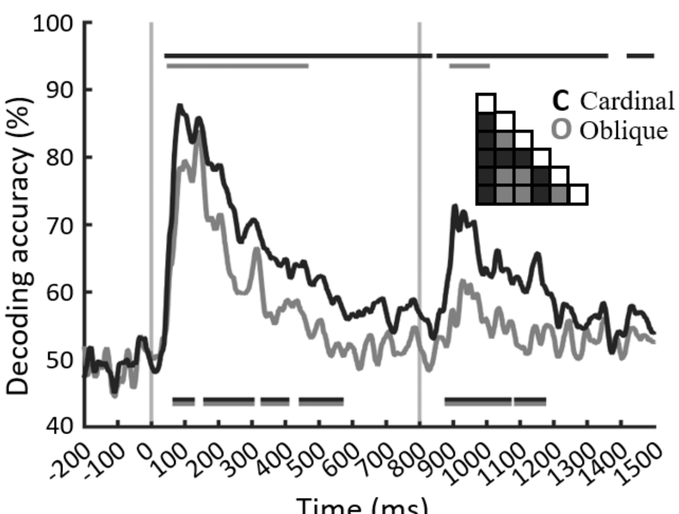
Pairwise pattern classification



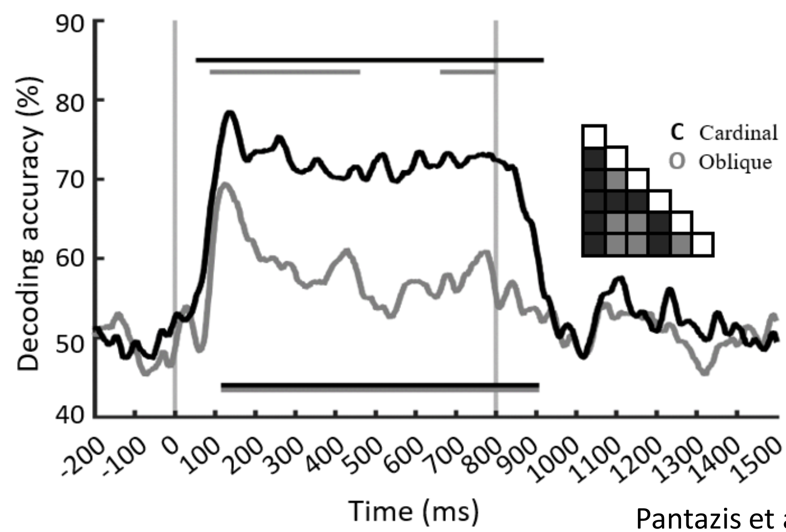
Time-resolved decoding matrix



Decoding orientation from *evoked* responses

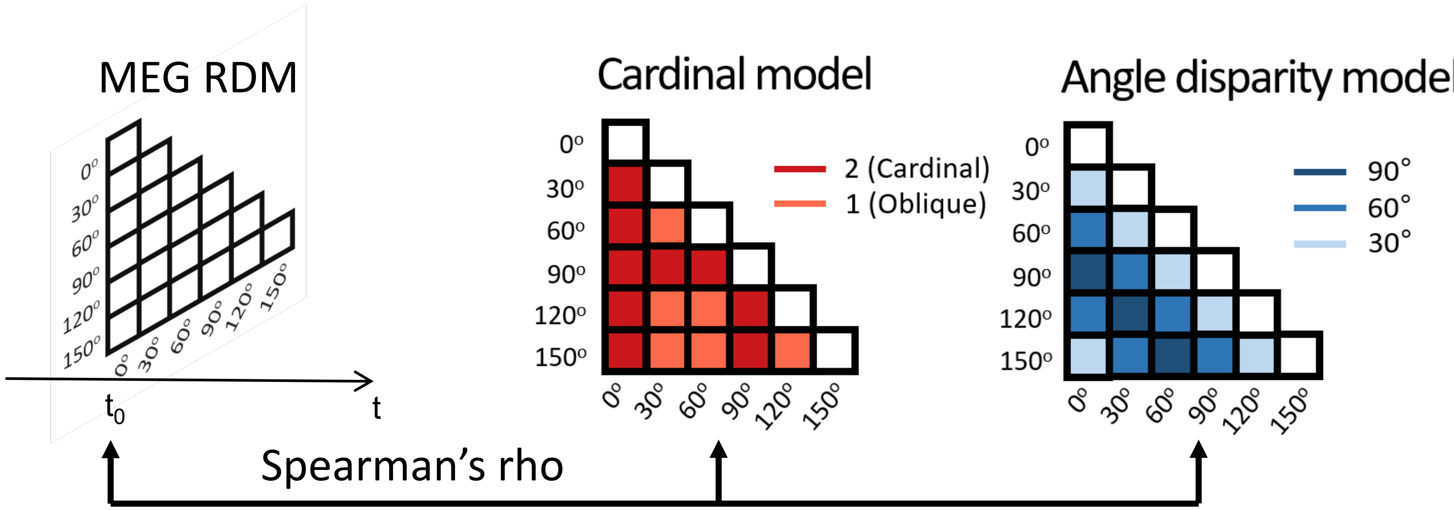


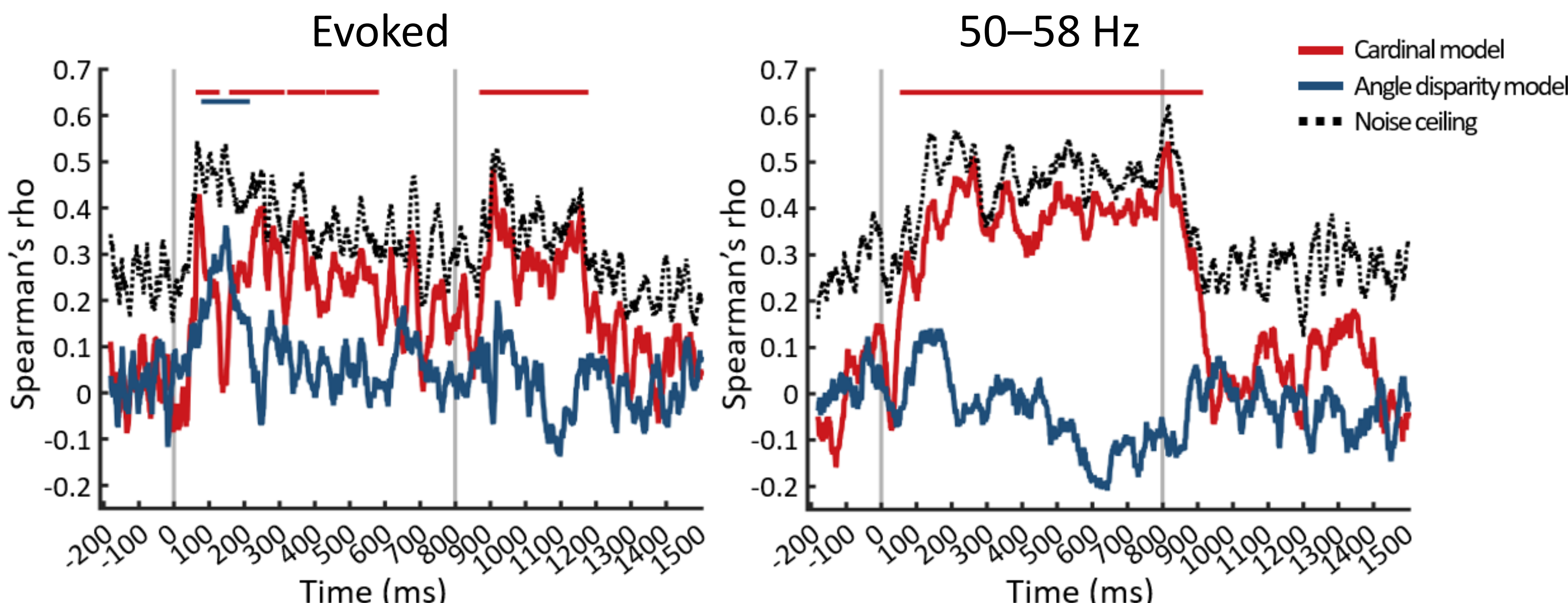
Decoding orientation from *50-58Hz* responses



Pantazis et al., 2017

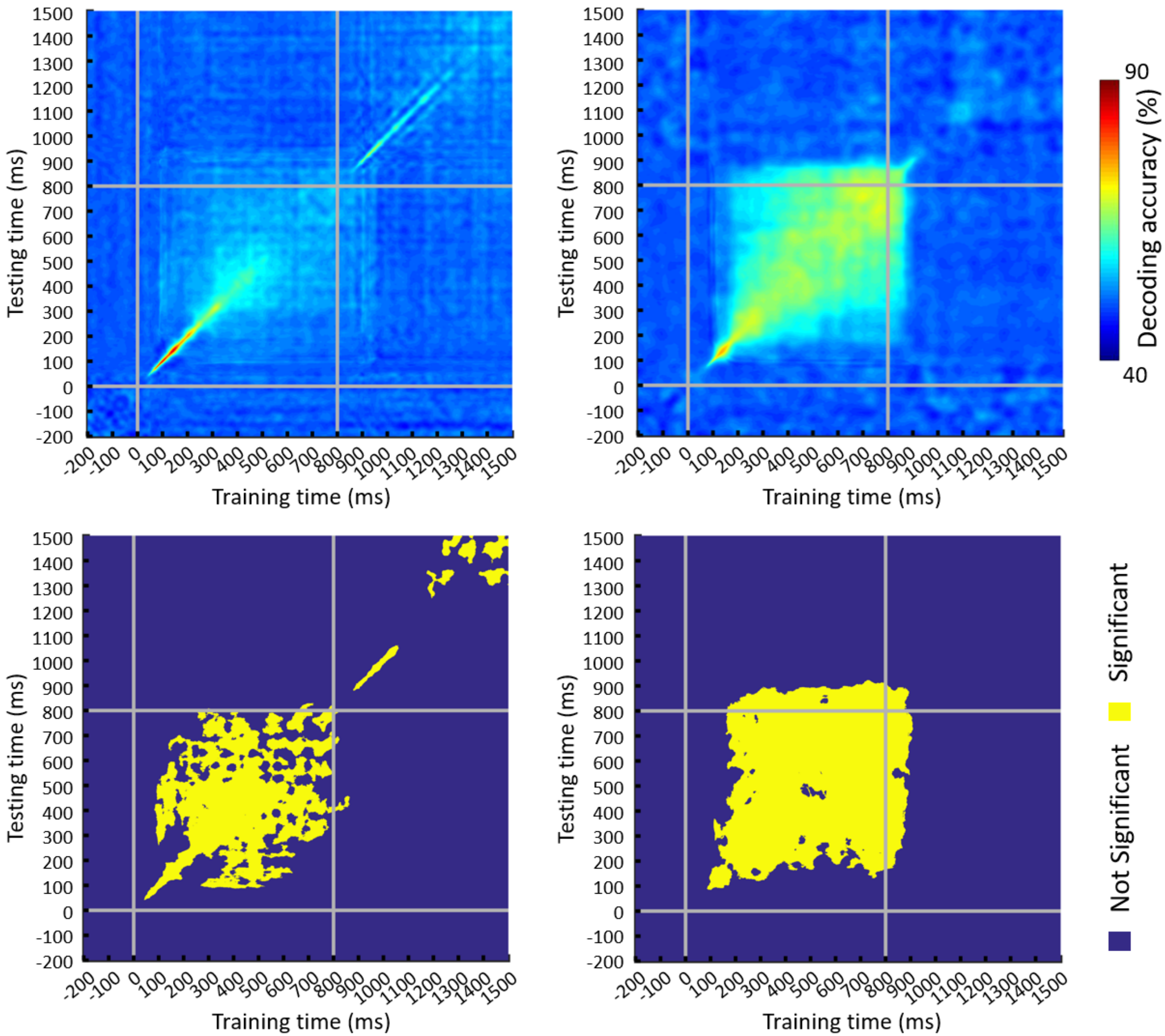
Representational similarity analysis: MEG vs. hypothesized models





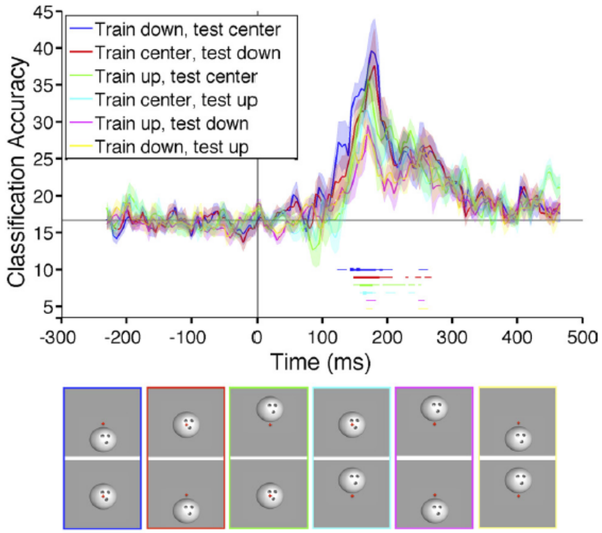
Pantazis et al., 2017

Temporal generalization of decoding

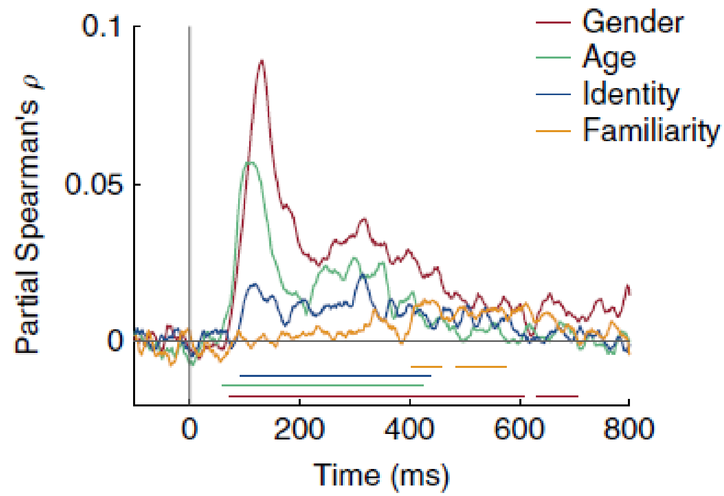


Pantazis et al., 2017

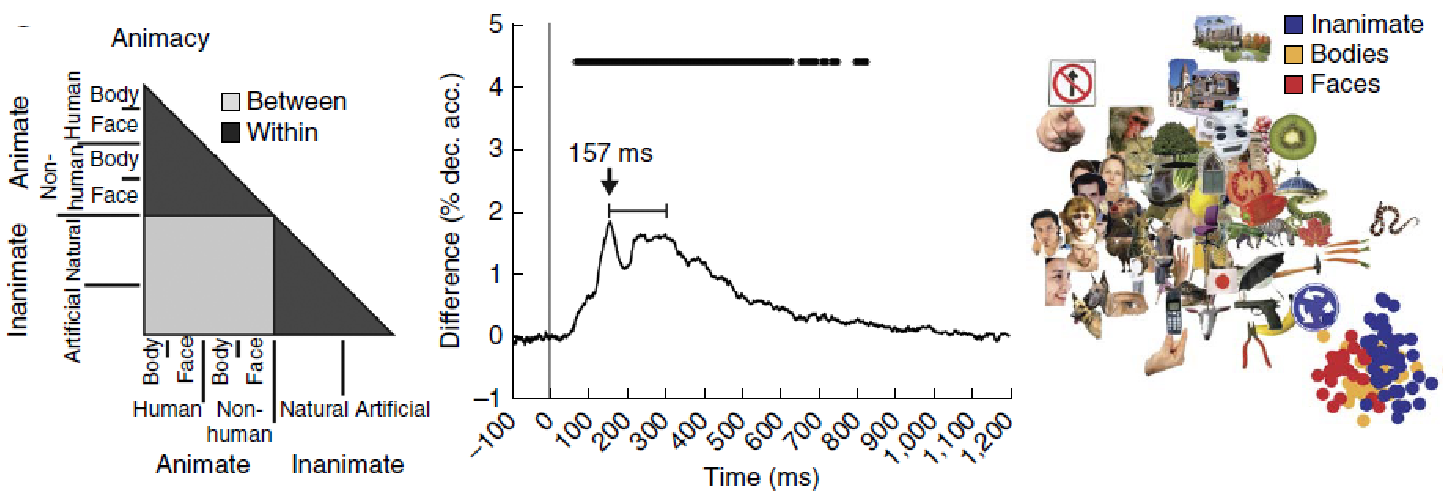
More decoding examples



Isik et al. 2013



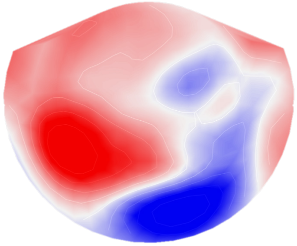
Dobs et al. 2019



Cichy et al. 2014

Conceptual issues

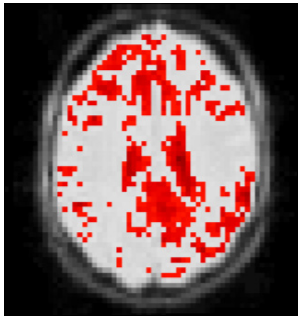
Why use decoding?



Multivariate pattern classification methods are powerful and robust.
(multivariate $\geq$ univariate sensitivity)

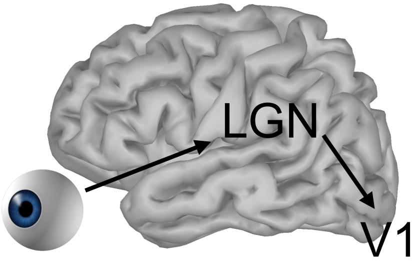
What is being decoded?

Information that is *represented* in brain signals is not necessarily *used* by the brain.



- ❑ 2006 Pittsburgh Brain Competition: Ventricles were the most informative region in the brain to decode humorous events; movement artifacts due to laughter

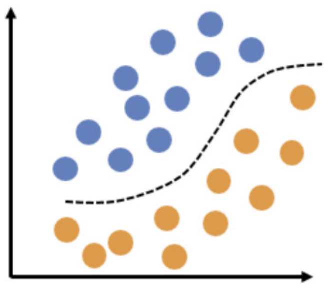
- ❑ Patterns of activity in retina can in principle decode all visual information with a sufficiently complex classifier (e.g. deep neural networks)



Tong and Pratte, 2012

Selection of a classifier

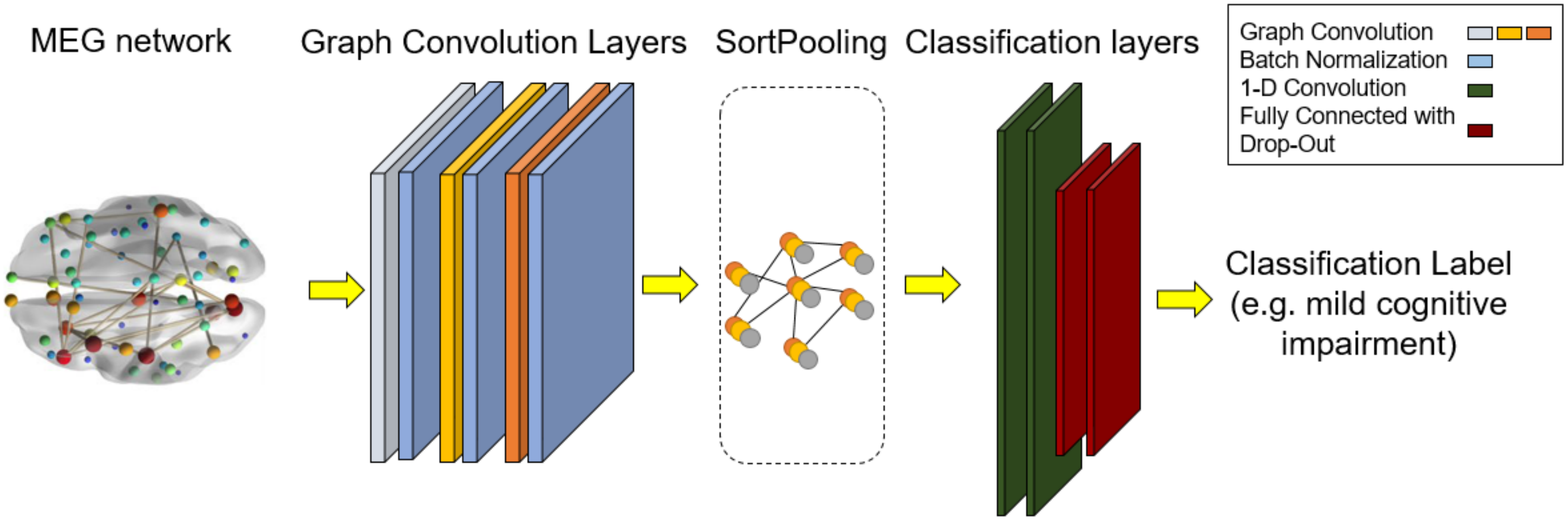
Nonlinear classifiers allow complicated and potentially more powerful decision boundaries to discriminate experimental conditions.



Predictions in real-world applications, such as the detection of Alzheimer's disease (higher decoding accuracies)

Example of non-linear classifier

AGCNN architecture for Alzheimer's disease progression assessment

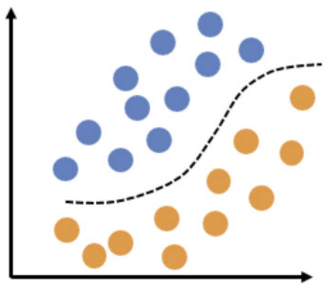


AGCNN classification accuracy: 90%

Li et al. MICCAI 2019

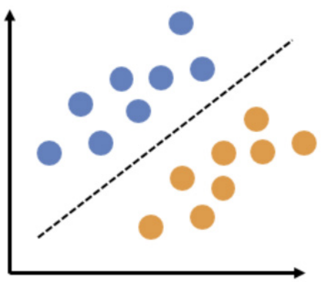
Selection of a classifier

Nonlinear classifiers allow complicated and potentially more powerful decision boundaries to discriminate experimental conditions.



Predictions in real-world applications, such as the detection of Alzheimer's disease (higher decoding accuracies)

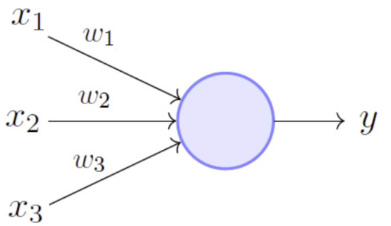
Linear classifiers restrict solutions to linear decision boundaries to discriminate experimental conditions



Understand the neural processes that carry discriminative information

A linear classifier can reveal the information that is <u>*explicitly*</u> represented in the brain

Amenable to a biologically plausible readout in a single step.
A single neuron that receives the pattern as a input has direct access to this information

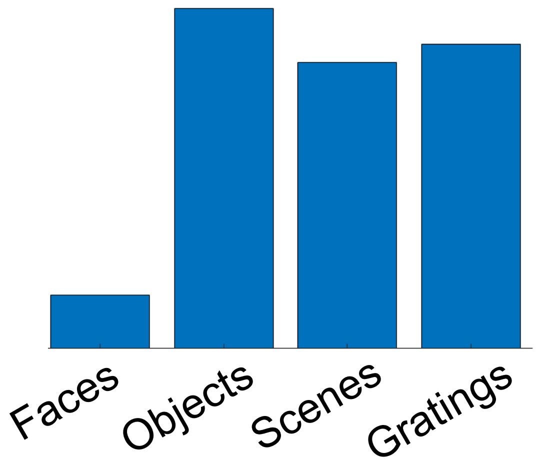


Interpreting decoding accuracies

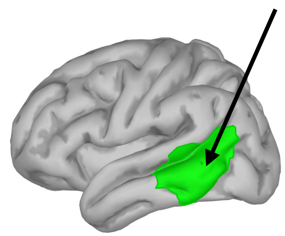
Keep decoding parameters constant when comparing decoding accuracies!

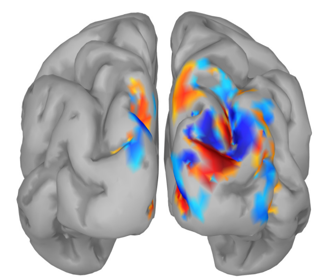
Classification performance depends on several factors: selection of the classifier, cross-validation scheme, degree of separation between the two classes, number of data samples, number and selection of variables to construct multivariate patterns, structure of noise and application of noise whitening

Activation vs. information imaging

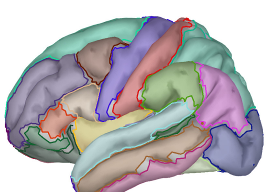


- Represents anything but faces
- Maximally informative about the presence of faces



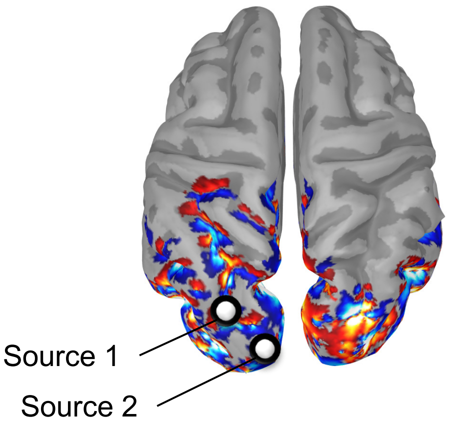


Directional vs. non-directional inference (A>B vs. A≠B)



Hebart et al., 2018

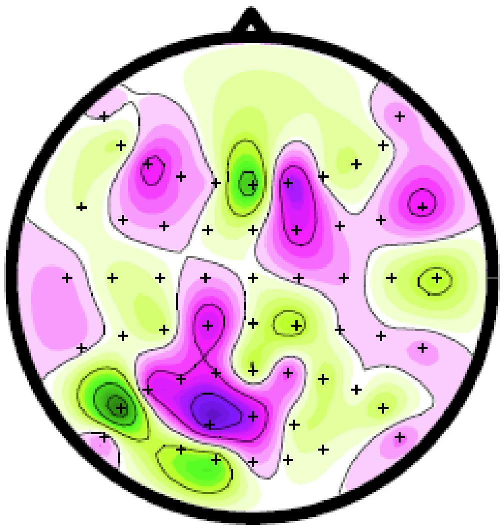
Interpreting decoding weights



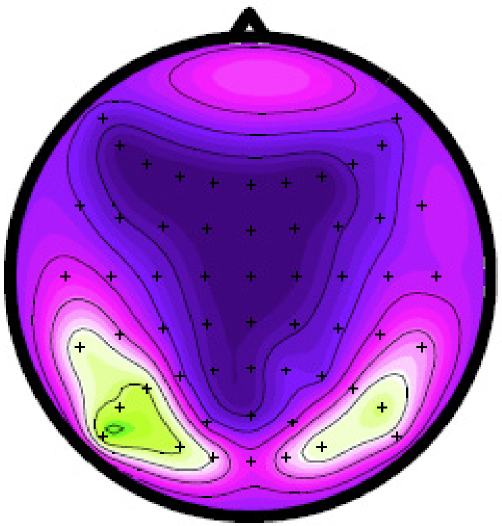
S_1 = signal + noise
S_2 = noise

Optimal weights: $W = [1 \quad -1]$

Weight map W



Activation pattern A



$$A = C_x W C_y^{-1}$$

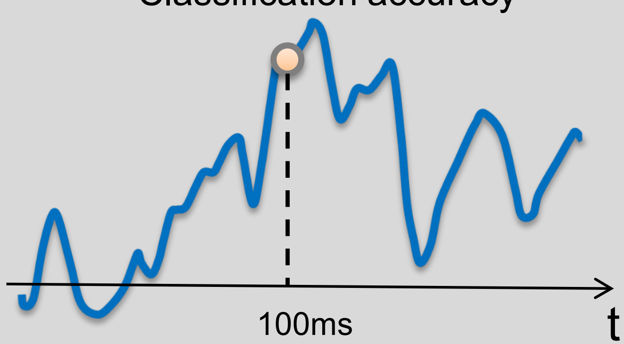
Haufe et al., 2014; Briebmann et al., 2013

What goes into the classifier? Single-stimulus decoding

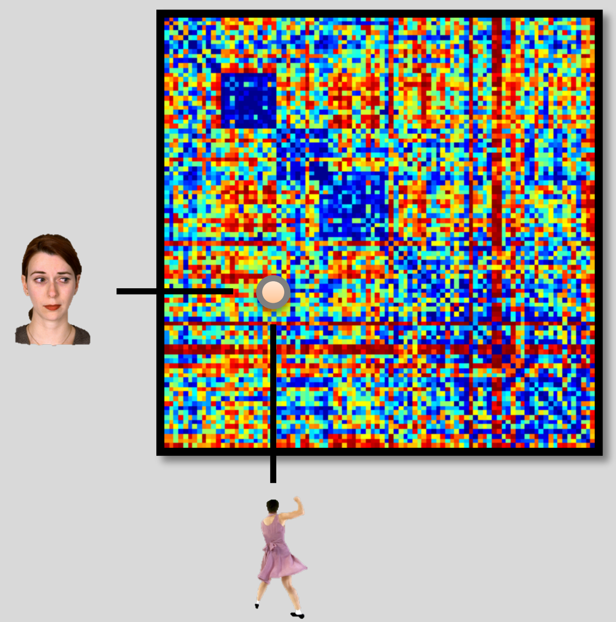
	Train	Test
Image A		
vs.		
Image B		

(cross-trial cross-validation)

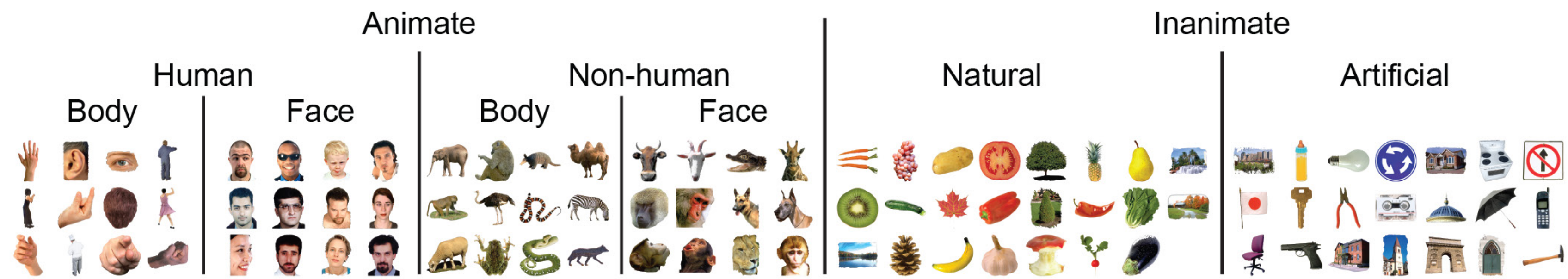
Decoding time series

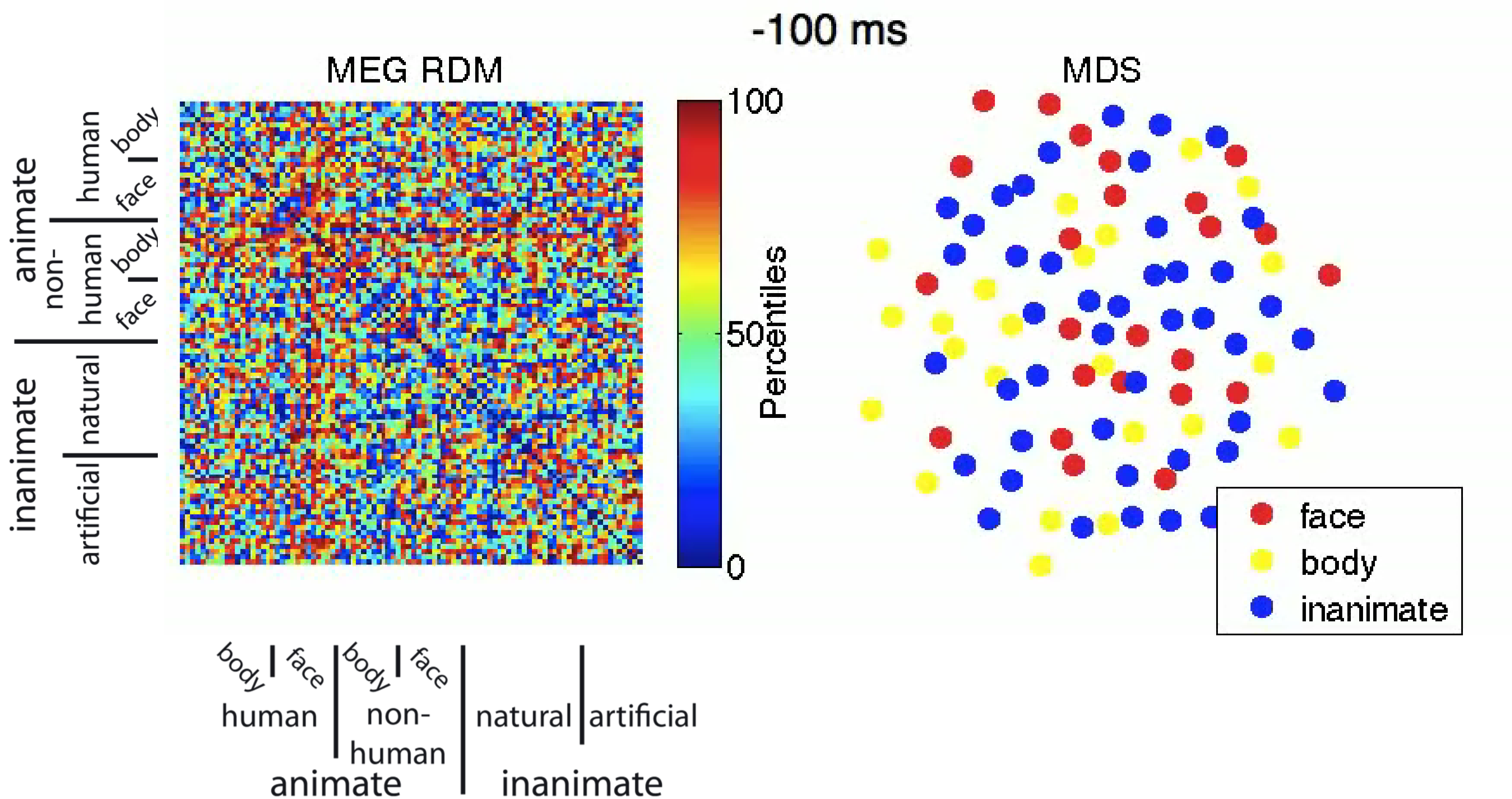


Representational dissimilarity matrix



Single-stimulus decoding - example





Condition decoding

	Training set	Testing set	Cross-validation scheme
Human Faces vs. **Human Bodies**			Cross-trial
Human Faces vs. **Human Bodies**			Cross-exemplar

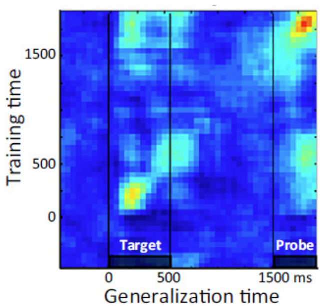
Cross-condition decoding

Human Faces vs. **Human Bodies**		 **Animal Faces** **Animal Bodies**	Cross-condition

Cross-time decoding

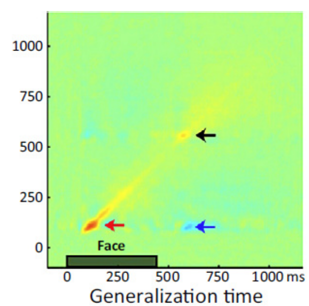
	Training set	Testing set	Cross-validation scheme
Factor A			
vs.			Cross-time
Factor B	 $t=t_1$	 $t=t_2$	

Visual categories



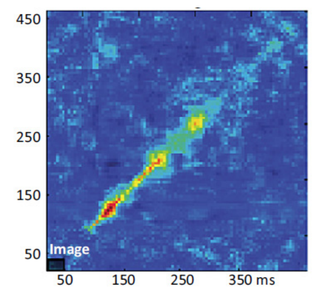
Meyers et al. 2008
Multi-unit recordings; IT cortex

Visual stimulus location



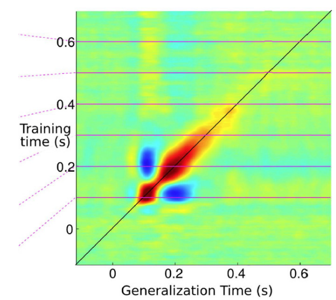
Carlson et al., 2011

Visual position invariance



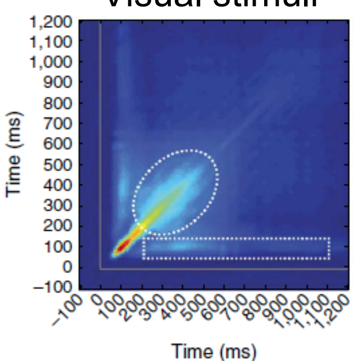
Isik et al., 2014

Standard vs. deviant tones



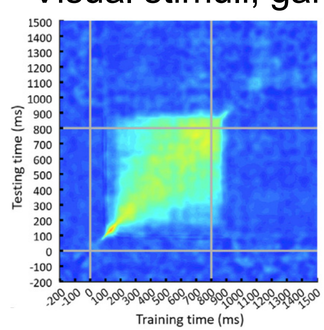
King et al., 2014

Visual stimuli



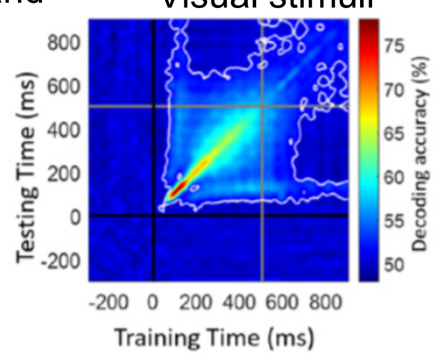
Cichy et al., 2014

Visual stimuli, gamma band



Pantazis et al., 2017

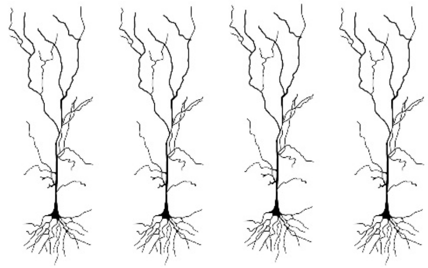
Visual stimuli



Mohsenzadeh et al., 2018

Source of decoding information

Cell recordings

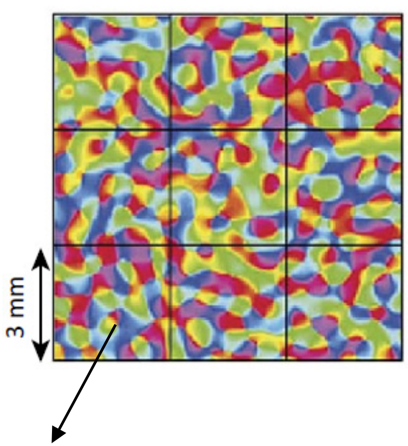


Population coding:
Groups of neurons jointly encode information about the outside world

Macroscopic measurements

Coarse-scale or fine-scale information? Orientation columns of about 800 μm in diameter

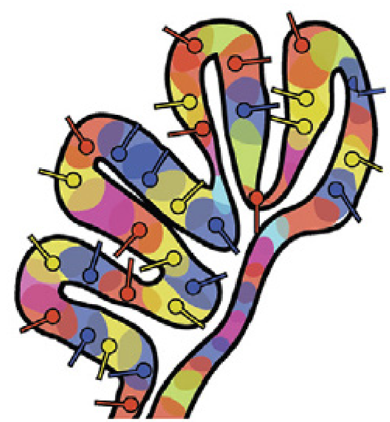
fMRI



fMRI voxel

Biased sampling account

MEG



Haynes and Rees, 2005; Kamitani and Tong, 2005

Cichy et al., 2015; Stokes et al. 2015

Good practices

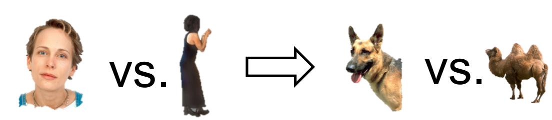
Information that is *represented* in brain signals is not necessarily *used* by the brain.

Good practices

✓ *Link decoding* performance ~ *behavioral* performance

✓ Investigate patterns across a variety of stimulus conditions (Representational similarity analysis)

✓ Show that decoding performance generalizes to novel and very different stimuli



✓ Use a linear classifier

Brain processes reflect a series of nonlinear computations. A linear classifier will capture the information processed at each step.