fMRI analysis. Brain decoding with machine learning.

Leila Reddy

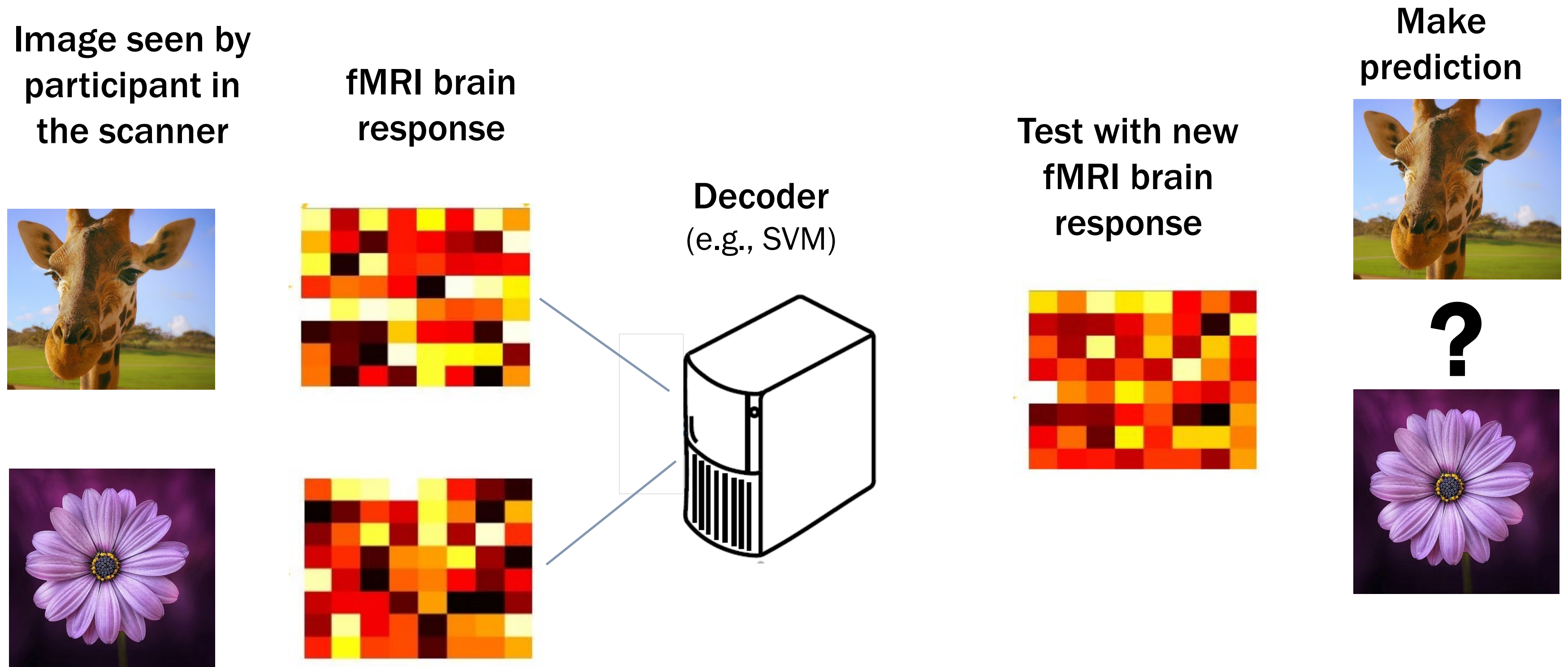
(leila.reddy@cnrs.fr)

CerCo, CNRS

Brain decoding

- Brain decoding: decode what people see, imagine, remember etc.

- A decoder is trained to learn a relationship between fMRI patterns and the “content” to be decoded (e.g., a seen image).



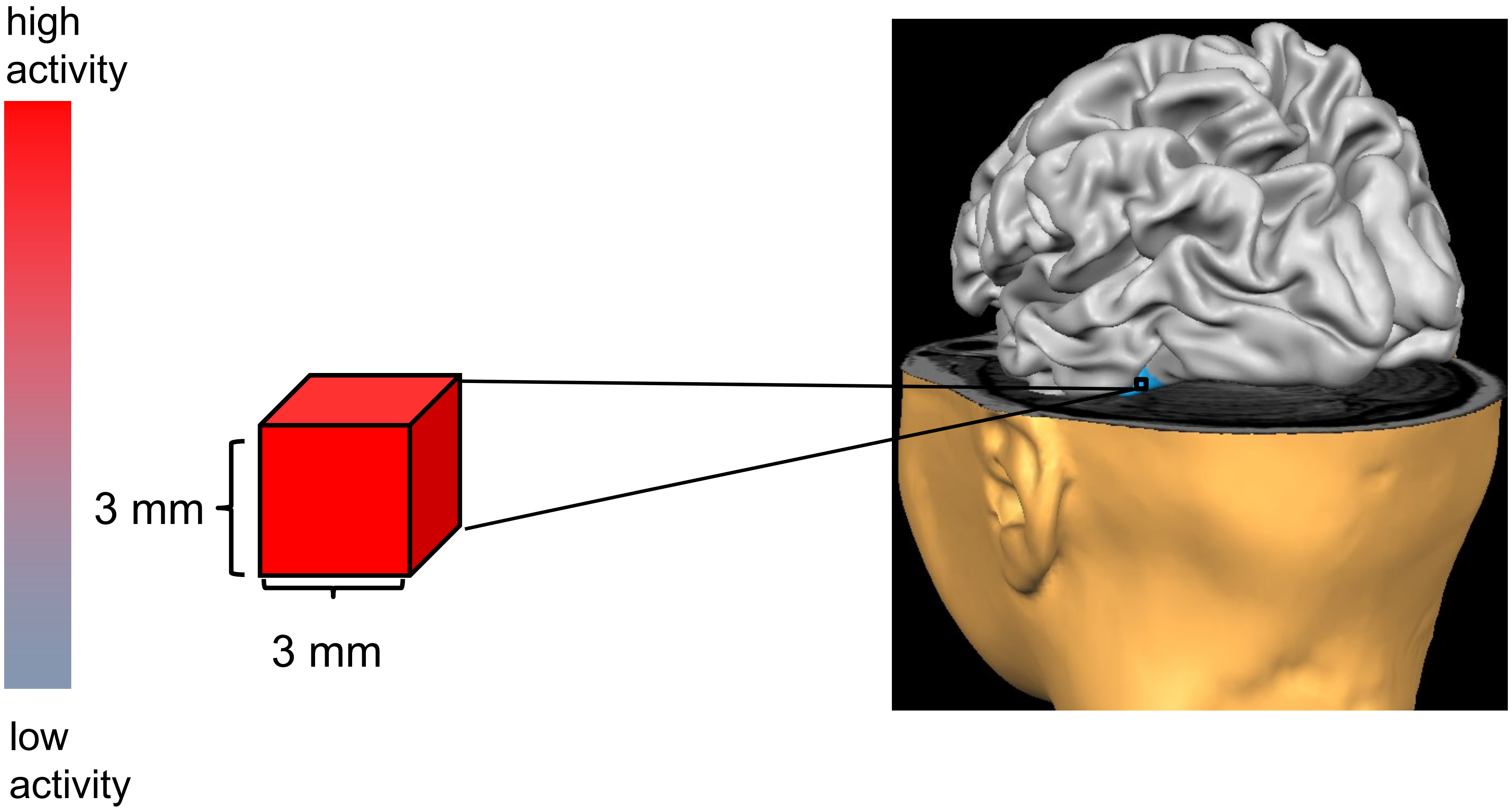
What is fMRI?

- Functional Magnetic Resonance Imaging.
- Takes pictures of the brain.
- Measures neural activity indirectly
- BOLD (blood oxygenation level dependent) signal:
    - Increased neural activity > increase in MR signal intensity (BOLD level)

What is a voxel?

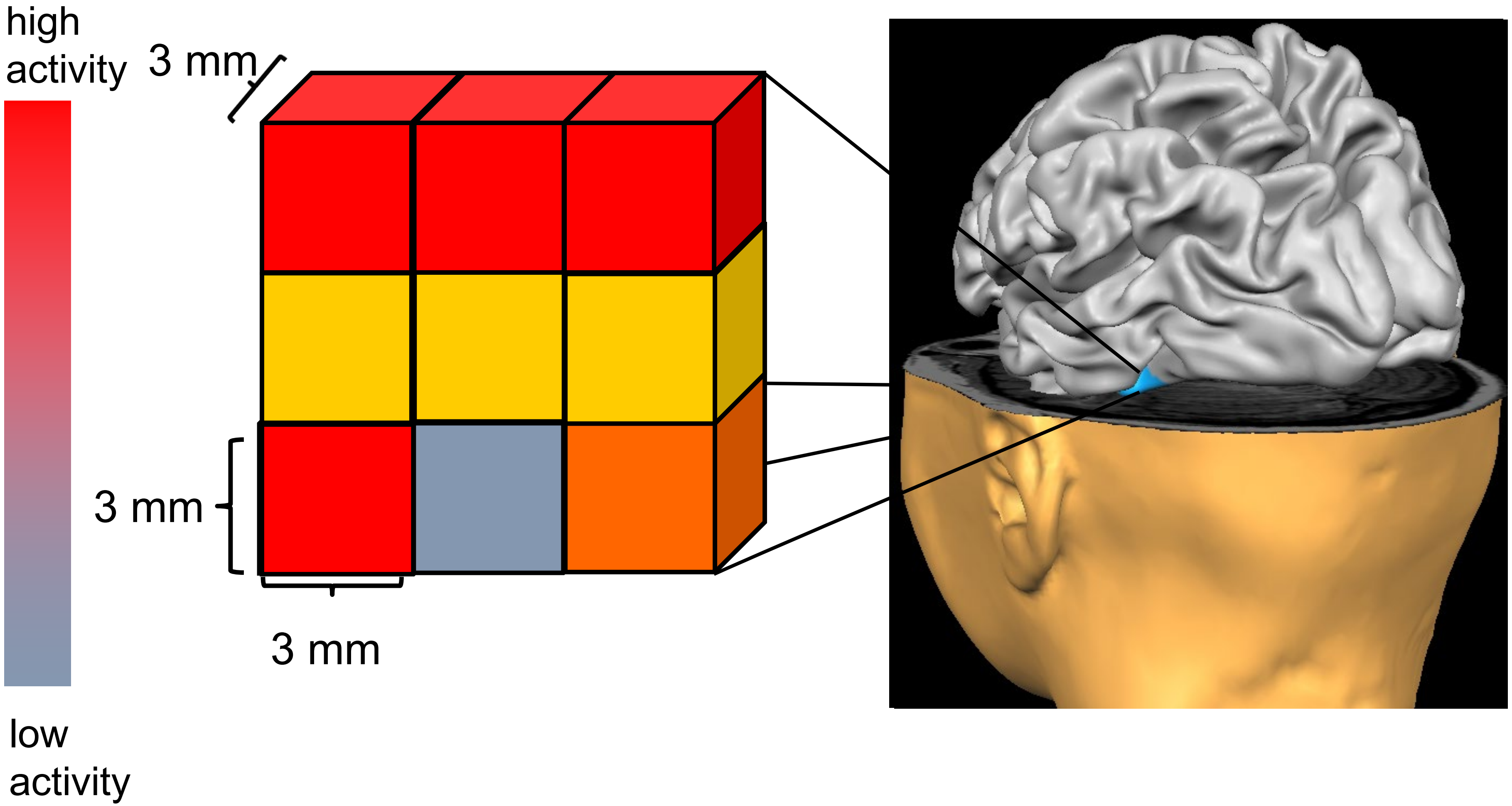
- fMRI /MRI take 3-dimensional pictures of the brain.
- fMRI doesn't provide detail at the level of a neuron.
- The fMRI unit is a voxel.
- One voxel represents a cube of cortex — a 3-D image analogous to the 2-D pixel of digital cameras.

Voxels and multi-voxel patterns of activity.



From Jody Culham

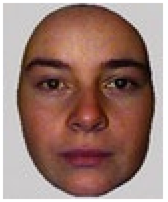
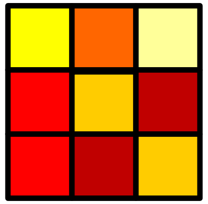
Voxels and multi-voxel patterns of activity.

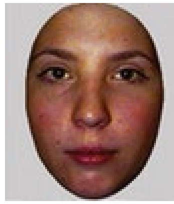
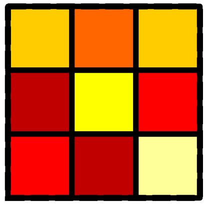


From Jody Culham

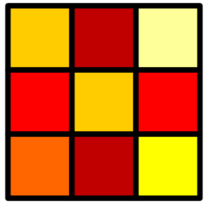
Standard fMRI Analysis

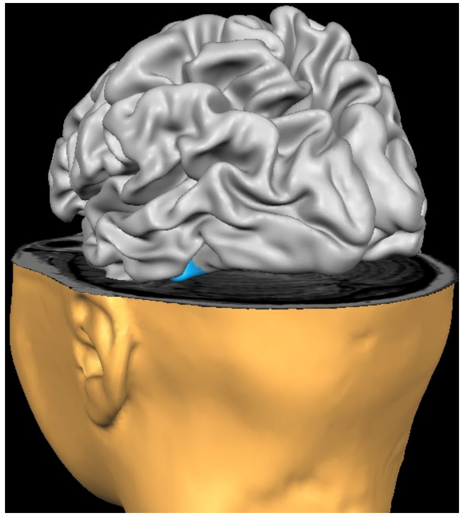
FACES

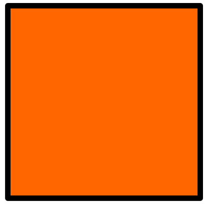
trial 1



trial 2



trial 3

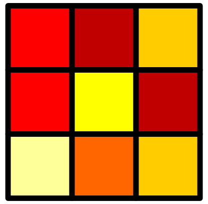


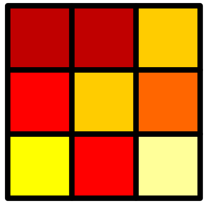


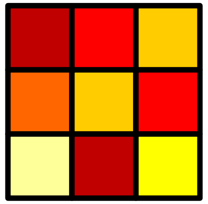


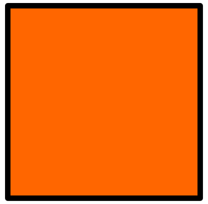
Average
Summed
Activation

HOUSES

trial 1



trial 2



trial 3


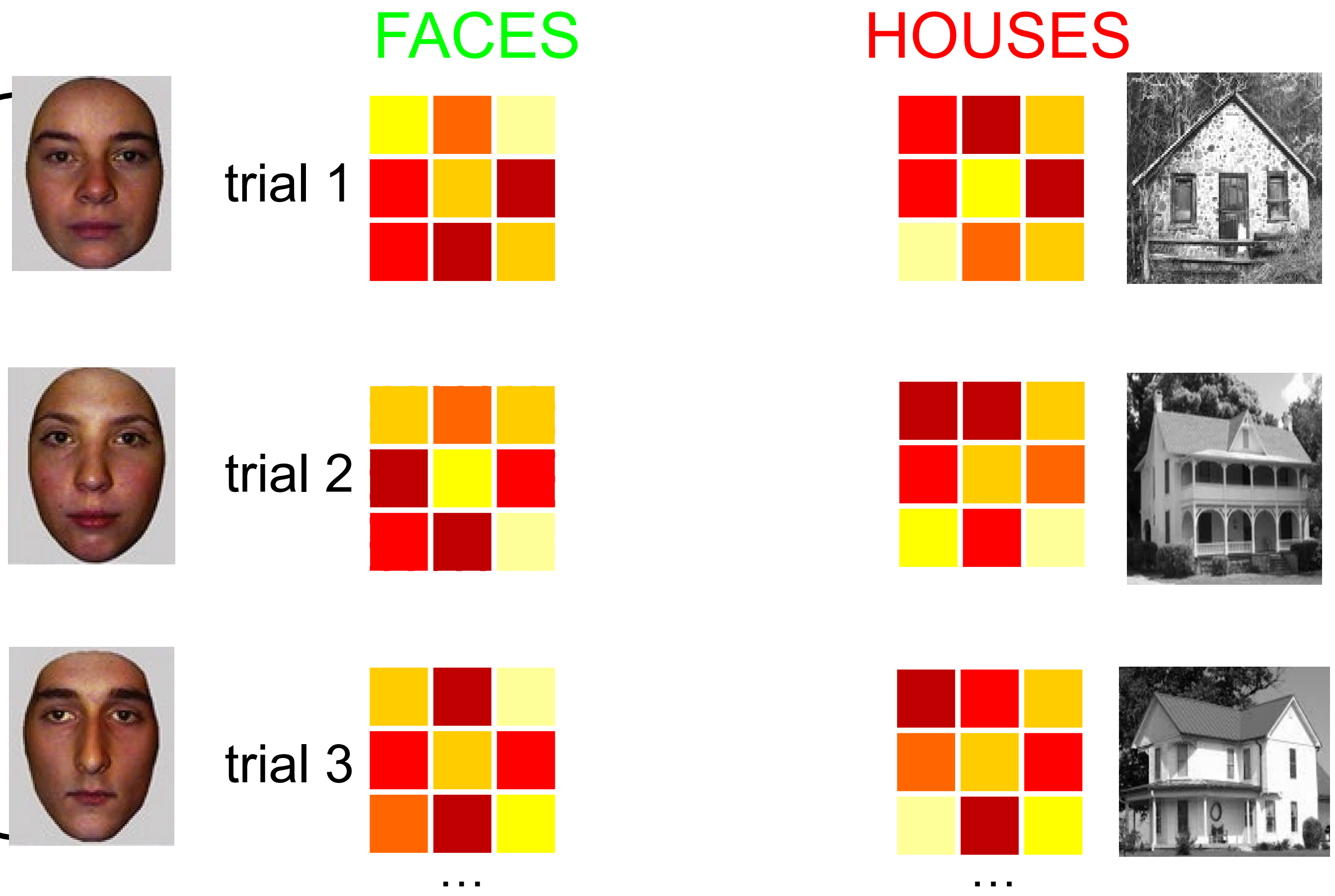
From Jody Culham

Multi-voxel patterns of activity.

- An fMRI response in a chunk of the brain can be thought of as a pattern of activity across voxels (millions of neurons).

- Patterns carry more information than the average across the pattern.

Multi-voxel pattern analysis (MVPA)



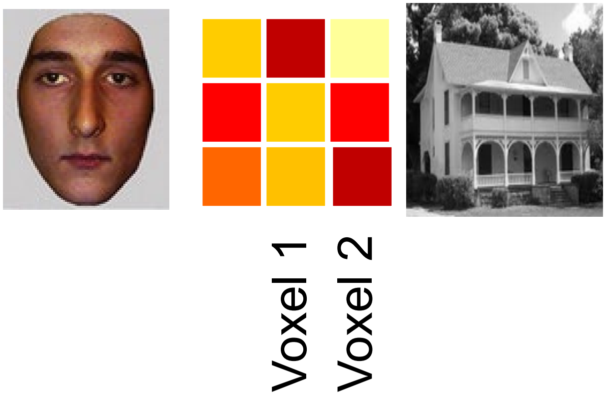
Training Trials

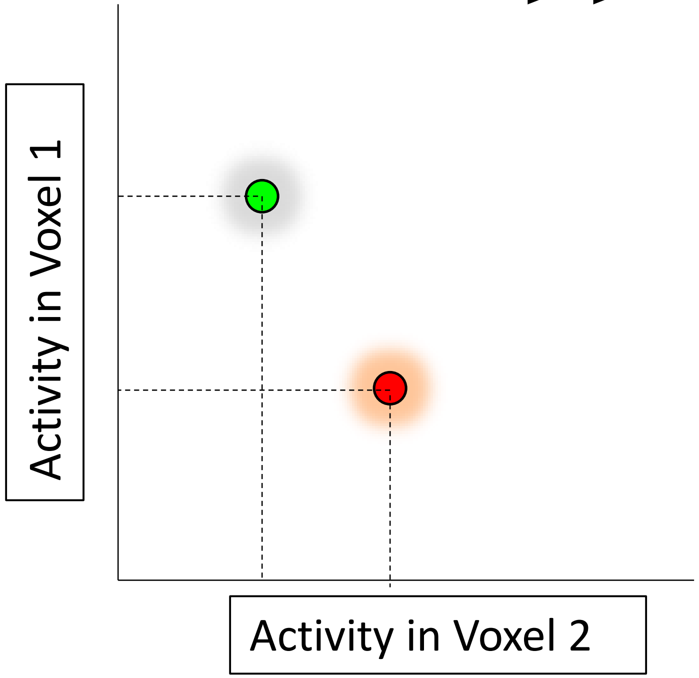
Test Trials
(not in training set)

Can an algorithm correctly "guess" trial identity better than chance (50%)?

From Jody Culham

Multi-voxel pattern analysis (MVPA)





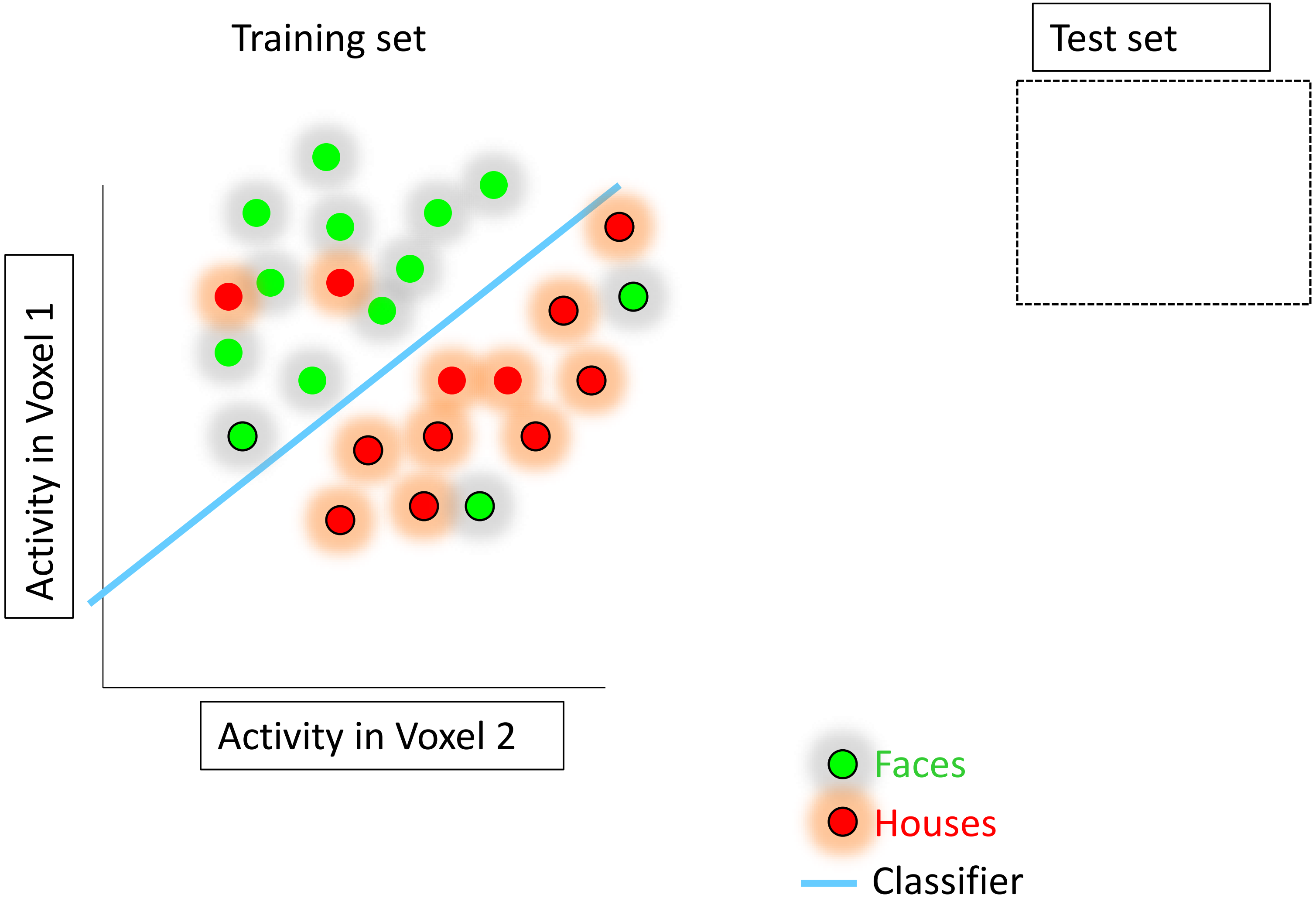
Each dot is one measurement (trial) from one condition (red circles) or the other (green circles)

Faces

Houses

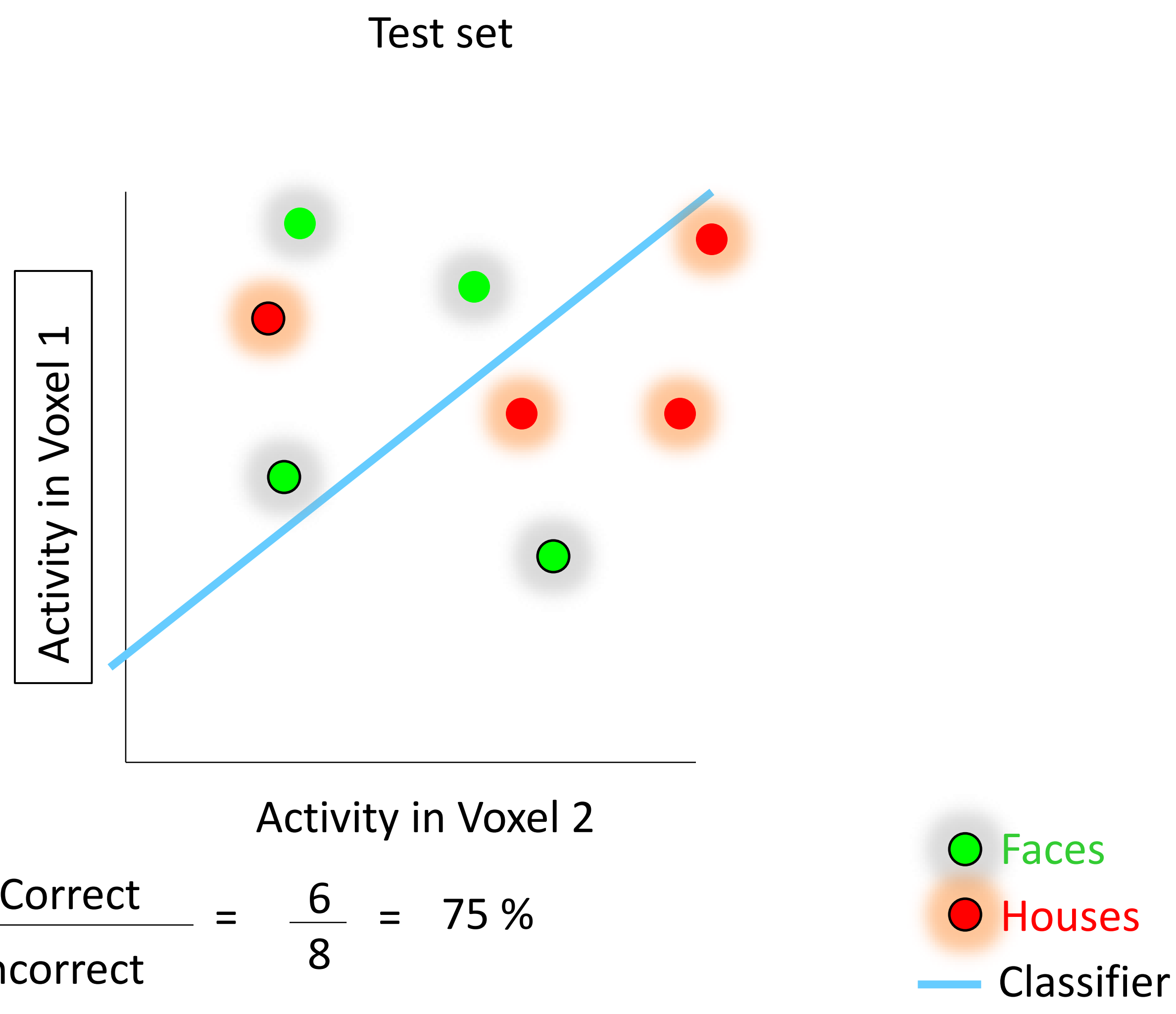
From Jody Culham

MVPA: training set



From Jody Culham

MVPA: test set

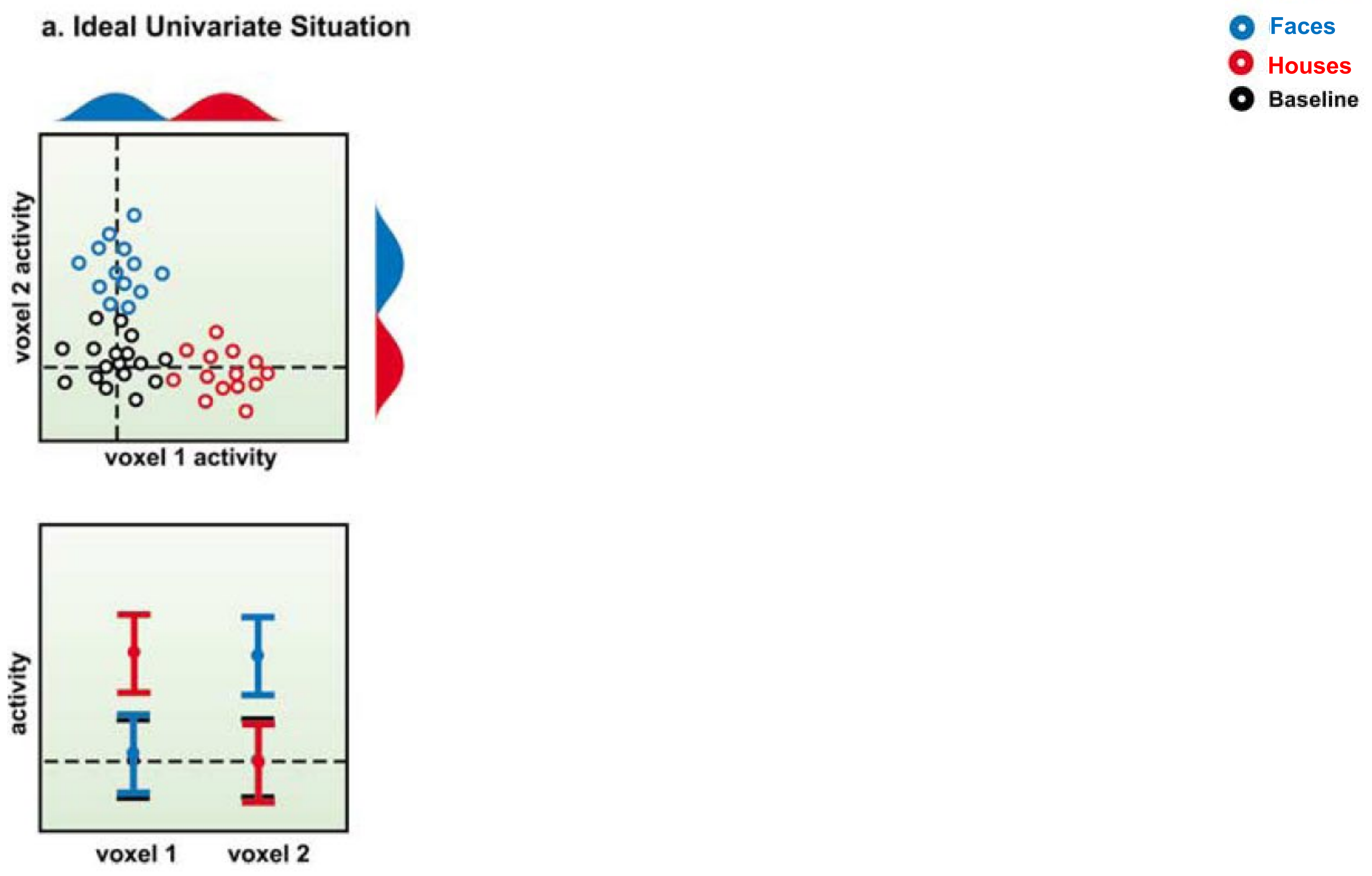


Classifier Accuracy = $\frac{\text{Correct}}{\text{Incorrect}}$ = $\frac{6}{8}$ = 75 %

From Jody Culham

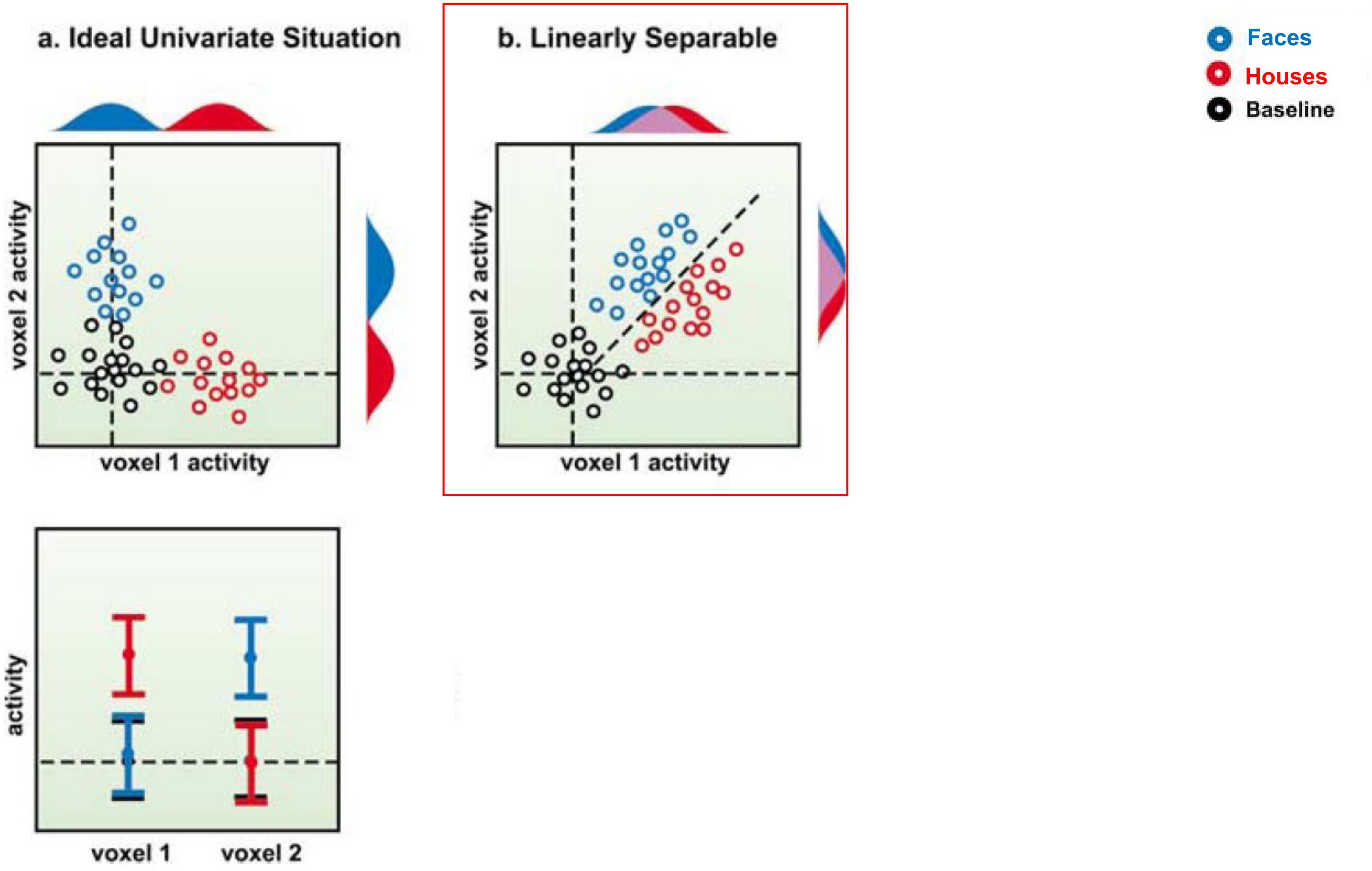
In a, the classifier can operate on single voxels because the response distributions are separable within individual voxels.

Cox and Savoy, 2003. From Jody Culham



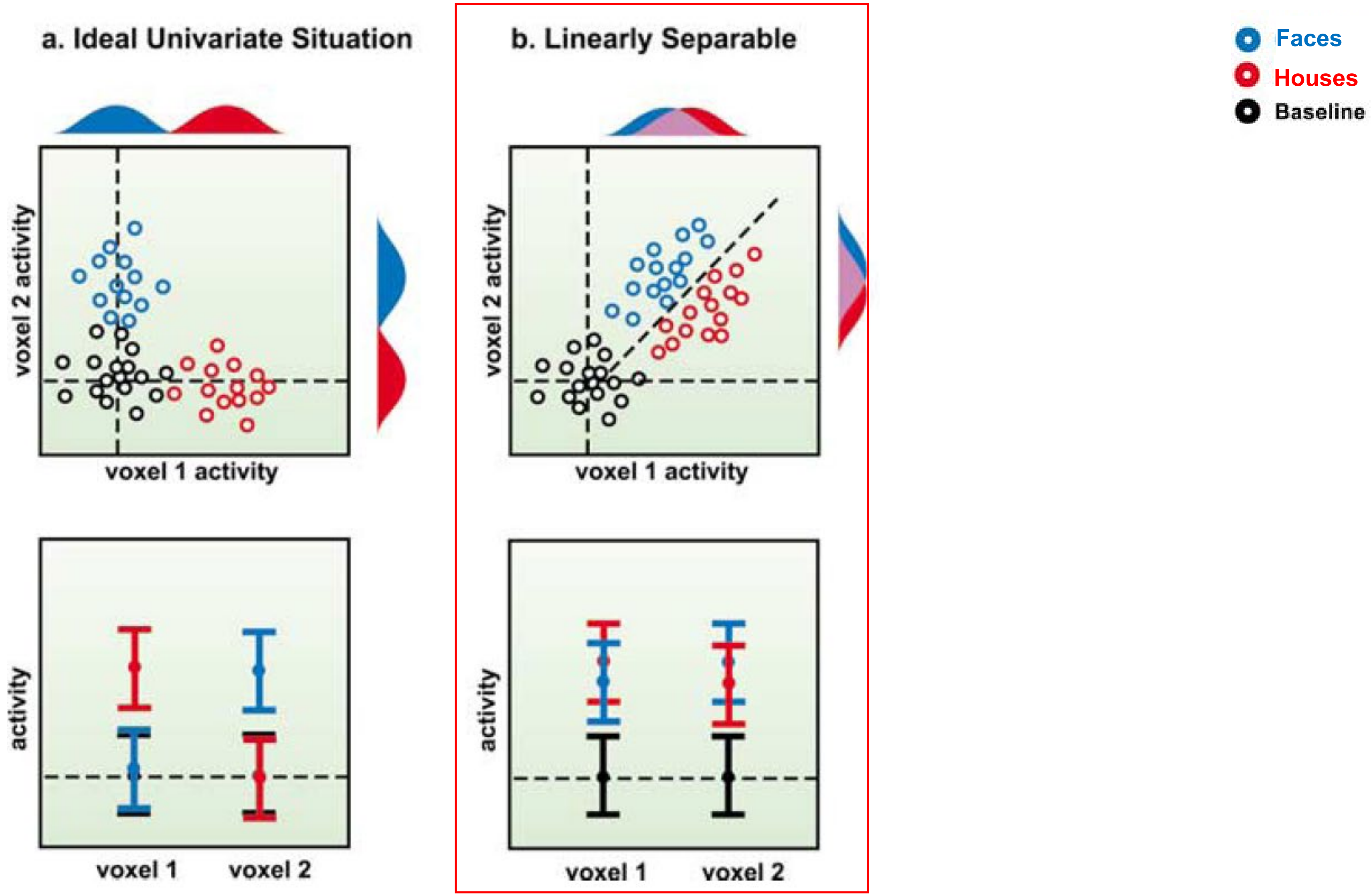
Also standard fMRI analysis can detect this difference

Cox and Savoy, 2003. From Jody Culham



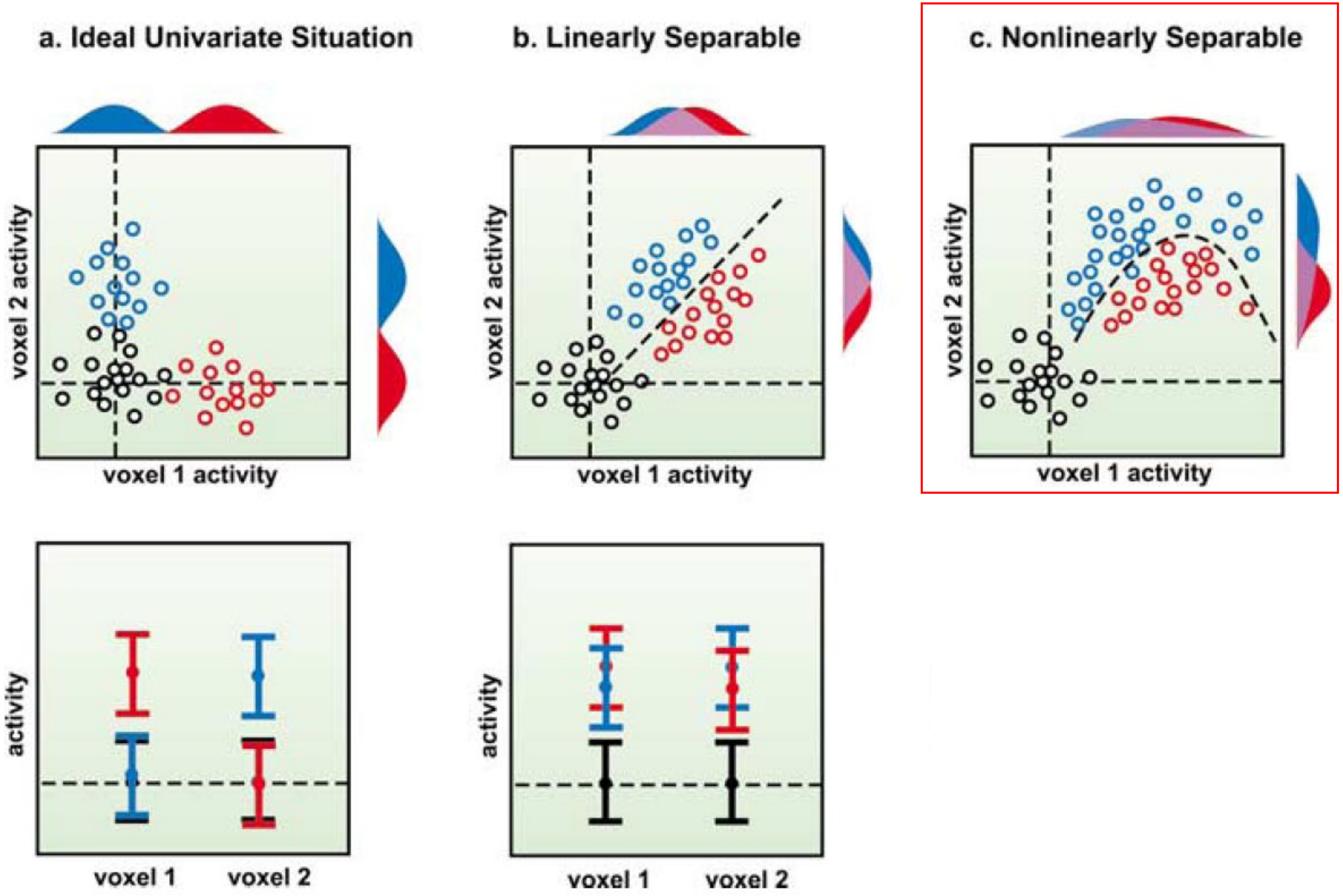
In b, the classifier cannot operate on single voxels because the response distributions are overlapping within individual voxels.

Cox and Savoy, 2003. From Jody Culham



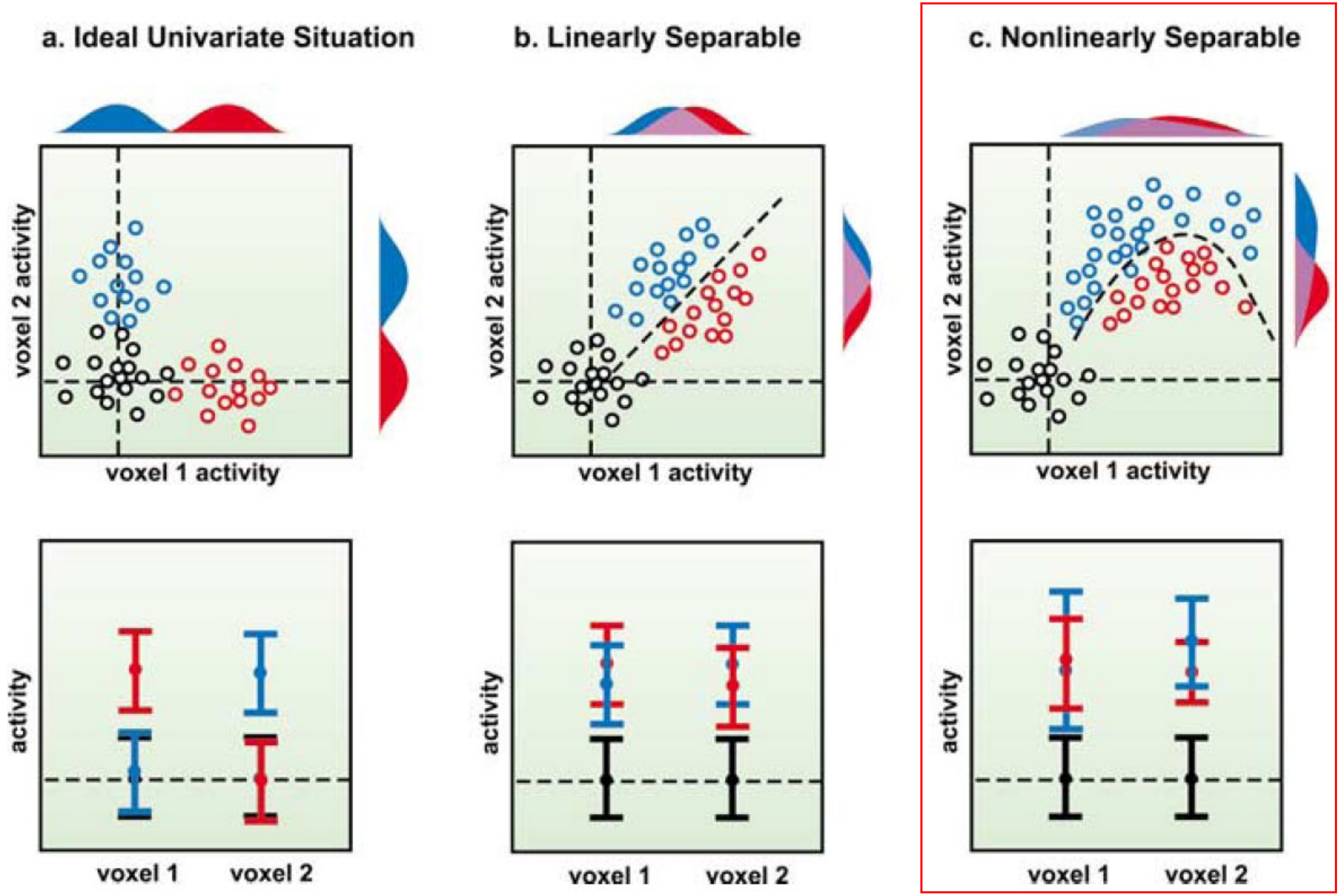
Standard fMRI analysis cannot detect this difference

Cox and Savoy, 2003. From Jody Culham



In c, the non-linear classifier draws decision boundaries other than straight lines

Cox and Savoy, 2003. From Jody Culham



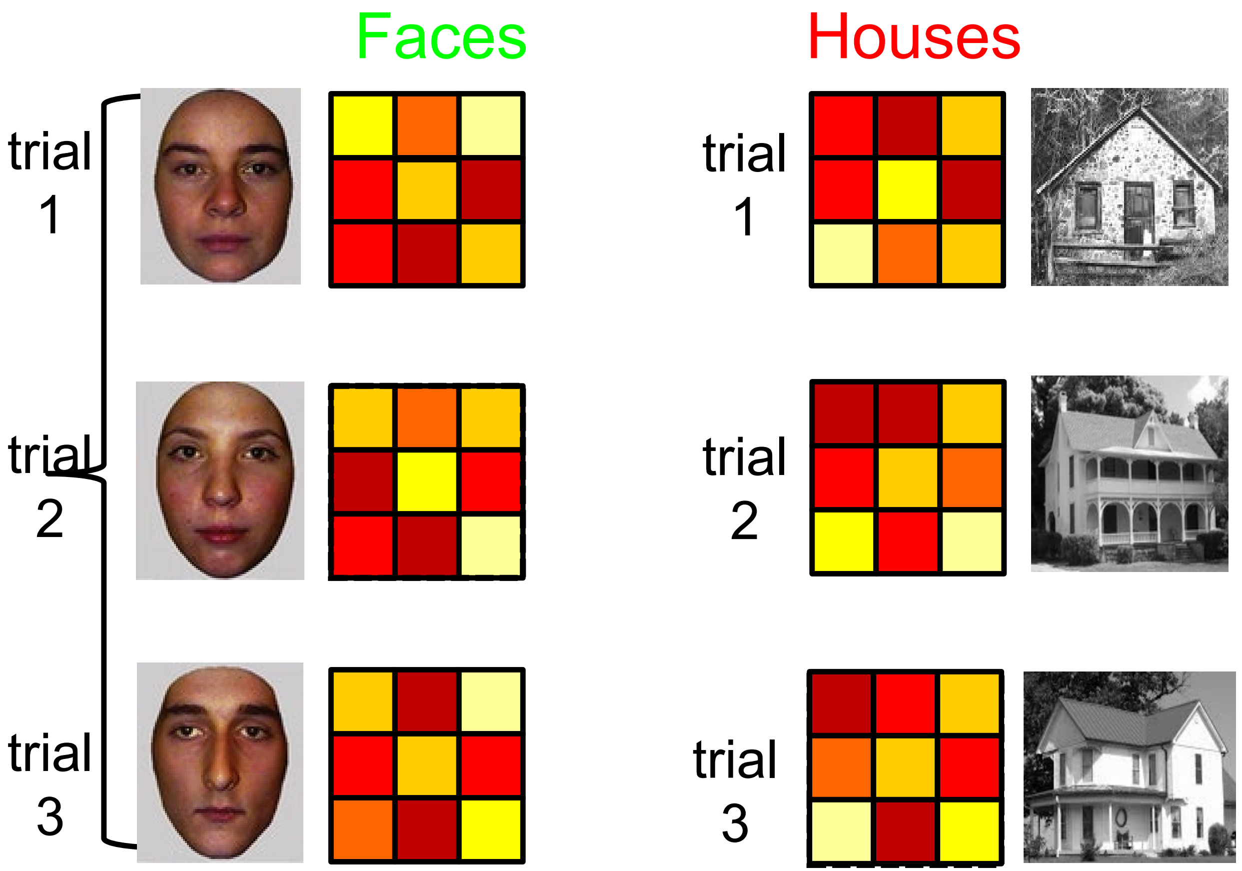
Standard fMRI analysis cannot detect this difference

Cox and Savoy, 2003. From Jody Culham

3 examples of fMRI decoding

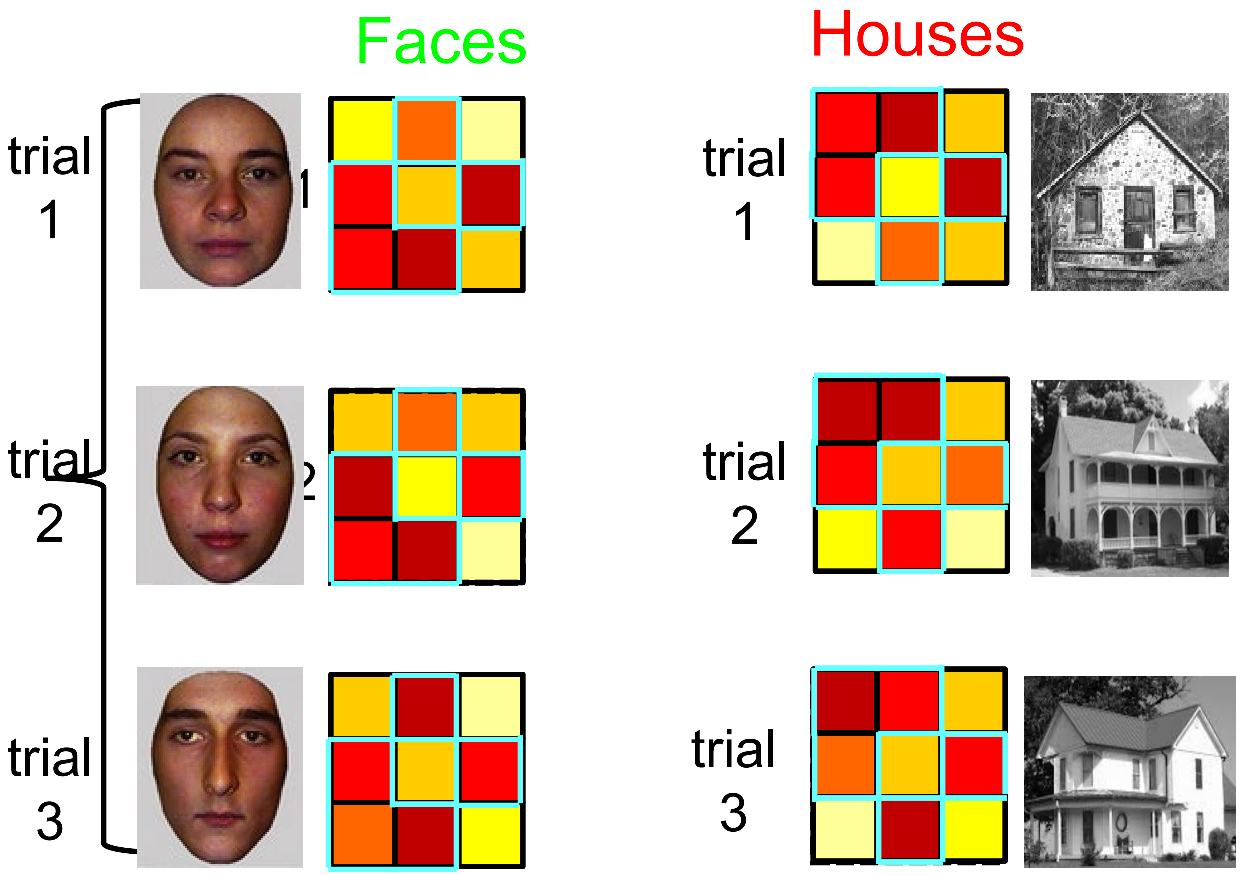
- Correlation-based
- Support Vector Machines
- Using DNN latent spaces
- There are others, but we don’t have time for them today.

Correlation-based MVPA



From Jody Culham

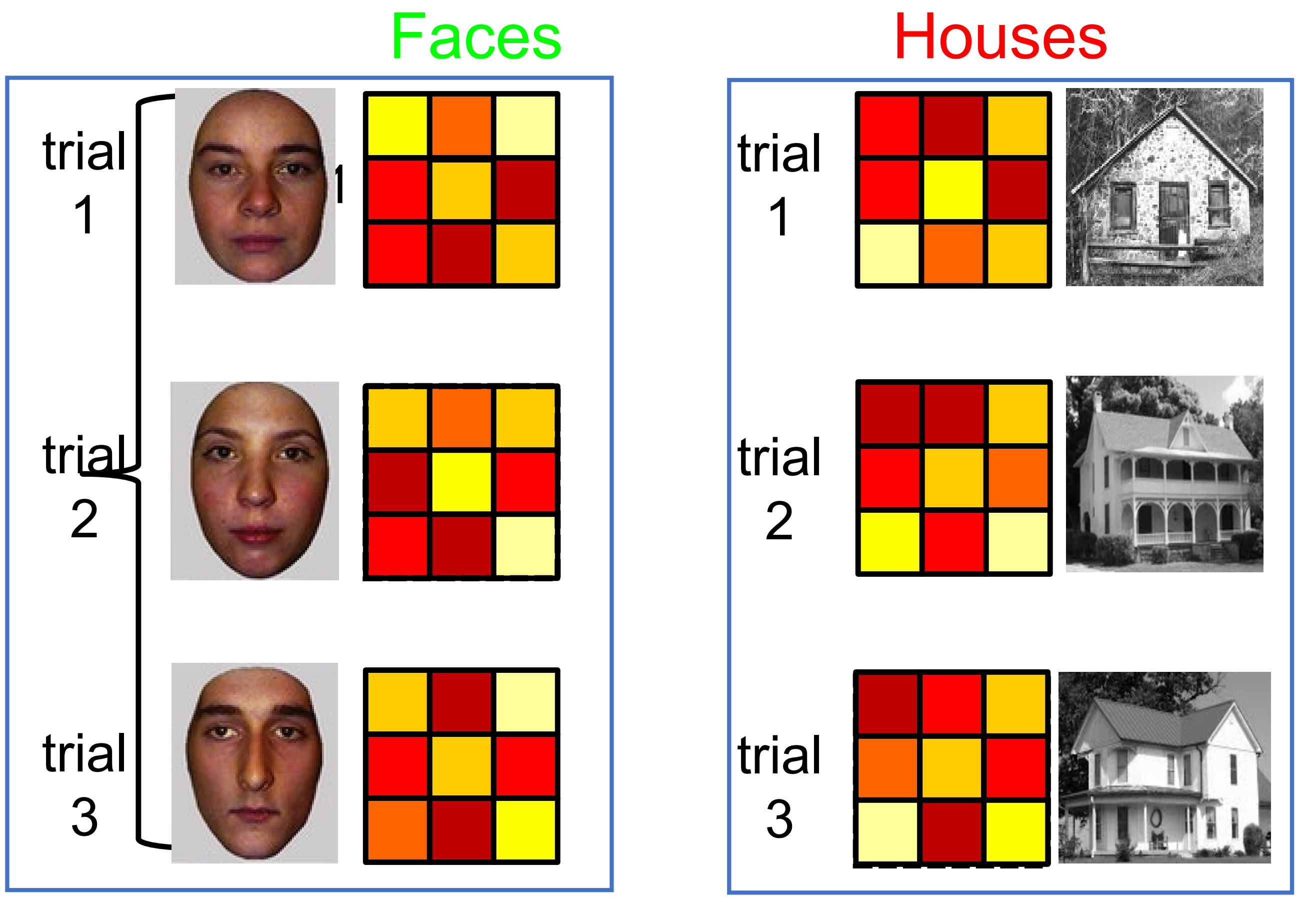
Correlation-based MVPA



The same category evokes similar patterns of activity across trials

From Jody Culham

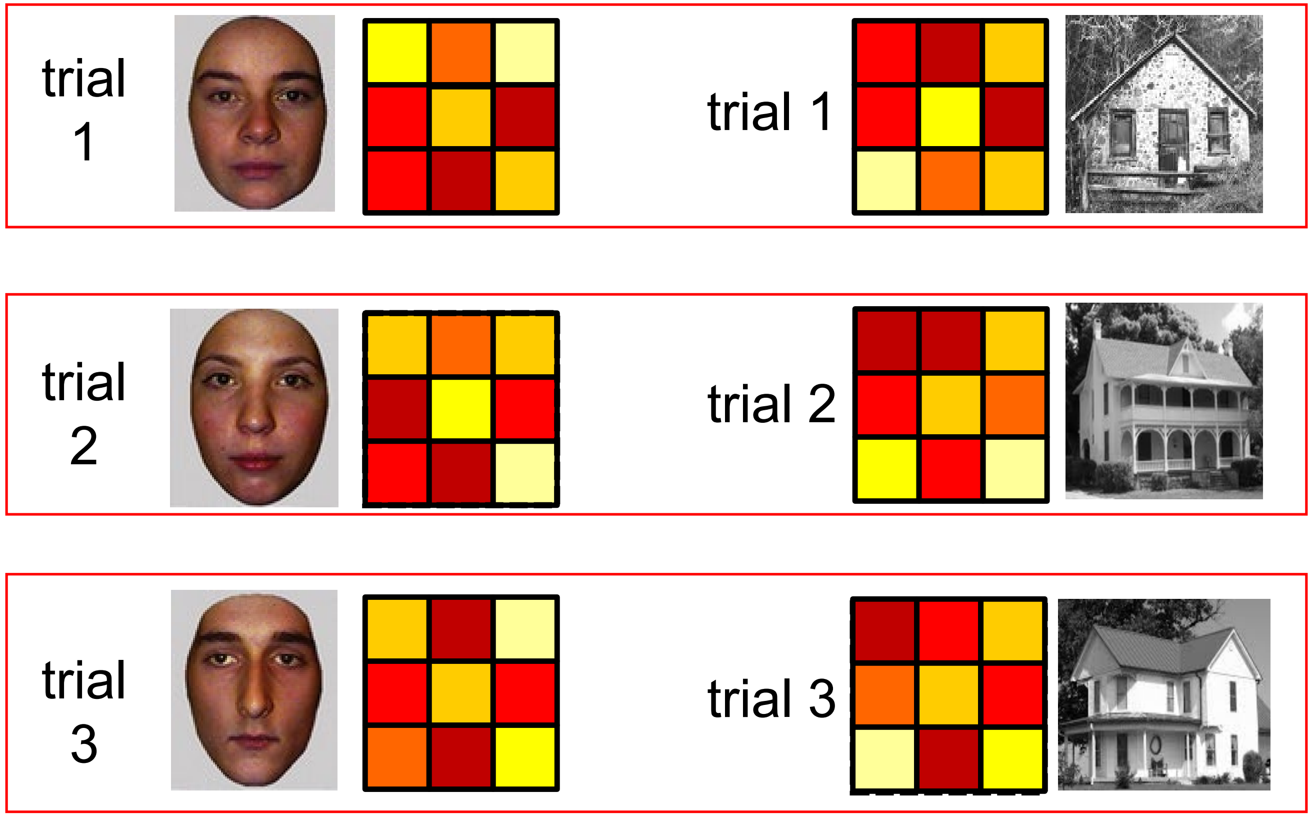
Correlation-based MVPA



Similarity Within the same category

From Jody Culham

Correlation-based MVPA



Similarity <u>Between different categories</u>

From Jody Culham

Correlation-based MVPA

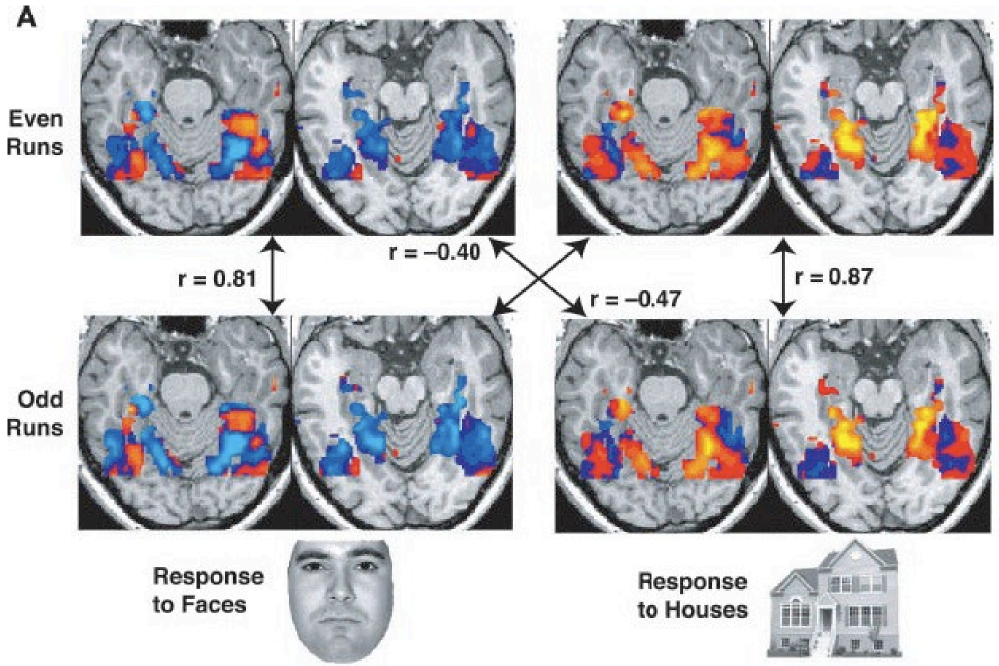
IF

Within-category similarity > Between-category similarity

Then the brain area contains information about faces and houses.

Correlation-based MVPA

Decoding category information

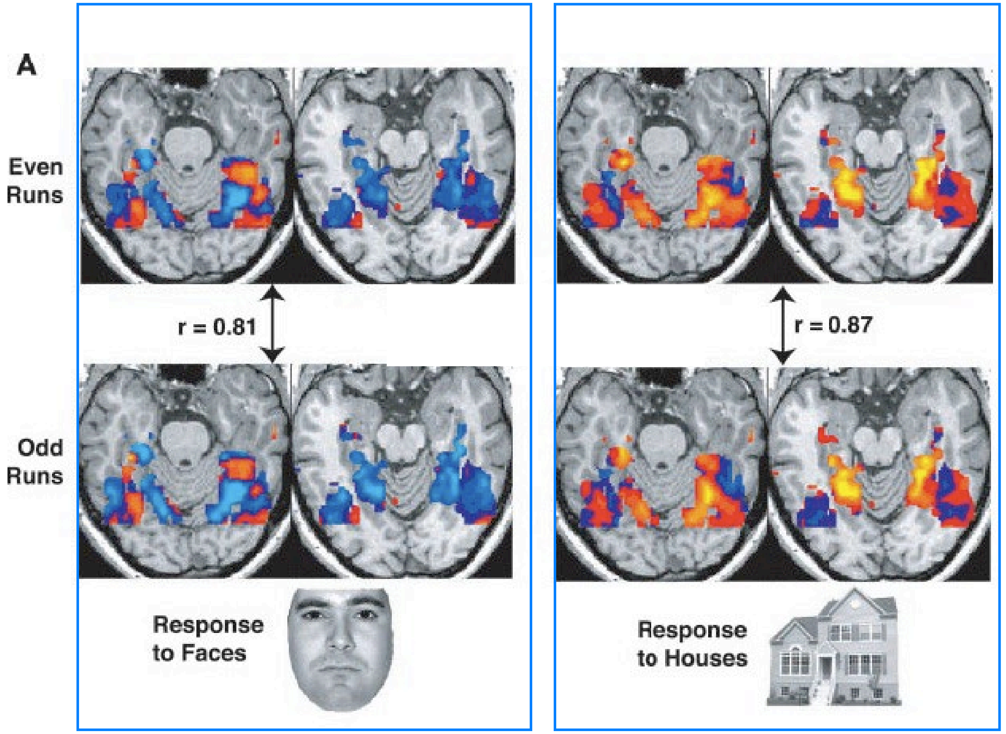


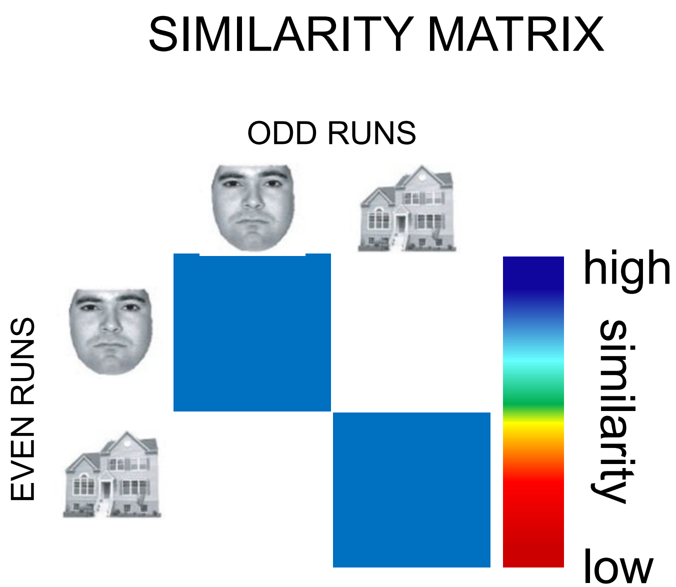
Haxby et al., 2001

From Jody Culham

Correlation-based MVPA

Decoding category information





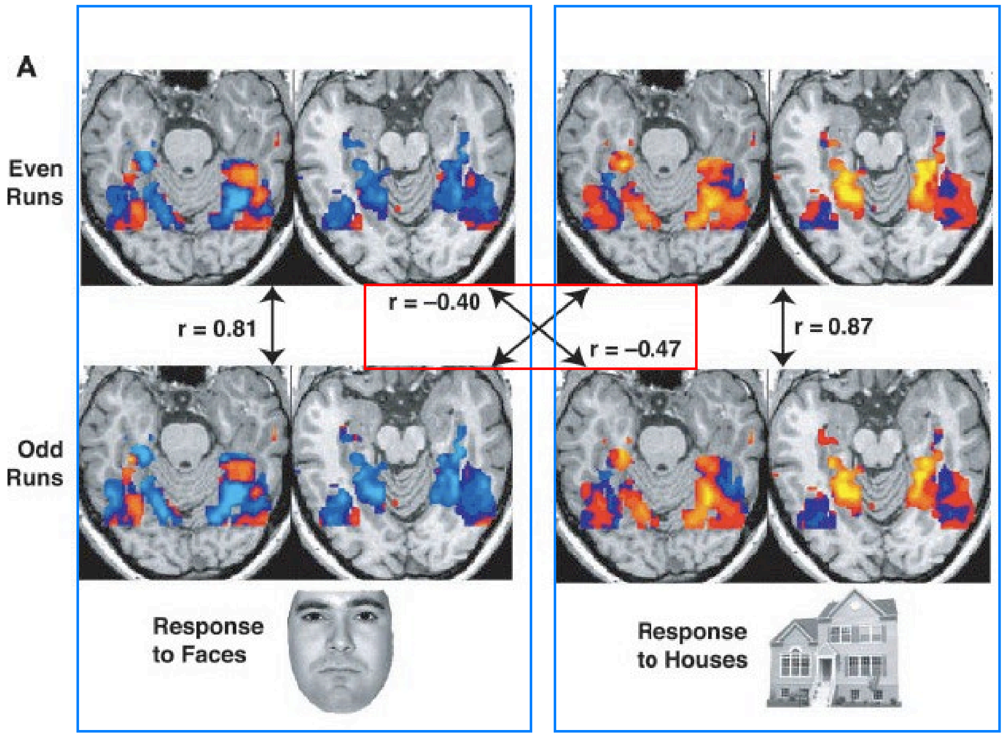
Within-category similarity

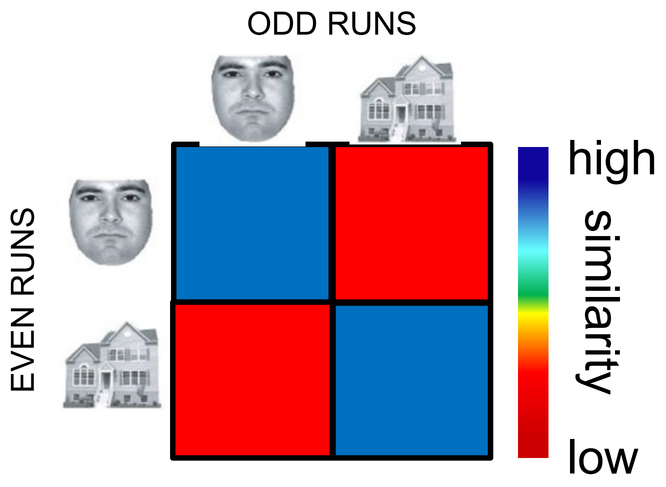
Haxby et al., 2001

From Jody Culham

Correlation-based MVPA

Decoding category information



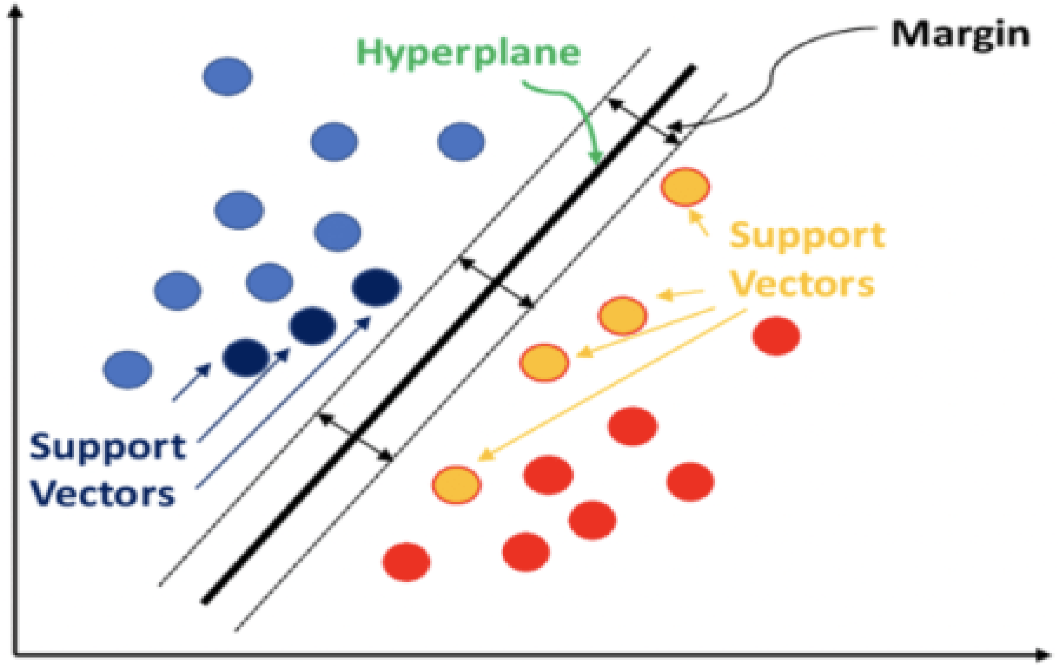


Between-category similarity

Haxby et al., 2001

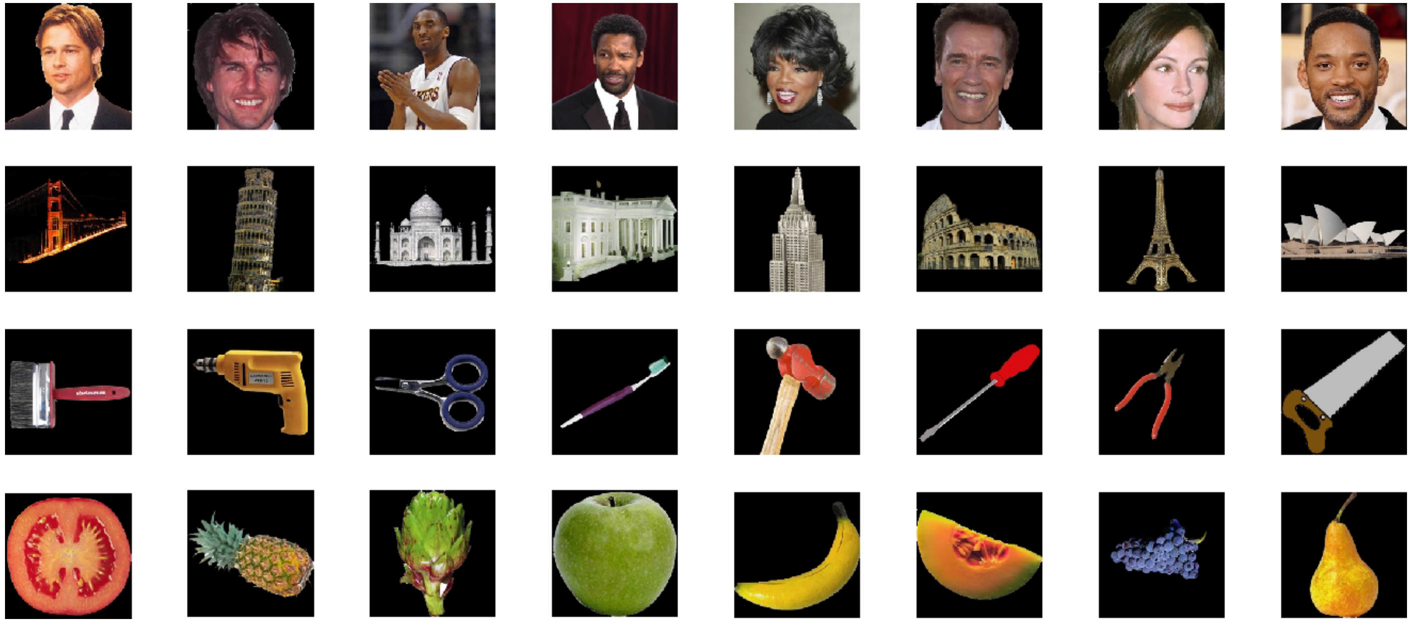
From Jody Culham

Support Vector Machines



Support Vector Machines

Decoding Perception and Imagery



Reddy et al., 2012

Support Vector Machines

Decoding Perception and Imagery

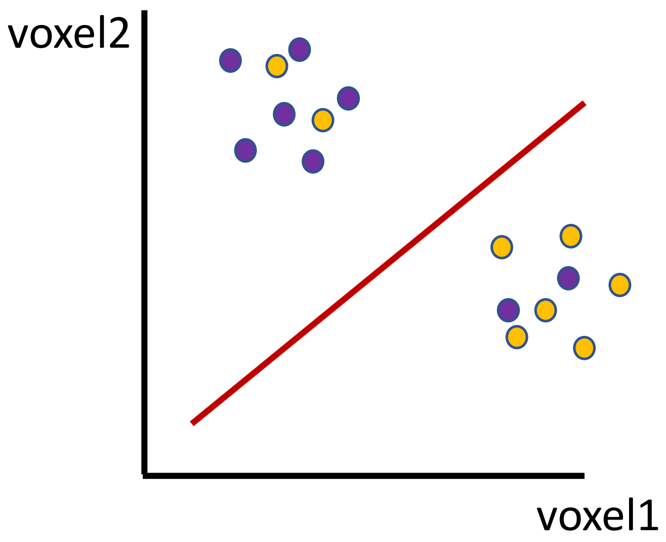
Reddy et al., 2012

Support Vector Machines
Decoding Perception and Imagery

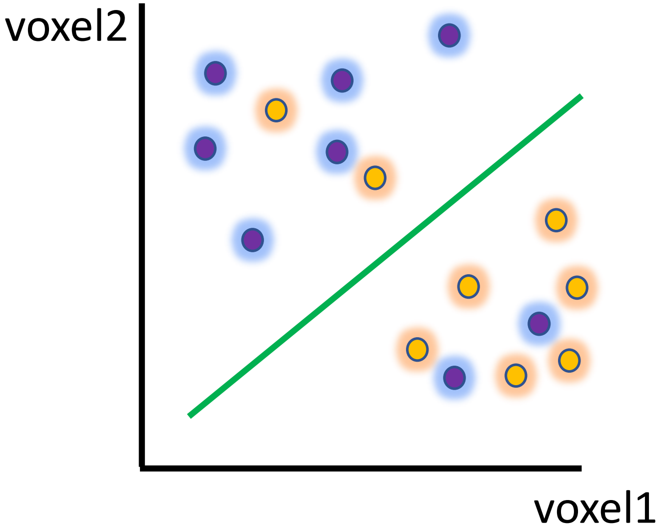
Question 1:
Is there information about perceived and imagined object categories?

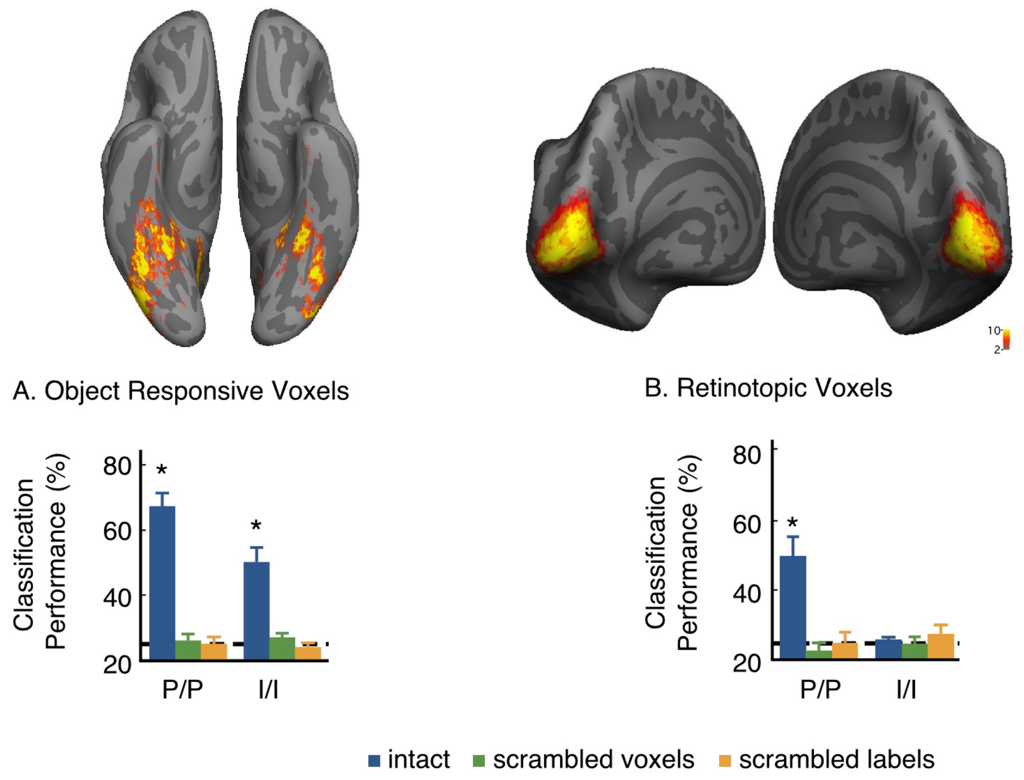
Reddy et al., 2012

Train Perception-Test Perception



Train Imagery- Test Imagery



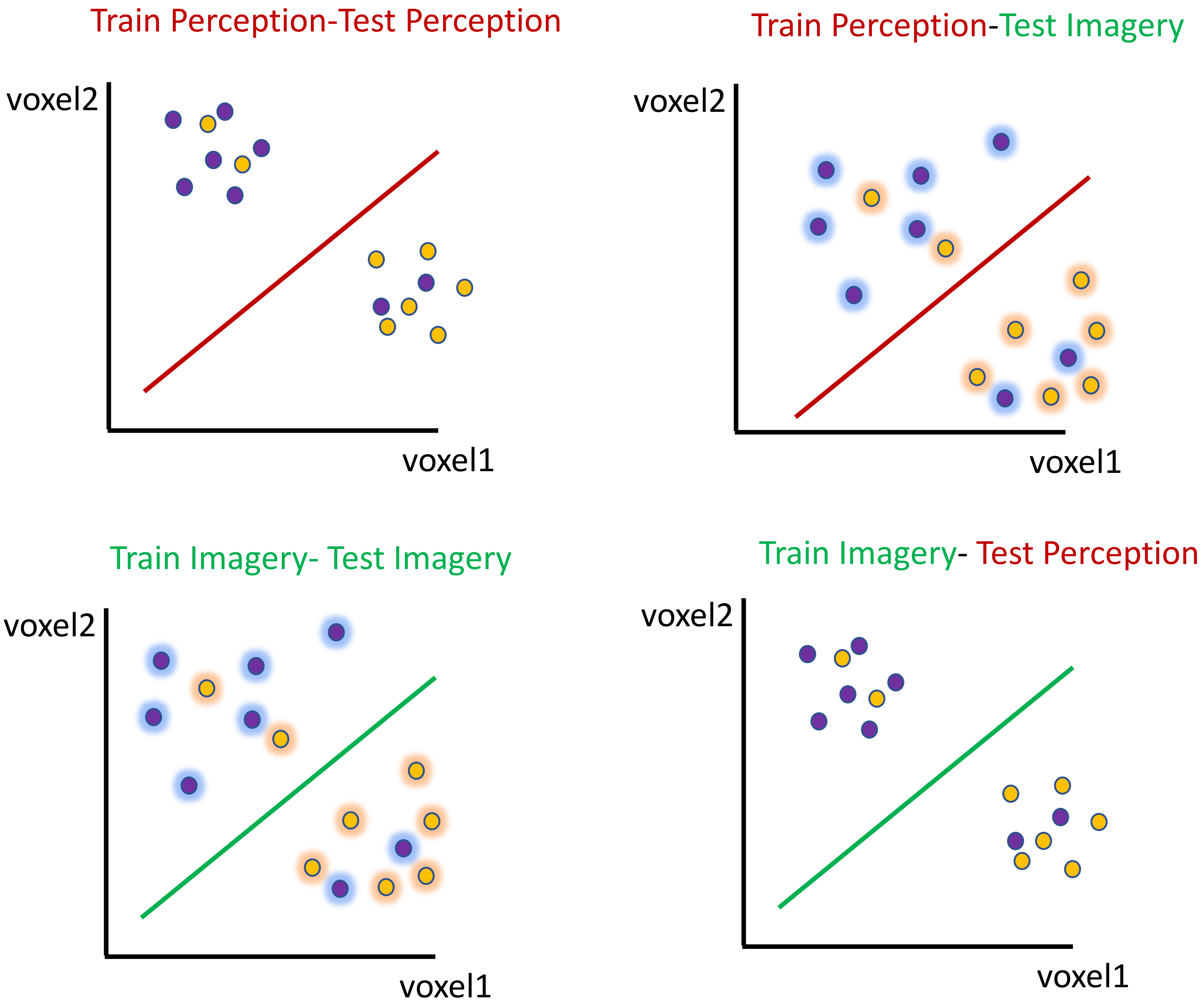


- Perception decoding from low- and high-level areas.
- Imagery decoding only in high-level regions.
- There is information about perceived and imagined objects in these regions.

Support Vector Machines
Decoding Perception and Imagery

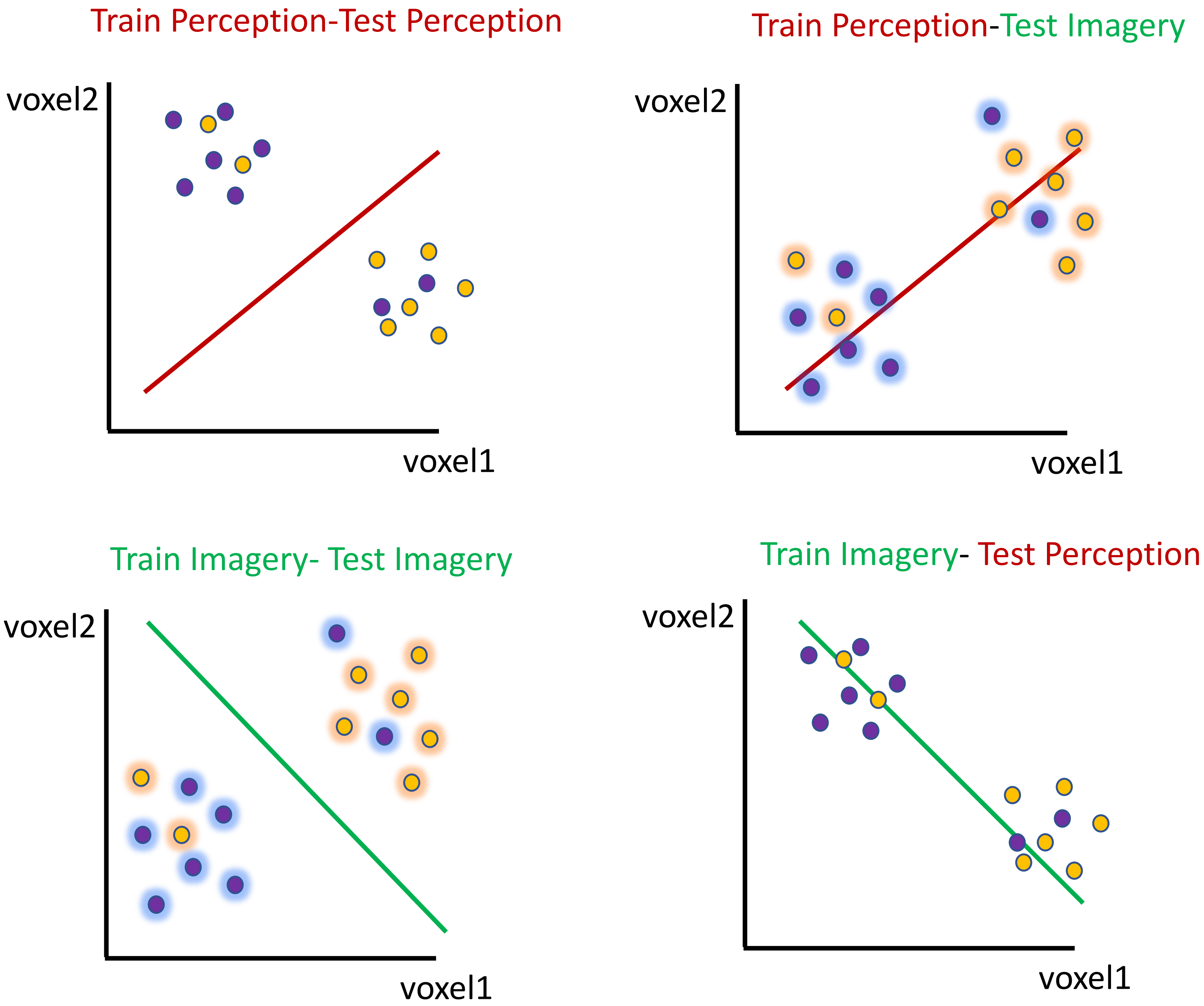
Question 2:
Do perceived and imagined object categories share common representations?

Reddy et al., 2012



Successful cross-decoding

- Perception and Imagery share common representations

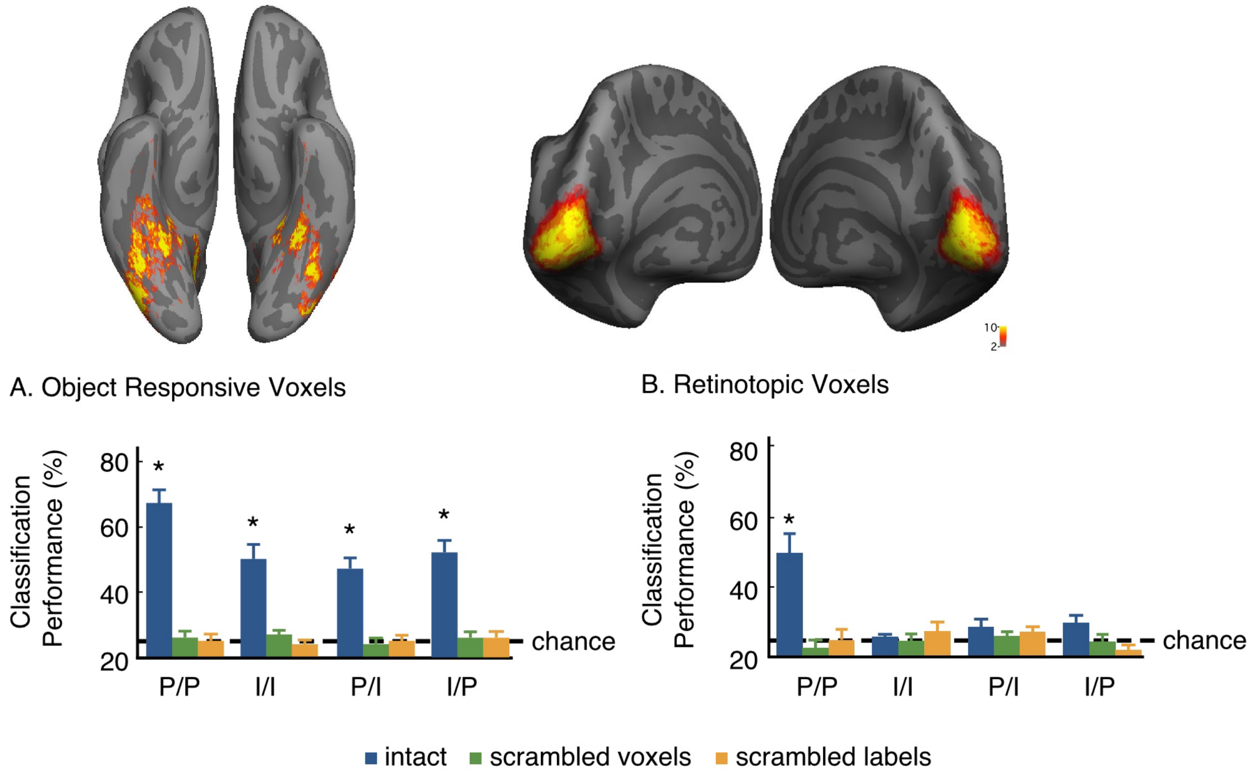


Unsuccessful cross-decoding

- Perception and Imagery do not share common representations

Support Vector Machines

Decoding Perception and Imagery



P/P: train and test on perception
I/I: train and test on imagery
P/I: train on perception, test on imagery
I/P: train on imagery, test on perception.

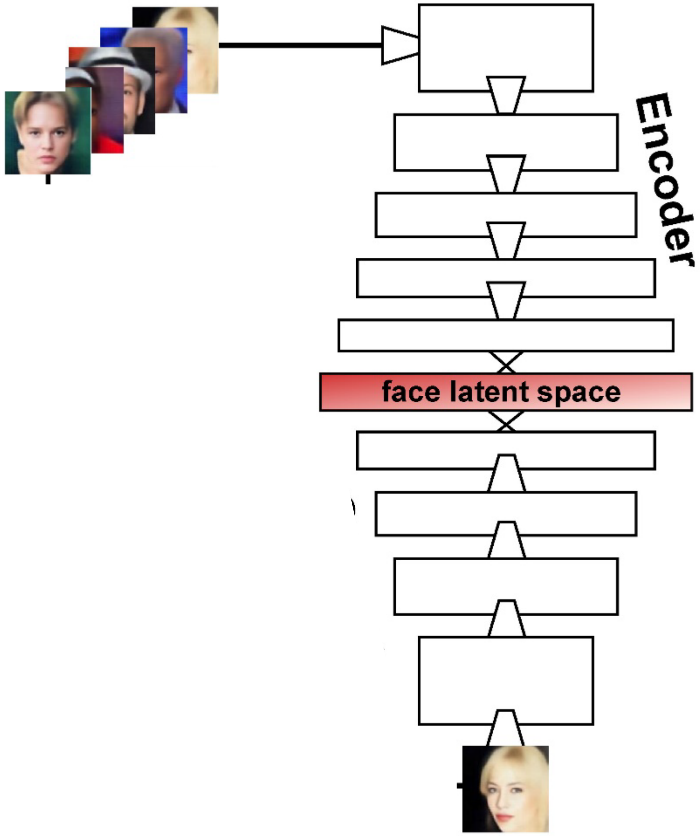
- Perception decoding from low- and high-level areas.
- Imagery decoding only in high-level regions.
- Perception and imagery share representations in high-level regions.

Reddy et al., 2012

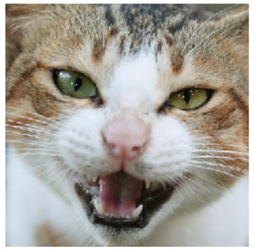
Decoding with DNN Latent Spaces

What is a latent space?

Decoding with DNN Latent Spaces



128x128x3
~50,000 dimensions

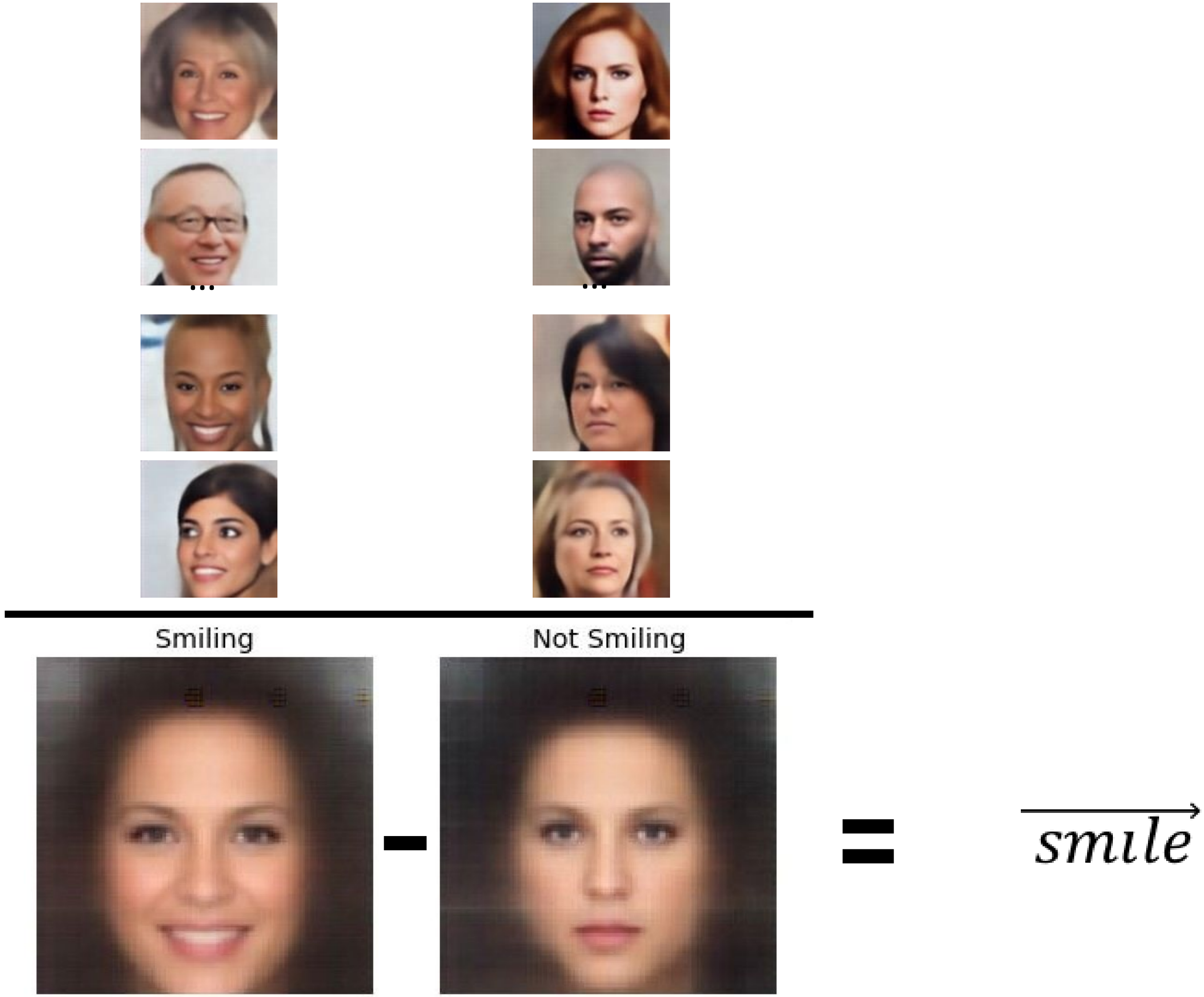


1024 latent dimensions

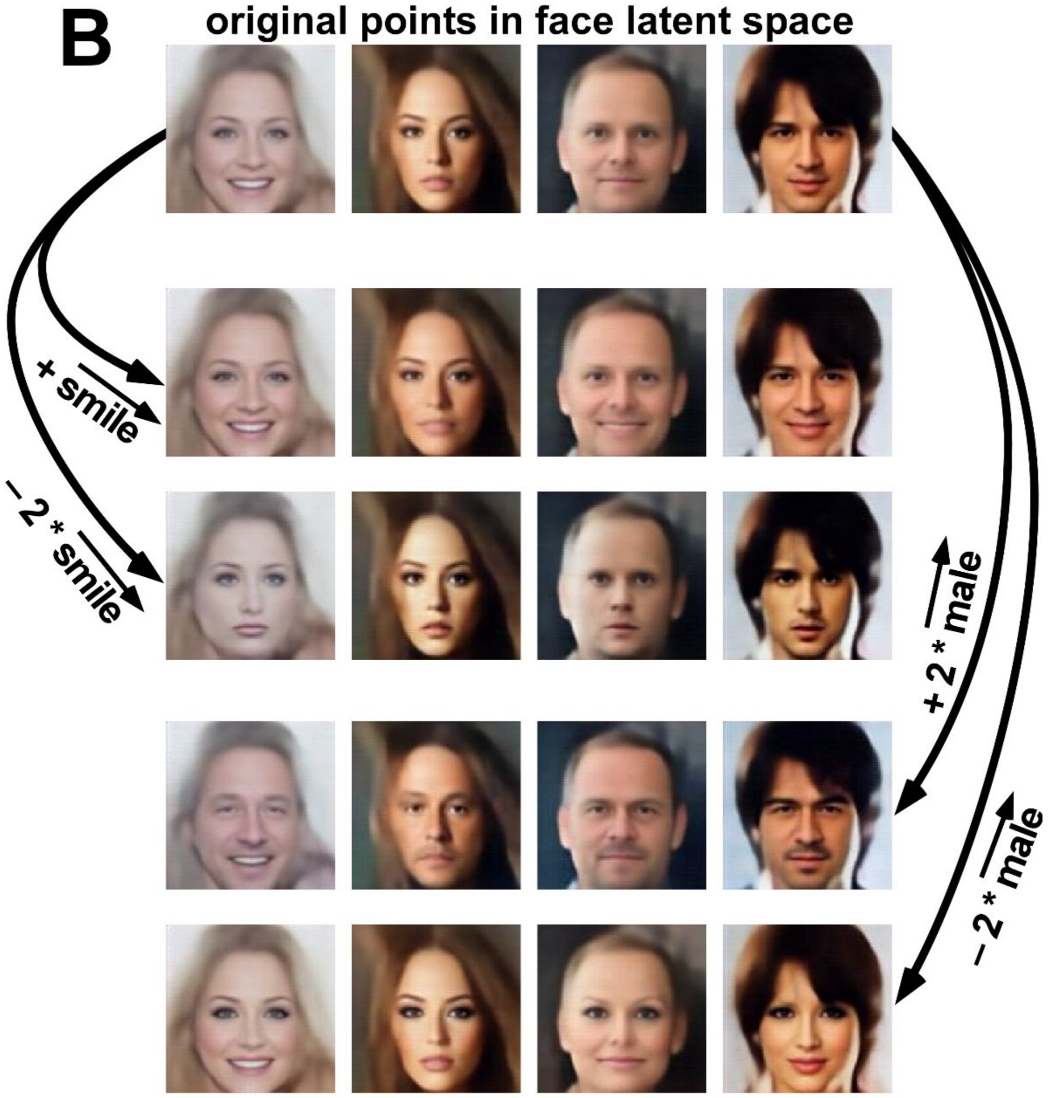
Face Latent Spaces

- The network translates an image into a code.
- It turns the code back into an image.
- The “code” defines a “**latent space**” of efficient dimensions.
- The latent space is an “internal representation” of the network.
- A point/vector in the latent space corresponds to a face.
- We can perform linear operations on these vectors and look at the faces that are generated.
- Do these vector operations make sense?

Face Latent Spaces



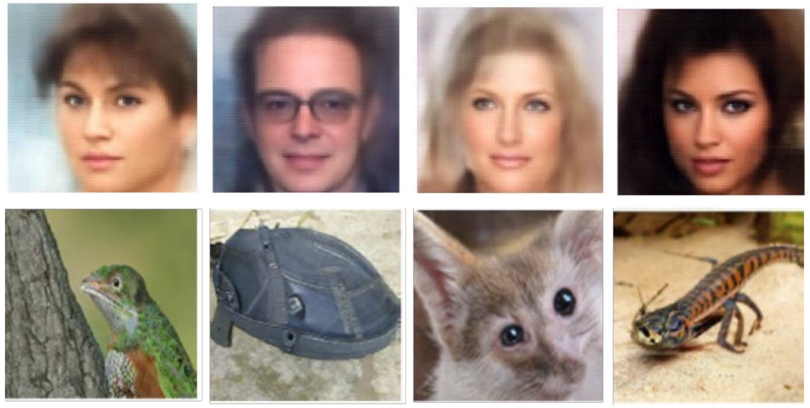
Face Latent Spaces

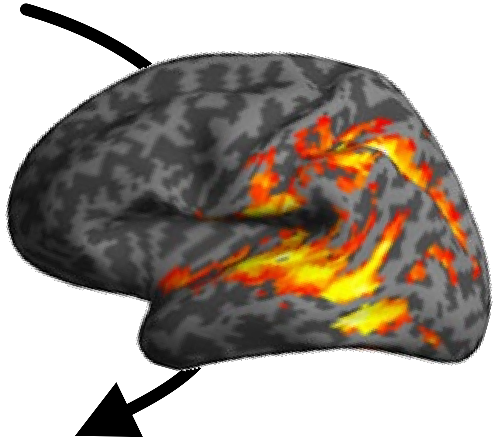


Decoding faces

images seen by subject in MRI scanner

images reconstructed from fMRI signals





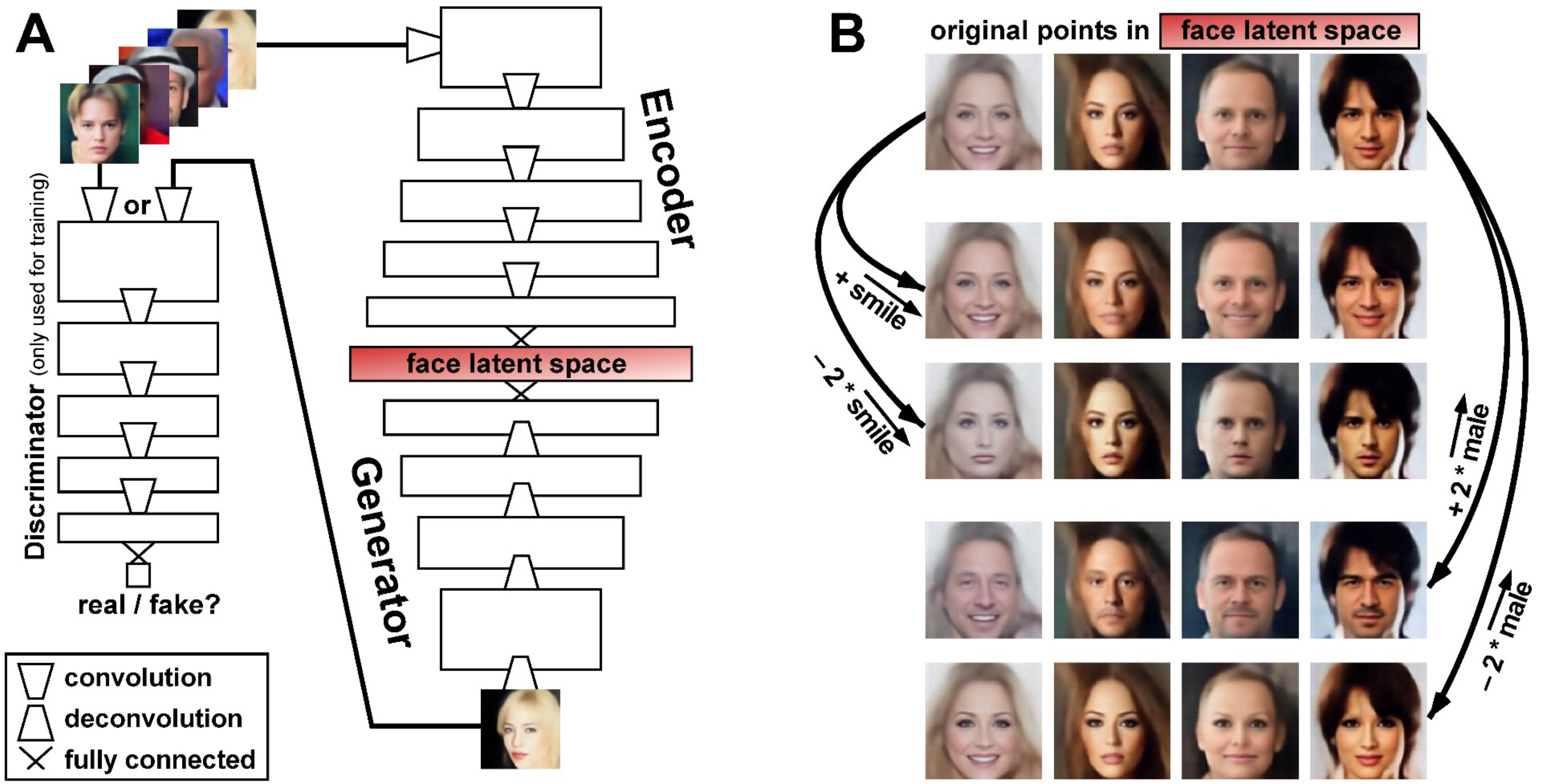
VanRullen & Reddy (2019) *Communications Biology*
Mozafari, Reddy & VanRullen (2020) *IJCNN'20*

Decoding faces – VAEGAN model

The network is trained on a celebrity dataset (200,000 images).

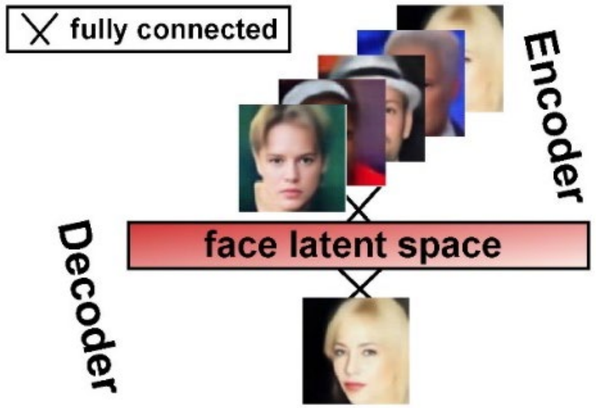
The encoder defines a "<u>latent space</u>" of 1024 dimensions.

After training, the weights are frozen and the discriminator network is dropped.



Larsen, A. B. L., Sønderby, S. K., Larochelle, H. & Winther, O. *ICML 2016*.

PCA as a comparison model

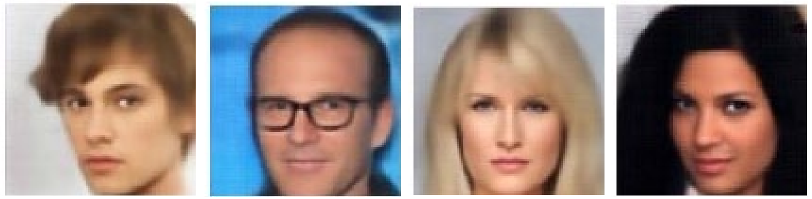


Faces are encoded into a latent space of 1024 principal components

➔ Previous state-of-the-art for face reconstruction from fMRI (Cowen et al., 2014; Lee et al., 2016.).

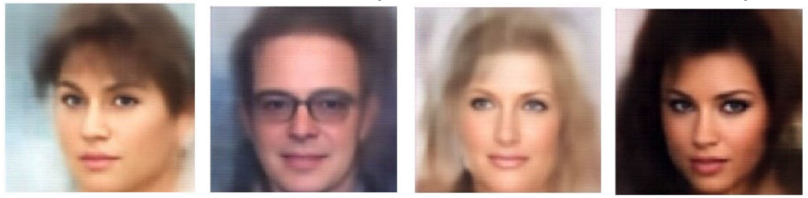
Face decoding via a GAN latent space

images seen by subject in MRI scanner

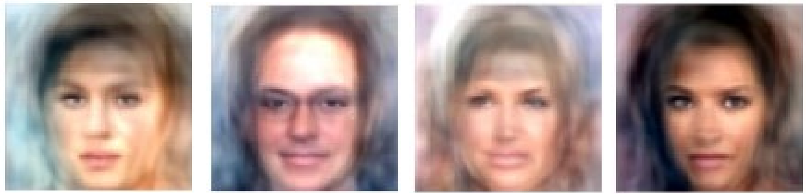


images reconstructed from fMRI signals

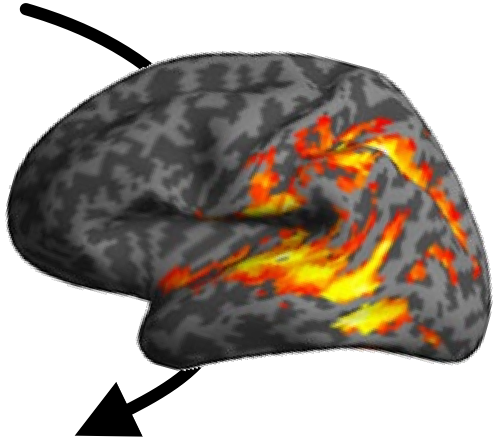
VAE-GAN model (1024 latent dims.)



PCA model (1024 latent dims.)



Cowen et al., 2014; Lee et al., 2016.

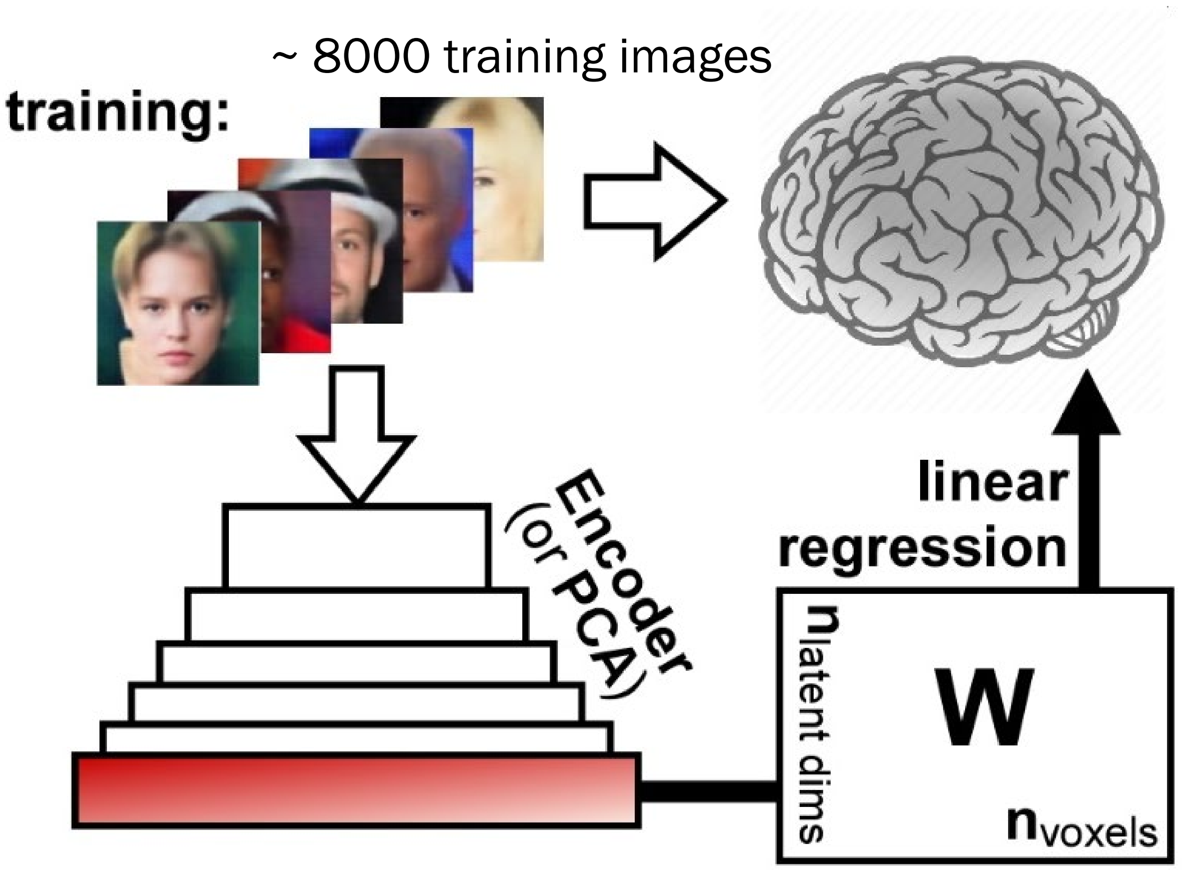


- 4 subjects
- ~8000 faces/subject
- 16 hours of scanning/subject

➔Freely available on OpenNeuro

VanRullen & Reddy, 2019. *Communications Biology*

Training a decoder



$$Y = XW \quad \Leftrightarrow \quad X^TY = X^TXW$$

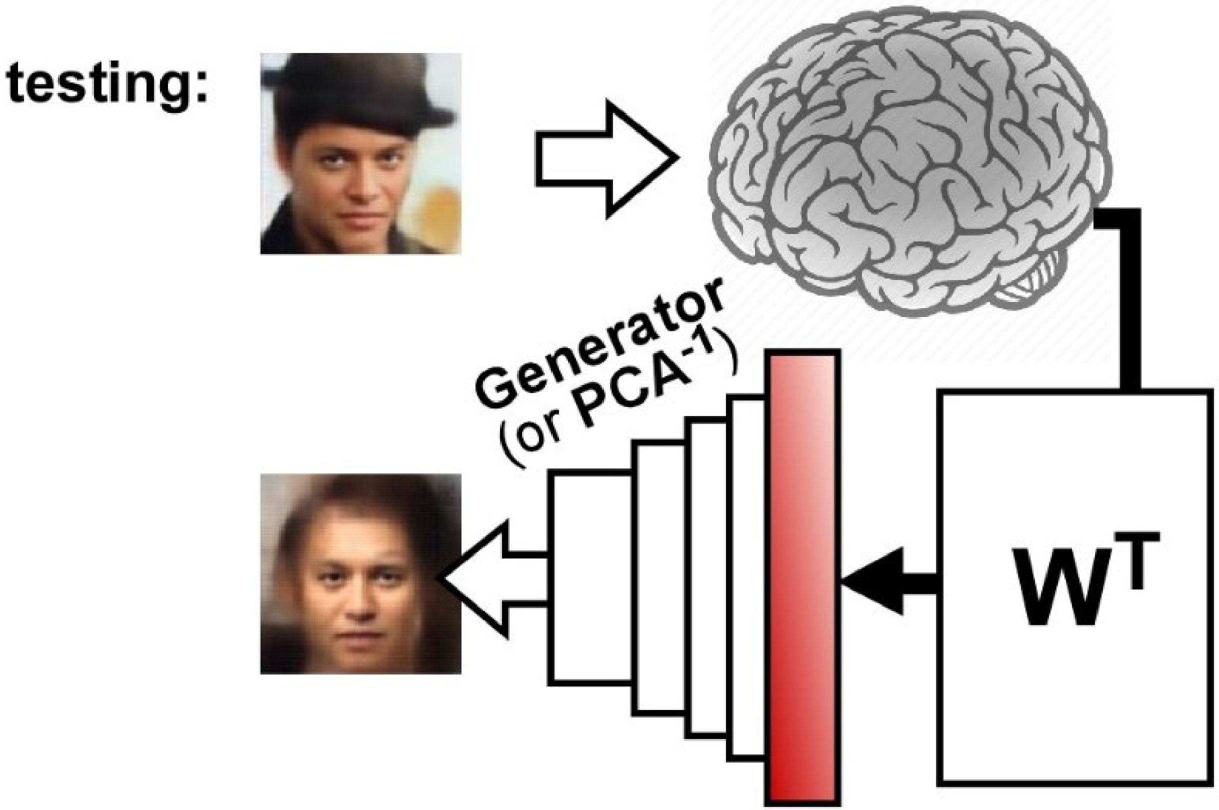
Y: (8000, n_{voxels}); X: (8000,1024); W: (1024, n_{voxels})

$$\boxed{W = (X^TX)^{-1}X^TY}$$

VanRullen & Reddy, 2019. *Communications Biology*

Testing the decoder

20 test images (x 45 repeats)



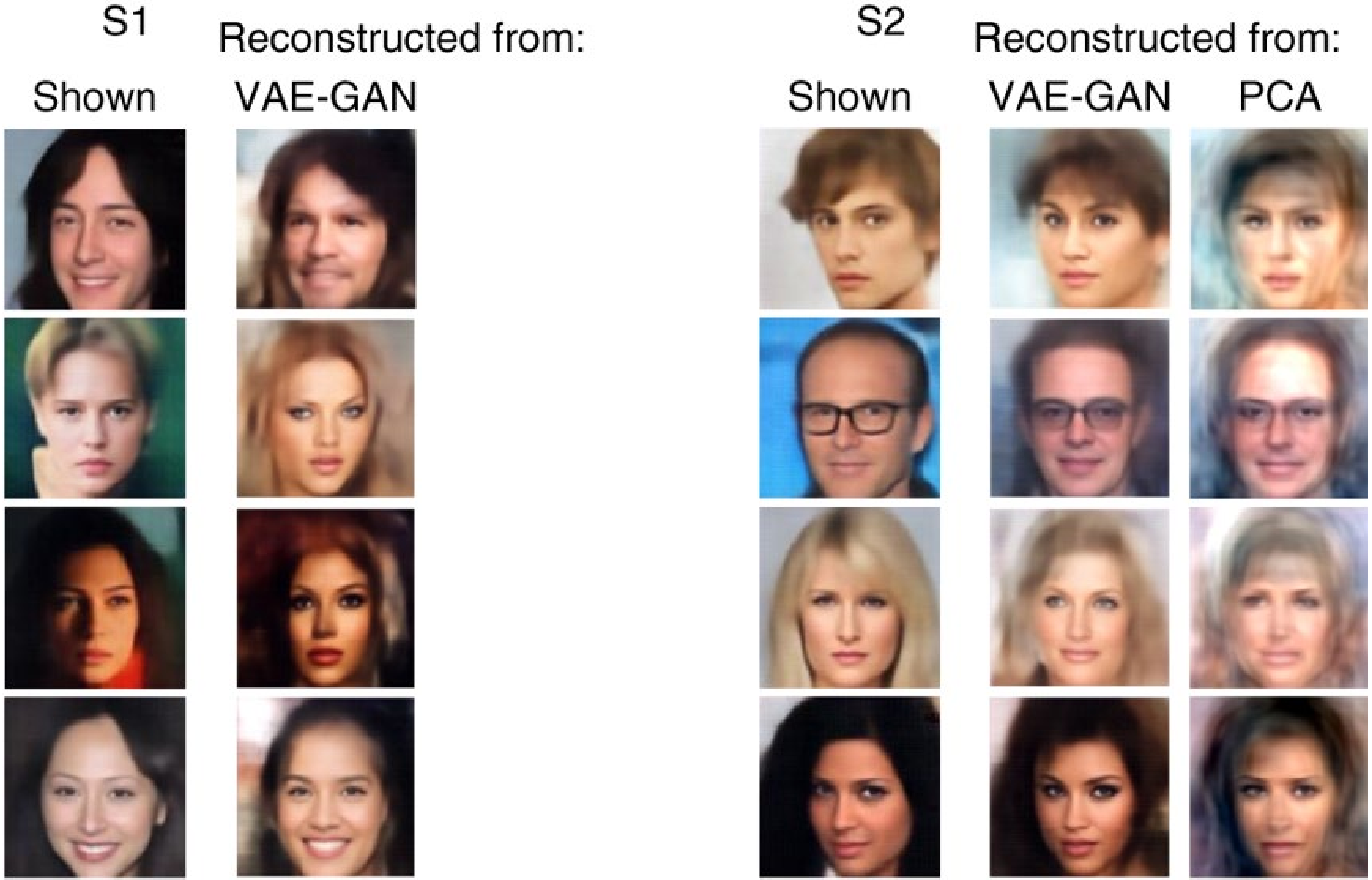
$$Y = XW \quad \Leftrightarrow \quad YW^T = XWW^T$$

(20,n_{voxels}) (20,1024) (1024,n_{voxels})

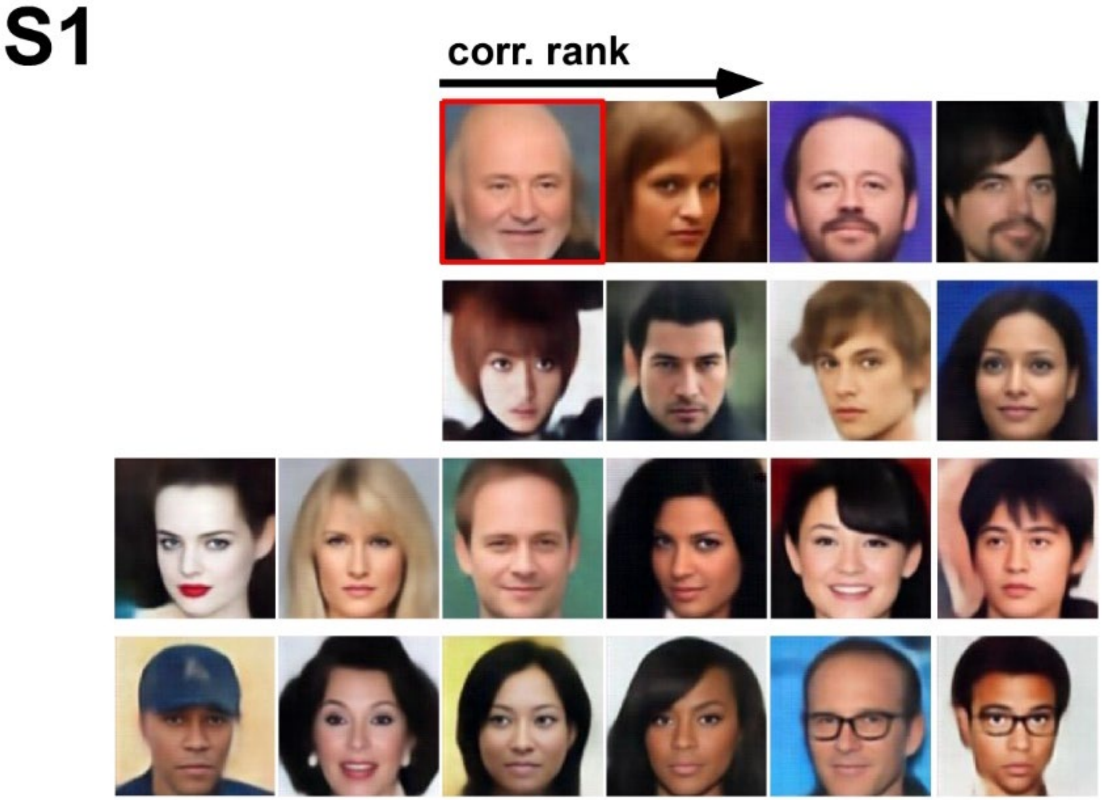
$$\boxed{X = YW^T(WW^T)^{-1}}$$

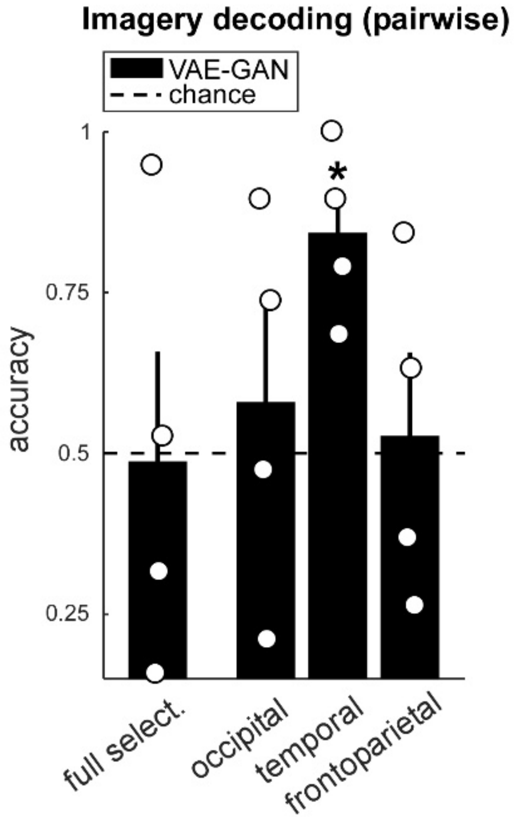
VanRullen & Reddy, 2019. *Communications Biology*

Face Reconstructions



Decoding mental imagery





VanRullen & Reddy, 2019. *Communications Biology*

Today we have seen…

- What do we mean by brain decoding
- Exploiting patterns of activity for decoding
- Different methods for decoding