University of Amsterdam

Symposium Computational cognitive neuroscience: What, why and how

Flavors of modelling in cognitive neuroscience

Lukas Snoek

Contents

- Aims of cognitive neuroscience
- Explanations in cognitive neuroscience
- Types of computational models
  - Stimulus-based vs. behavior-based
- Techniques to relate models to the brain
  - Encoding models
  - Representational similarity analysis (RSA)

Let’s take a step back …

What do we want to know at our department (brain & cognition)?

Adapted from Gazzaniga et al., 2014

Aim of brain & cognition (cogn. neuro.)

- First and foremost: understand **behavior and cognition**
  - We're psychologists, not biologists
- … but also use the **brain** in our explanations
  - We're the department of **brain** & cognition
  - Also: because this is what gives rise to cognition and behavior, after all

In other words...

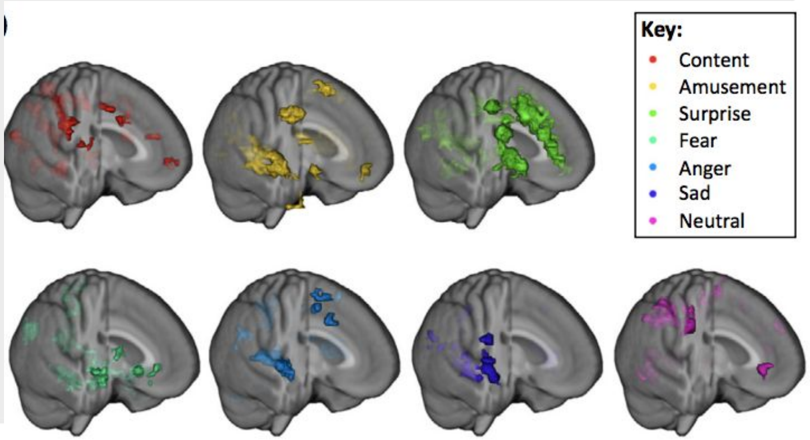
- Closing **Leibniz' gap**: how does the brain give rise to psychological phenomena?
- Explaining psychological effects (explananda) by functional analysis (decompose effect into underlying properties and mechanism)
  - Constrained/influenced by brain measurements

So, how have we gone about using the brain in our explanations?

Kragel & LaBar, 2016

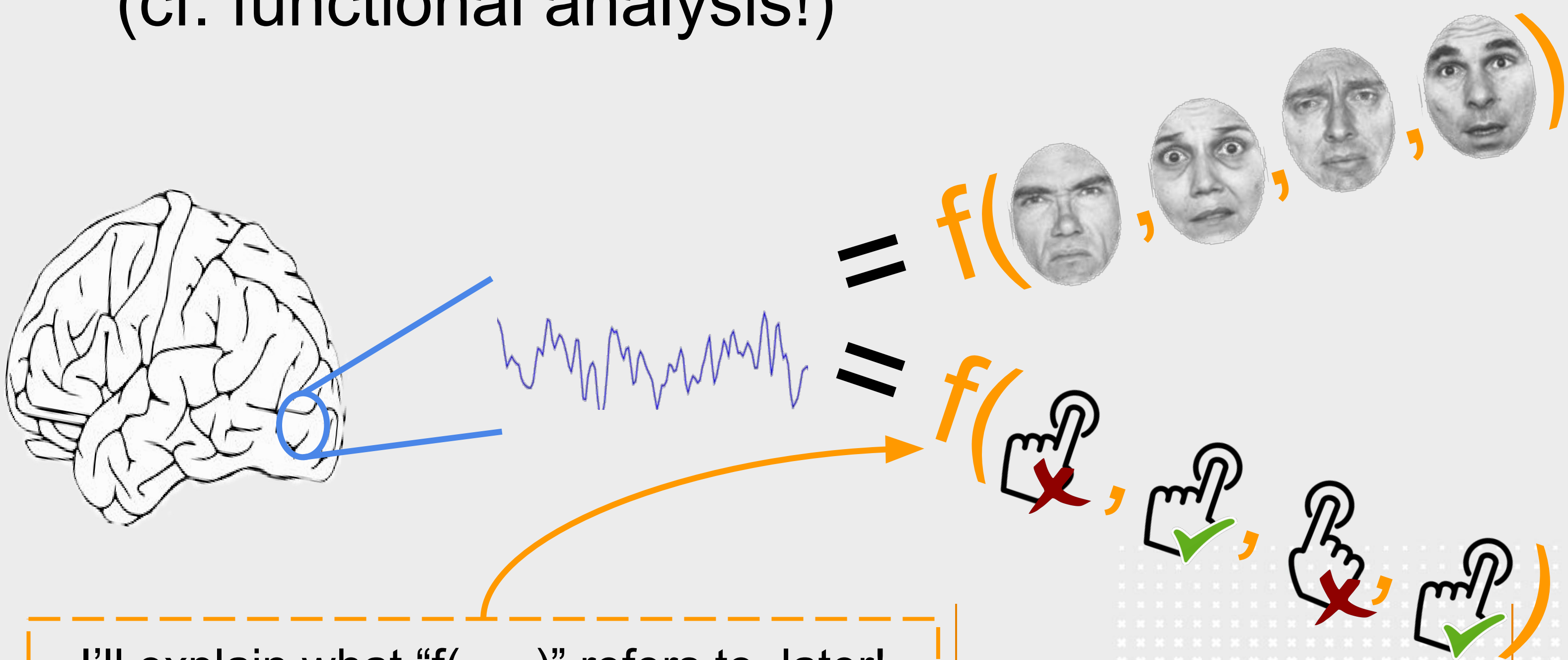
Explanations in cognitive neuroscience

- One way we use the brain is to **validate** our theories and hypotheses about behavior and cognition
- For example, the theory that there are 6 categorical emotion categories has been "verified" with fMRI



Explanations in cognitive neuroscience

- In other words, we model brain measurements as a **function of stimulus or behavior properties** (cf. functional analysis!)

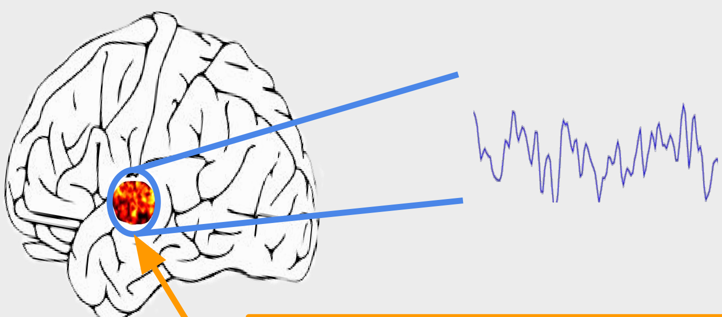


I'll explain what "f(…)" refers to, later!

Explanations in cognitive neuroscience

- Implicit assumption:

  "If we can accurately predict the responses of all neurons in an area, we have captured the computations up to that area." (Kriegeskorte & Kievit, 2013)



= f(anger)

Given that there is a (specific) region for anger, anger must be a basic, real concept represented/computed by the brain

So, what is this “computational” or “model-based” cognitive neuroscience, then?

Based on Kay, 2017

Every analysis is a model!

- *All analyses that relate stimuli/cognition/behavior to the brain are (computational) models!*
  - Including standard EEG/MEG/fMRI analyses
  - Difference faces / houses, unattended / attended — all models!

Based on Kay, 2017

Every analysis is a model!

- All analyses that relate stimuli/cognition/behavior to the brain are (computational) models!
- But models differ in their **computational details** and **sophistication**

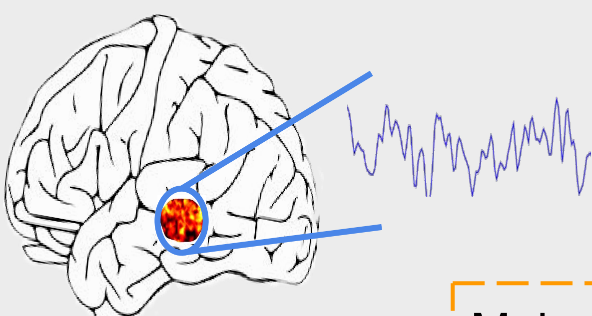
Based on Kay, 2017

Every analysis is a model!

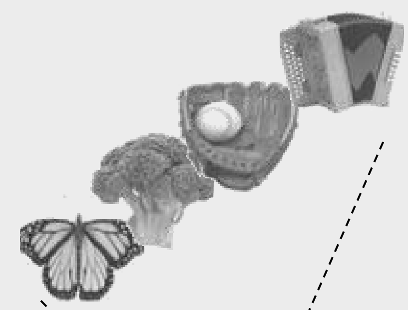
- All analyses that relate stimuli/cognition/behavior to the brain are (computational) models!
- But models differ in their **computational details** and **sophistication**
- We can specify models based on “folk psychology”, computationalism, or connectionism, which all differ in how “computationally sophisticated” they are

Folk psychological models

- Very common in cognitive neuroscience
  - e.g., research on “object representation”



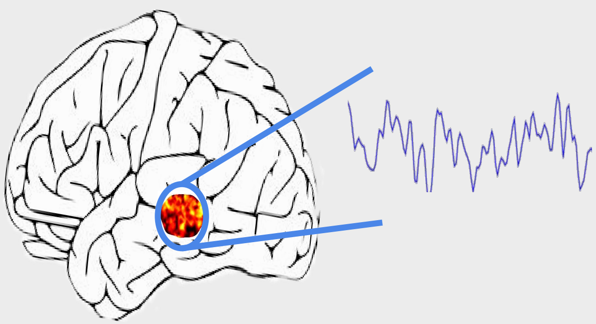
= **f(**object category**)**



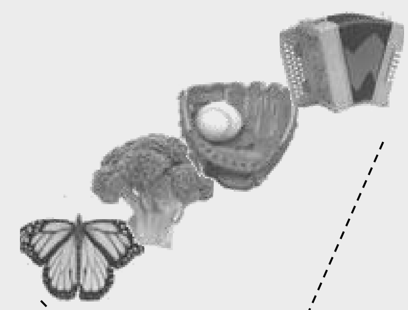
Majority of the traditional “brain mapping” studies: face vs. house, attended vs. unattended, etc.

Computationalism-inspired models

- Gaining popularity in cognitive neuroscience
  - e.g., research on “object representation”

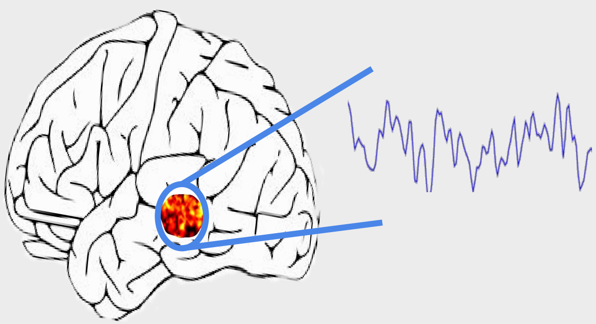


= f(δ(object, exemplar))

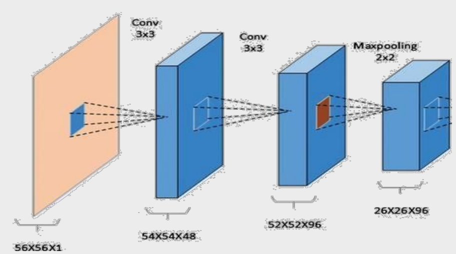


Connectionism-inspired models

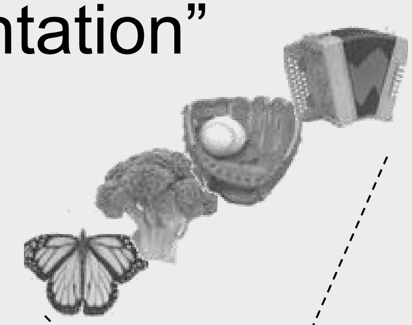
- Revival in cognitive neuroscience due to regained interest in artificial neural networks
  - e.g., research on “object representation”



= f(

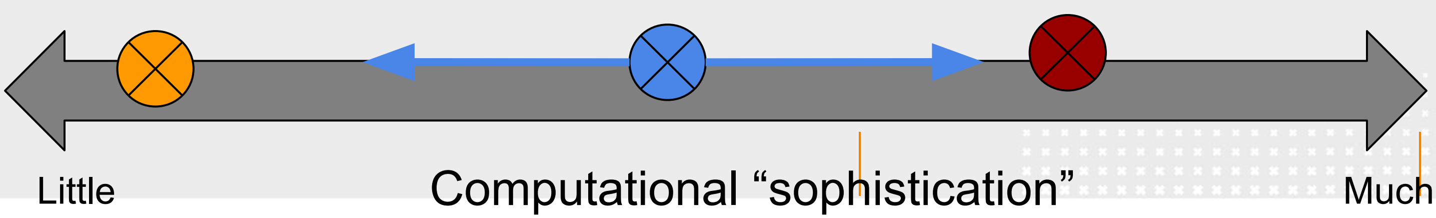


)



Every analysis is a model

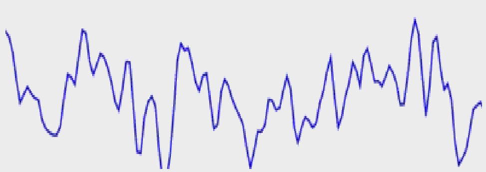
- All analyses that relate stimulus or behavioral properties to the brain are "computational" or "model-based"
- Models can be inspired by **folk psychology** concepts, **computational** principles, or **connectionist** (ANN) properties



How to decide on your model?

It’s all in the “features”

- A model should reflect **your hypothesis** of how the brain represents/processes the stimulus/task
- Your hypothesis can be summarized in one or multiple “**features**”, i.e., whatever goes inside the function:

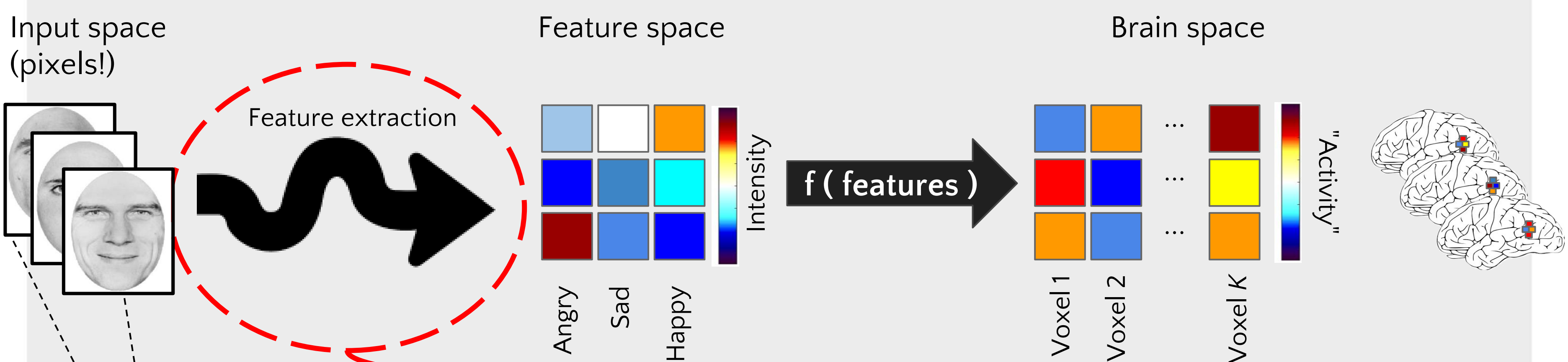


$= f(\ldots)$

Based on Naselaris & Kay (2011)

Input space, feature space, brain space

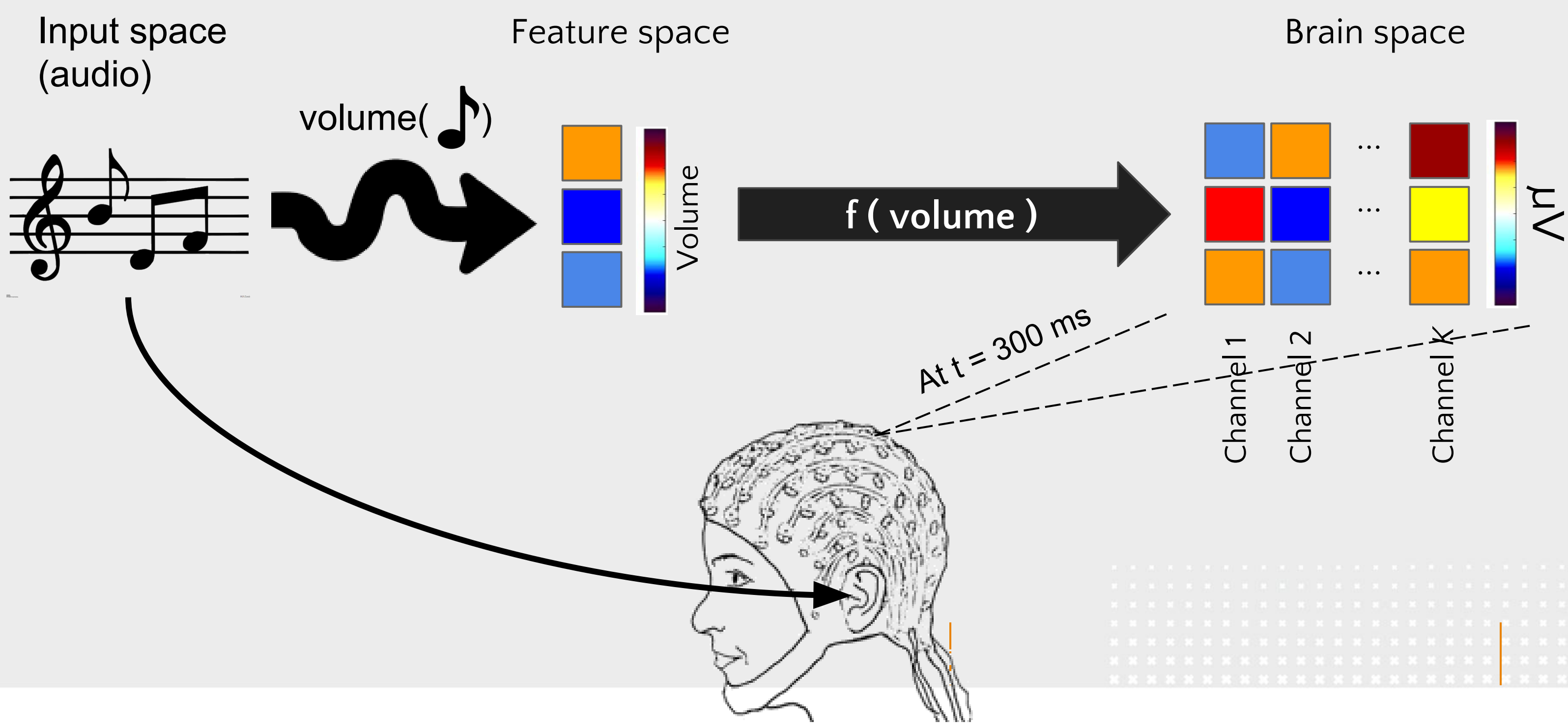
(Note: this is an example of a “folk psychology” model!)



How we infer/compute/define our features represents our hypothesis (model) of how the brain does this!

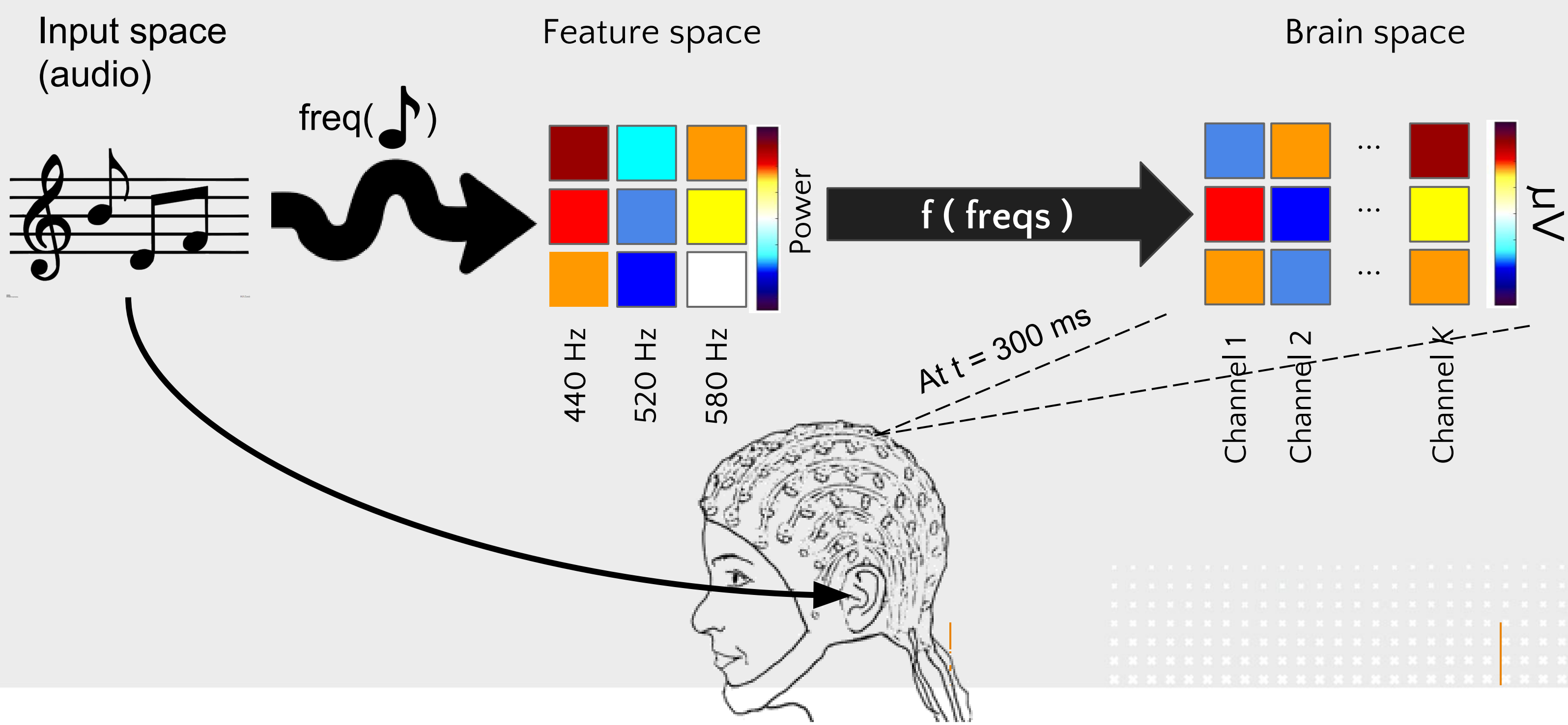
Another example:

How does the brain process auditory information over time?



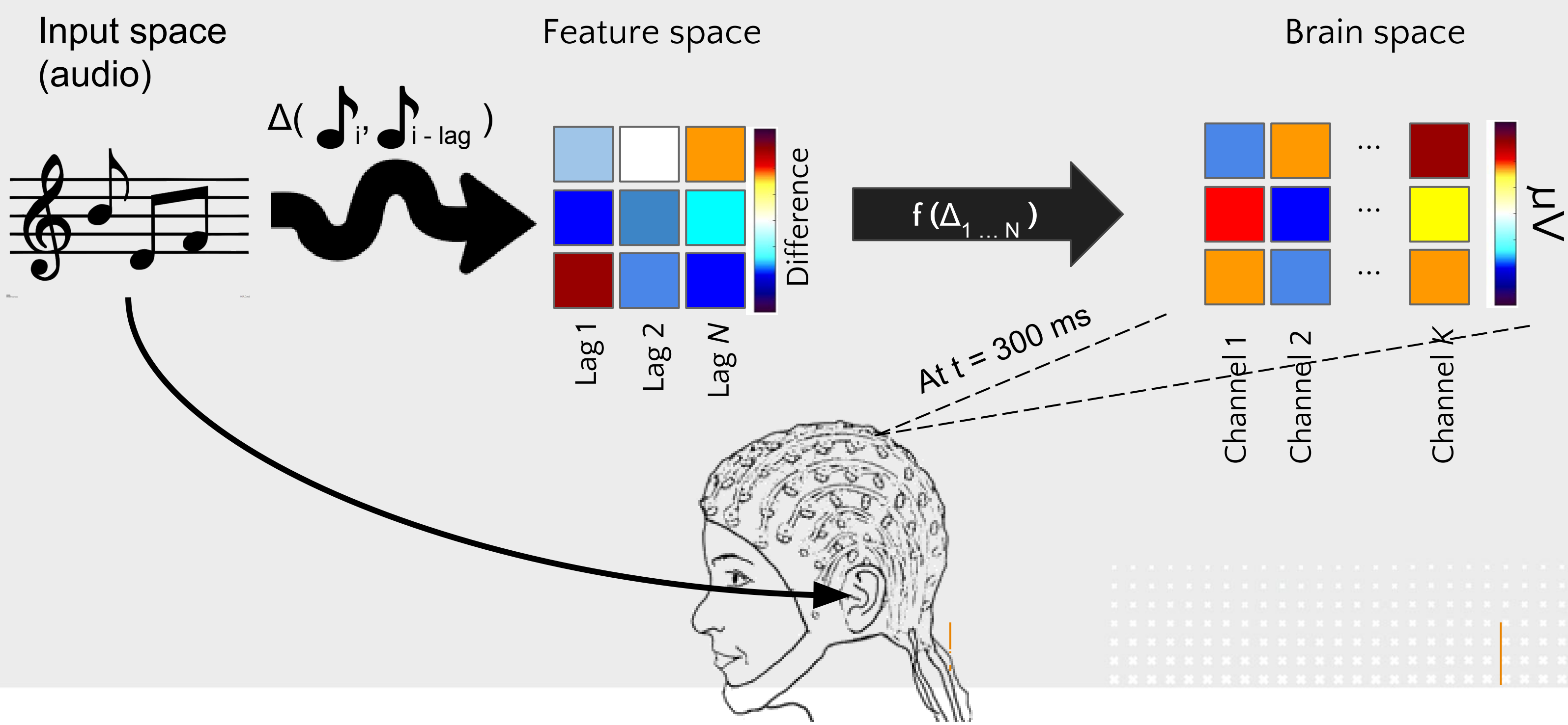
Another example:

How does the brain process auditory information over time?



Another example:

How does the brain process auditory information over time?

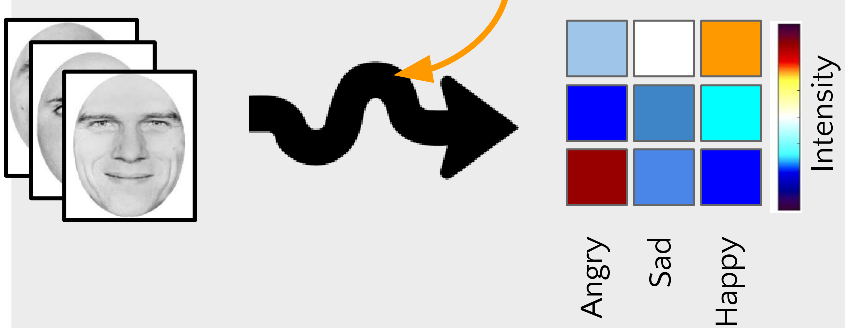


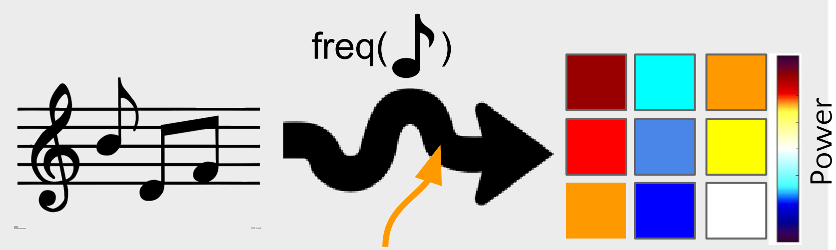
Conceptual vs. computational

Relative!

- Conceptual models do not *compute* (but "hard code") features
- Computational models *do* compute features (and thus specify a mechanism)

"Oracle", "epiphenomenal", "folk psychology models





$$\mathbf{F}(n) = \sum_{k=0}^{N-1} f(k)\, e^{-j2\pi nk/N}$$

Interim summary

- All analyses are computational, but the relative complexity depends on how you define/compute/extract your **features**
- Your model on which you base your features should reflect your **hypothesis** about the corresponding brain mechanism

From parameter testing → prediction

- Traditional analyses often focus on **parameter inference** (is my factor significantly different?)
  - ERPs and contrast-analyses in fMRI
  - $y_{\mu V,\, t\, =\, 300} = X_{freq\, =\, 400}\boldsymbol{\beta_1} + X_{freq\, =\, 500}\boldsymbol{\beta_2} + \varepsilon$
- Computational models (often) focus on **model fit** primarily ("How accurate is my model?")
  - Prediction, $\hat{y} = X_{freq\, =\, 400}\mathbf{b_1} + X_{freq\, =\, 500}\mathbf{b_2}$
  - Model fit = correlation(y, $\hat{y}$)

Model comparison

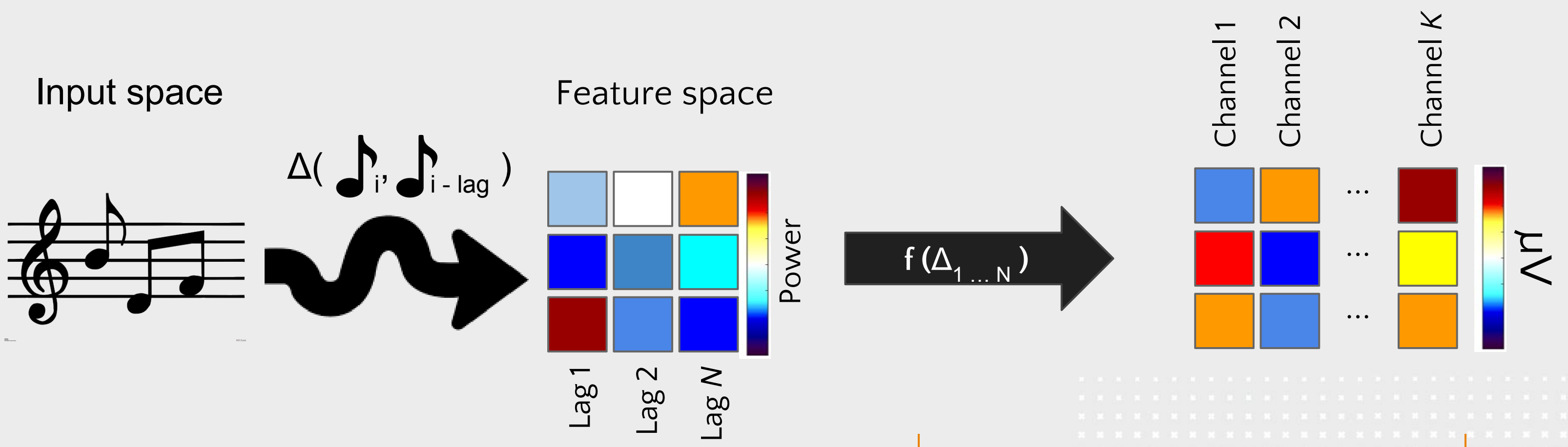
- Focusing on model fit allows for easy model comparison
- Is my **volume**-model better than my **frequency**-model in explaining activity during auditory processing?
  - correlation($\hat{y}$, f(X_{volume})) > correlation($\hat{y}$, f($X_{frequency}$)) ???
- Compare across **time** (EEG/MEG) or **space** (fMRI):
  - e.g., frequency model best at 70-150 ms, but frequency better at 150-300

What are the major ‘flavors’ of computational models in cognitive neuroscience?

University of Amsterdam

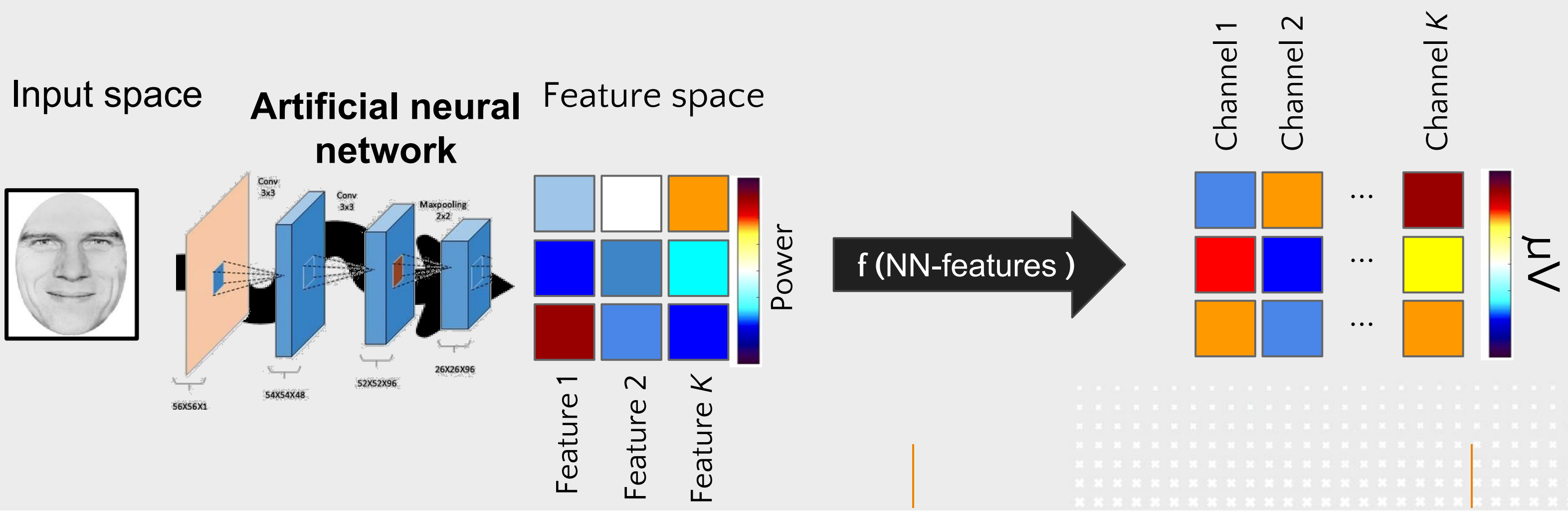
Flavors of computational models

- **Stimulus**-based models
  - Bread-and-butter of sensory neuroscience



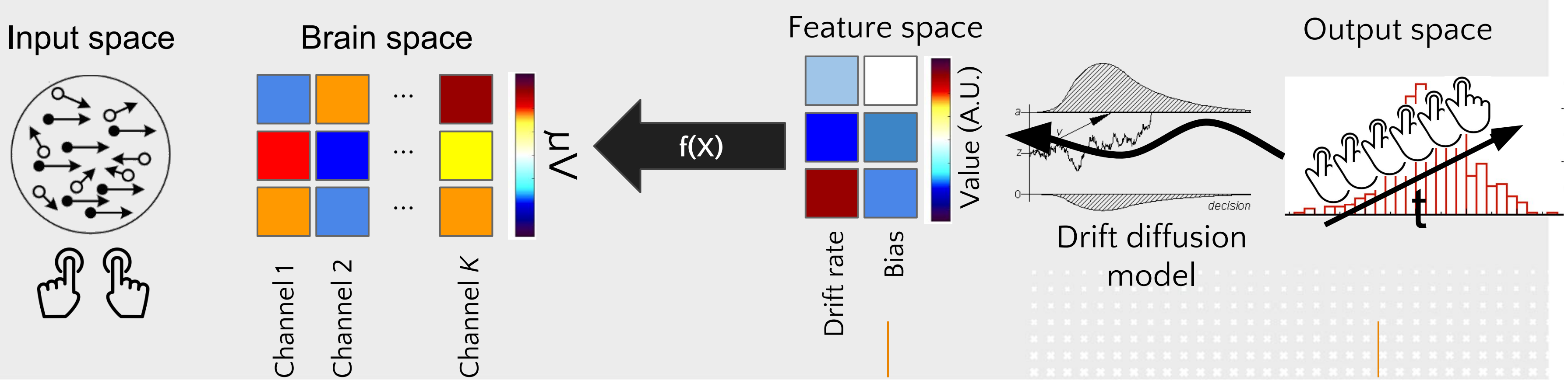
Flavors of computational models

- **Stimulus**-based models
  - Bread-and-butter of sensory neuroscience



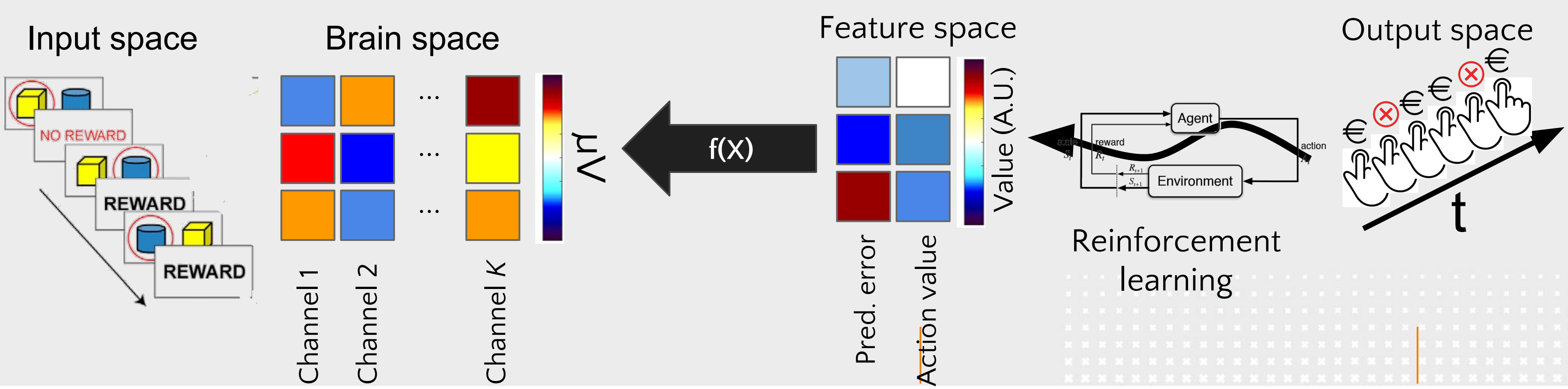
Flavors of computational models

- Stimulus-based models
- **Behavior**-based models
  - Often inspired by existing cognitive models



Flavors of computational models

- Stimulus-based models
- **Behavior**-based models
  - Often inspired by existing cognitive models



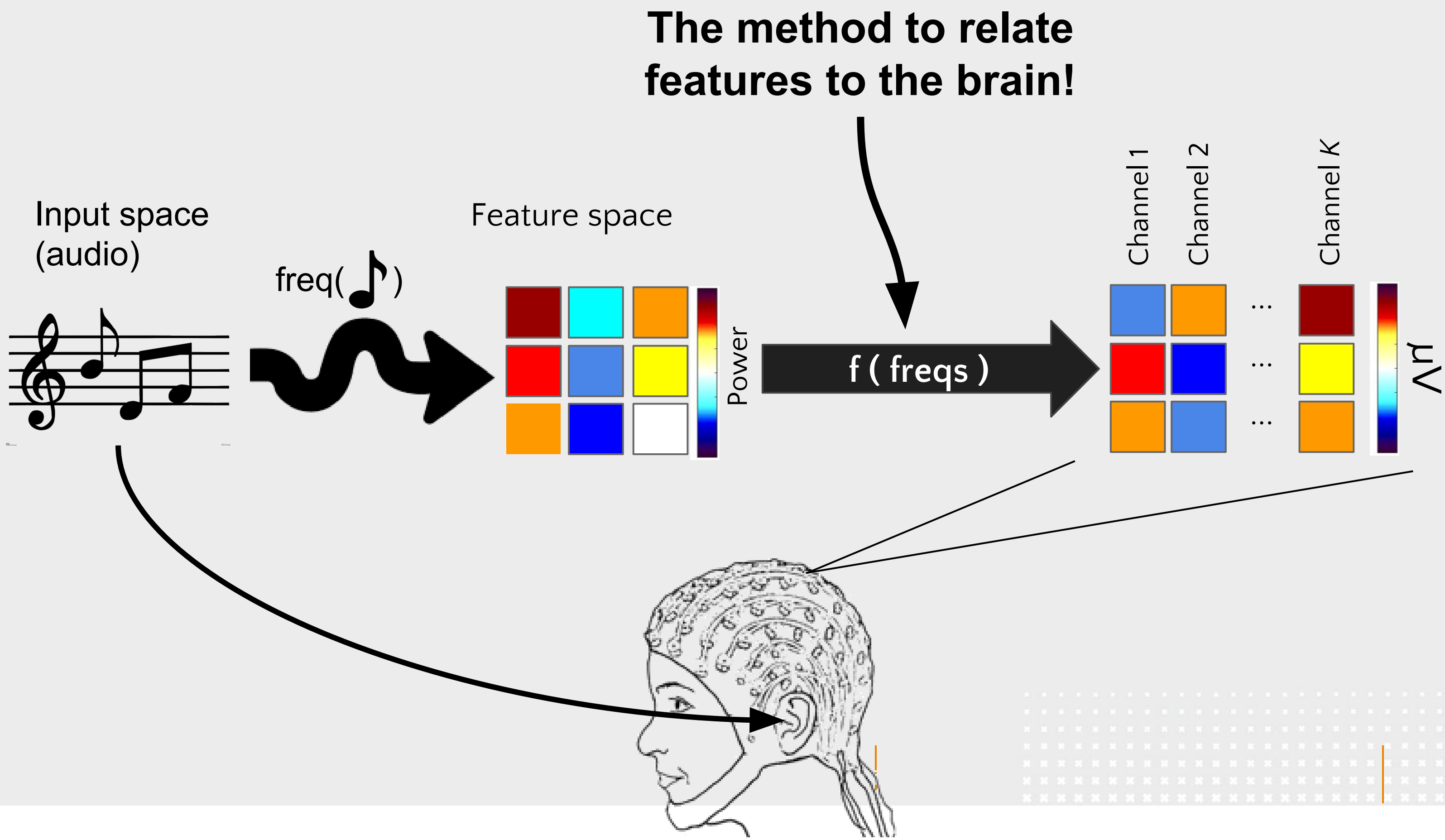
Interim summary

- The strength of computational research is often not due to state-of-the-art **methods**, but due to smart **operationalizations of features**
  - Derived from existing cognitive/behavioral models or artificial neural networks

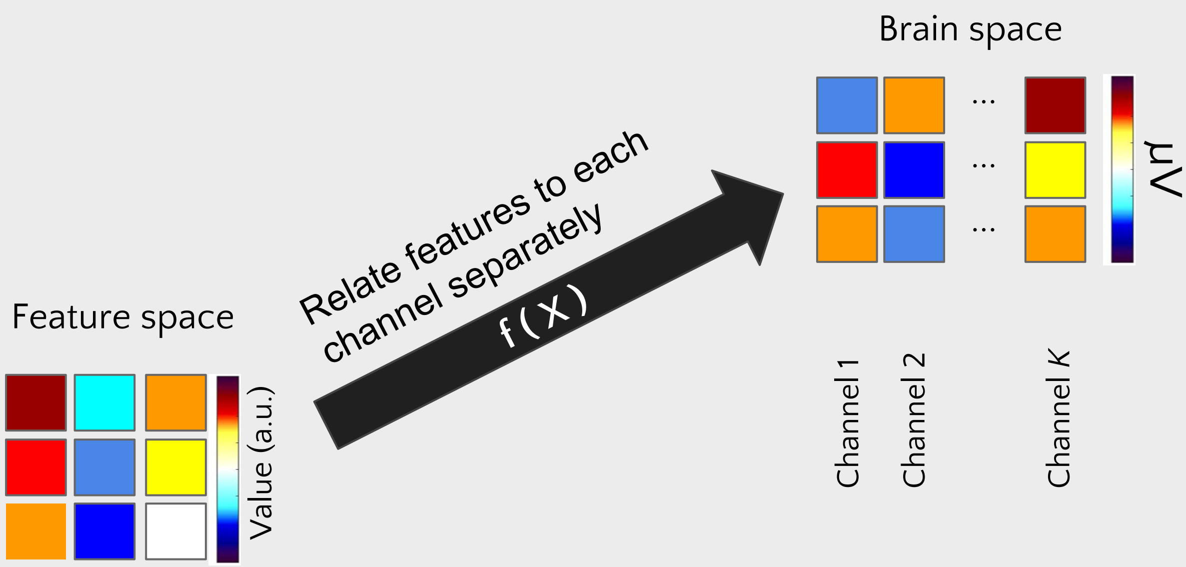
Methods for relating features to brain measurements

or: what on earth do you mean with **f(** … **)**?

What on earth is "f()"?

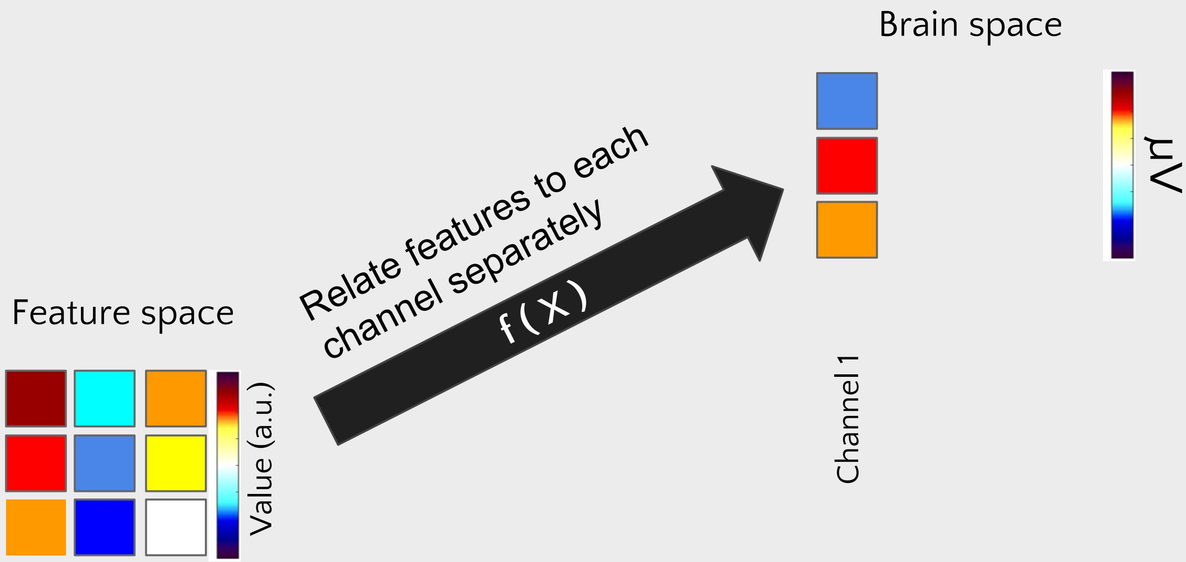


Two types of methods



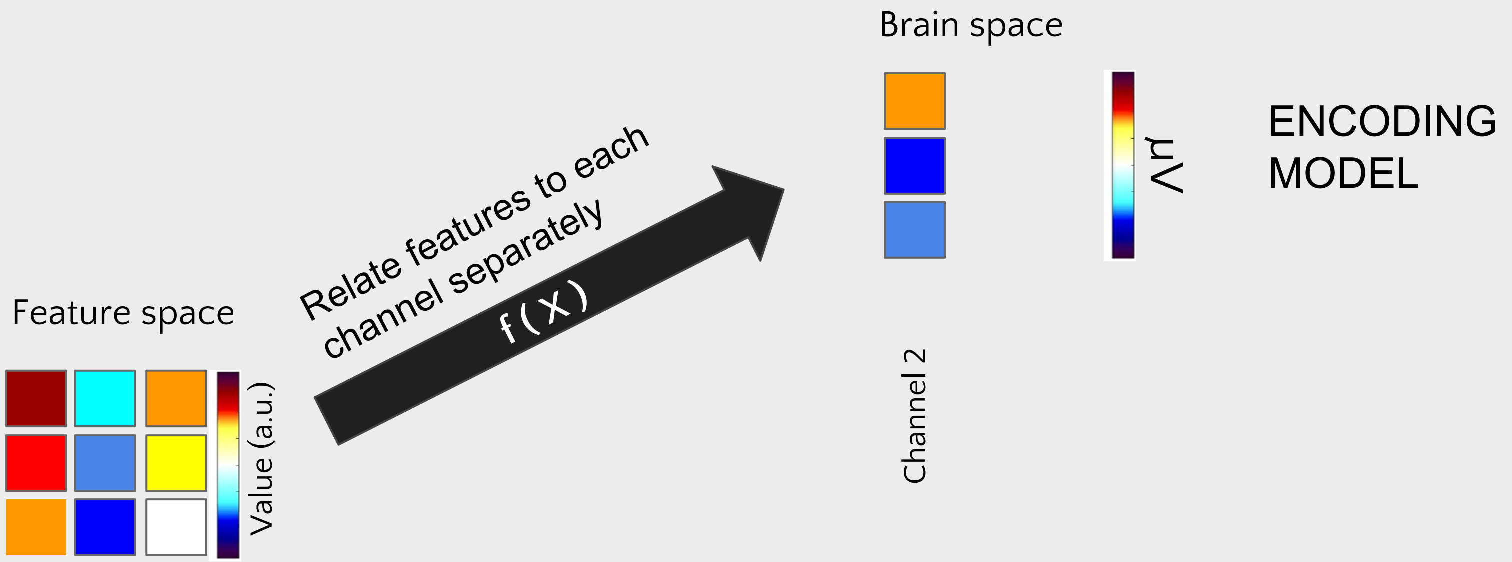
ENCODING MODEL

Two types of methods

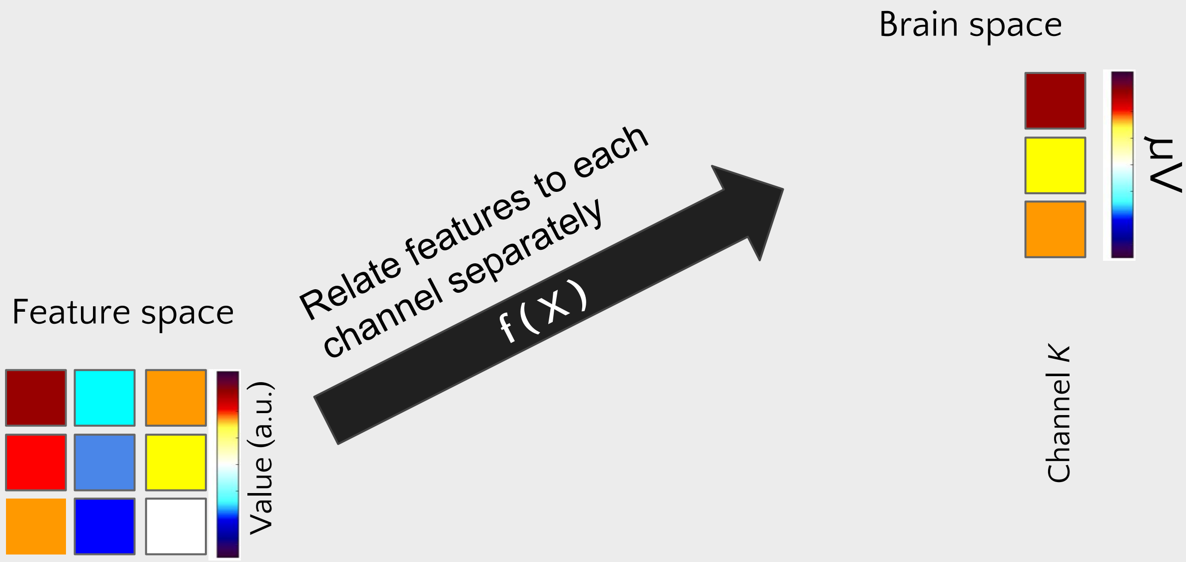


ENCODING MODEL

Two types of methods

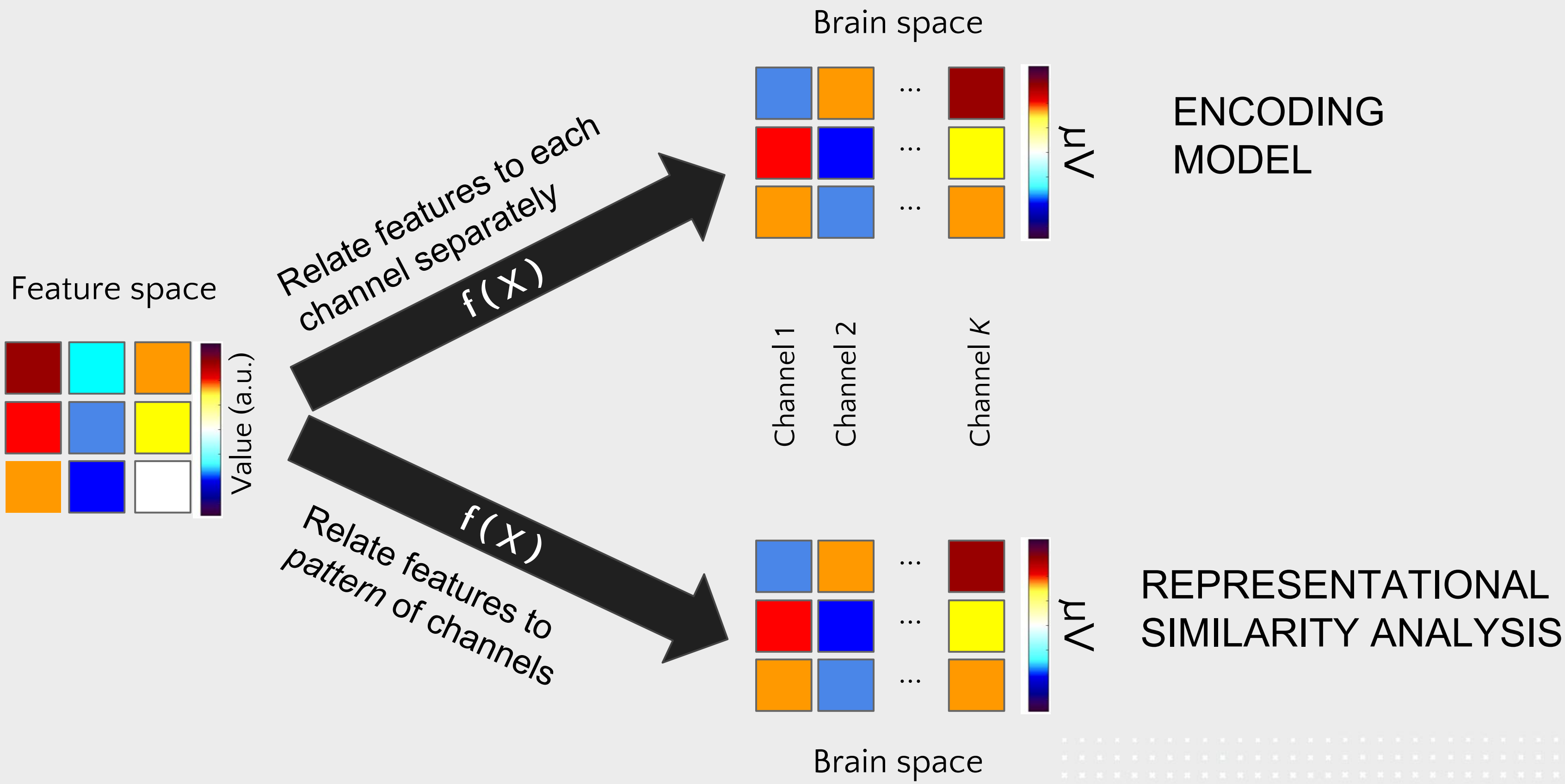


Two types of methods



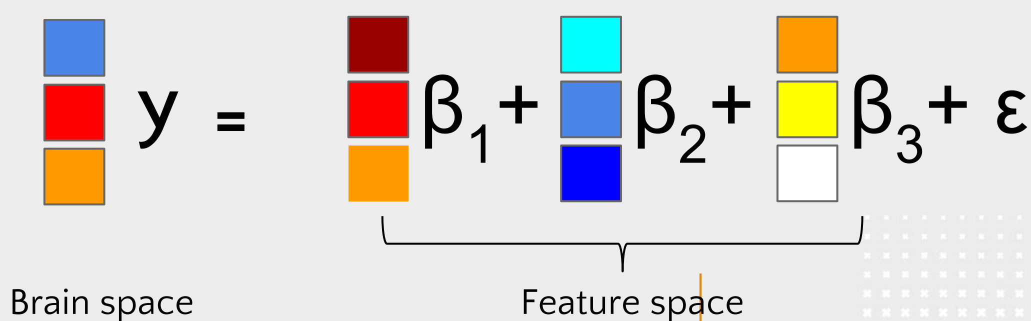
ENCODING MODEL

Two types of methods



Encoding models

- Encoding models relate features (X) to brain measurements (y), often using linear regression models
  - The good ol' GLM: $y = X\beta + \varepsilon$



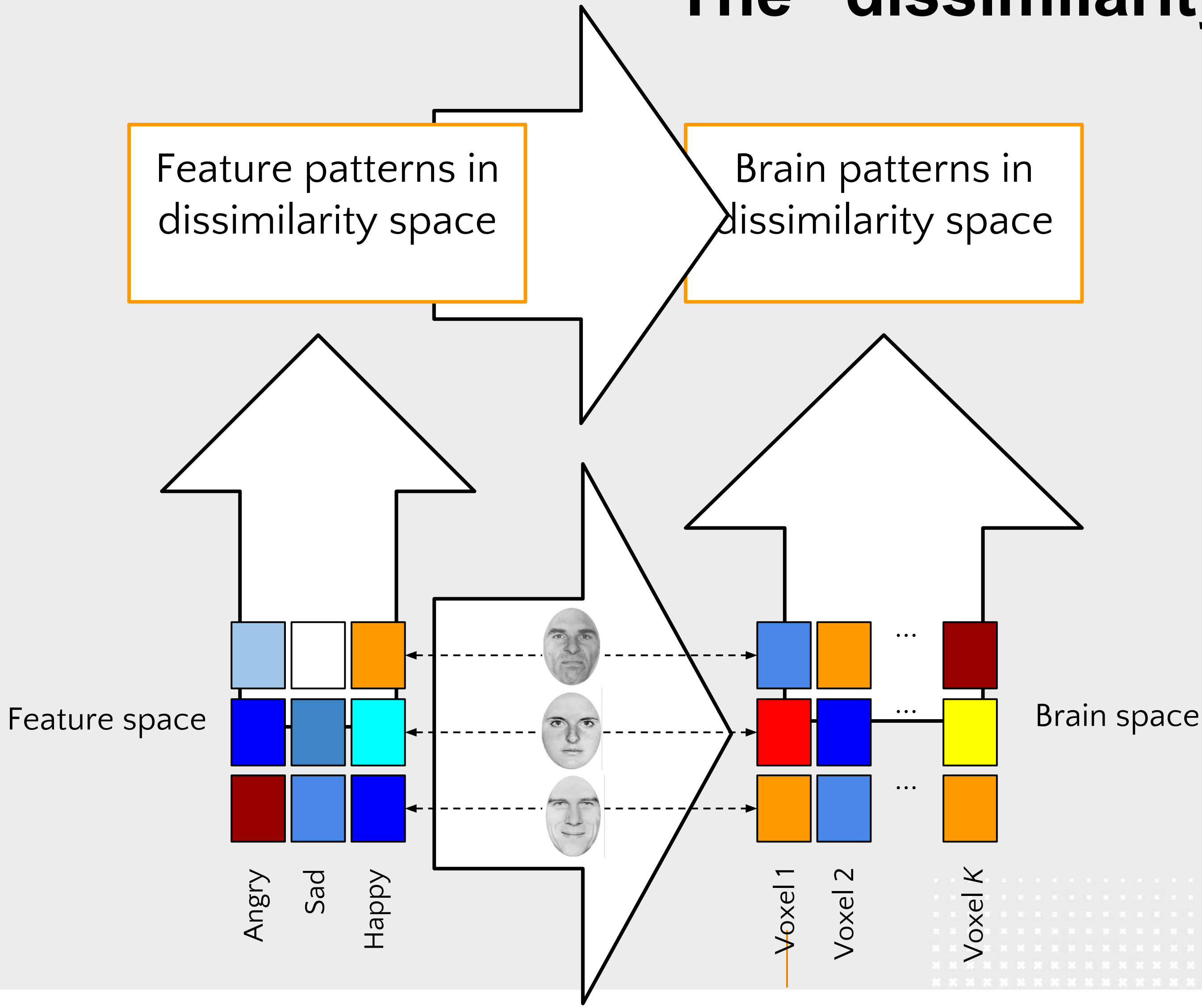
Regularization

- Sometimes, machine learning tricks ("**regularization**") are used to deal with many features (e.g., artificial NNs: K > 100,000!)
  - "Ridge regression", "LASSO"
  - Think of it as slightly extended GLMs

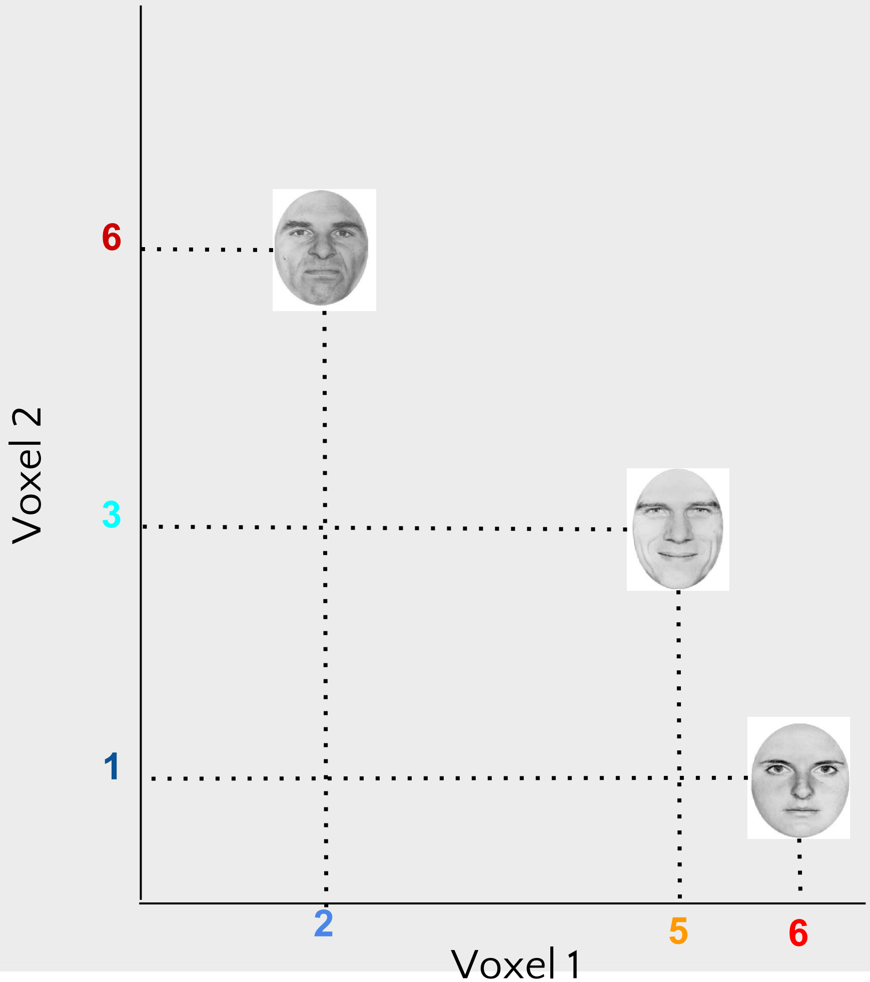
RSA

- Encoding models relate features to **single** voxels/channels
- RSA relates features to **patterns** of voxels/channels, using the “dissimilarity trick”

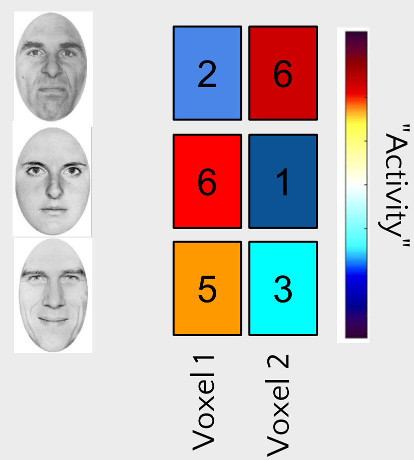
The "dissimilarity trick"



Dissimilarity space



"Dissimilarity" is the distance of points in N-dimensional space



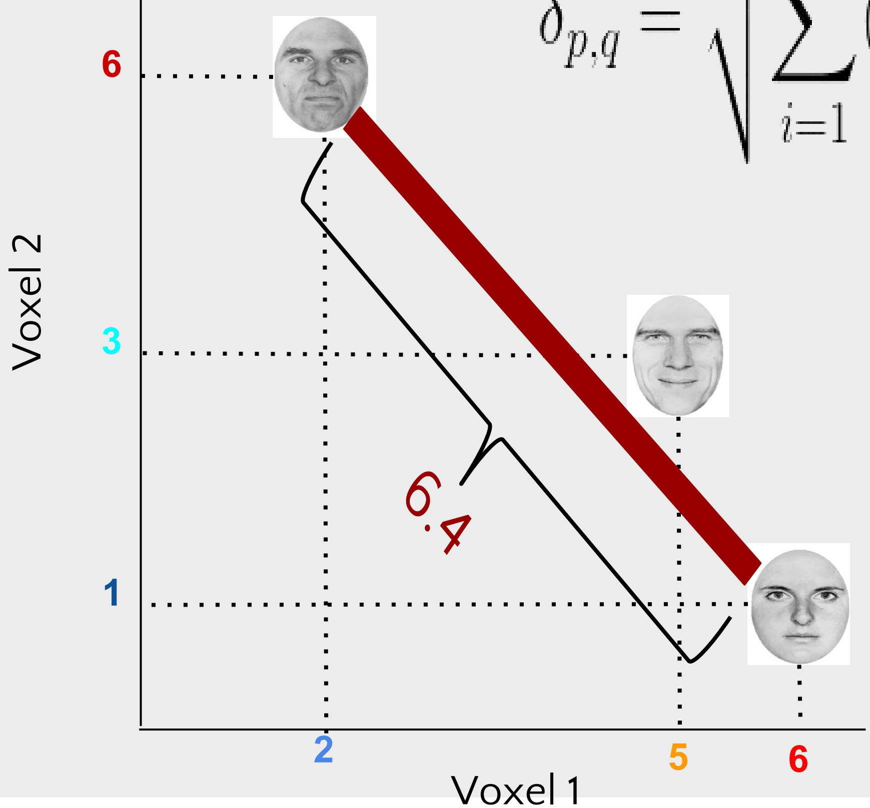
Note: the visualization is in 2-dimensional space, but usually K > 2!

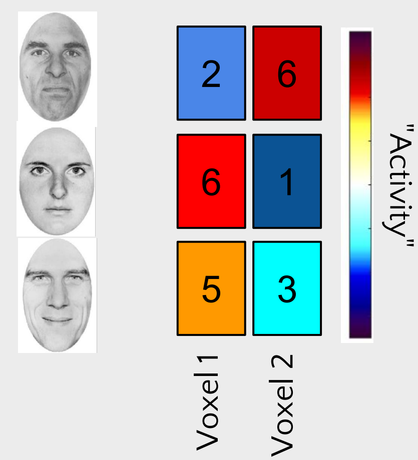
Dissimilarity space

$$\delta_{(\text{face}_1,\text{face}_2)} = \sqrt{(2-6)^2 + (6-1)^2} = 6.4$$

$$\delta_{p,q} = \sqrt{\sum_{i=1}^{K} (p_i - q_i)^2}$$

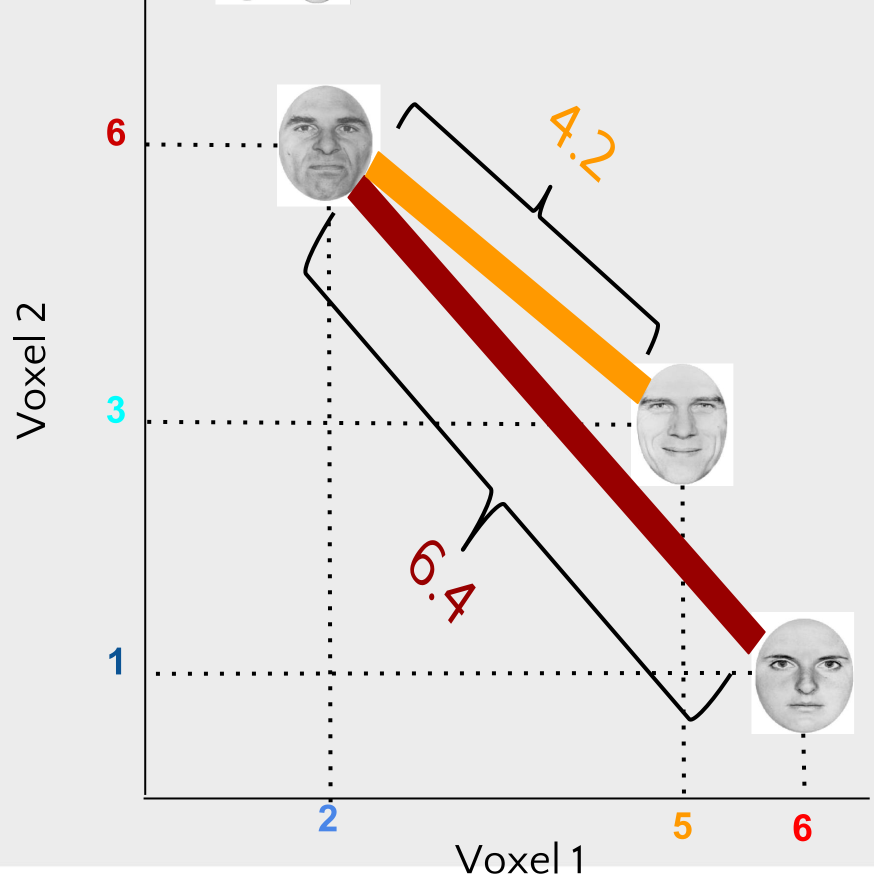
Euclidean distance

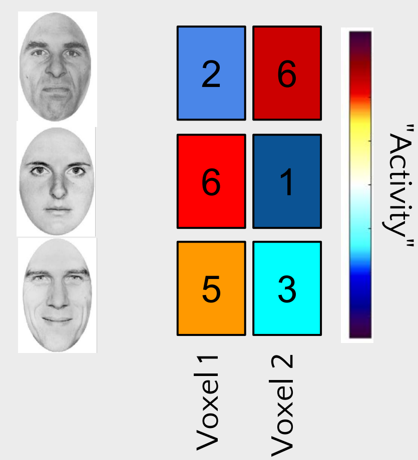




Dissimilarity space

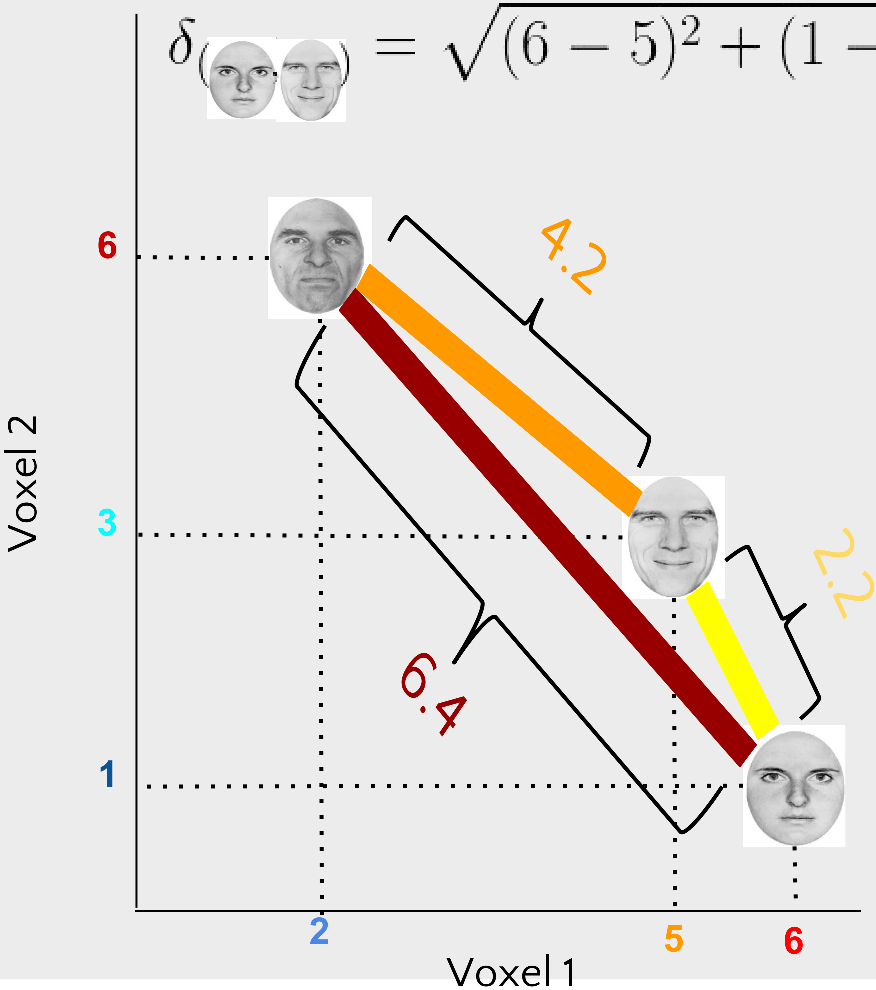
$$\delta = \sqrt{(2-5)^2 + (6-3)^2} = 4.2$$

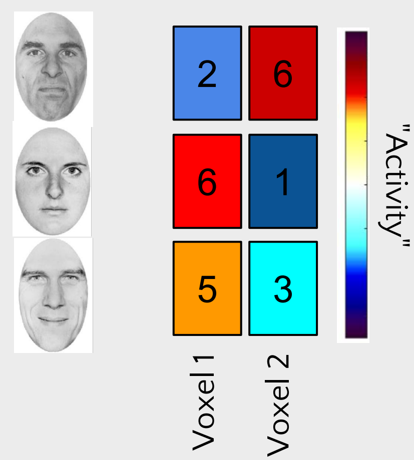




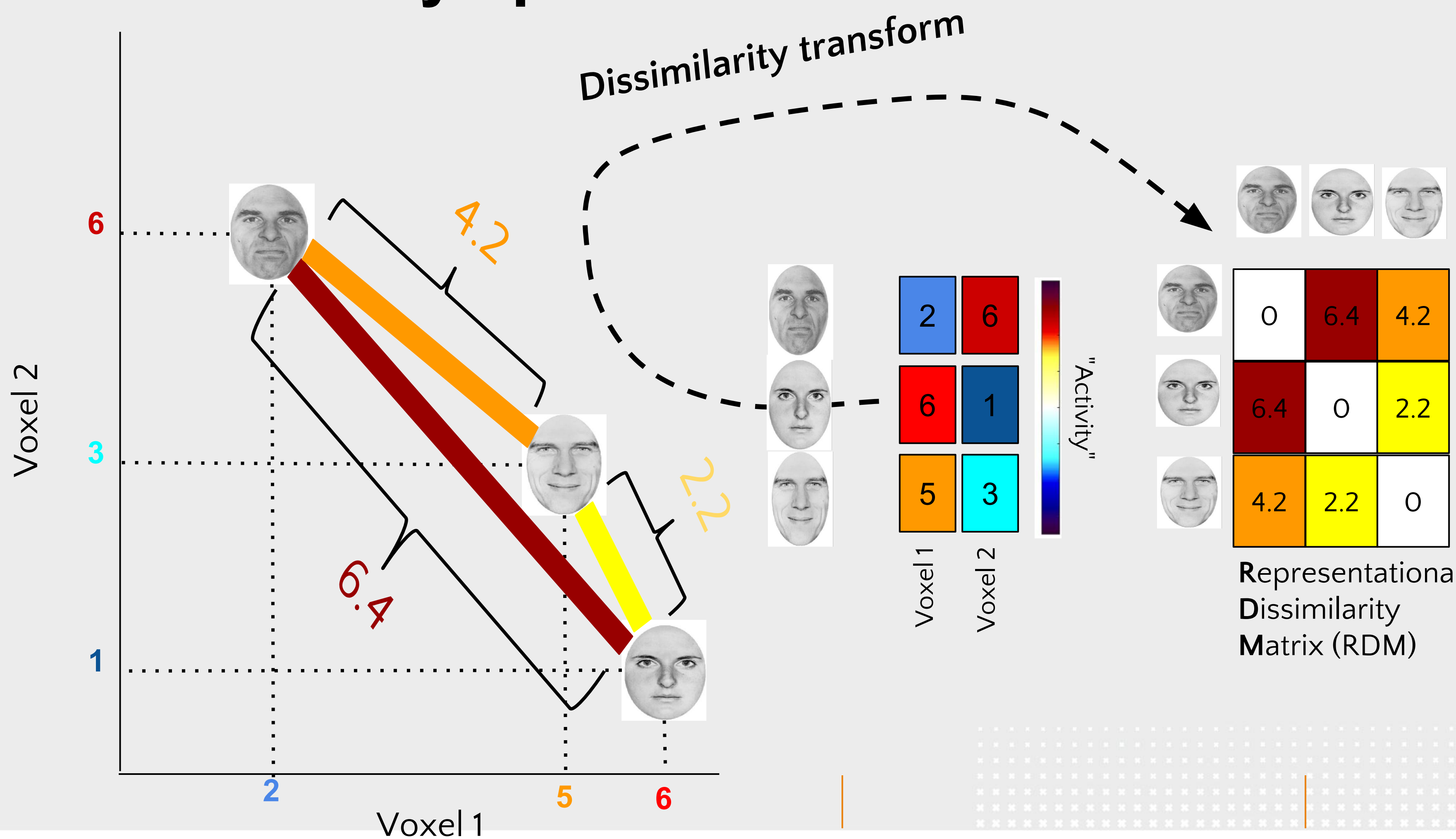
Dissimilarity space

$$\delta_{(\cdot,\cdot)} = \sqrt{(6-5)^2 + (1-3)^2} = 2.2$$





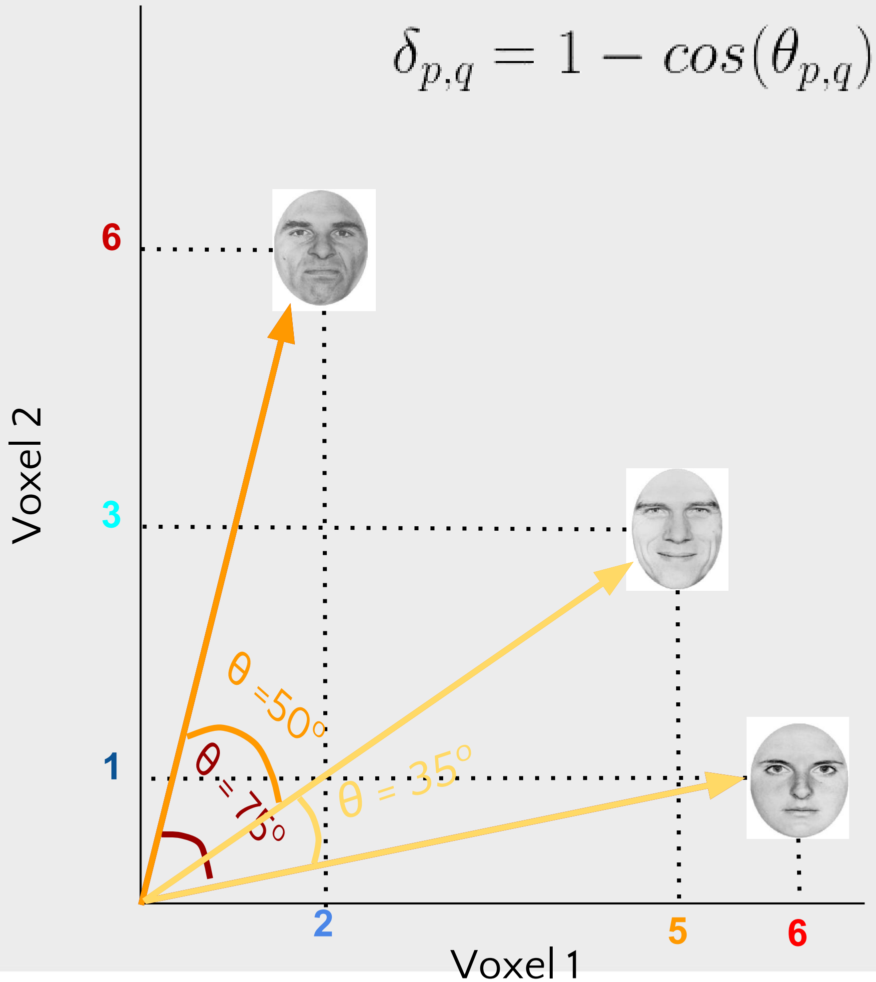
Dissimilarity space

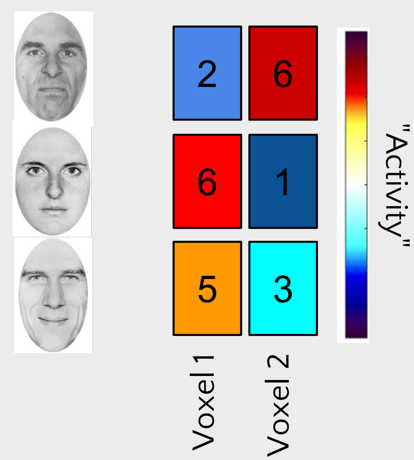


Dissimilarity space

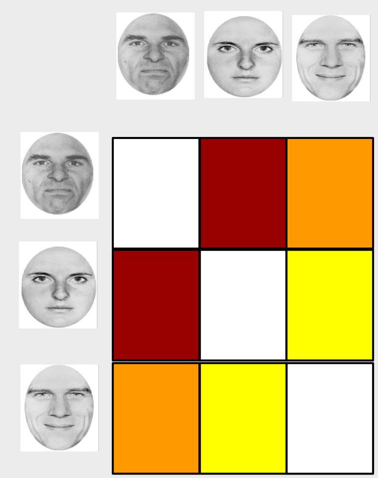
$$\delta_{p,q} = 1 - cos(\theta_{p,q})$$

Cosine distance



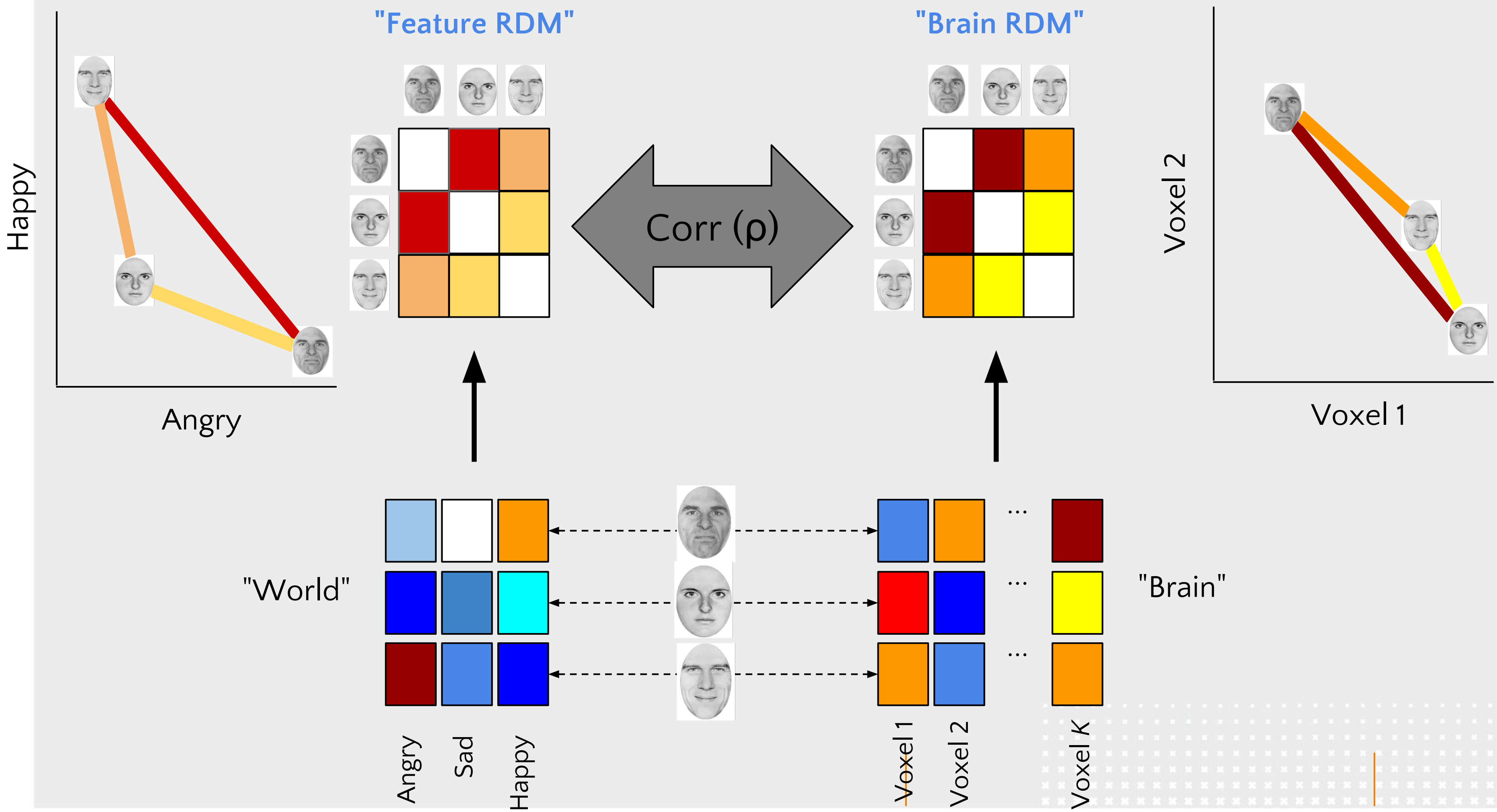


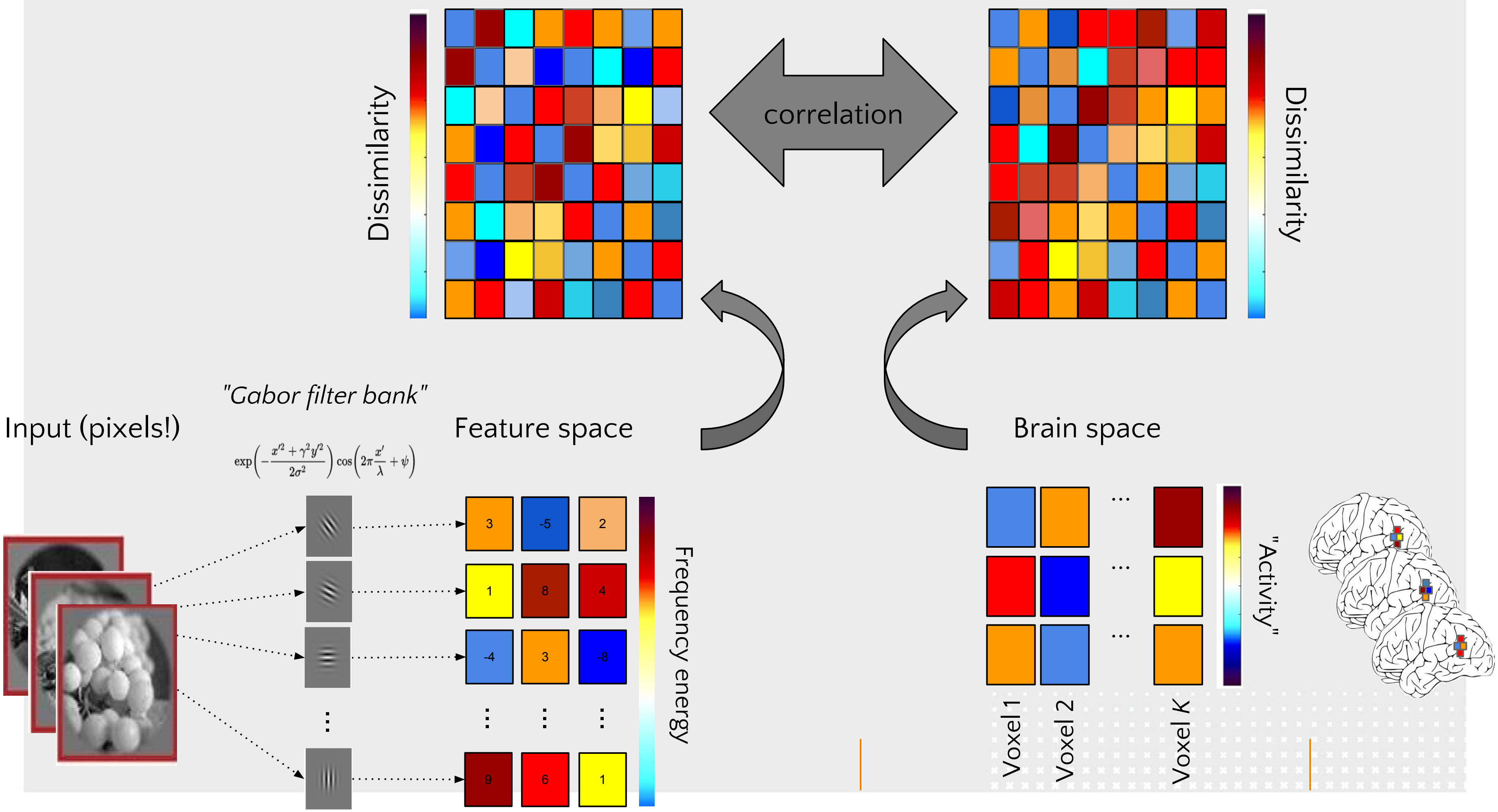
1 – cos(x)



Representational **D**issimilarity **M**atrix (RDM)

RSA


"Feature RDM"
"Brain RDM"
Corr (ρ)
Happy
Angry
Voxel 2
Voxel 1
"World"
"Brain"
Angry
Sad
Happy
Voxel 1
Voxel 2
...
Voxel K


Dissimilarity
correlation
Dissimilarity
"Gabor filter bank"
Input (pixels!)
$\exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$
Feature space
3
-5
2
1
8
4
-4
3
-8
9
6
1
Frequency energy
Brain space
"Activity"
Voxel 1
Voxel 2
Voxel K

Summary

- Encoding models and RSA both relate features to brain measurements
  - Encoding models: features → single units
  - RSA: features → patterns
- Choice depends on how your feature is encoded in the brain
  - Spatial frequency → single channels
  - Object category → patterns (?)

Questions?