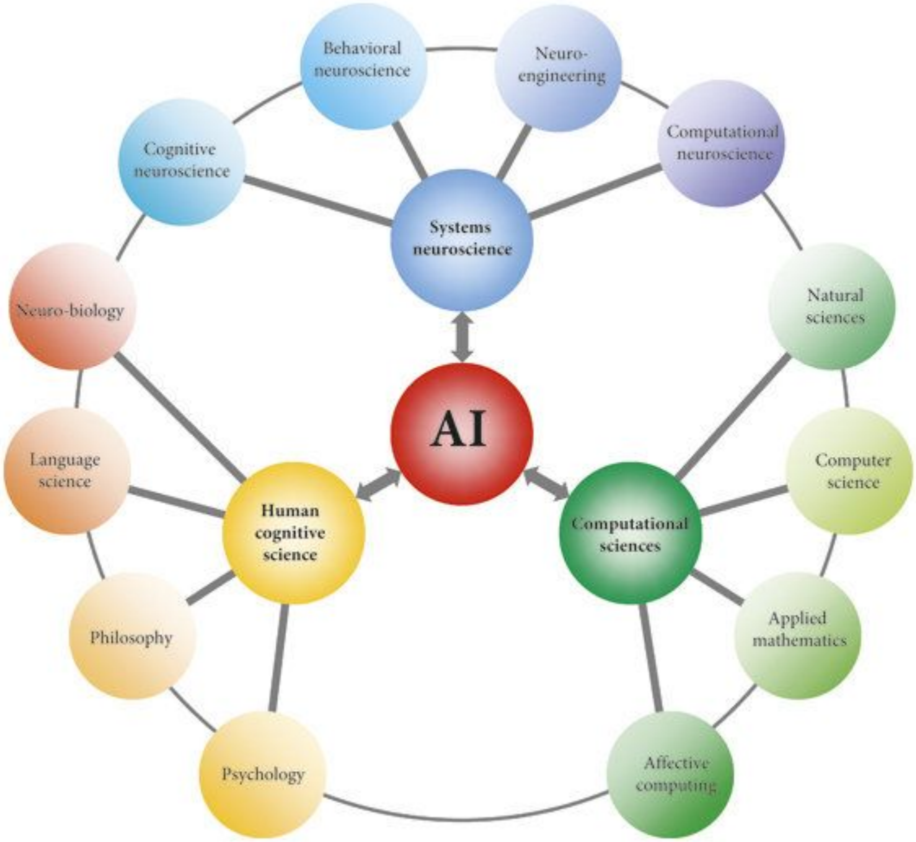
Applications of AI: Cognitive Science

Naiti Bhatt
University of Edinburgh

Roadmap

- ➔ Cognitive Science
- ➔ Measuring cognition: Neuroimaging
- ➔ Algorithmic tools for modeling cognitive systems

Cognitive Science


Behavioral neuroscience
Neuro-engineering
Cognitive neuroscience
Computational neuroscience
Systems neuroscience
Natural sciences
Neuro-biology
AI
Language science
Computer science
Human cognitive science
Computational sciences
Philosophy
Applied mathematics
Psychology
Affective computing

What is the difference between the brain and the mind?

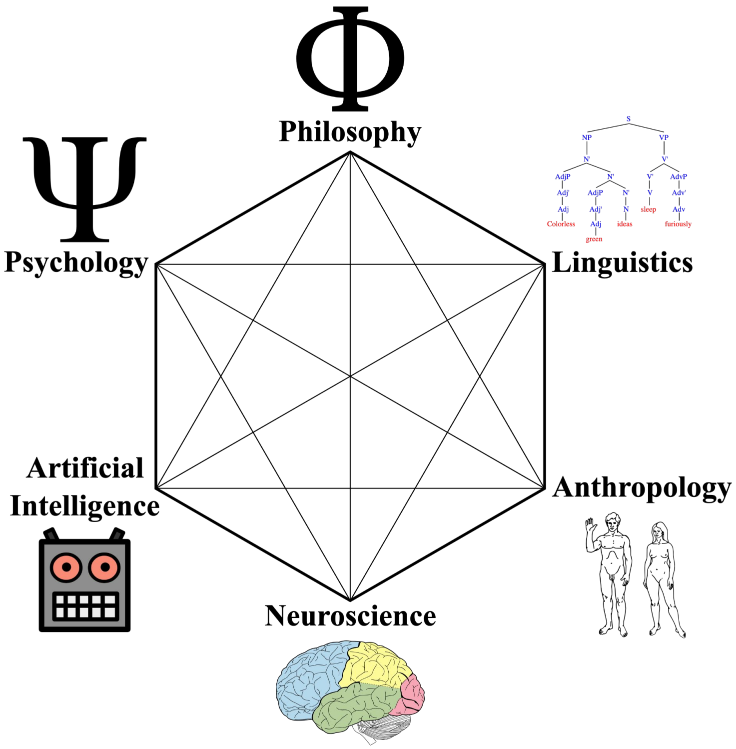
Brain versus Mind

Brain	**Mind**
Anatomical Organ	Hypothetical Construct
Vehicle for thoughts	Origin of thoughts
Composed of neurons	Composed of phenomena
Involuntary operator of consciousness	Volitional source of consciousness
Hardware of the human computer	Software of the human computer

What is cognitive science?

The study of the human mind and brain, focusing on how **the mind** represents and manipulates **knowledge** and how mental representations and processes are realized in **the brain**.

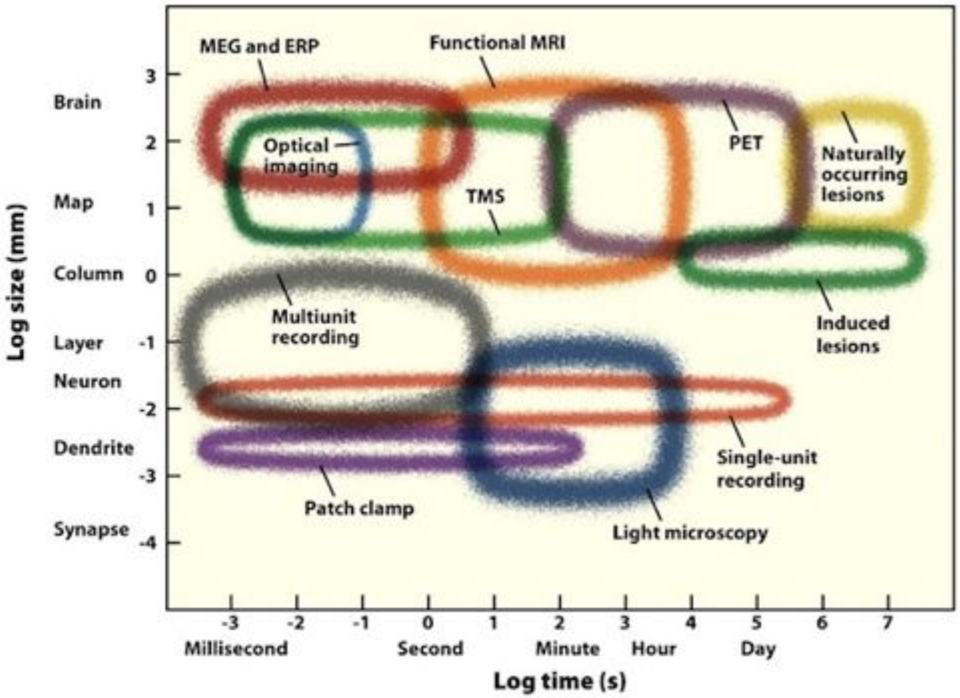
Quintessentially Interdisciplinary!



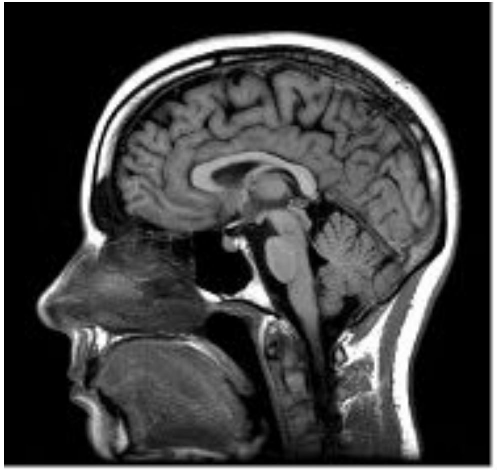
How do we do research in cognitive science?

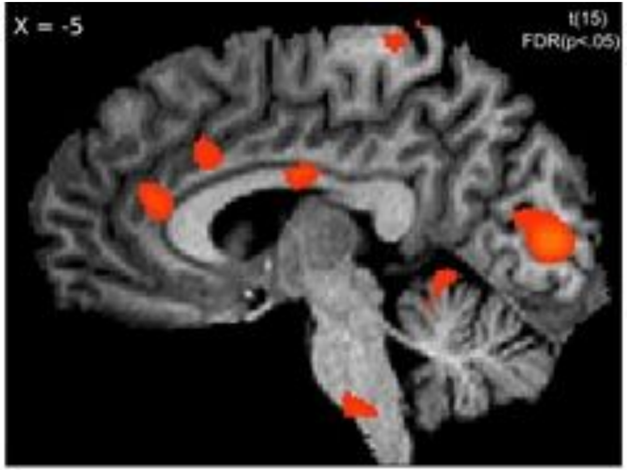
Measuring cognition: Neuroimaging

Loads of different neuroimaging methods!

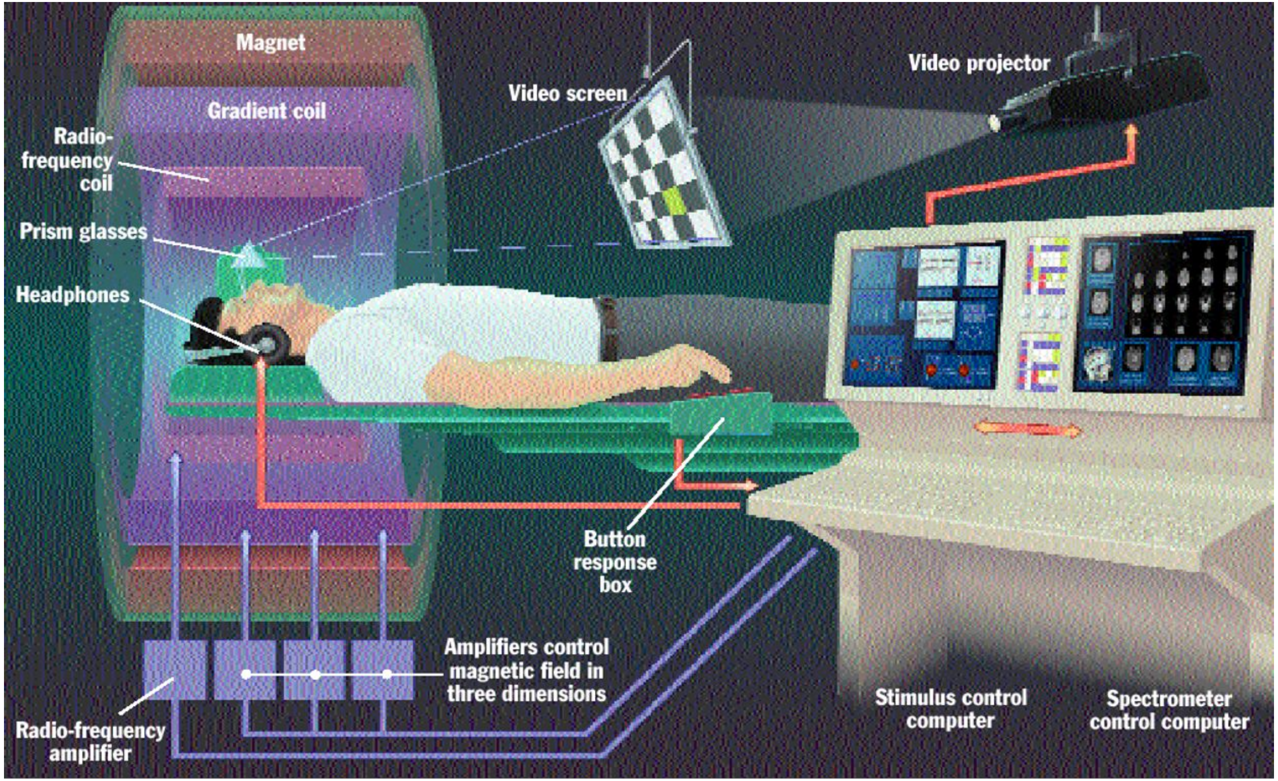


Magnetic Resonance Imaging



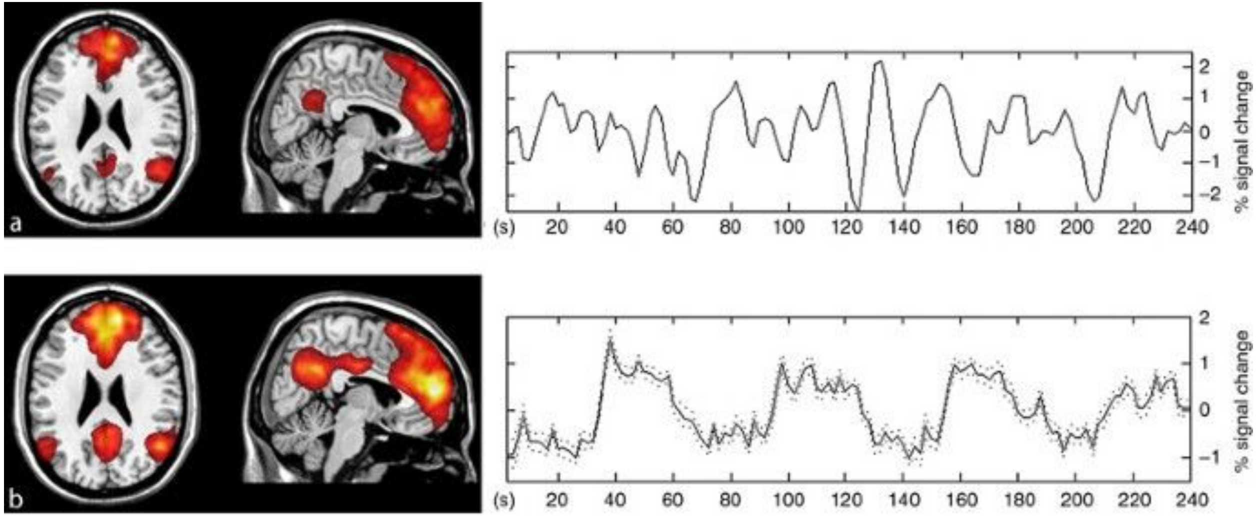


fMRI Setup


Magnet
Gradient coil
Radio-
frequency
coil
Prism glasses
Headphones
Video screen
Video projector
Button
response
box
Amplifiers control
magnetic field in
three dimensions
Radio-frequency
amplifier
Stimulus control
computer
Spectrometer
control computer

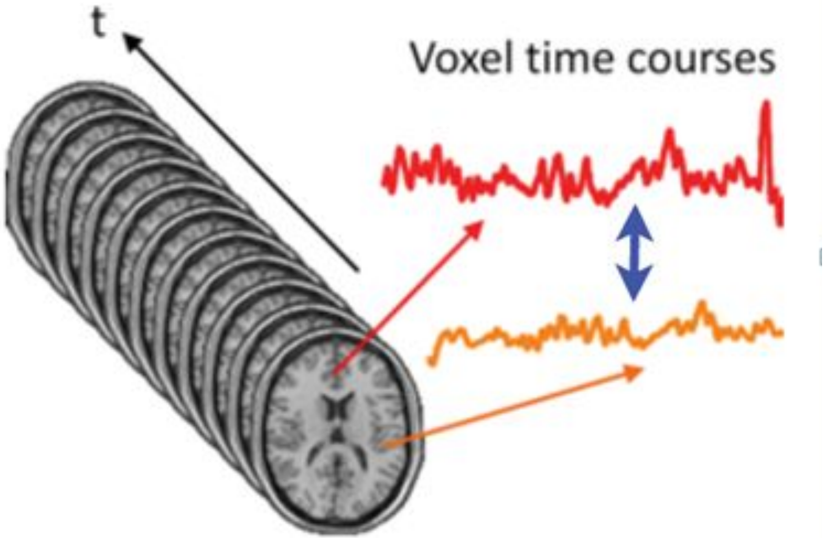
fMRI: **F**unctional **M**agnetic **R**esonance **I**maging

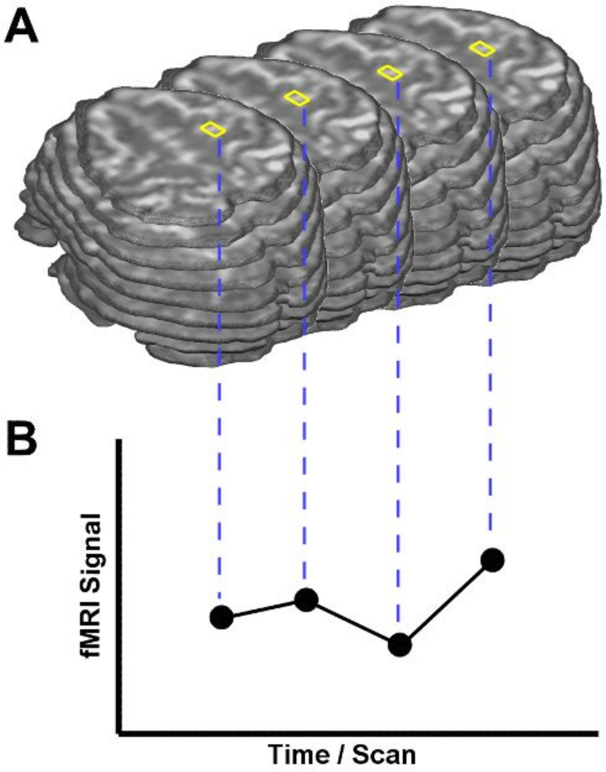
- Noninvasive and indirect method of measuring neural 'activity'
- Uses flow of oxygenated blood as a proxy
- Measure voxel-wise



Voxel (**Vo**lumetric Pi**xel**)

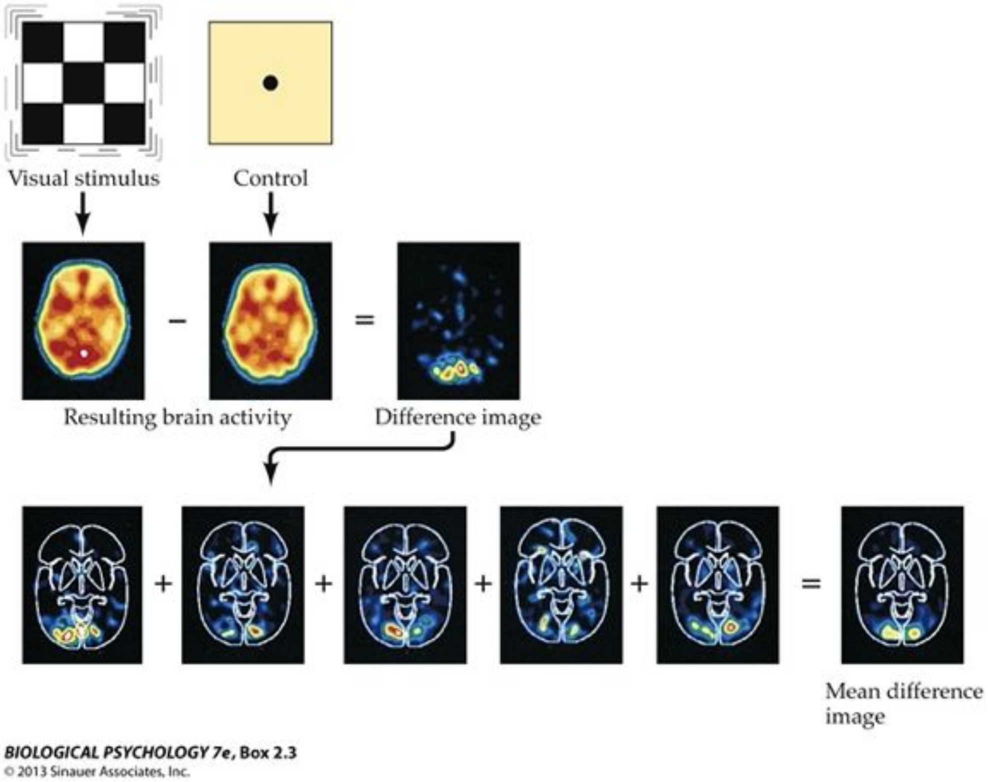
Usually, 2-3mm^3



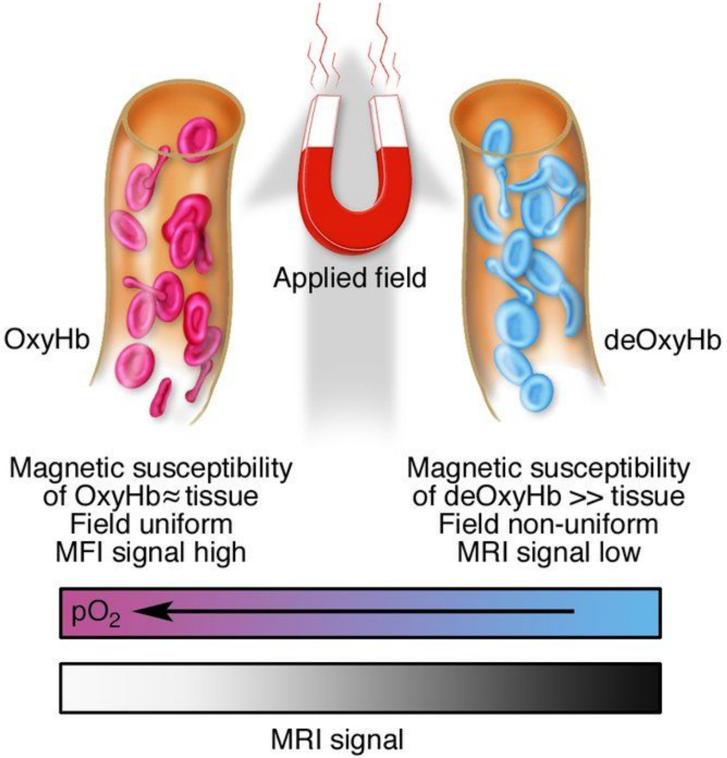


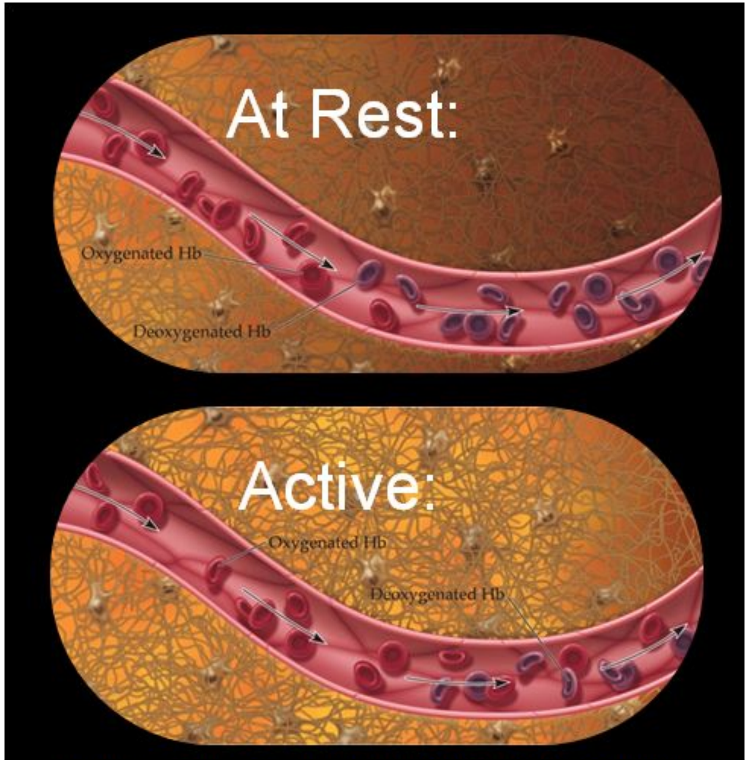
Final result: Activation Maps

Stimulus - Control = Activation



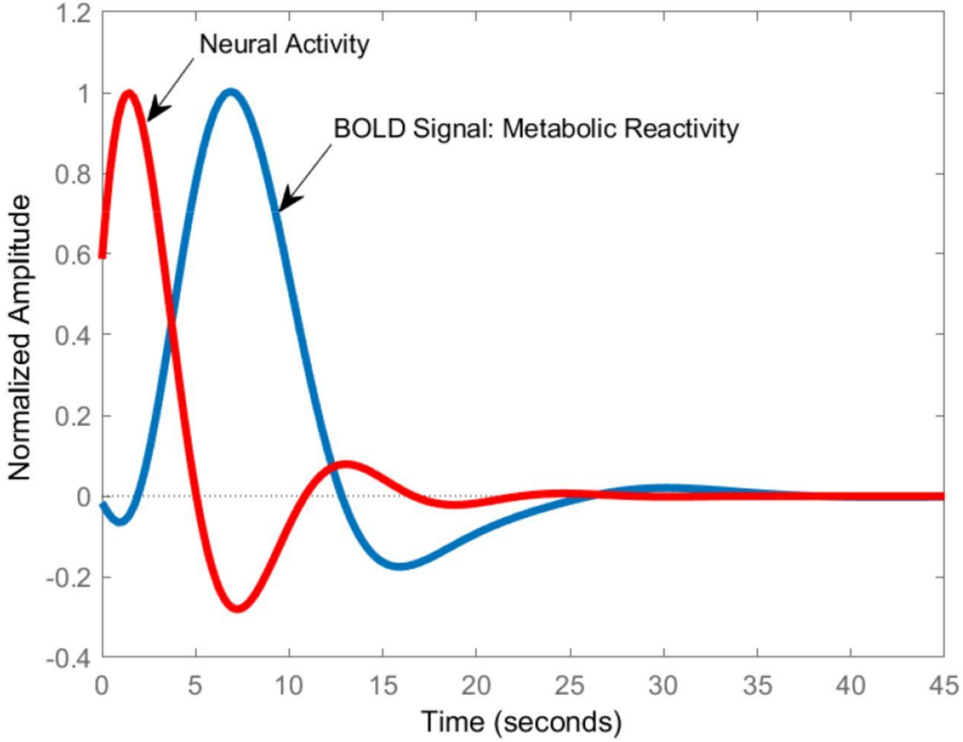
BOLD (**B**lood **O**xygen **L**evel **D**ependent) Response



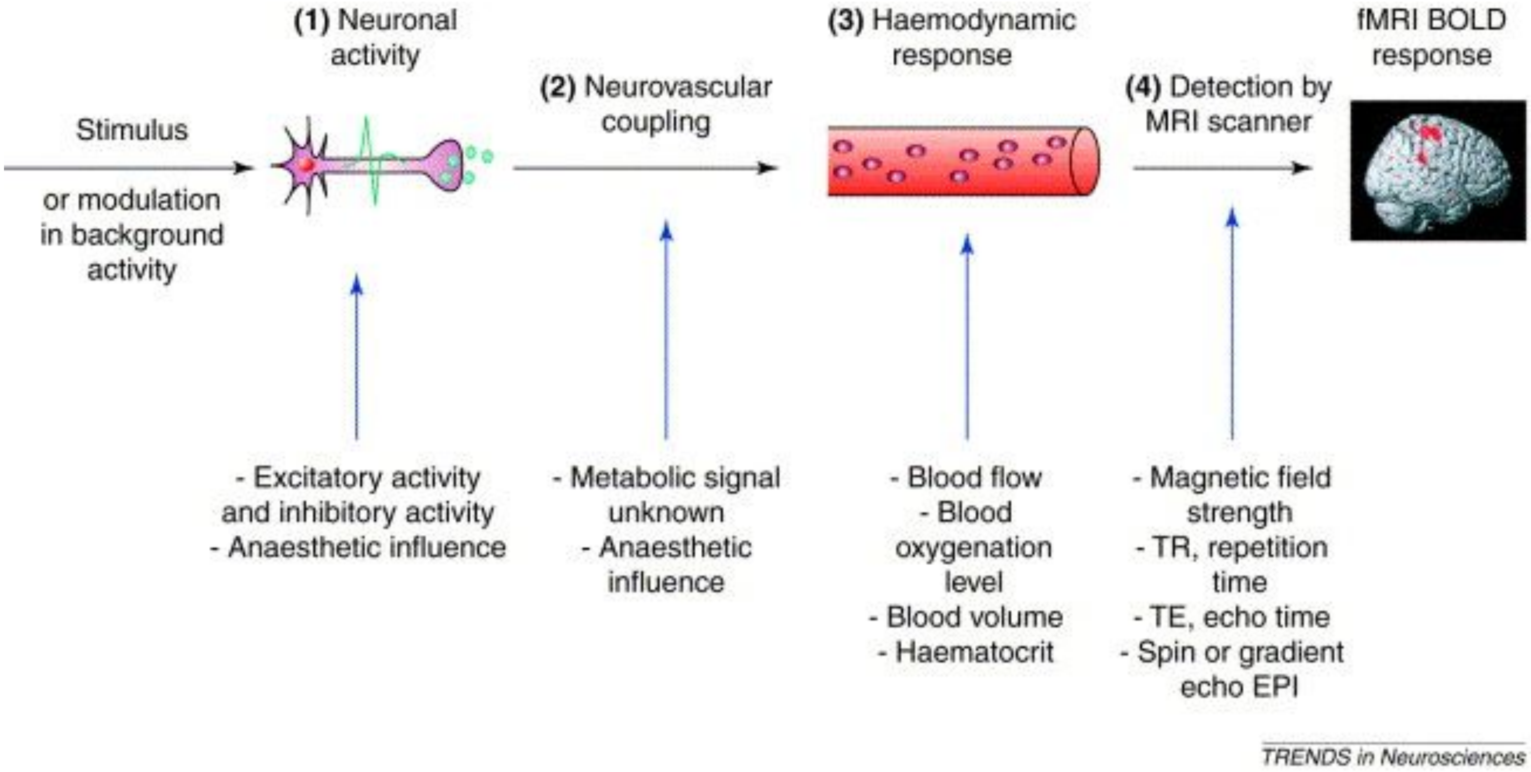


Hemodynamic **R**esponse **F**unction (HRF)

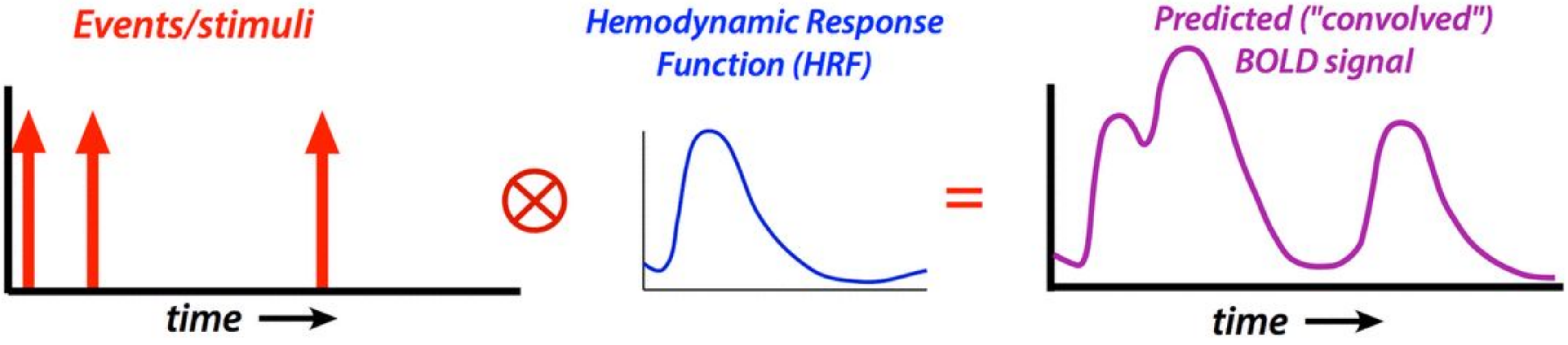
- Model of BOLD response
- Oxygenated blood follows neural activity
- Slower response gives time for measurement



Overview of the fMRI BOLD Signal



Predicting responses using convolution



Two approaches to analyzing fMRI data

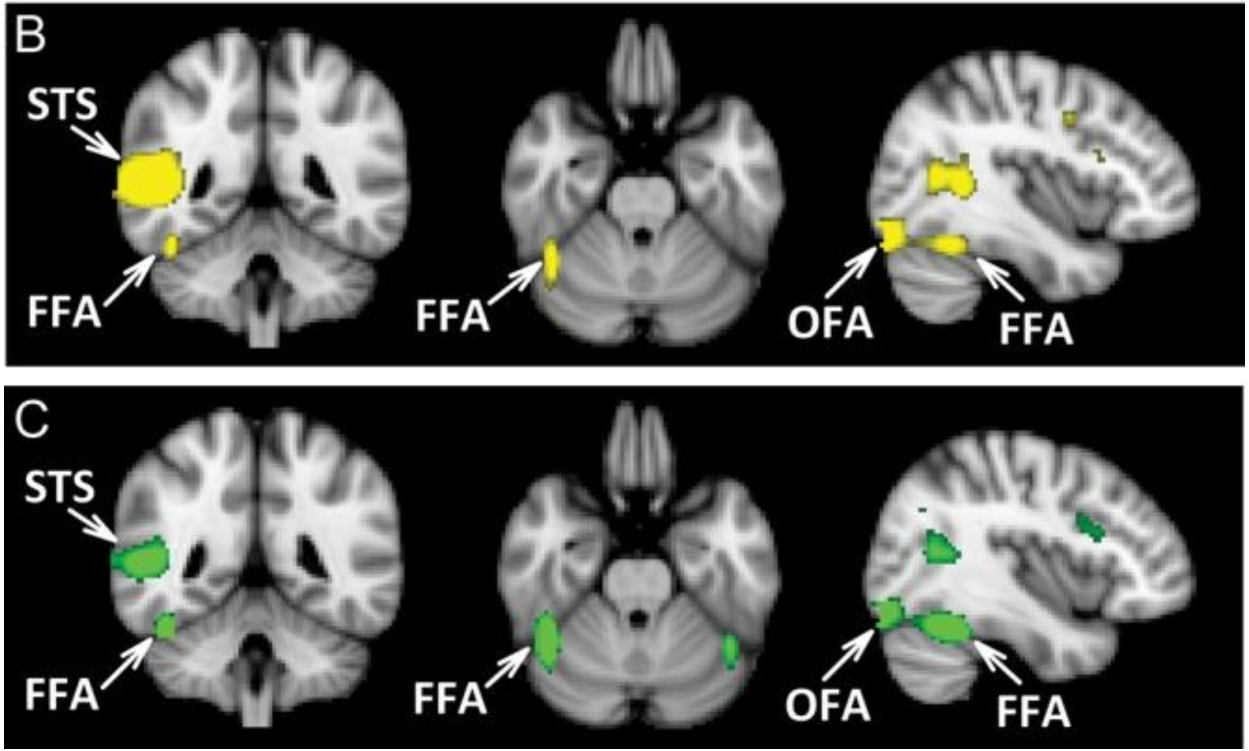
1. “Whole-brain” or “voxelwise” approach

2. Region-of-interest approach

1. “Whole-brain” or “voxelwise” approach

- For every voxel in the scanned volume (typically but not necessarily the whole brain), do a statistical test to look for expected differences in activation
    - e.g., ANOVA
        - main effect of stimulus category (faces vs. hands)
        - main effect of direction (left, direct, right)
        - interaction

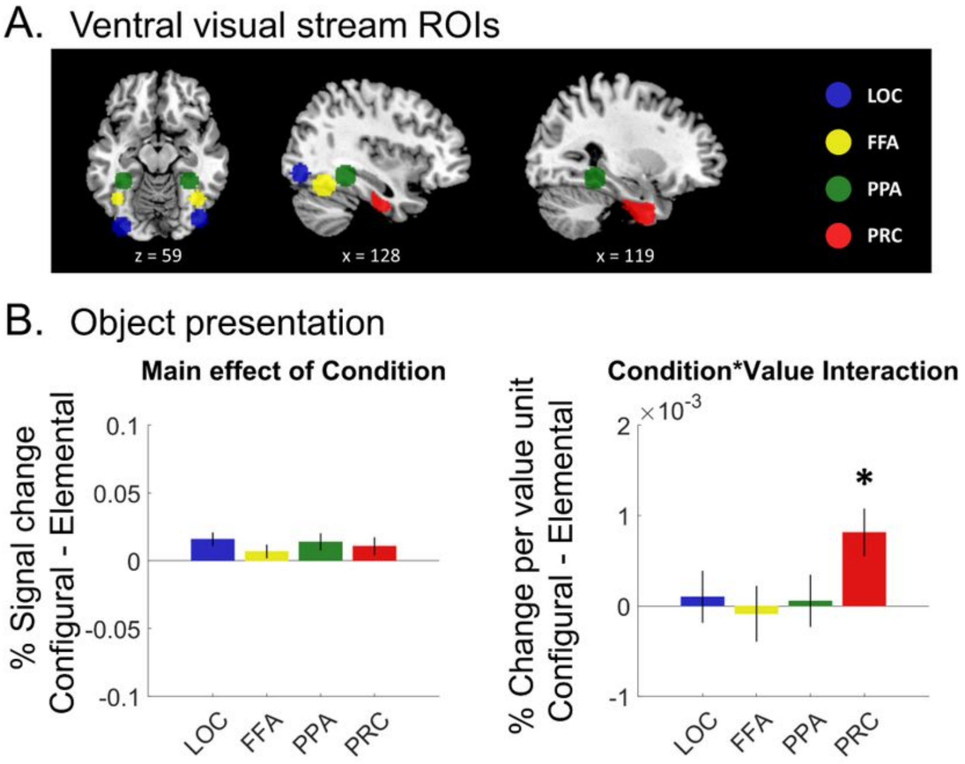
1. "Whole-brain" or "voxelwise" approach



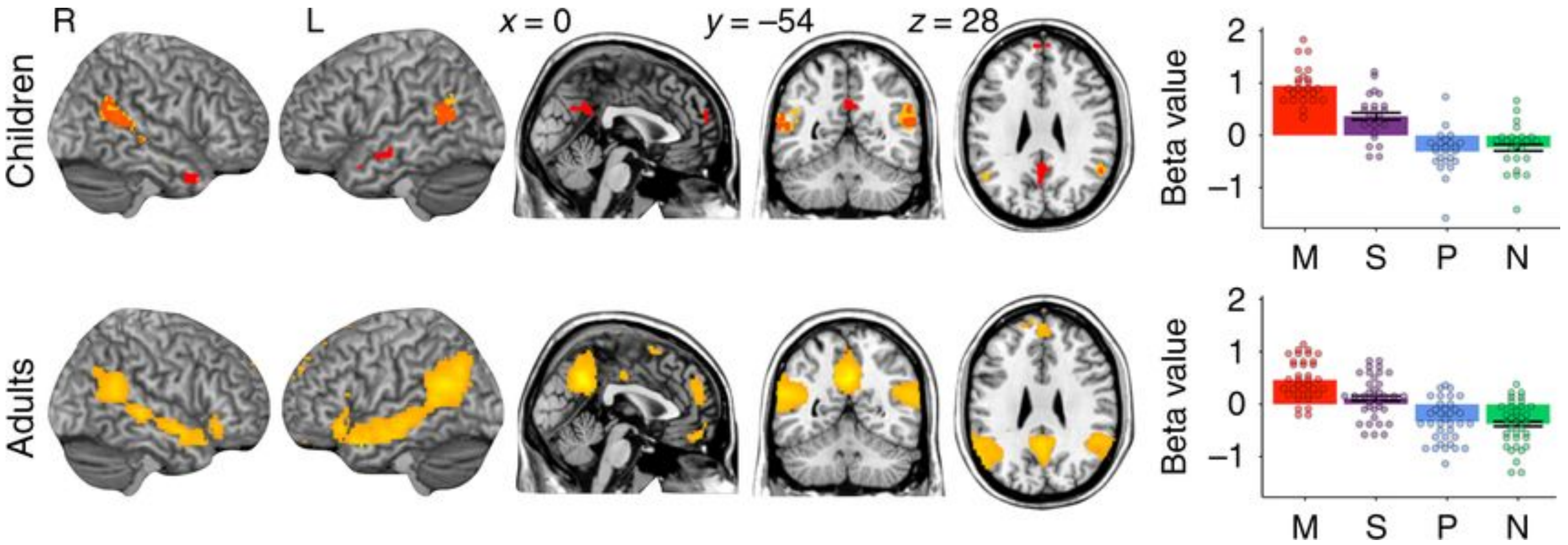
2. Region-of-interest approach

- Conduct separate “localizer” scan(s) to identify expected brain regions based on typical activation differences
  - e.g., Define FFA by a contrast of faces vs. other visual stimuli
- Regions can be identified in individual participants or across a group average
- Activation levels for different conditions can be extracted from main experiment data and tested

2. Region-of-interest approach



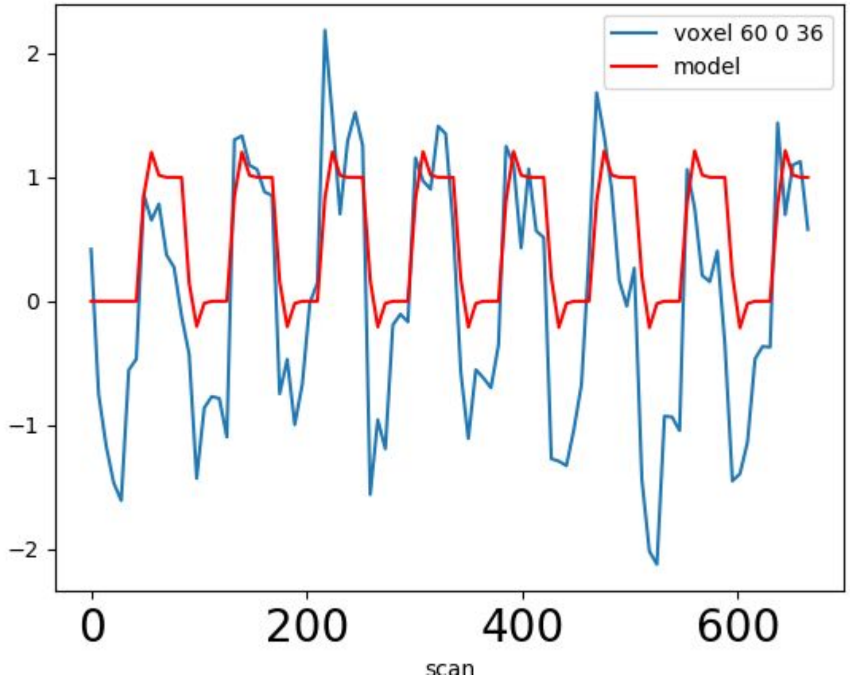
Using fMRI to study theory of mind development



fMRI Summary

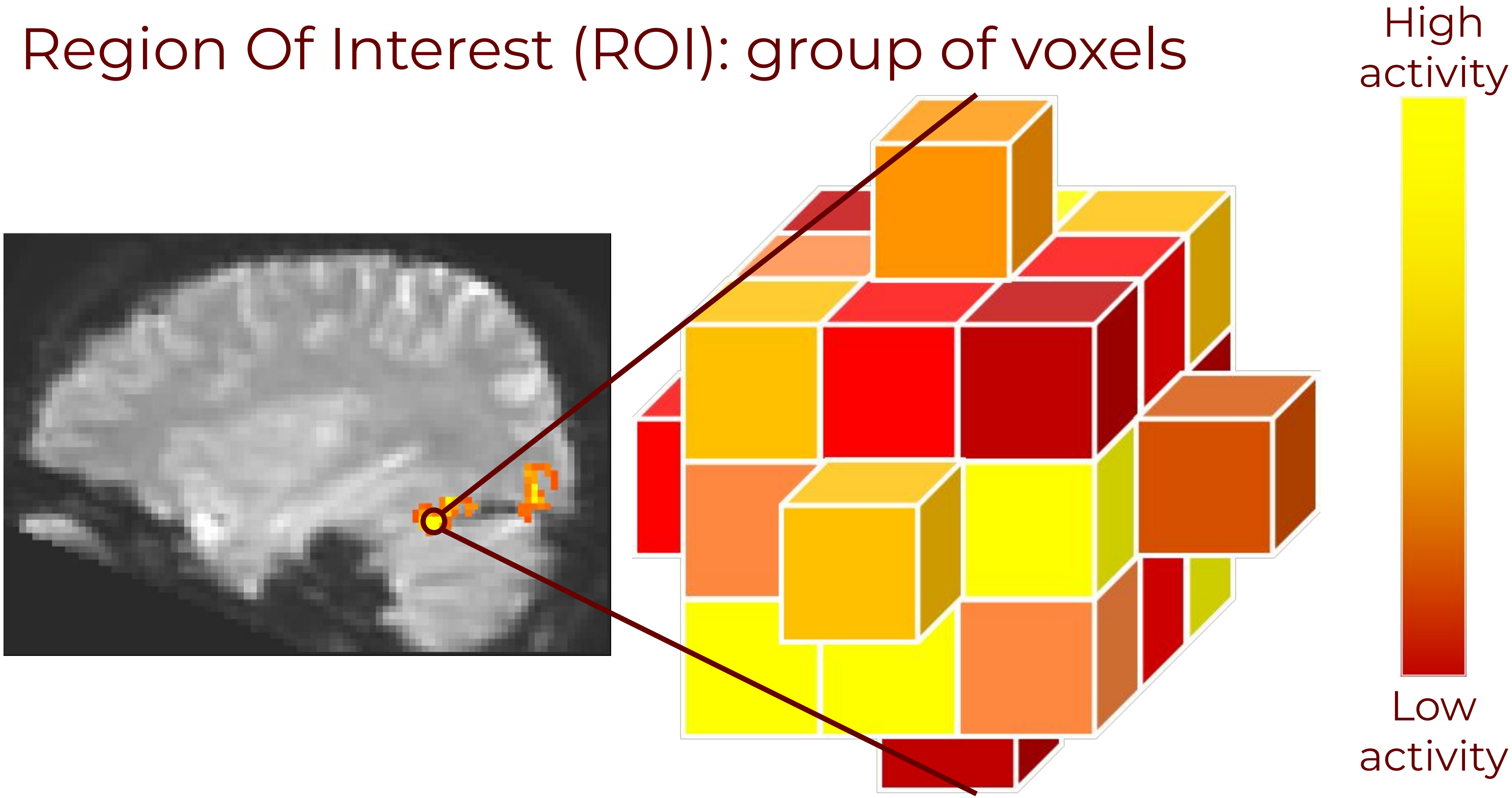
Measuring neural responses to external stimuli

These data are really multidimensional!



Algorithmic modeling tools

Region Of Interest (ROI): group of voxels



Methods of Analysis

Univariate Analysis

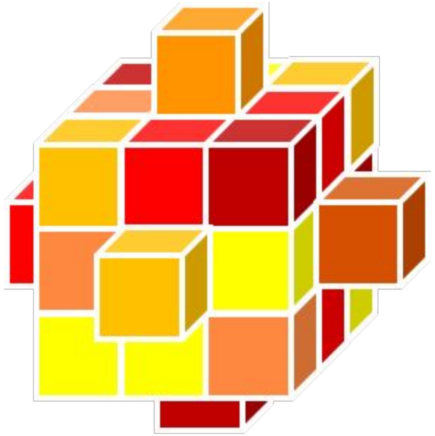
- look at one variable
- level of activation across a region
- average data from each voxel
- works well for strong differences

Multivariate Analysis

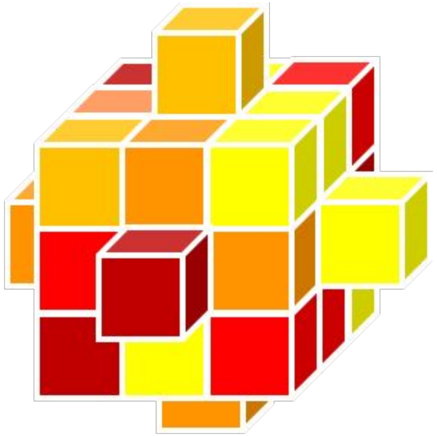
- look at multiple variables
- pattern of activation within a region
- preserve data from each voxel
- works well for weak effects across many voxels

Voxel Pattern Information

Condition 1
e.g., Faces



Condition 2
e.g., Hands



Two regions might have similar average **levels** of activation but different spatial **patterns** of activation

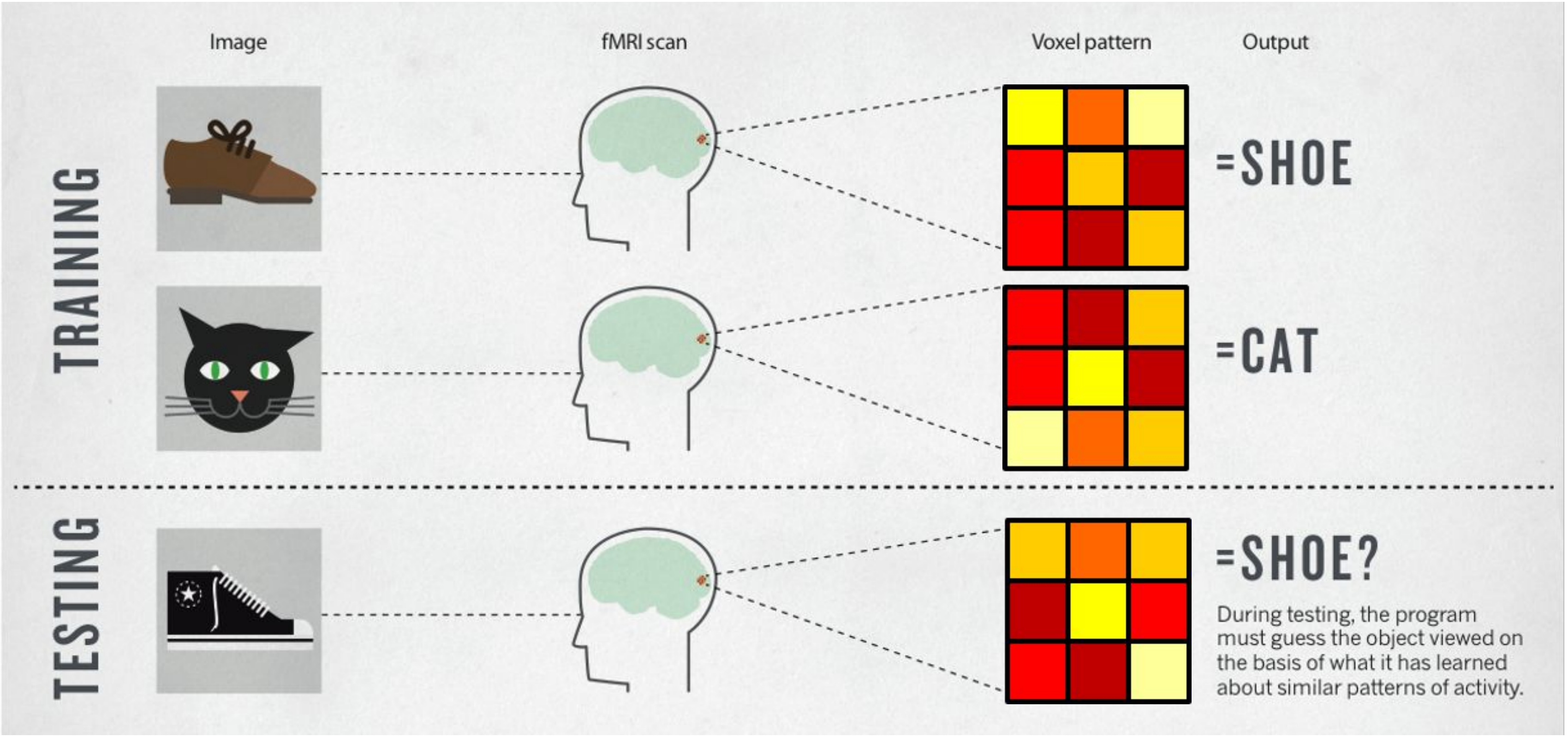
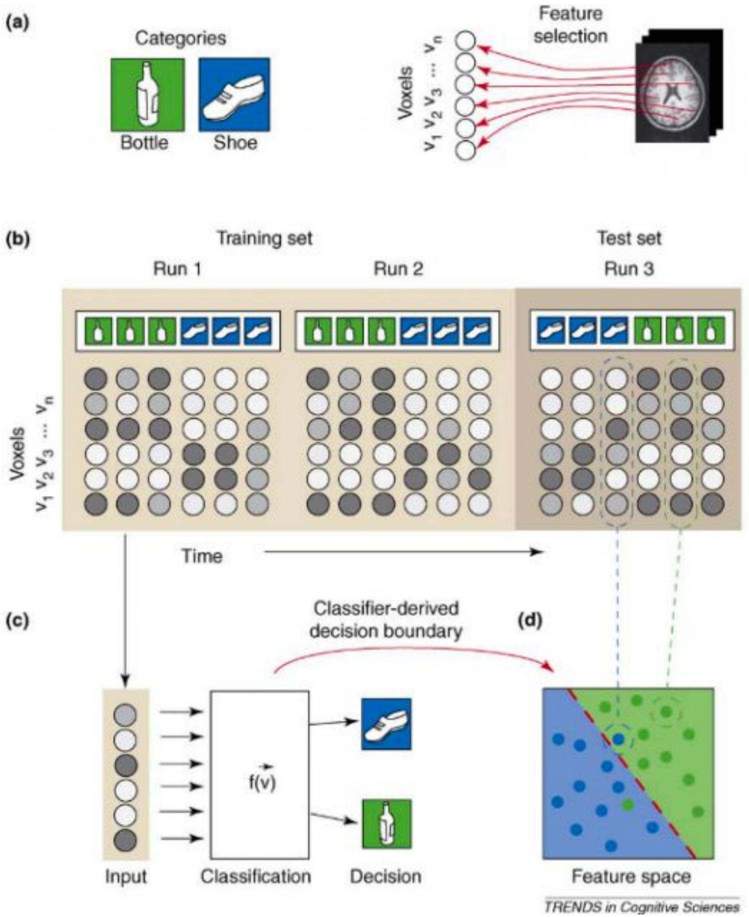


Image
fMRI scan
Voxel pattern
Output
TRAINING
=SHOE
=CAT
TESTING
=SHOE?
During testing, the program must guess the object viewed on the basis of what it has learned about similar patterns of activity.

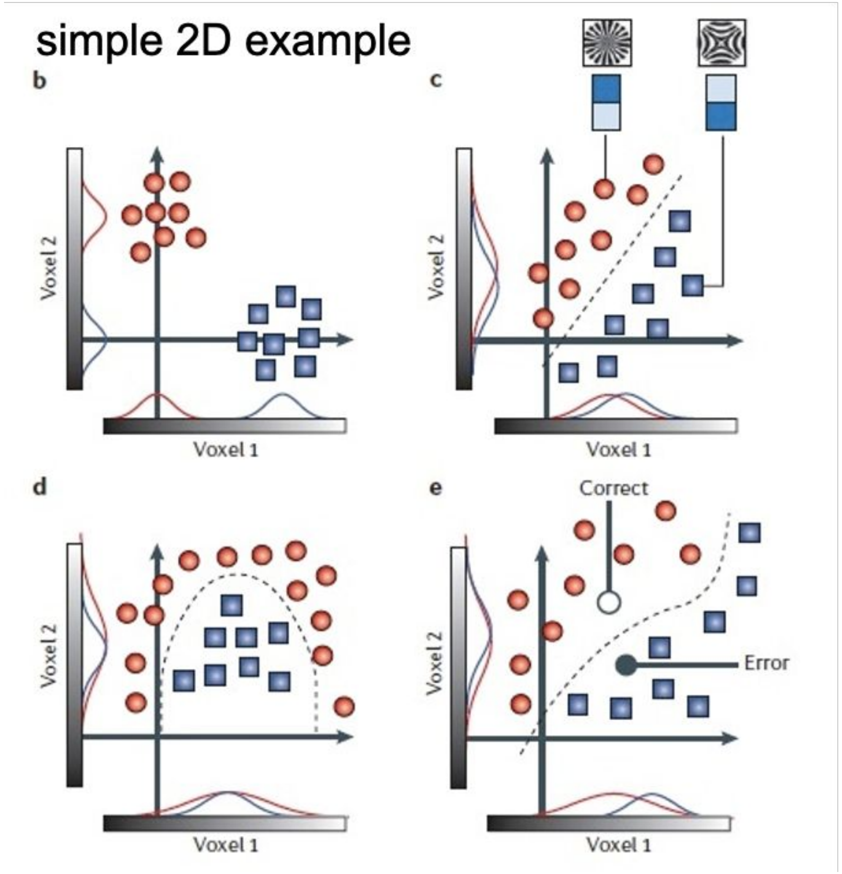
Multi-Voxel Pattern Analysis

Supervised classification method where a classifier attempts to capture the relationships between **spatial pattern of fMRI activity** and **experimental conditions**.



Classifier can act on single voxels. Conventional fMRI analysis would detect the difference.

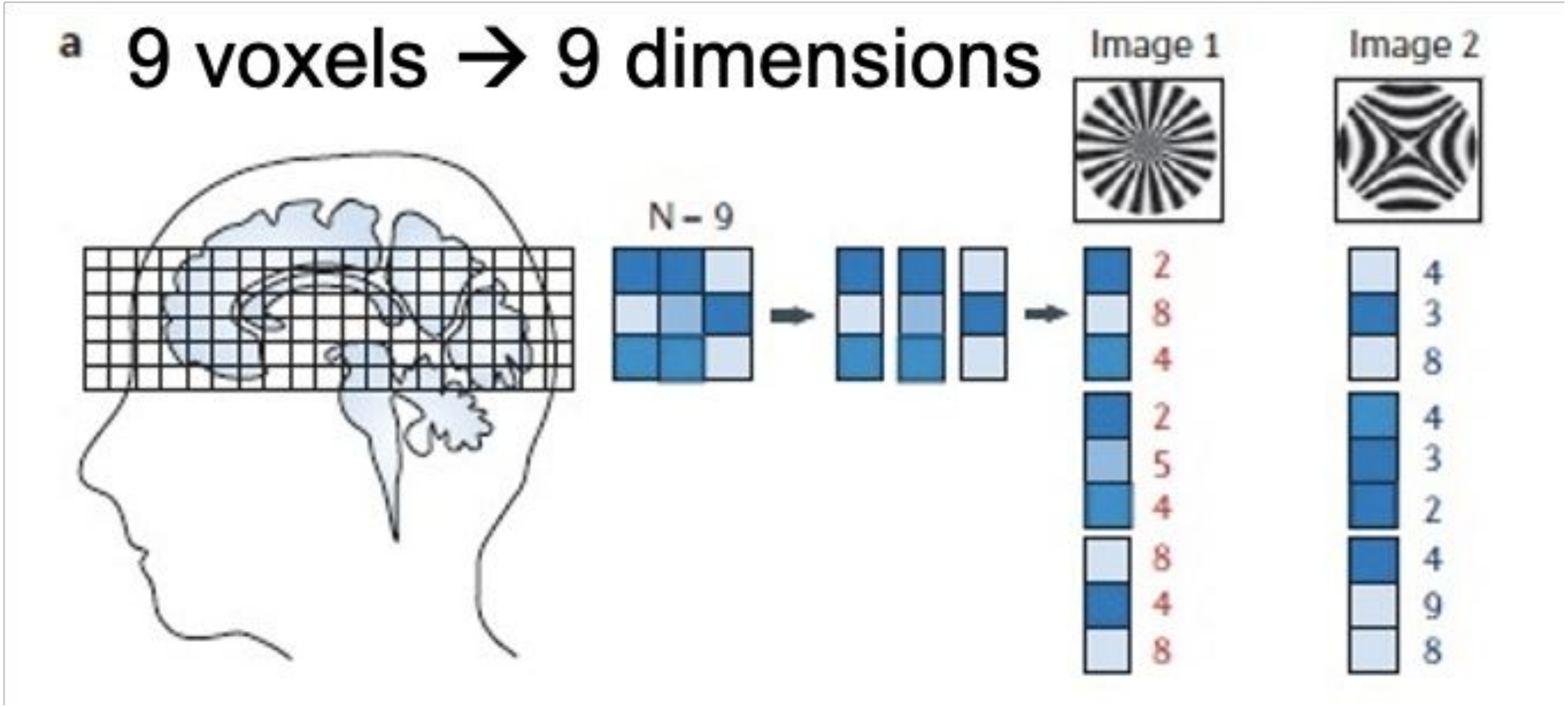
Classifier would require curved decision boundary



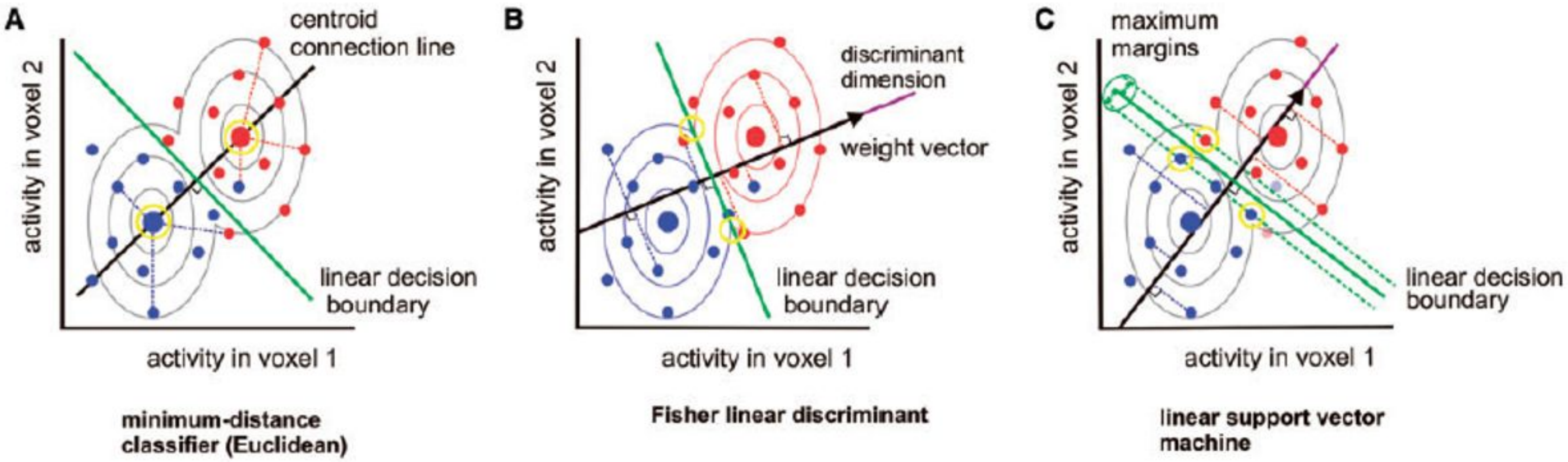
Classifier can not act on single voxels because distributions overlap

Classifier can act on combination of voxels using a linear decision boundary

White and black circles show examples of correct and erroneous classification in the test set


a
9 voxels → 9 dimensions
Image 1
Image 2
N = 9
2
8
4
2
5
4
8
4
8
4
3
8
4
3
2
4
9
8

Linear support vector machine (SVM) is the most common approach to decide how to draw the line.

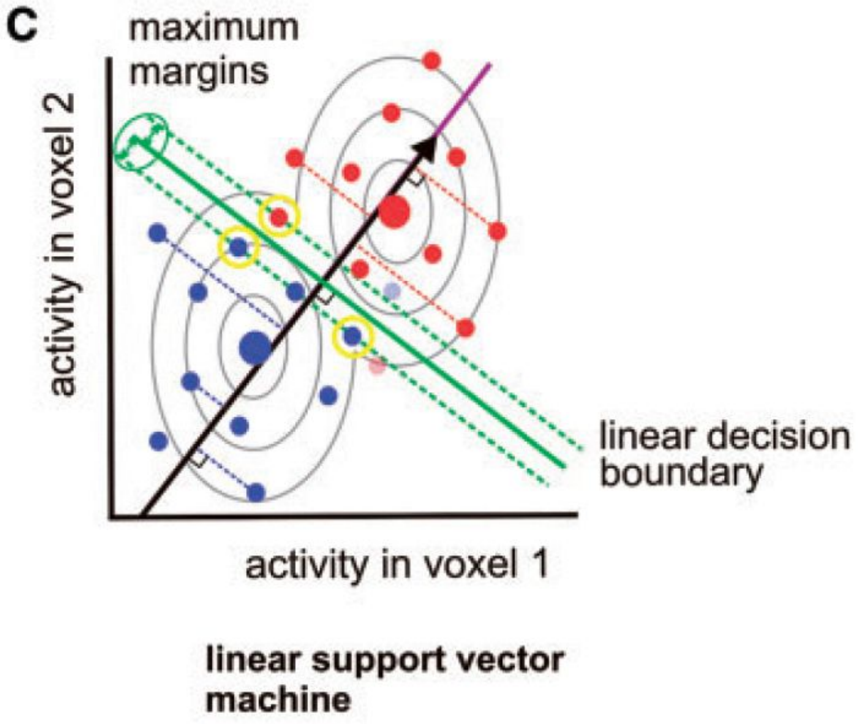


Support Vector Machine (SVM)

SVM finds a linear decision boundary that discriminates between two sets of points

Constrained to have a the largest possible distance from the closest points on both sides.

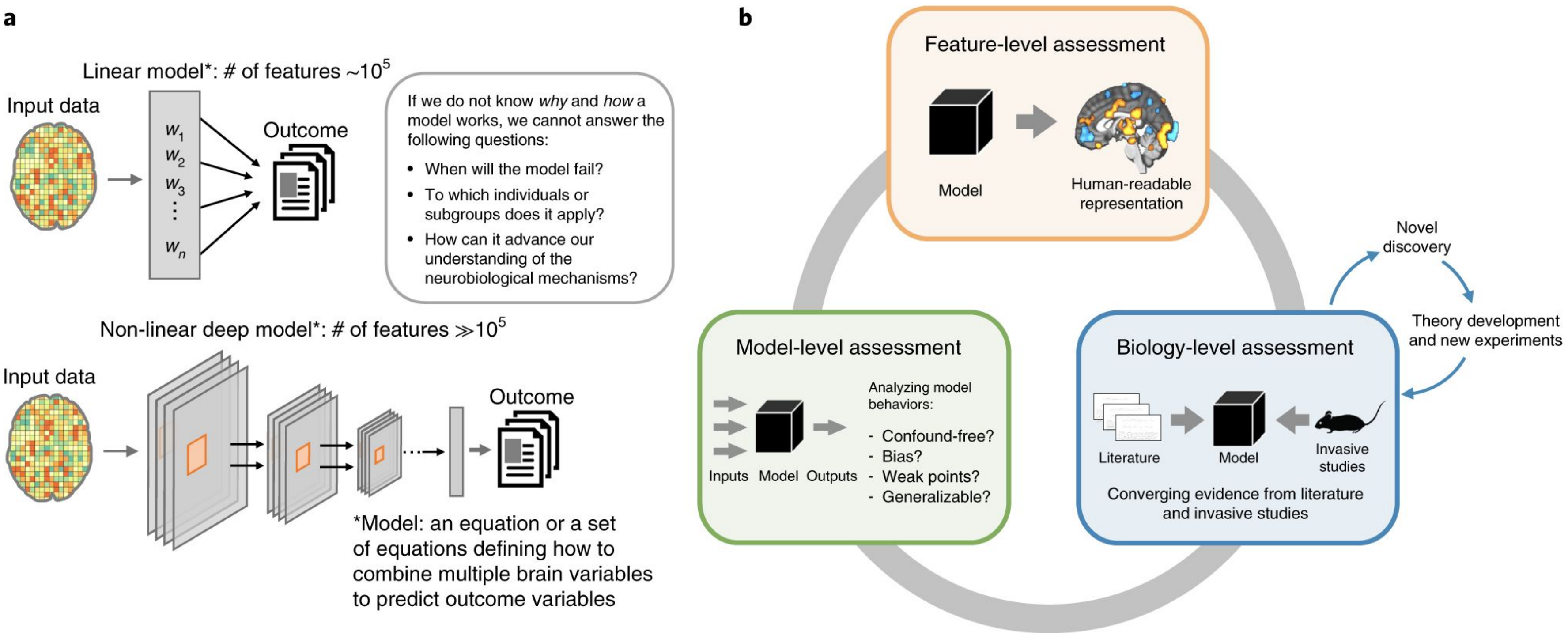
Response patterns closest to the decision boundary (yellow circles) that defined the margins are called “support vectors”.

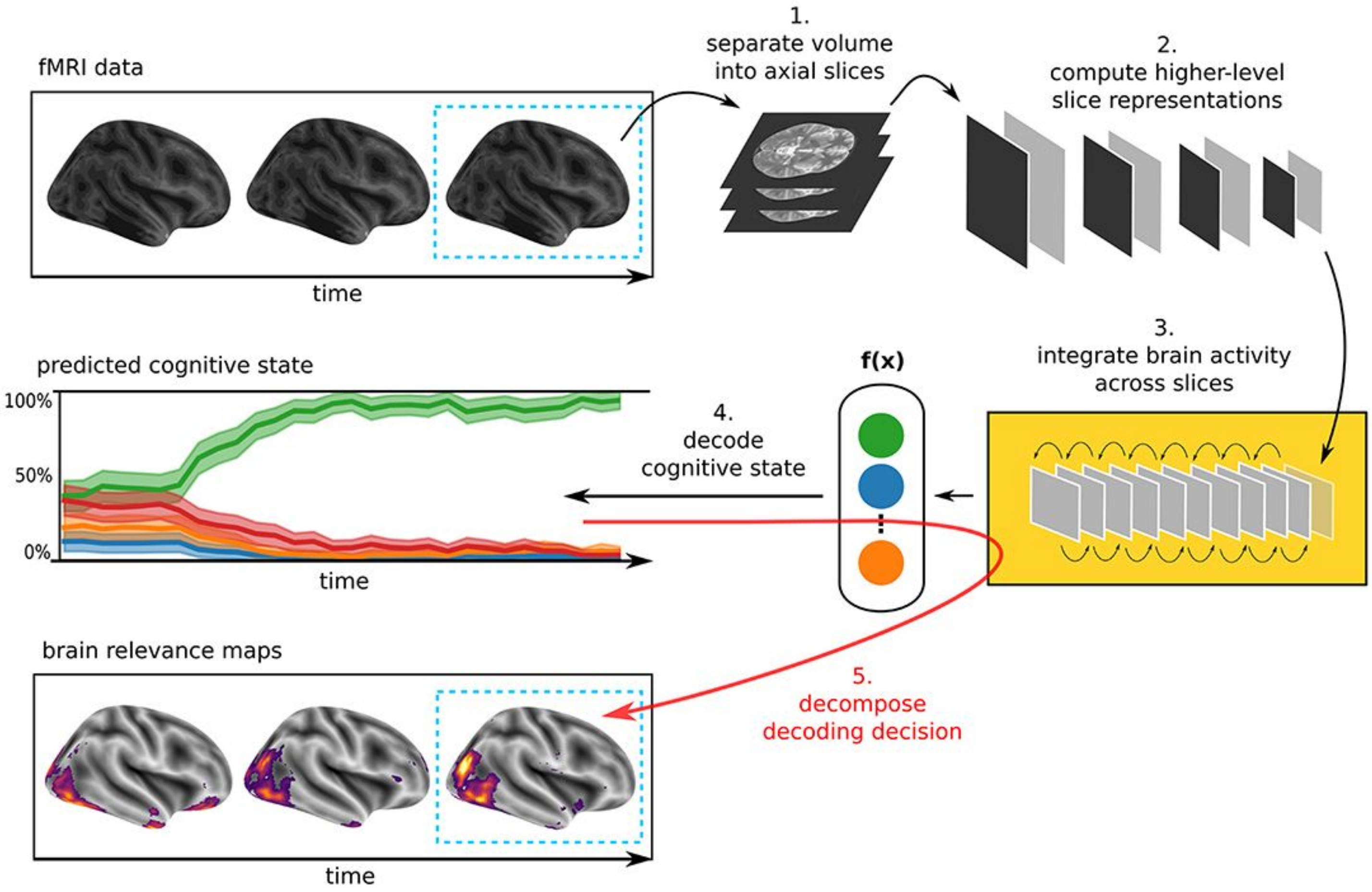


Steps to use classifiers

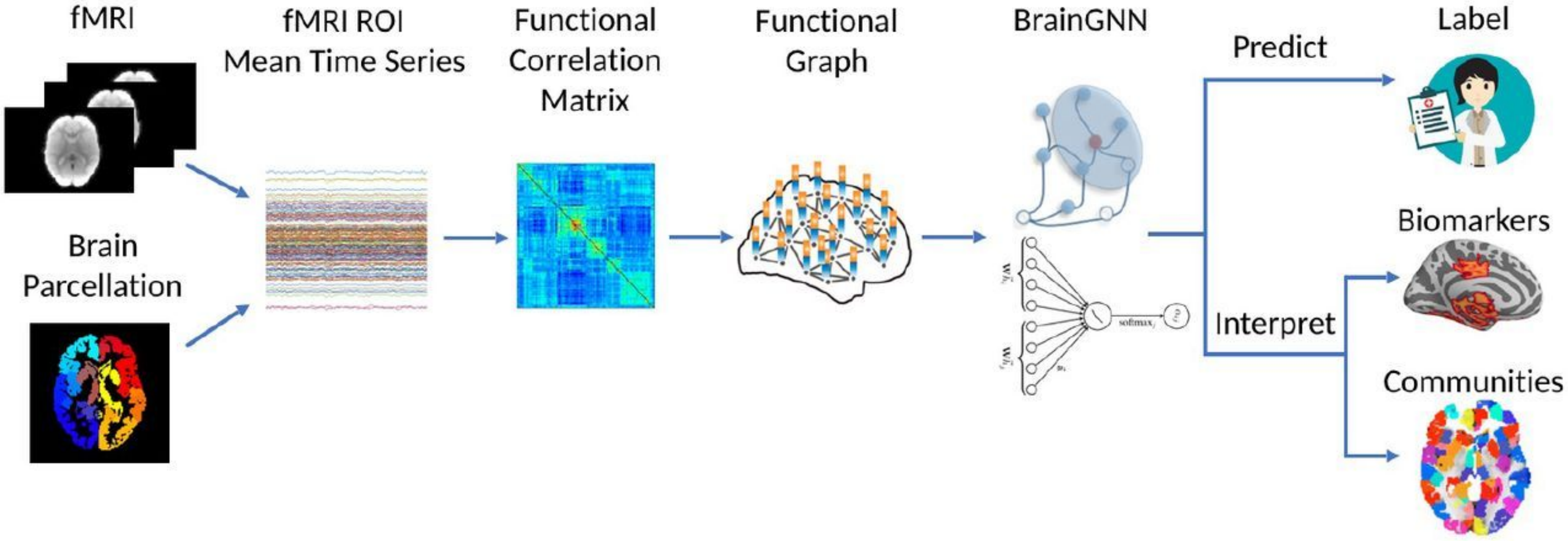
1. Get an estimate of activation for each voxel in ROI
2. Break data into independent training and test sets
3. Train classifier on training set, test on test set to determine accuracy of classification for each subject
4. Test significance of classification

Advanced Algorithms


a
Linear model*: # of features ~10^5
Input data
w1
w2
w3
wn
Outcome
If we do not know why and how a model works, we cannot answer the following questions:
• When will the model fail?
• To which individuals or subgroups does it apply?
• How can it advance our understanding of the neurobiological mechanisms?
Non-linear deep model*: # of features ≫10^5
Input data
Outcome
*Model: an equation or a set of equations defining how to combine multiple brain variables to predict outcome variables
b
Feature-level assessment
Model
Human-readable representation
Novel discovery
Theory development and new experiments
Model-level assessment
Inputs
Model
Outputs
Analyzing model behaviors:
- Confound-free?
- Bias?
- Weak points?
- Generalizable?
Biology-level assessment
Literature
Model
Invasive studies
Converging evidence from literature and invasive studies



fMRI data
time
1.
separate volume
into axial slices
2.
compute higher-level
slice representations
3.
integrate brain activity
across slices
f(x)
4.
decode
cognitive state
predicted cognitive state
100%
50%
0%
time
5.
decompose
decoding decision
brain relevance maps
time


fMRI
fMRI ROI
Mean Time Series
Functional
Correlation
Matrix
Functional
Graph
BrainGNN
Predict
Label
Brain
Parcellation
Biomarkers
Interpret
Communities

Appendix

fMRI Experimental Design