Kanwisher Lab Intro Tutorial #1

Nancy Kanwisher
Brain and Cognitive Science Department, MIT
McGovern Institute for Brain Research, MIT

Other Useful Resources:

My course on The Human Brain is online at MIT's Open Course Ware:
https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-13-the-human-brain-spring-2019/

Basics of fMRI Methods:

dan Blank's lectures on the GLM here:
https://www.youtube.com/watch?v=qqKm3EayUWY
https://www.youtube.com/watch?v=883bzPU6ndU&t=3s
https://www.youtube.com/watch?v=vsZez1MI4P4&t=2s
https://www.youtube.com/watch?v=qDiNg9-LDDg&t=3s
https://www.youtube.com/watch?v=BjDRsCC-8p8&t=1s

Rebecca Saxe's Amazing Bootcamp:
https://cbmm.mit.edu/fmri-bootcamp

1

Outline

1. Big Picture Overview
   What we think we have found & what questions we want to answer

2. BASIC fMRI methods:
   the fMRI signal and its limitations
   anatomical variability and the fROI method
   example: FFA

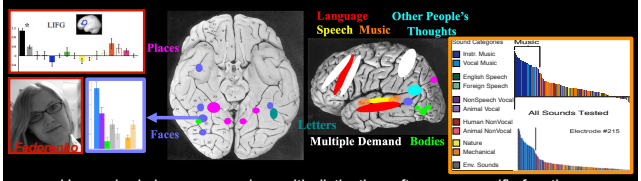
3. Alternate views about the ventral visual pathway
   other overlapping topographies
   Haxby's critique

4. Multiple Voxel Pattern Analysis (MVPA) and Representational Similarity Analysis (RSA)

5. ANNs and how we can compare them to Brains

2

Functional Distinct Regions in the Human Cortex

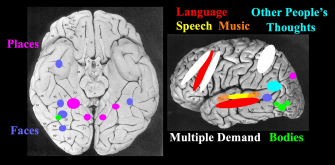


- Human brain has many regions with distinctive, often very specific functions
  Including regions selectively engaged by faces, places, bodies, etc
  & even regions specific for abstract distinctively human functions like
  language ≠ thought
  music
  thinking about each other's thoughts
- Not all cortical regions are functionally specific (the Multiple Demand system)!
- Of course each of these regions is part of a broader network.
- This set of regions can be taken as an initial rough sketch of the human mind.

Cool.
But mostly this picture reveals a vast landscape of questions……

3

Questions



- Do these regions play a specific *causal* role in behavior?
- What other mental functions get their own private patch of cortex?
- How does all this systematic organization arise in development?
  (or over evolution)
- What computations go on in each region?
- Why is this a good way to design a brain?

4

Functional Magnetic Resonance Imaging (fMRI)

The best spatial resolution available for measuring neural activity noninvasively in the whole human brain.

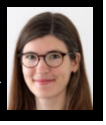
Because this is a blood-flow based signal….

"BOLD" (blood oxygenation level dependent) signal:
Increased neural activity >
Increased local blood flow more than compensates for O_2 use >
decrease in deO_2Hb concentration>
increase in MR signal intensity (deO_2Hb is paramagnetic)

5

Important caveats about the BOLD fMRI signal:

- Because the BOLD signal is based on blood flow,
  temporal resolution is bad (6 second delay, ~1 second res*.)
- Spatial resolution is limited:
  standard fMRI: ~.5 million neurons per voxel
  *#$@?!
- Cannot measure absolute amounts of activity/metabolism,
  only *differences* between two conditions.
- How would you use it to ask, for example…..



*Laura Lewis, new faculty member in EECS, is improving this!

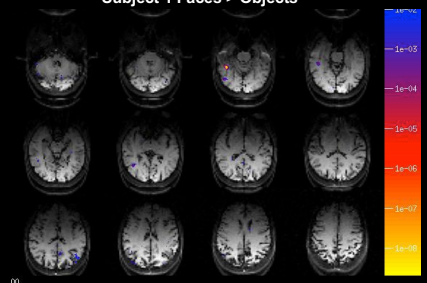
6

Do Any Brain Regions Respond Selectively to Faces?

7

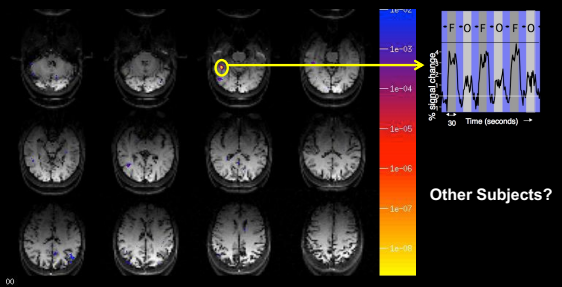
Do Any Brain Regions Respond Selectively to Faces?

Subject 1 Faces > Objects



8

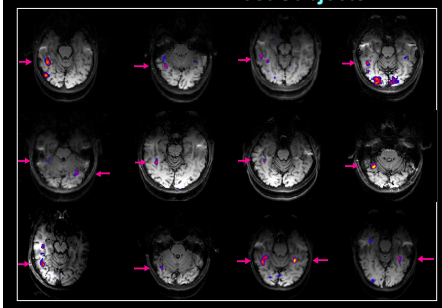
Do Any Brain Regions Respond Selectively to Faces?



Other Subjects?

9

Similar Activation for Faces > Objects In Most Subjects



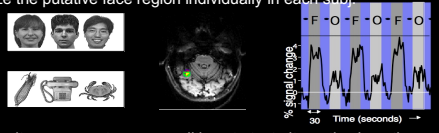
Cool, but…
Insufficient to prove this region is specific for faces.
Could respond to…
- anything animate
- anything roundish
- anything interesting
- etc.

To claim face specificity, we need to rule these out.
How, given anatomical variability?
The fROI method…

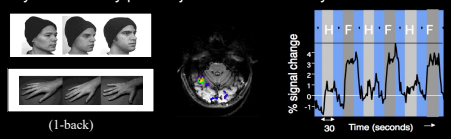
10

Functional Region of Interest (fROI) Method:

1. Localize the putative face region individually in each subj:



2. Measure its response to new conditions to test alternative hypotheses.
FFA: ~~any human body part~~? ~~"just attention"~~? ~~Only front views?~~

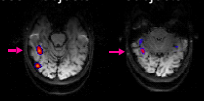


(1-back)

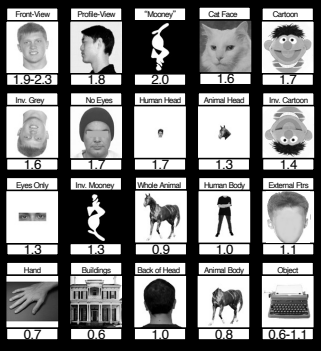
11

Fusiform Face Area

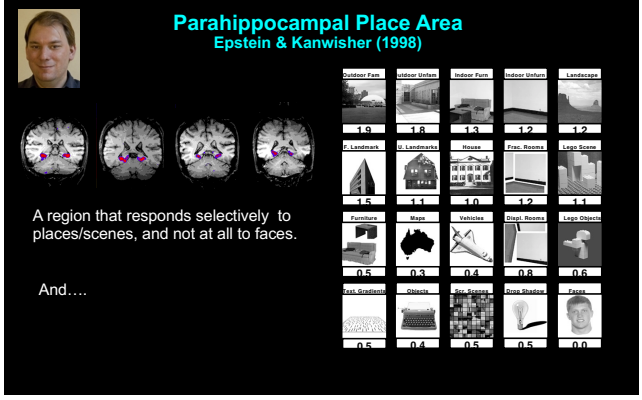
Faces > objects in 2 subjects:



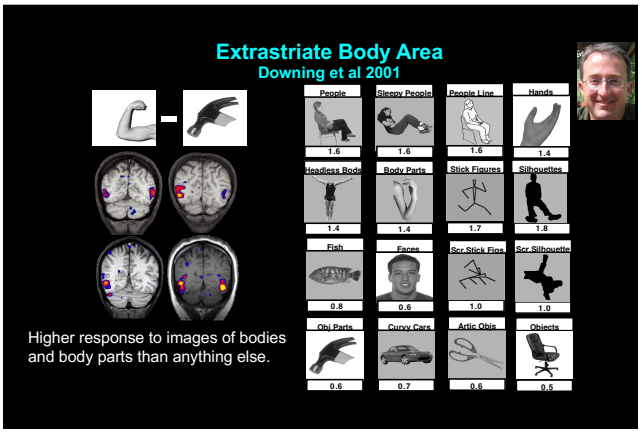
This region is found in a similar location in ~ every normal subject.
Responds more to faces>
anything else yet tested.
Does this mean that
this region is innate?
this region is *necessary* for face recognition?
Does it tell us how face recognition *actually works*?



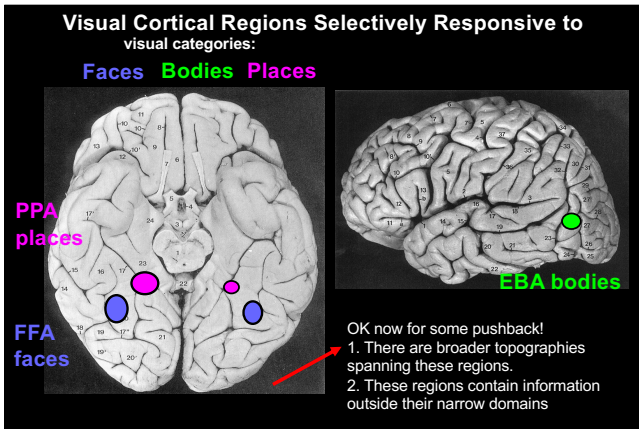
12



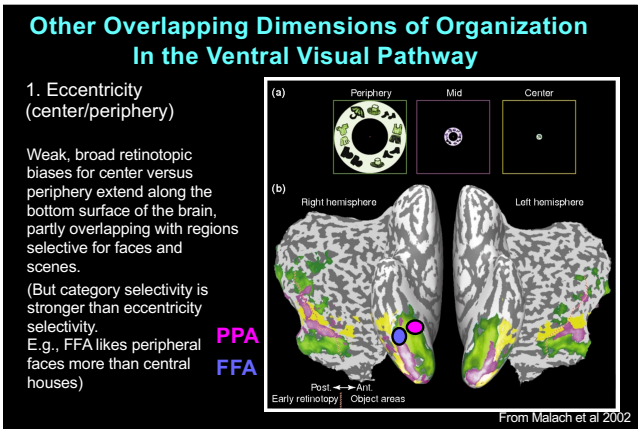
13



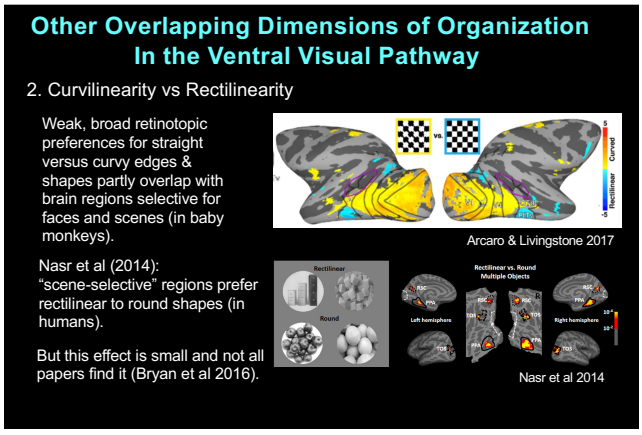
14



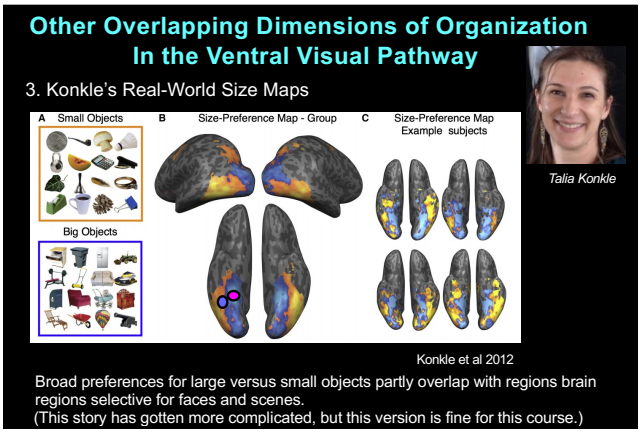
15



16



17



18

Other Overlapping Dimensions of Organization In the Ventral Visual Pathway

So, at least three different dimensions are coded along the same broad swath of brain where we find face and scene selectivity:

1. Eccentricity preferences
2. Rectilinearity/Curvilinearity preferences
3. Preferences for Real-World Size

So, things are more complicated than a few simple patches of category selectivity.

Many important, ongoing questions arise here:

E.g.: Might some "category-selective" regions really be showing *just* preferences for these dimensions?

That would be a confound that provides an alternative account,

e.g.: is the FFA just a "round things" area?

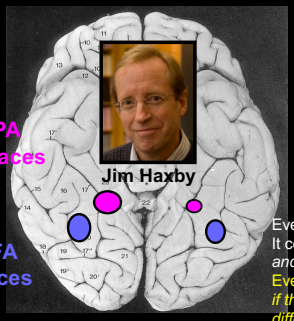
No, but the weak roundness preference is worth thinking about.

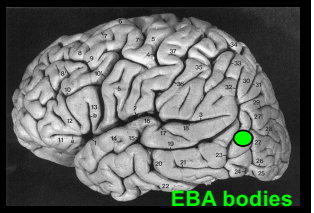
But the most serious challenge to the category selectivity comes from Haxby, so let's go over that….

19

Visual Cortical Regions Selectively Responsive to visual categories:

Faces Bodies Places





Even if FFA responds weakly to chairs & cars, It could still *represent information about chairs and cars.*

Even low responses could hold information *if the pattern of response across voxels is different for each category.*

The gist of the idea….

20

NK's Intuitive Version of Haxby's Central Point

Is visual information coded by big peak responses in the cortex, like this?

Haxby's (2001) experiment asks if this is true in the brain.

Or, can information be coded in the cortex by a *distributed pattern of weaker responses*, like this (Haxby)?

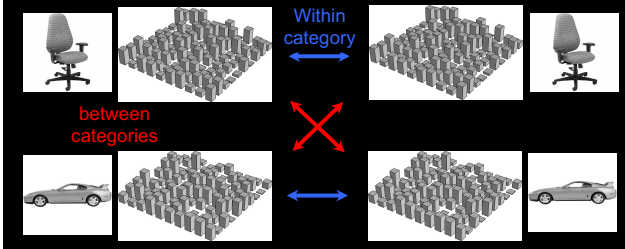
If the weak responses are fairly consistent within an object class, yet *consistently different* between object classes, then those weak responses contain information.

So, who says the peak responses are particularly important?
And who says peak regions represent only their preferred category?

21

Does the FFA hold Information about Nonfaces (e.g., cars versus chairs)?

1. Collect fMRI response to chairs and cars, for each voxel in FFA.
2. Repeat in same subject.
3. Now ask: is the pattern more similar within a category… than between



What does this method reveal about FFA?

If r(Within) > r(Between) the region contains info. distinguishing cars & chairs!

22

Does the Pattern of Response Across Voxels in the FFA Contain Information about Nonfaces? YES!

Haxby et al (2001): yes
"Regions such as the …. 'FFA' are not dedicated to representing only ... human faces.. but, rather, are part of a more extended representation for all objects".

Spiridon & Kanwisher (2002): no

O'Toole, Haxby et al. (2005): not very much
"preferred regions for faces & houses are not well suited to object classifications that do not involve faces and houses, respectively."

Reddy & Kanwisher (2007): uh, a little

Does that mean we should throw out the whole idea that the brain contains specialized processors that do specific things?
Discuss with your neighbors….

23

How do we Reconcile these Findings?

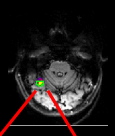
Cosmides & Tooby
Findings from fMRI: selective responses of FFA, PPA, EBA
E-stim of FFA (Japanese patient)
TMS to OFA and EBA (Pitcher)
Patients with prosopagnosia
Haxby et al (2001)

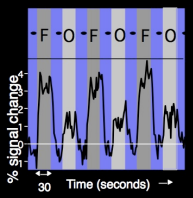
Does that mean we should throw out the whole idea that the brain contains specialized processors that do specific things?
Discuss with your neighbors….

24

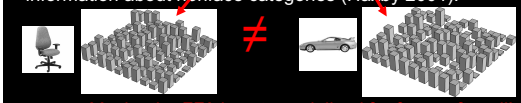
Haxby's Challenge to Face Selectivity of FFA

Faces > Objects





So what the FFA responds more to faces -
The *pattern* of response across voxels in the FFA contains information about nonface categories (Haxby 2001):



Maybe the FFA is not specialized for faces after all!
Also: this introduces a new method: MVPA

25

Multiple Voxel Pattern Analysis (MVPA)

Question: Can activity in region X distinguish A (cats) vs B (dogs)?
Scan ppl looking at cats & dogs.
Split data into: Odd runs data >> Even runs data

dogs odd

cats odd

dogs even

Pattern of resp in X

r(Within dogs)	r(Between)
r(Between)	r(Within cats)

cats even

Pattern of resp in X

If r(Within) > r(Between)
the region contains info. distinguishing dogs & cats!

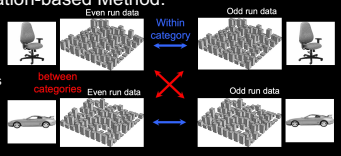
Two versions of MVPA…

26

Two Versions of Multiple Voxel Pattern Analysis

I. Correlation-based Method:

If r(within) > r (between) the region can distinguish cars from chairs

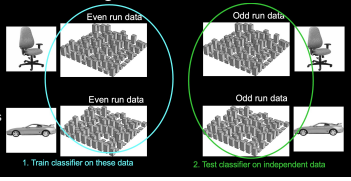


Same data, same idea, different math

Both are ways to ask: what information does neural response contain?

II. ML- based Decoding/Classification Method:

If classification accuracy > chance the region can distinguish cars from chairs

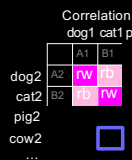


A generalization of this method….

27

RSA: Characterizing Representational *Spaces*

Correlation Method r (within)> r(between)



What is this?

MVPA: Can this region discriminate btwn class A (eg dog) vs class B (eg cat)?
worth knowing but binary so somewhat impoverished/arbitrary
Enter: Representational Similarity Analysis (RSA)
the *pattern of correlations across multiple conditions.*
You don't have to chose an arbitrary binary classification
a richer, more general characterization of representational *spaces*
Better yet: you can compare this RSA matrix to an RSA matrix based on a totally different kind of data (w/ same row and column labels)…..

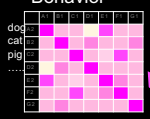
28

Relating Different RSA Matrices to Each Other

Having measured representational spaces in each data type we can now…

Compare them to each other!

Behavior

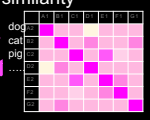


e.g. similarity ratings from ten subjs, a *cognitive* rep. space

r (behav – fMRI)

r (behav – phys)

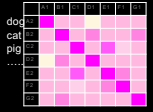
Neurophysiological similarity



e.g. from responses of 150 neurons

r (fMRI – phys)

fMRI Voxels

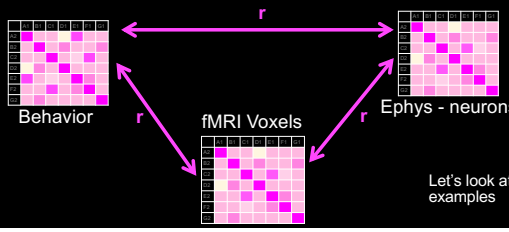


e.g. MVPA correlation across 200 voxels

29

Why this is so totally cool

1. It enables us to compare representational spaces
   across subjects (no spatial registration required!)
   across species
   & methods: 150 neurons (monkeys) to 200 voxels (humans)
2. It probes representations in a more comprehensive way.
   not just one or 2 discriminations (MVPA), but *lots* (RSA)
3. Key requirement: need same stimuli across data sets.

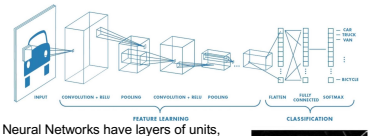


Let's look at some examples

30

What are Convolutional Neural Networks (CNNs)

see also https://poloclub.github.io/cnn-explainer



Deep Neural Networks have layers of units, each producing an output that is a nonlinear function of the input from some of the units in the previous layer.

What does 'convolutional' mean?
It means you apply the same "filter" to each patch of an image to extract each feature. For example….

"Strong" Edge Detect

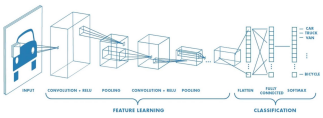
Loosely inspired by the primate visual system: (similar edge detectors all over the visual field in V1)

A really nice explainer on convolution: https://www.youtube.com/watch?v=8rHTiUzyZA

How you "train" the network:
Adjust the strength of every connection ("weight") in the entire network to optimize performance on a given task (e.g., recognizing objects despite image variation and clutter), over millions of examples using "backprop". (A big change from the days of hand-engineering networks.)

31

How can CNNs inform Mind & Brain?



1. Since 2012, CNNs trained on object recognition (like Alexnet) suddenly *work*, giving us:
   the first image-computable models of visual recognition.
   these now provide *possible* models of how obj rec works in humans/monkeys!
   So, are they in fact good models of obj rec in humans & monkeys?
   That is, do they fit data from humans/monkeys?
   How would we find out?
   How can we compare a brain to a bunch of software? (apples and oranges!)
   Two main ways…

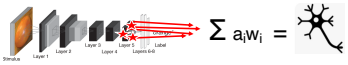
32

"Encoding Models": How well Can an ANN Predict a Neural Response?

Jim DiCarlo, MIT

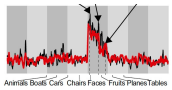
Dan Yamins, Stanford

2. Measure "response" of ANN already trained on object recognition to same images.



1. Measure response of neuron to natural images.

Yamins & DiCarlo 2014



YES!!!

3. Fit a ANN-based model to brain responses to a subset of images.
4. Does this model predict brain responses to *held-out images*?

Evidence for (some kind of) similarity of CNNs and brains

This method is what was used for your Assignment #3. Specifically…..

Yamins & DiCarlo, 2014

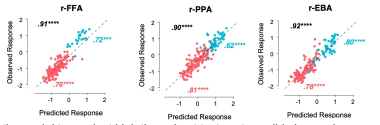
33

Will ANN-Based Encoding Models of FFA, PPA, and EBA Accurately Predict their Response?

Ratan Murty
Georgia Tech

Do the same thing with fMRI!*
Namely:
1. Scan each participant while they view 185 images,
2. Measure magnitude of e.g. FFA response to each image.
3. Find ANN weights that when summed fit response of FFA

Then test predictions on held out images.
Does this work? ***Yes!***



The other method for comparing ANNs to brains you already know…

Use the model to conduct high throughput tests not possible in people, or to pilot test new hypotheses,like you guys did for Assignment 3. we now routinely do this in my lab, totally new.
Also, can use GANs to generate "best" images

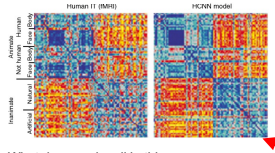
**Paper assigned for this week.
Understand the overall logic.
Don't worry about every detail.

Murty, Bashivan, et al 2021

34

Use RSA to Compare Layers of a Network to Regions of the Brain

1. Measure fMRI response of each voxel in human IT (=ventral pathway) to each of ~100 stimuli;
2. RSA: make representational dissimilarity matrix (RDM):

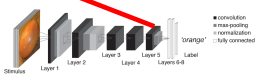


4. Make a representational dissimilarity matrix for a particular layer of CNN using the same stimuli

Fit is not perfect, but often quite good. Further, "higher" layers of the network better match "higher" layers of the visual system…

What does each cell in this matrix mean?
What can you see here?
What does the dark blue square mean?

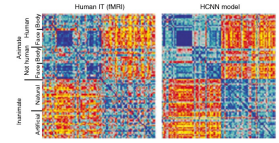
3. Feed same stimuli to a deep net, and "record" responses of each "neuron".



Khaligh-Razavi and Kriegeskorte, 2014

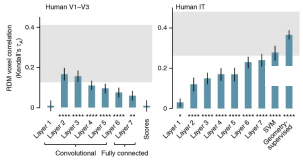
35

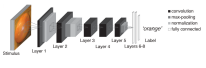
Use RSA to Compare Layers of a Network to Regions of the Brain



Fit is not perfect, but often quite good. Further, "higher" layers of the network better match "higher" layers of the visual system…

And "lower" layers of the visual system…





Khaligh-Razavi and Kriegeskorte, 2014

36