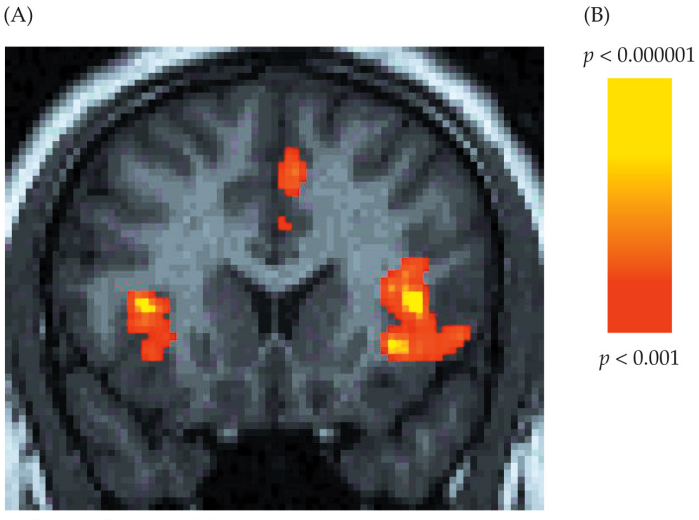
ANALYSIS APPROACHES PART ONE

1

Subtraction and "blobology"

- A simple approach: Response(X) > Response(Y) ?
- It has limitations, but there is nothing intrinsically wrong with this approach
- There are many important and interesting statistical and anaylysis issues that already arise under this approach.



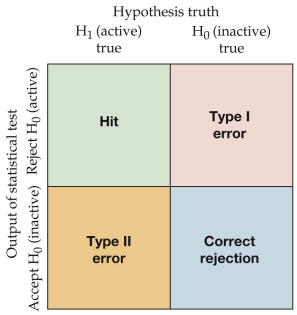
FUNCTIONAL MAGNETIC RESONANCE IMAGING 3e, Figure 10.1

2

Null Hypothesis Significance Testing (NHST)

- **Null hypothesis** (There is no interesting effect)
- **Significance** (How unlikely is the null hypothesis?)
- **Testing** (Implies a dichotomy or that a choice is made)

- The meaning of a *p*-value
- Statistical inference (An inference based on NHST)
- Type 1 error (There is no effect, but you think there is)
  - Also known as a "false positive"
- Type 2 error (The effect is there, but you fail to detect it)
  - Can view as a "miss" (like in signal detection theory)
  - Concept of power (assuming the effect exists, how likely will we detect it?)



3

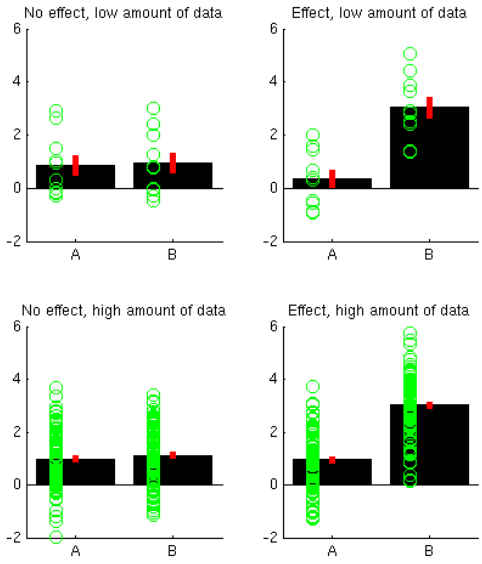
Null Hypothesis Significance Testing (NHST)

- **Effect size** is the magnitude of a phenomenon and does NOT change with the amount of data collected
- Examples: correlation coefficient, regression coefficient (beta weight), difference in means

4

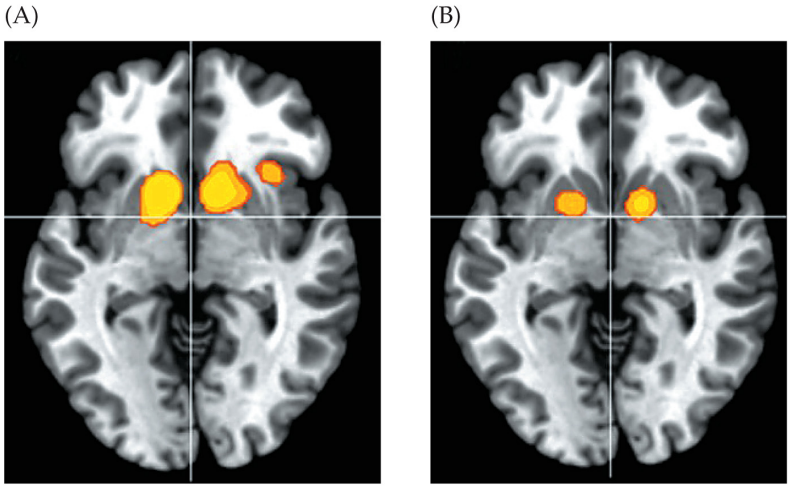
Null Hypothesis Significance Testing (NHST)

- Effect size vs. *p*-value
  - *P*-values depend on the amount of data
  - Idea of a statistically significant but tiny effect
  - *P*-values induce an artificial dichotomy
- A *p*-value is a single number, but we really need two numbers for full information
  - We need to understand both the response and the error on the response



5

Spatial smoothing will affect results



FUNCTIONAL MAGNETIC RESONANCE IMAGING 3e, Figure 8.29
© 2014 Sinauer Associates, Inc.

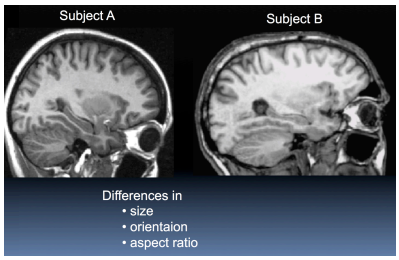
The location of activated regions *appears* to have changed.

Smoothing is useful for averaging out noise, but it will blur the signal, too.

6

Group-level analysis

- Two different types of goals:
  - How to establish that an effect generalizes across subjects within a group
  - How to compare two different groups
    - (Beware of confounding variables (e.g. group differences in head motion, vascularization, BOLD percent signal change)
- How do we anatomically register subjects?
  - Classically, people have used volume-based approaches (Talairach, MNI)
  - Newer surface-based approaches are more accurate (but are specific to cortex)
  - Alternatively, can perform functional localizers in individual subjects



slide courtesy of D. Van Essen

7

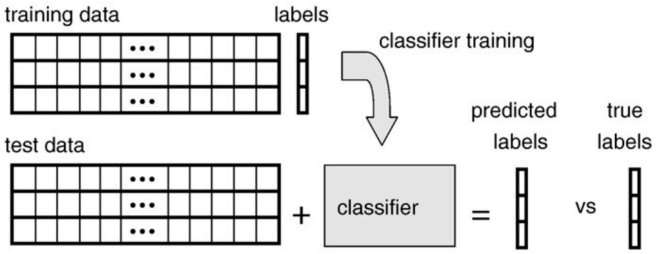
Subtraction approach: a summary

- Approach: "determine whether the BOLD response is higher in one experimental condition than another"
- This is a simple approach that can provide valuable insight
- There are important analysis issues that should be considered (e.g., effect size, smoothing, group spaces)

8

Multivariate pattern analysis (MVPA)

- A classifier is trained to distinguish between experimental conditions based on brain activity



from Pereira, *NeuroImage*, 2009

Fig. 2. A classifier is learned from the training set, examples whose labels it can see, and used to predict labels for a test set, examples whose labels it cannot see. The predicted labels are then compared to the true labels and the accuracy of the classifier–the fraction of examples where the prediction was correct–can be computed.

9

Multivariate pattern analysis (MVPA)

- A classifier is trained to distinguish between experimental conditions based on brain activity
  - E.g., show that activity in a region can be used to predict which of two conditions occurred with 70% accuracy
- Multivariate in a statistical sense (f(weights•voxels) = 0 or 1)
- Increased statistical power compared to "mass-univariate" approach
- Directionality is flipped
- Cross-validation becomes critical
- Classifiers can be fancy "machine learning" techniques (e.g. SVM) but can also be simple ones (e.g. LDA, nearest-prototype)

10

Multivariate pattern analysis (MVPA)

- Often thought to be accessing "fine-scale activity patterns" but this is not necessarily true in all cases
- It does allow abstraction away from particular activity patterns found in individual subjects (i.e. the units are now % correct(!))

- The concept of searchlight
  - Often used in conjunction with MVPA
  - Idea: search the whole volume, considering small groups of voxels at a time
  - It is essentially a way to perform voxel selection (i.e., not whole-brain, not 1 voxel, but somewhere in between)
  - Can also be viewed as a way to regularize whole-brain classifier weights

11

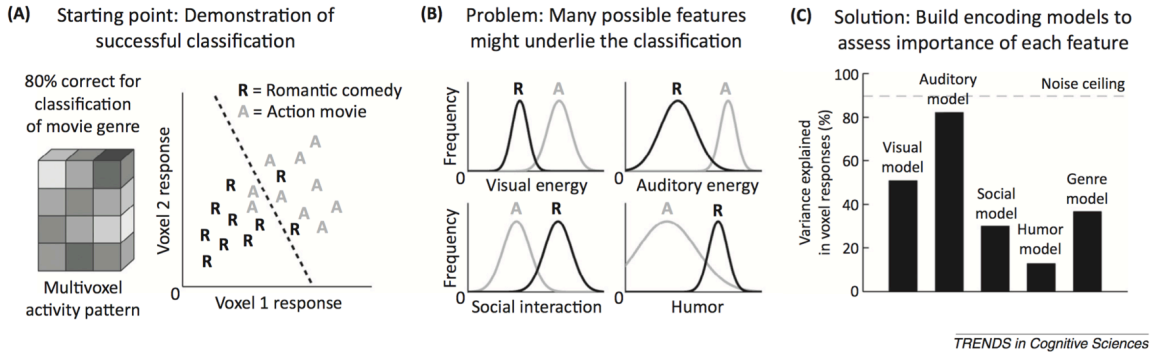
Multivariate pattern analysis (MVPA)

- Although MVPA is commonly characterized as vastly different from and superior to subtraction, this is probably an overstatement. Consider the following:
  - Subtraction can reveal fine-scale activity patterns (as long as you don't smooth)!
  - MVPA loses the sign of the effect! This is not ideal.
  - You can use one voxel to perform classification; thus, classification is not inherently multivariate...
  - Subtraction and MVPA are similar in that they are both methods for establishing whether a given experimental manipulation modulates brain activity.

12

Multivariate pattern analysis (MVPA)

- And there are some deeper limitations of MVPA…



Naselaris and Kay, *TICS*, 2015

13

MVPA approach: a summary

- Approach: "determine whether the BOLD response (distributed across several voxels) is reliably different for different experimental conditions"
- Compared to subtraction, this approach may provide more statistical power
- But somewhat more complicated to implement and more difficult to interpret

see also Etzel et al., *NeuroImage*, 2013; Haufe et al., *NeuroImage*, 2014

14

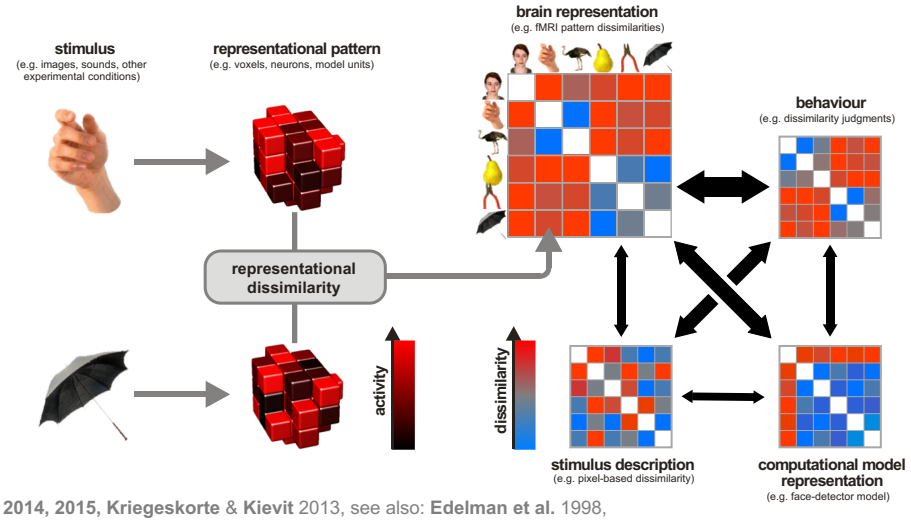
Representational similarity analysis (RSA)

- Create a matrix quantifying the dissimilarity of activity patterns for different experimental conditions
- For example, if **X** is voxels x conditions, a representational dissimilarity matrix (RDM) can be computed as conditions x conditions where element (i,j) is one minus the correlation between the ith and jth columns of **X**
- Note that this is essentially just MVPA on steroids (i.e., classification between all pairs of conditions)

- Primary uses:
  - As a method to visualize data
  - As a method to compress/summarize data
  - As a method to compare data from different modalities (Kriegeskorte et al., *Neuron*, 2008)

15

Representational similarity analysis (RSA)



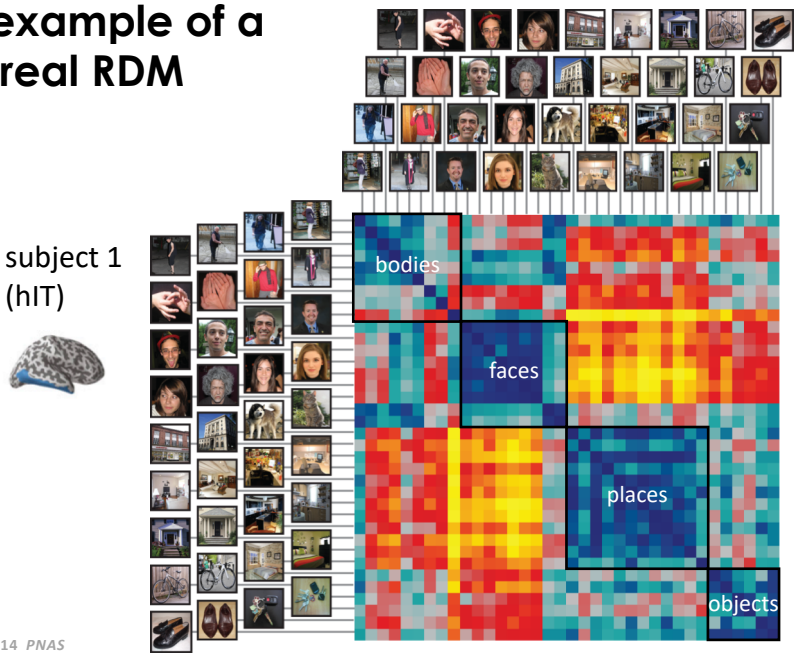
Charest et al. 2014, 2015, Kriegeskorte & **Kievit** 2013, see also: **Edelman et al.** 1998, **Laakso & Cottrell** 2000, **Op de Beeck** et al. 2001, **Haxby** et al. 2001, **Aguirre** 2007, **Kriegeskorte** et al. 2008

Slide adapted from Ian Charest

16

An example of a real RDM

subject 1
(hIT)



Charest et al. 2014 *PNAS*

Slide adapted from Ian Charest

17

Multidimensional scaling

A classic technique to help visualize the content of a dissimilarity matrix

Arrange dots in 2D to optimally approximate the full dissimilarity structure

- **bodies**
- **faces**
- **places**
- **objects**

Slide adapted from Ian Charest

18

Representational similarity analysis (RSA)

- Some choices to make:
  - What units should the data matrix be in? (e.g. % BOLD, z-score units, t-values?)
  - What distance metric should be used? (e.g. Euclidean, cosine, correlation?)
- The choices WILL affect the results and the interpretation that can be given

19

RSA approach: a summary

- Approach: "use a single 2D matrix to characterize how different the BOLD response is for different experimental conditions"
- Abstracts away from, and compresses, the data
- This can be seen as a good thing (e.g., now we can compare data from different modalities) or a bad thing (e.g., an RDM cannot be used to predict the actual level of brain activity in any given voxel)

20