# Learning Virtual Sensors for Cognitive States (10-702 Lecture 1/13/03)

Tom M. Mitchell

Rebecca Hutchinson, Marcel Just, Sharlene Newman, Francisco Pereira, Radu Niculescu, Xuerui Wang

**Paper draft available: www.cs.cmu.edu/~tom/nips02.ps**

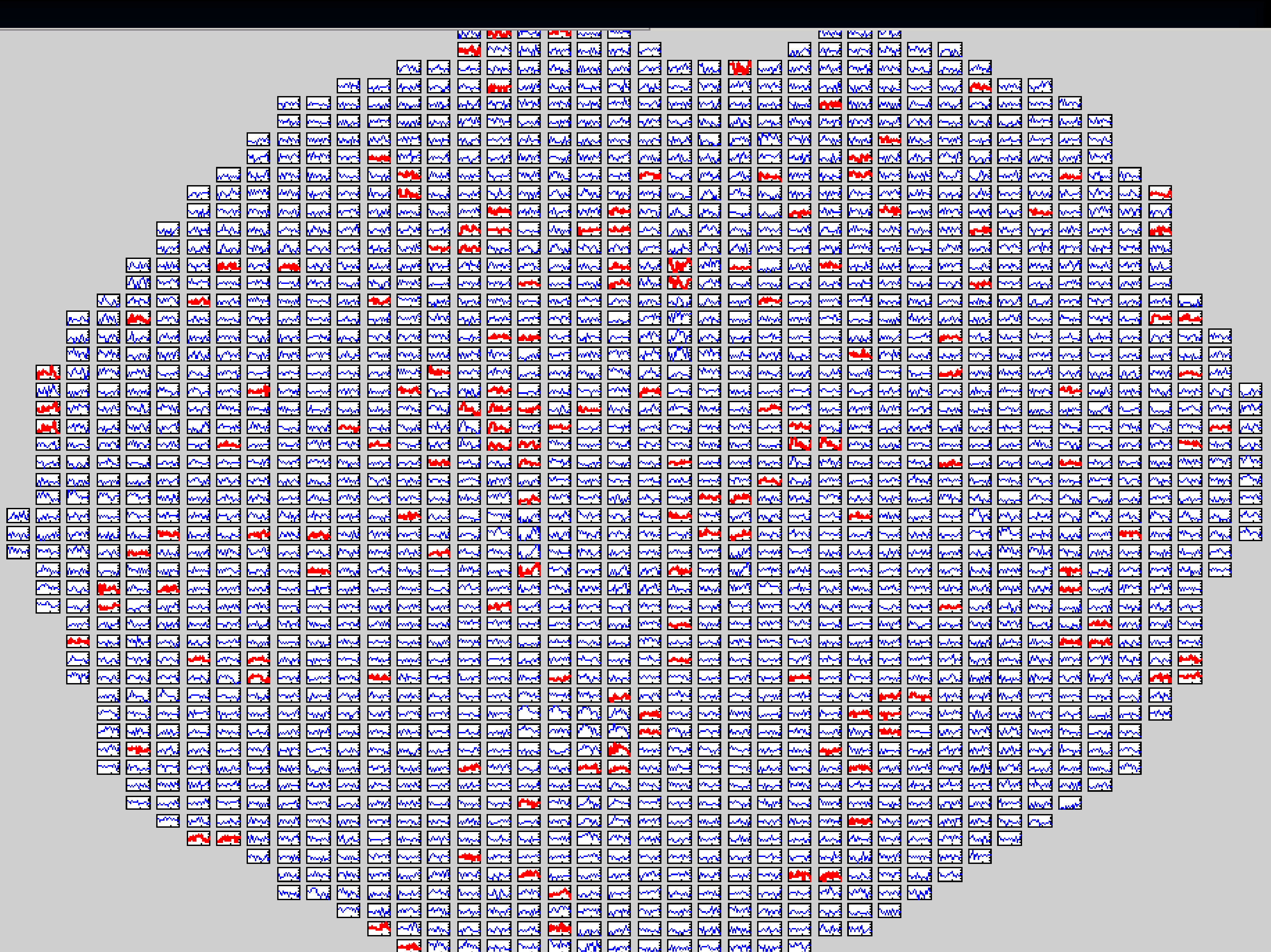
# Can we train classifiers to decode instantaneous cognitive state?

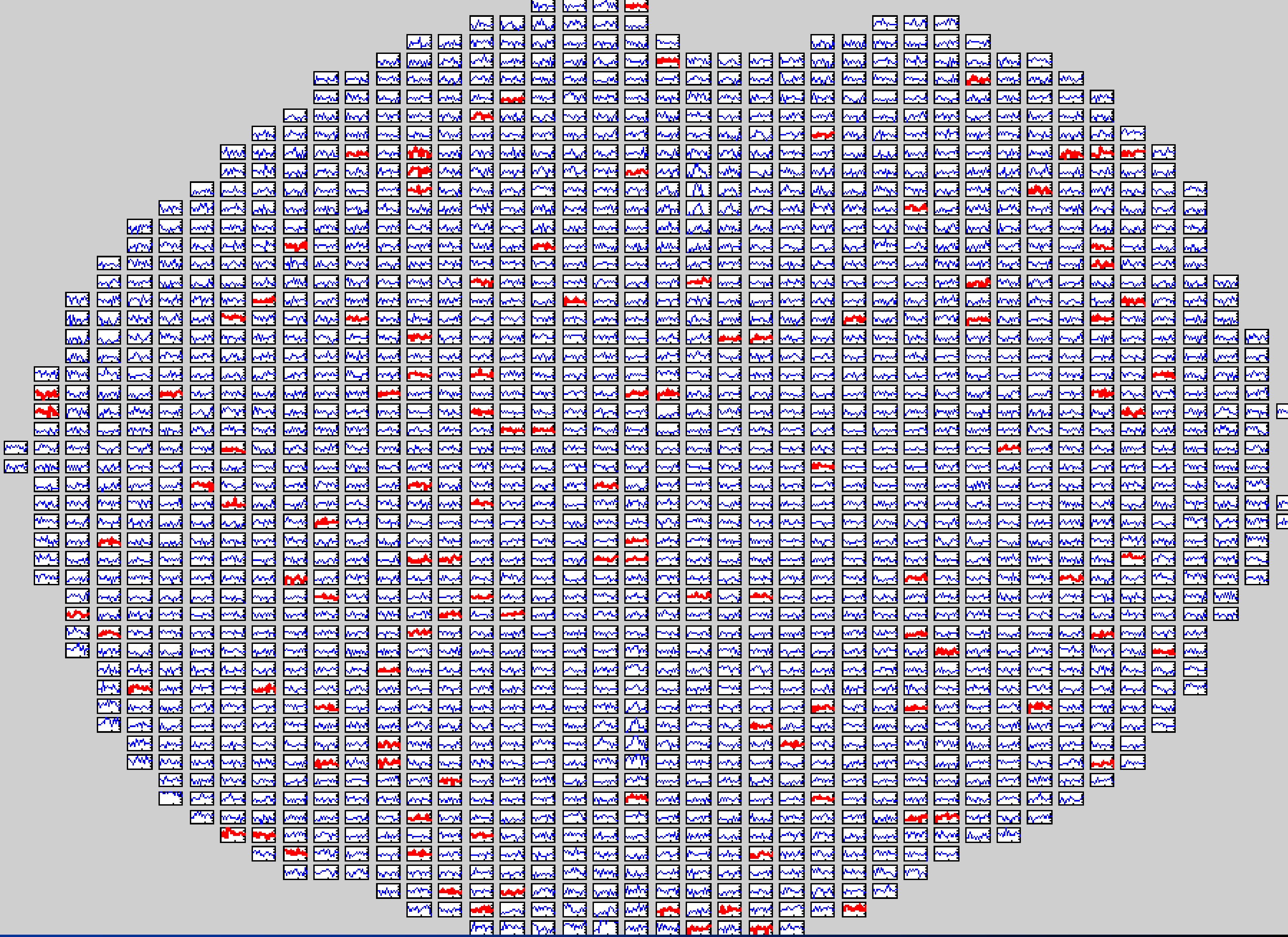
- → virtual sensors for cognitive states
- much analysis of *average* fMRI response
- little consideration of this question!

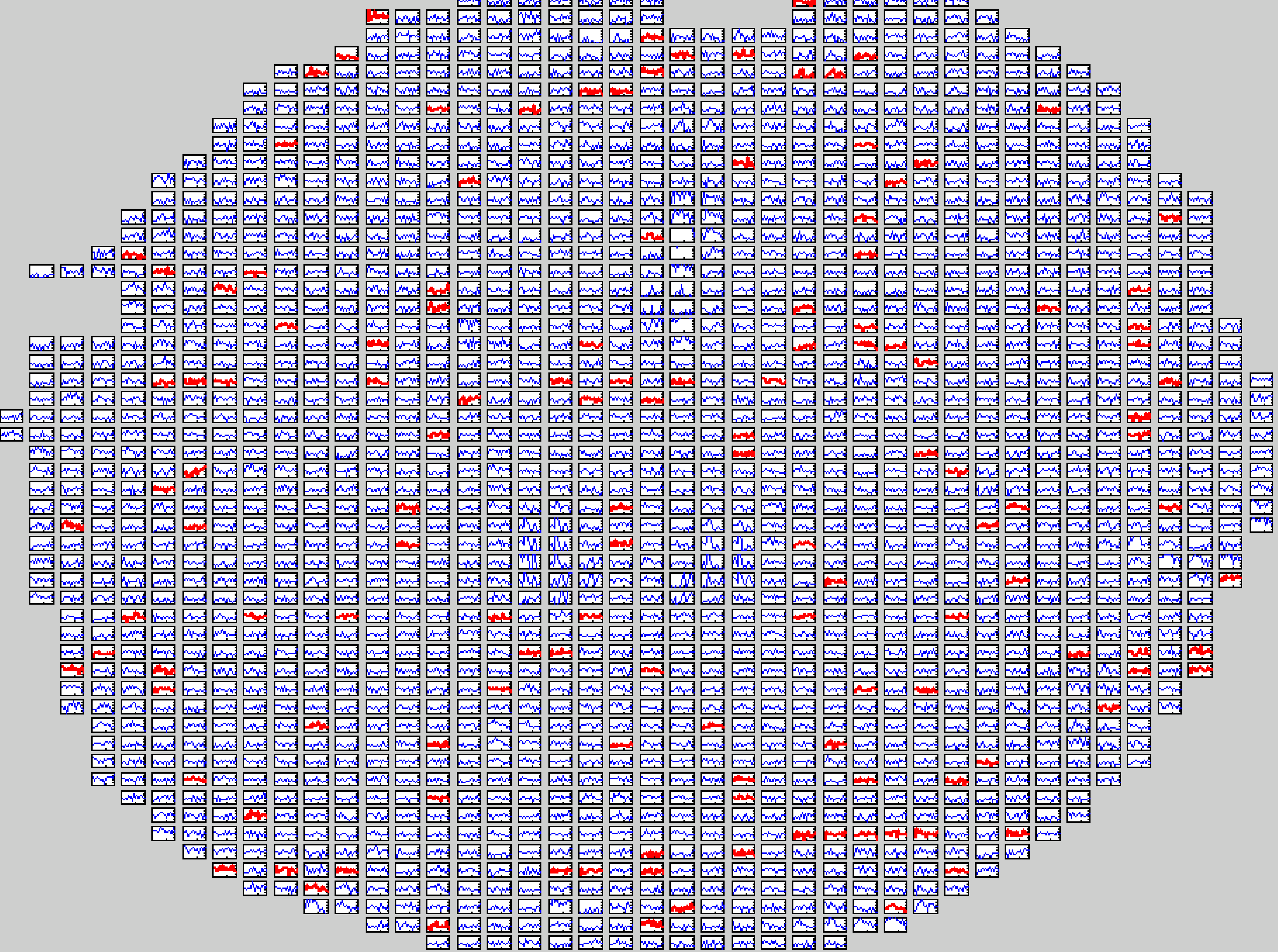
# Classifiers for Cognitive States

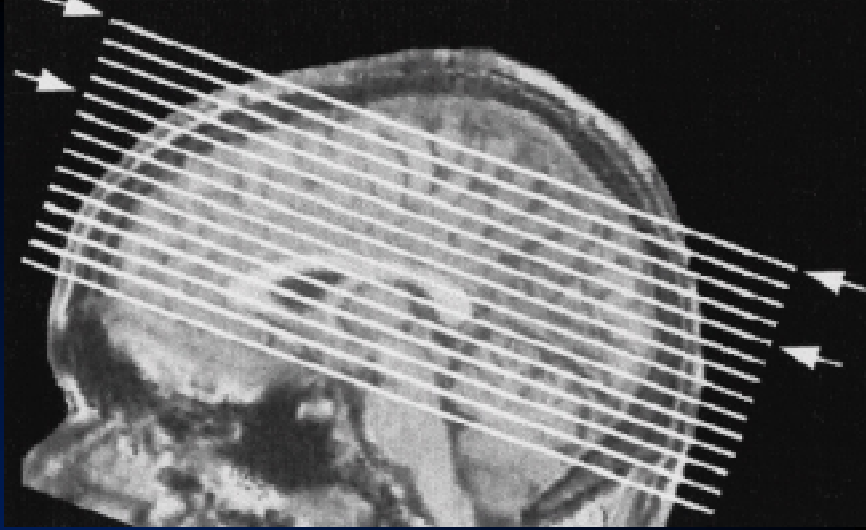
Difficult!

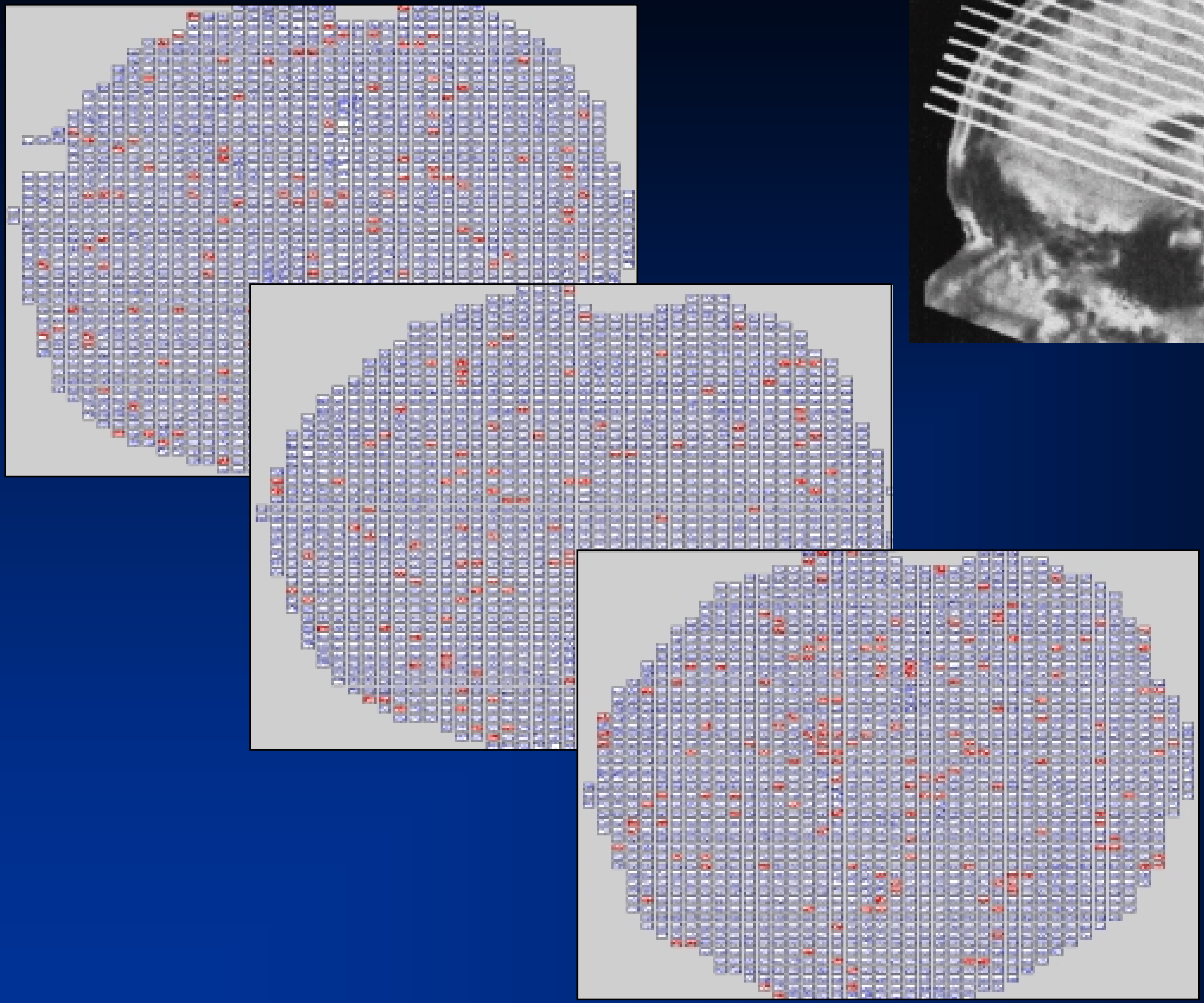
- Data very noisy
- High dimensional
- Sparse training data











• • •

# Approach

- Learn fMRI(t,…,t+k) → CognitiveState
- Classifiers:
  - Gaussian Naïve Bayes, SVM, kNN
- Feature selection/abstraction
  - Select subset of voxels (by signal, by anatomy)
  - Select subinterval of time
  - Average activities over space, time
  - Normalize voxel activities
  - …

# Study 1: Word Categories

[Francisco Pereira]

- Family members
- Occupations
- Tools
- Kitchen items
- Dwellings
- Building parts
- 4 legged animals
- Fish
- Trees
- Flowers
- Fruits
- Vegetables

# Word Categories Study

- Ten neurologically normal subjects
- Stimulus:
  - 12 blocks of words:
    - Category name (2 sec)
    - Word (400 msec), Blank screen (1200 msec); answer
    - Word (400 msec), Blank screen (1200 msec); answer
    - …
  - Subject answers whether each word in category
  - 32 words per block, nearly all in category
  - Category blocks interspersed with 5 fixation blocks

# Training Classifier for Word Categories

Learn fMRI(t) → word-category(t)

- fMRI(t) = 8470 to 11,136 voxels, depending on subject

Feature selection: Select n voxels

- Best single-voxel classifiers
- Strongest contrast between fixation and some word category
- Strongest contrast, spread equally over ROI's
- Randomly

Training method:

- train ten single-subect classifiers
- Gaussian Naïve Bayes → P(fMRI(t) | word-category)

# Results

Classifier outputs ranked list of classes

Evaluate by the fraction of classes ranked ahead of true class

- 0=perfect, 0.5=random, 1.0 unbelievably poor

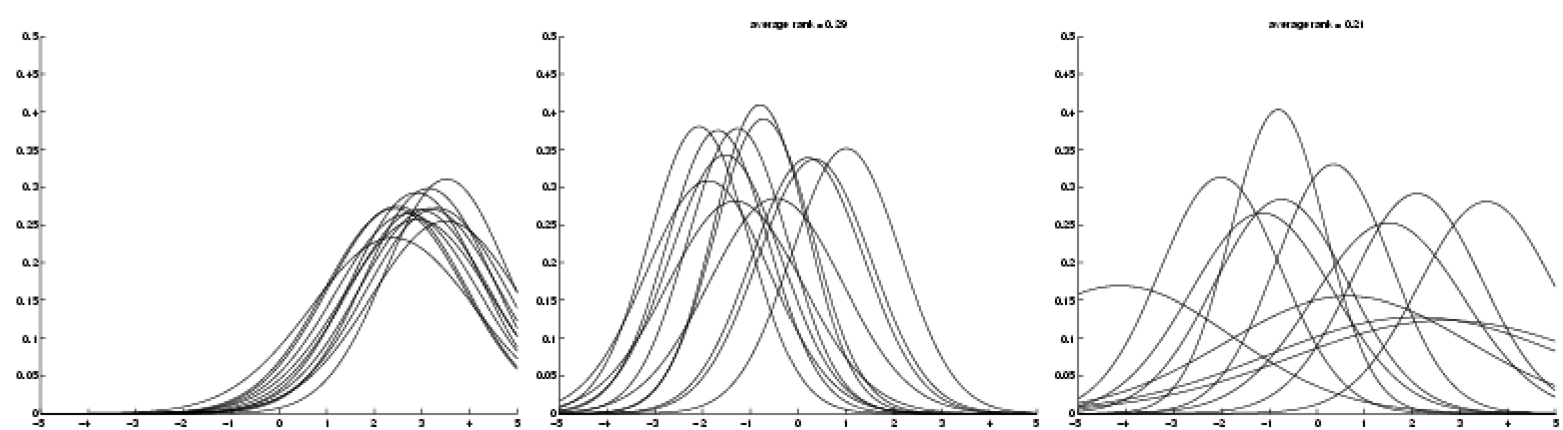
| *Experiment* | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *J* |
|---|---|---|---|---|---|---|---|---|---|---|
| 12 classes | 0.13 | 0.17 | 0.04 | 0.12 | 0.06 | 0.069 | 0.2 | 0.04 | 0.14 | 0.05 |
| 6 classes | 0.45 | 0.52 | 0.4 | 0.5 | 0.42 | 0.38 | 0.52 | 0.35 | 0.50 | 0.29 |

Try abstracting 12 categories to 6 categories

e.g., combine “Family Members” with “Occupations”

# Impact of Feature Selection

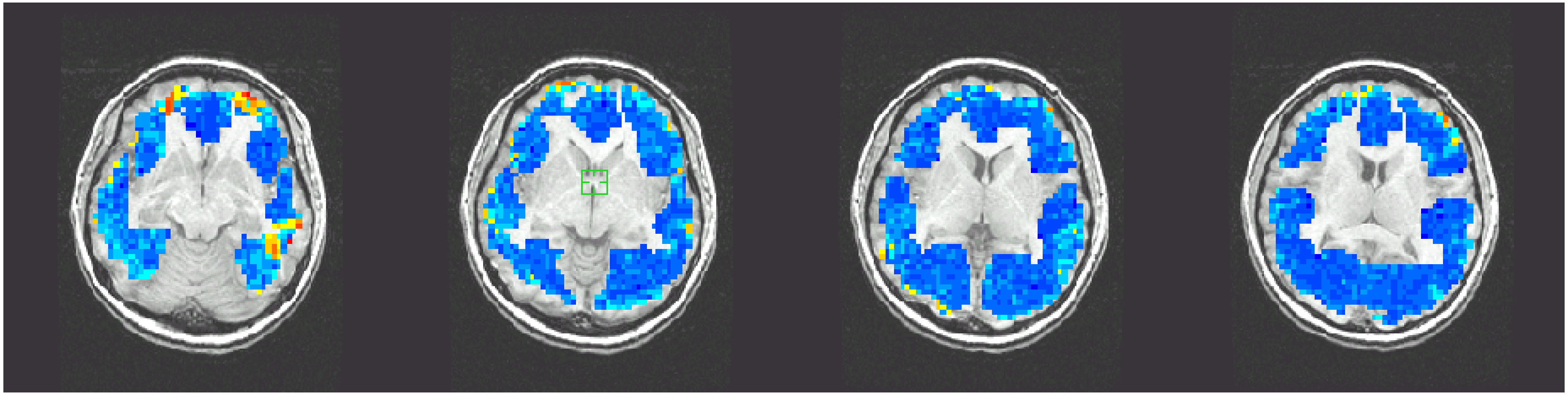
| *Feature Selection Method* | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *J* |
|---|---|---|---|---|---|---|---|---|---|---|
| Feature Average Rank | 0.11 | 0.18 | 0.04 | 0.12 | 0.065 | 0.085 | 0.20 | 0.047 | 0.13 | 0.054 |
| Activity p-value | 0.049 | 0.042 | 0.024 | 0.07 | 0.029 | 0.032 | 0.049 | 0.02 | 0.078 | 0.013 |
| Activity p-value per ROI | 0.13 | 0.15 | 0.051 | 0.12 | 0.058 | 0.068 | 0.13 | 0.039 | 0.15 | 0.042 |
| Random | 0.14 | 0.18 | 0.065 | 0.13 | 0.072 | 0.082 | 0.22 | 0.052 | 0.15 | 0.059 |
| Use all voxels | 0.13 | 0.17 | 0.044 | 0.12 | 0.058 | 0.069 | 0.2 | 0.043 | 0.14 | 0.05 |



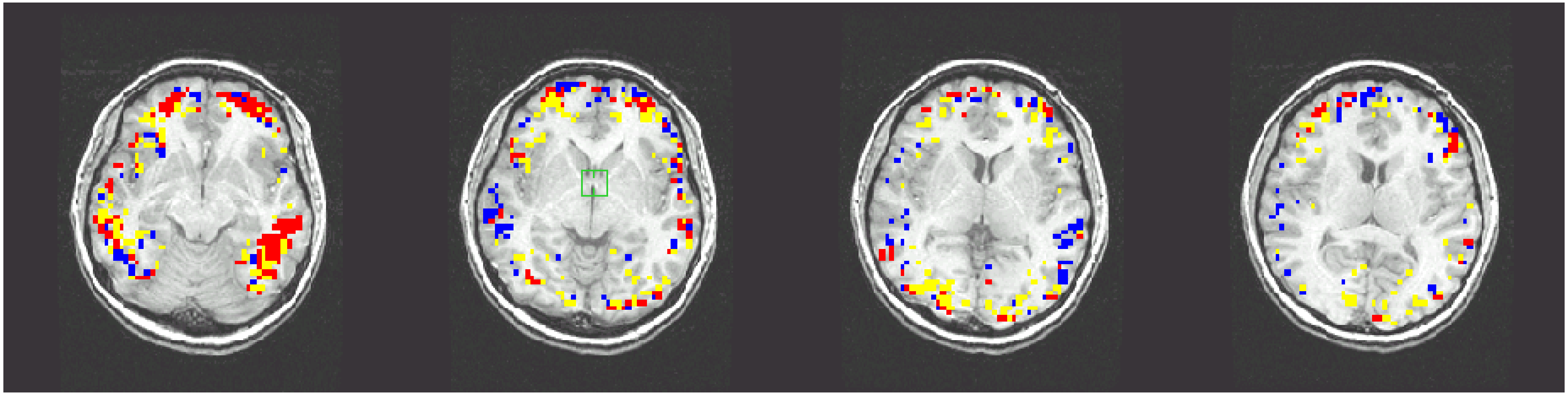
(a) Voxel in 1200 most active only

(b) Voxel in 1200 best discriminators only

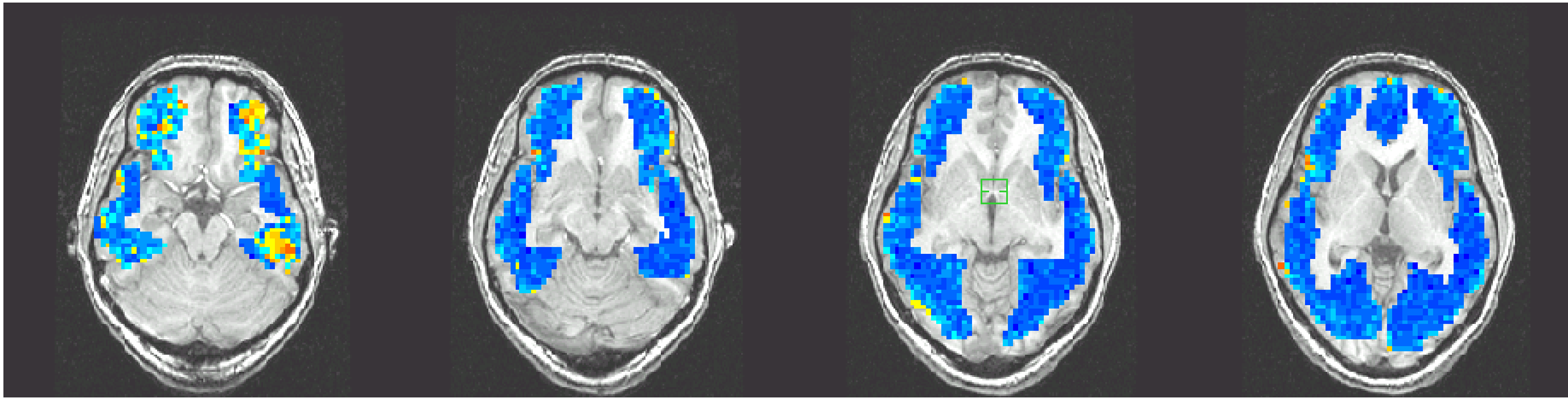
(c) Voxel in both groups



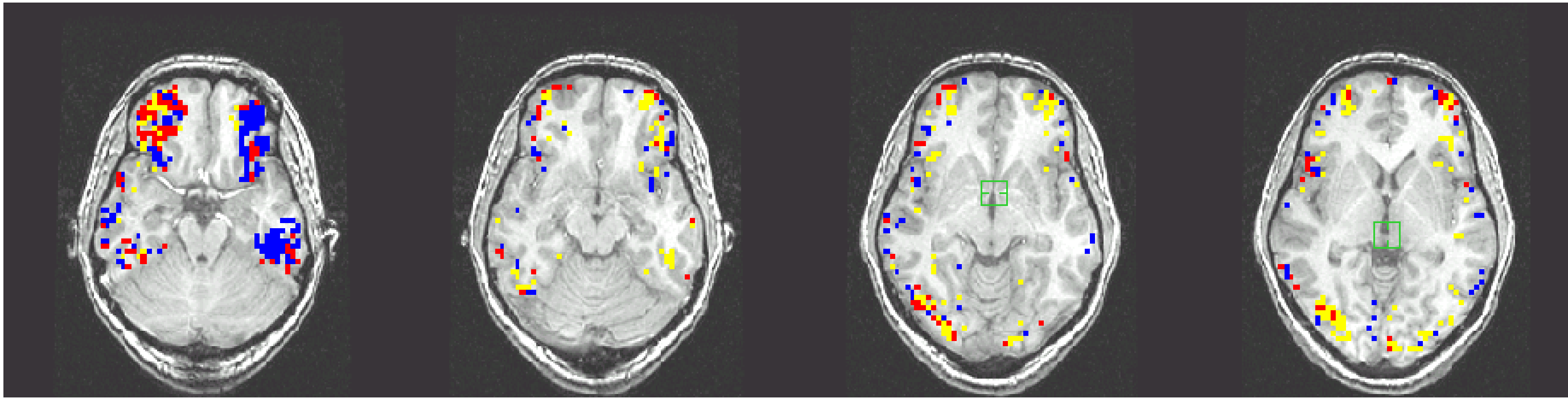
(a) Distribution of discriminating voxels for subject 1 (the more red the better)



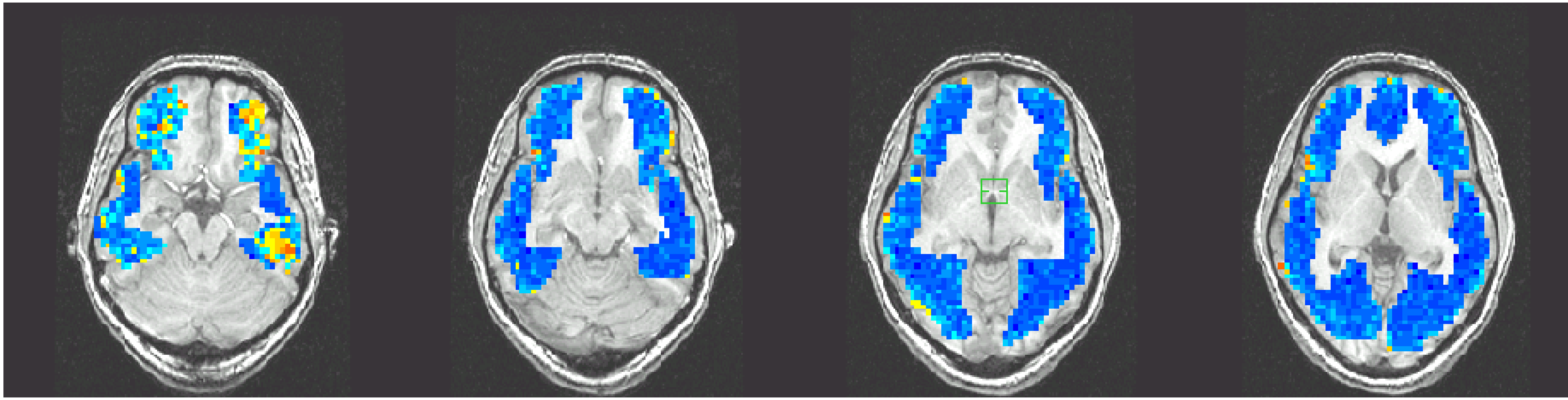
(b) Overlap (red) of the active (yellow) and discriminating (blue) 1200 voxel subsets for sub. 1



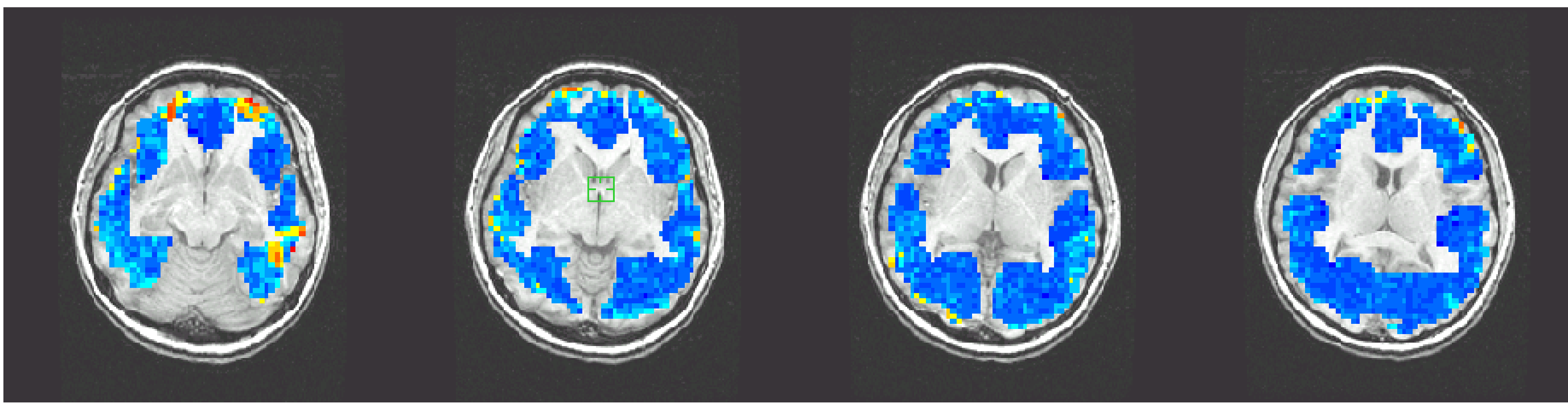
(c) Distribution of discriminating voxels for subject J (the more red the better)



(d) Overlap (red) of the active (yellow) and discriminating (blue) 1200 voxel subsets for sub. J

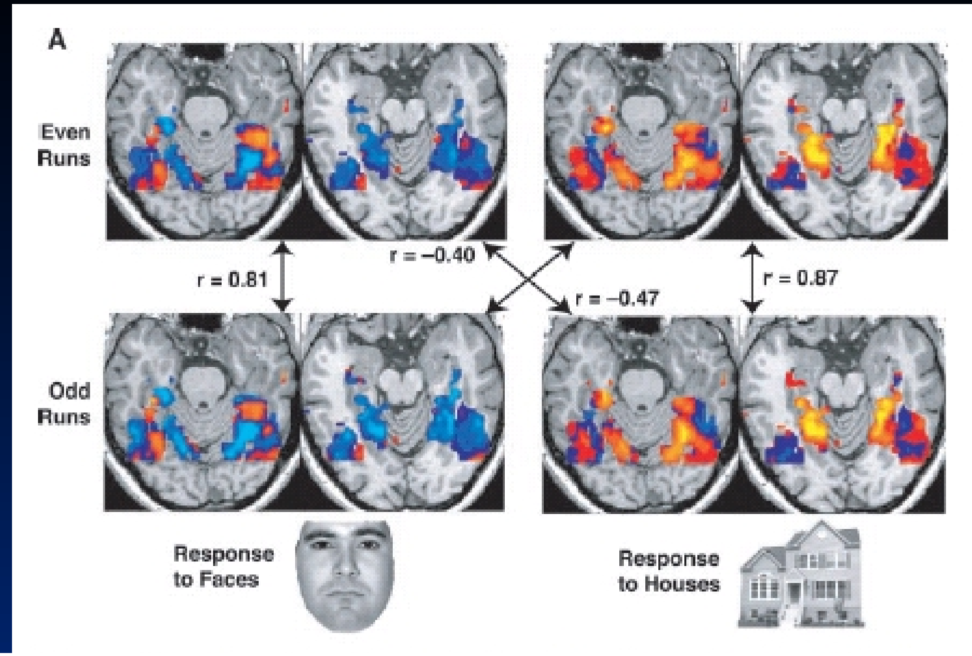


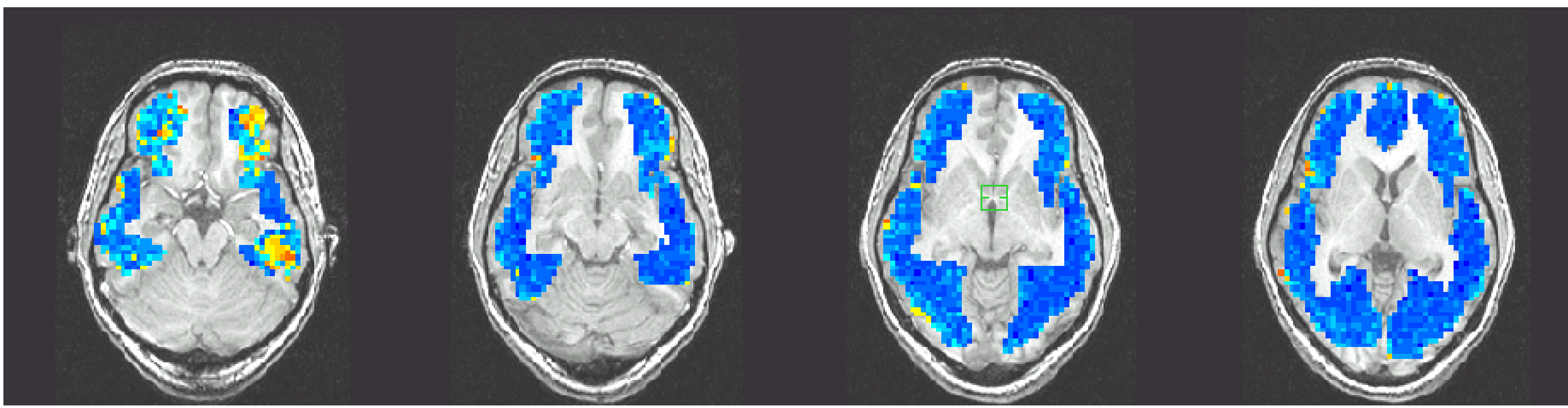
(c) Distribution of discriminating voxels for subject J (the more red the better)



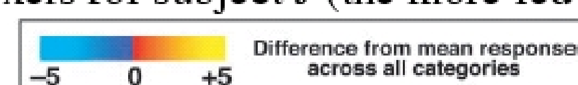
(a) Distribution of discriminating voxels for subject I (the more red the better)

# [Haxby et al., 2001]



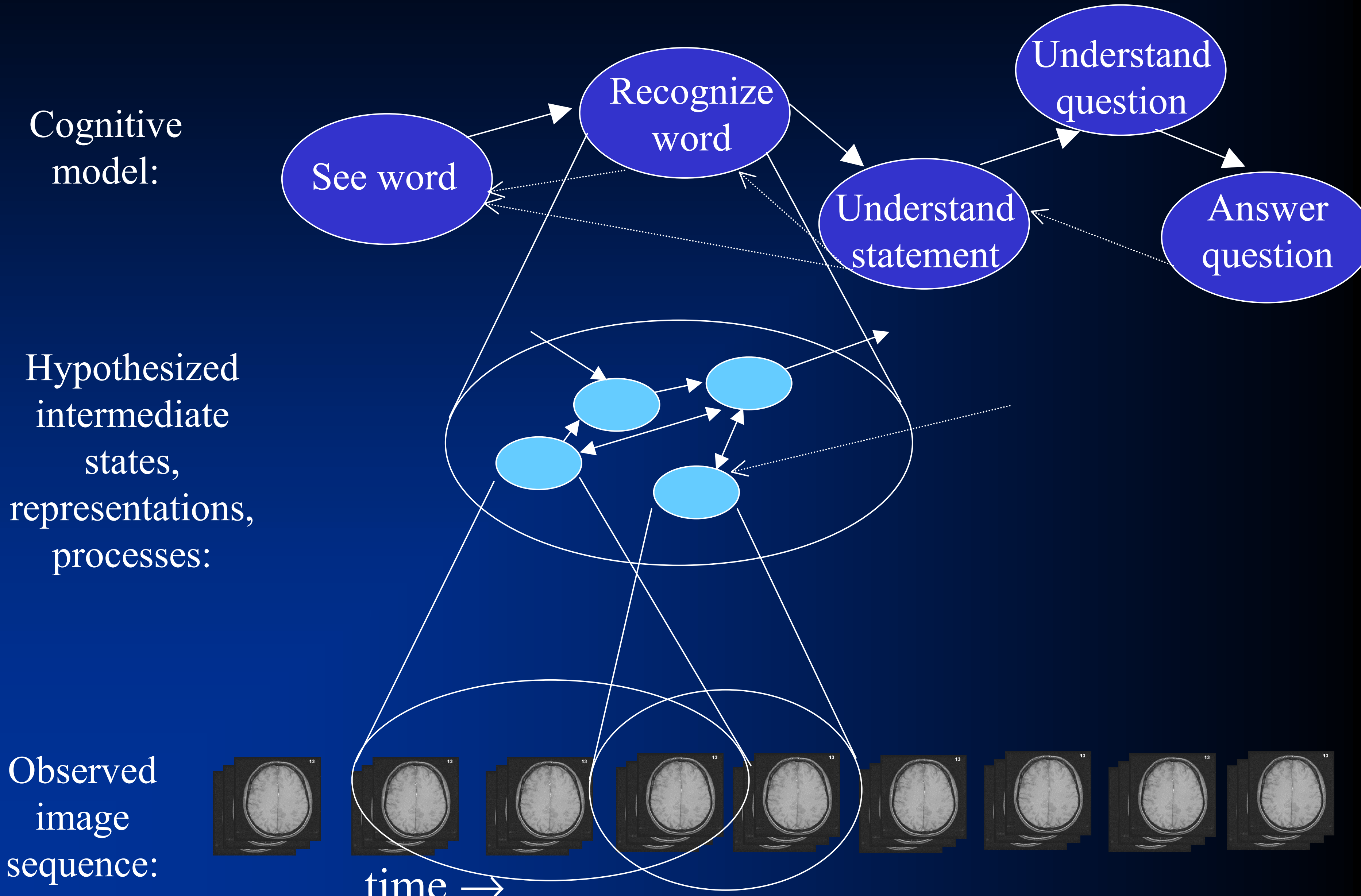
(c) Distribution of discriminating voxels for subject J (the more red the better)



# Study 1: Summary

- Able to classify single fMRI image by word category block
- Feature selection important
- Is classifier learning word category or something else related to time?
  - Accurate across ten subjects
  - Relevant voxels in similar locations across subjs
  - Locations compatible with earlier studies

# What We'd Like



# Challenge: virtual sensors to track sequence of cognitive states

a=6,…  3x+a=2