

# Inferring mental states from imaging data: OpenfMRI and the Cognitive Atlas

Russell Poldrack

Department of Psychology

Stanford University



# Two kinds of cognitive neuroscience questions

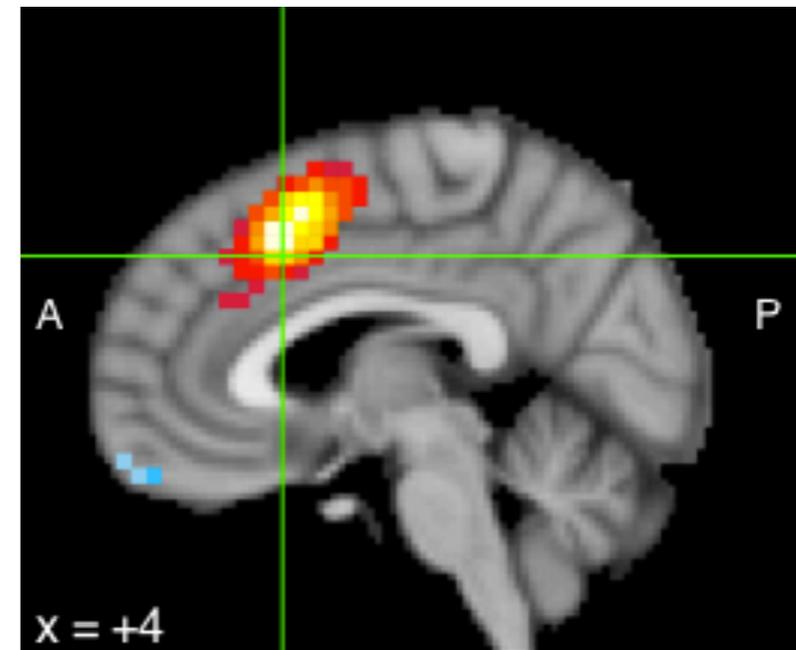
1. What are the neural correlates of mental process X?
2. What does area Z do?

# Neural correlates

Manipulate some  
mental process

Observe associated  
brain activation

working memory  
maintenance



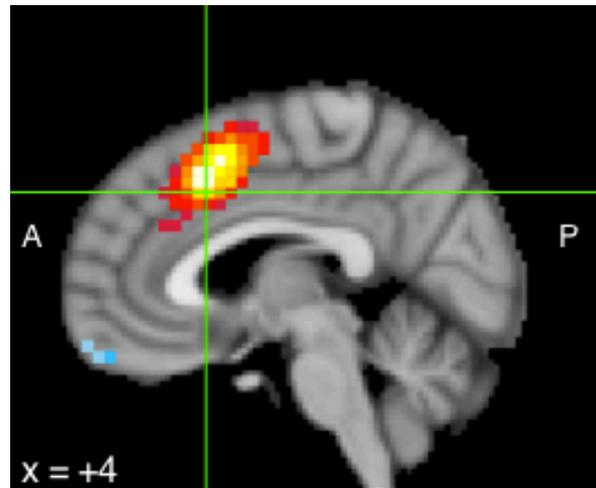
working memory is sufficient to activate ACC

~~working memory is necessary to activate ACC~~

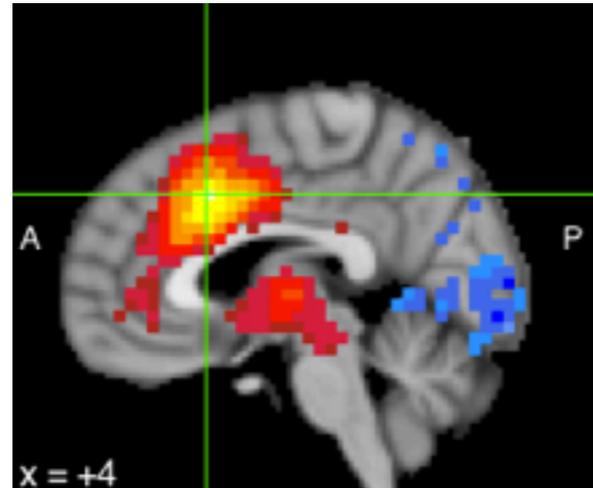
~~ACC is necessary or sufficient for working memory~~

# What does the ACC do?

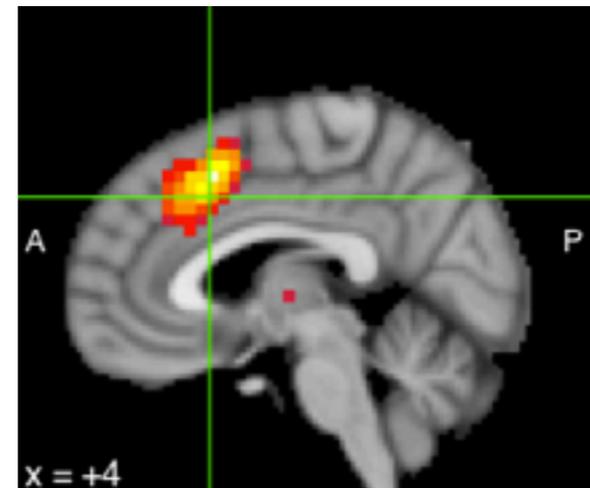
maintenance



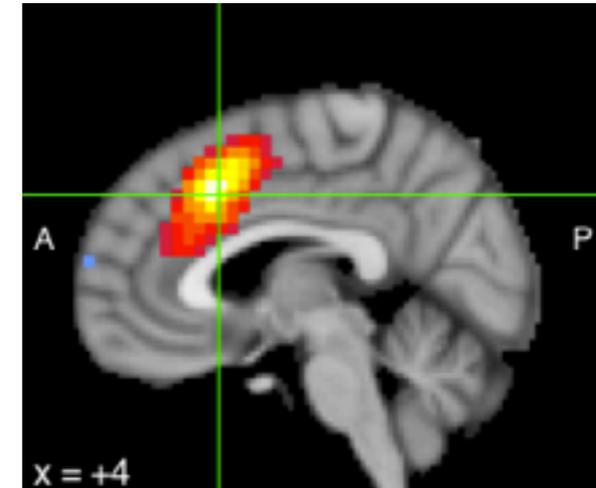
pain



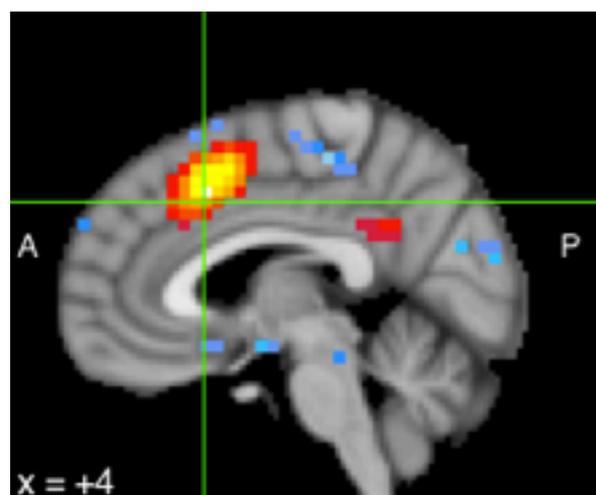
phonology



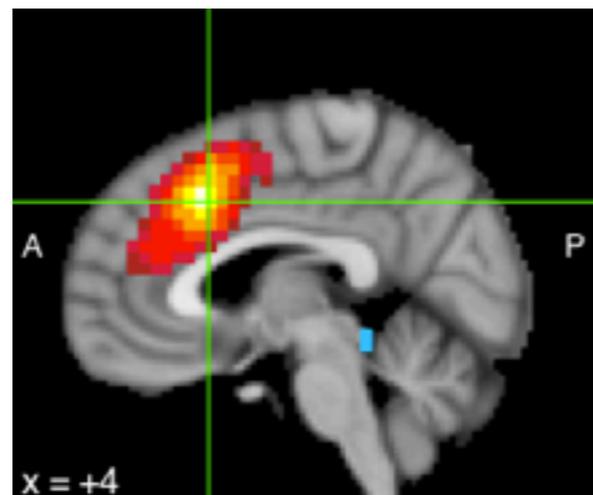
interference



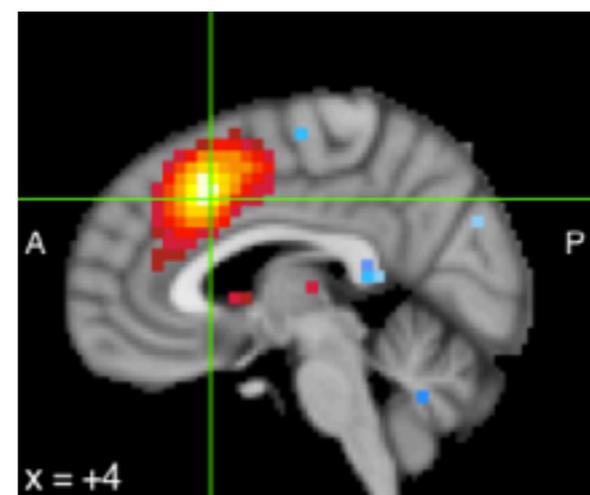
difficulty



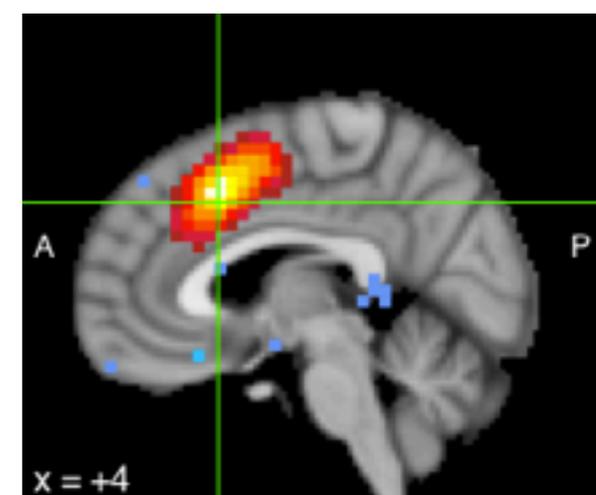
conflict



errors



attention

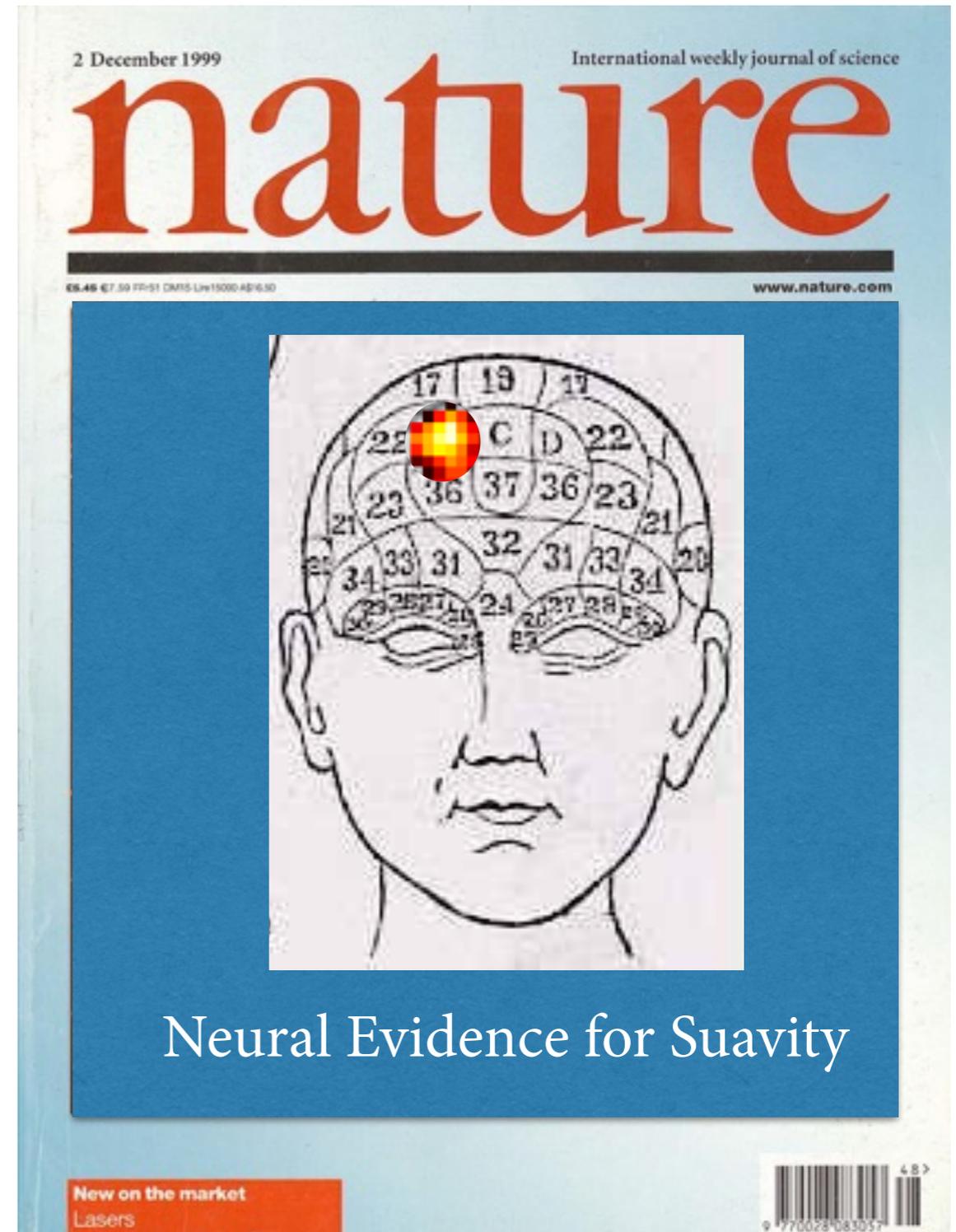
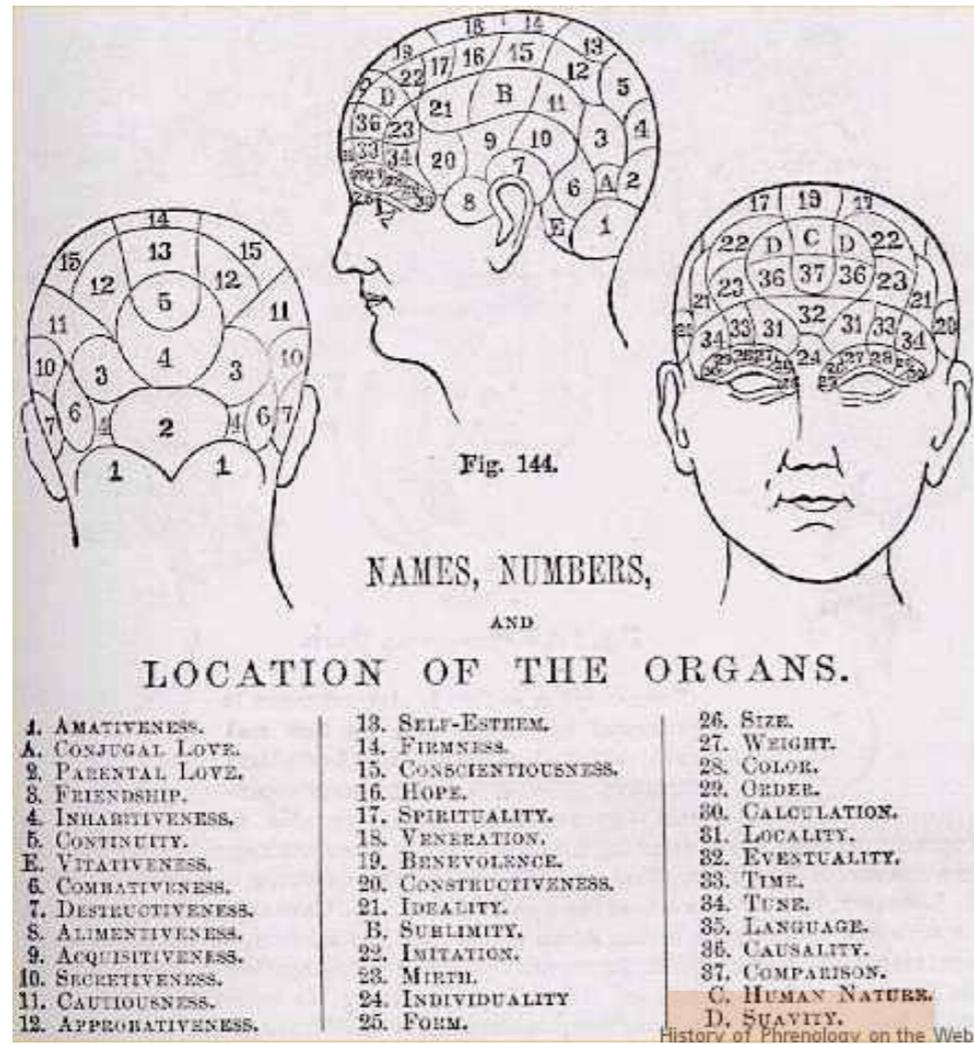


forward inference Z estimated using neurosynth.org

# Some alternatives

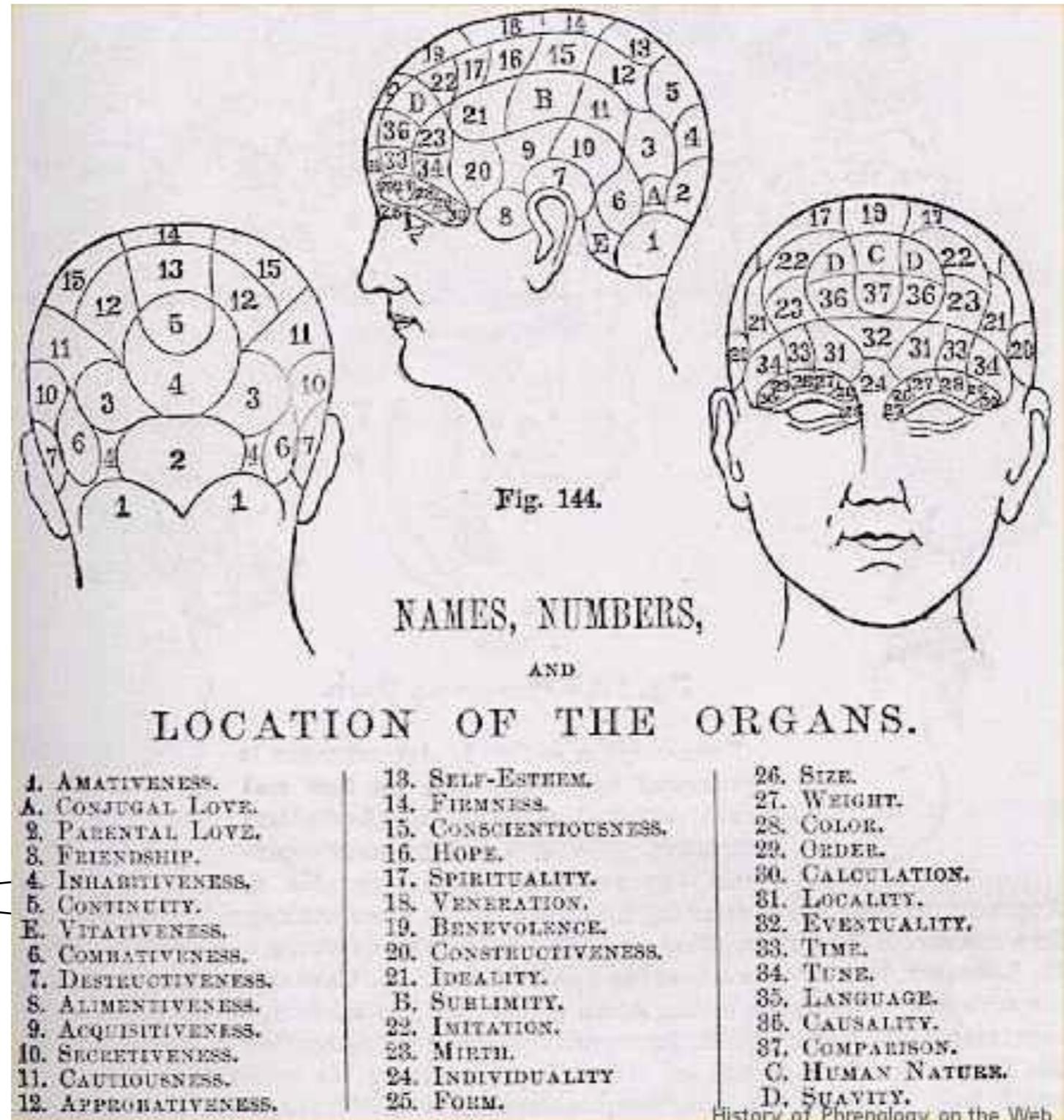
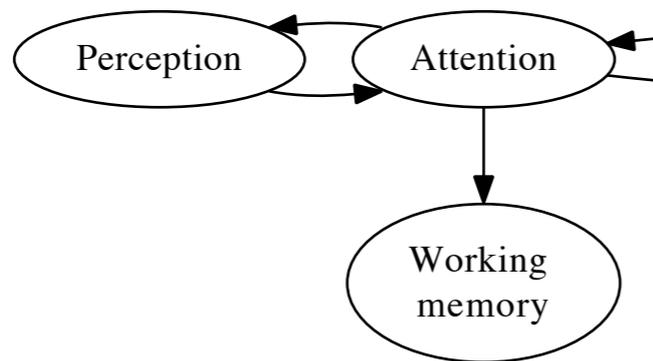
- There is some confound driving all of these (such as autonomic arousal or breathing)
- These are all truly distinct functions performed by subsets of neurons in the ACC
- These are all truly distinct functions subserved by ACC in different neural contexts
- These are not truly distinct functions
  - We are chopping up mental function in the wrong way
  - Thought experiment: What if the phrenologists had fMRI?

# What if the phrenologists had fMRI?



# Mapping cognition

- What are the atoms of the mind?
- How are they related to one another?



Habit  
arning

llection



a collaborative knowledge base characterizing the state of current thought in Cognitive Science.

CONCEPTS / 695

TASKS / 512

DISORDERS / 214

COLLECTIONS / 20

ABOUT

BLOG

# Welcome to Cognitive Atlas

The Cognitive Atlas is a collaborative knowledge building project that aims to develop a knowledge base (or ontology) that characterizes the state of current thought in cognitive science. The project is led by Russell Poldrack, Professor of Psychology at Stanford University. Development of the project was supported by grant RO1MH082795 from the National Institute of Mental Health.

## Sign In

Registered users may edit and contribute to the Cognitive Atlas

your email address

\*\*\*\*\*

Keep me logged in

SIGN IN

[Forgotten password? »](#)  
[Request a contributor account »](#)

Recently updated mental

### CONCEPTS

- *defiance*
- *irritability*
- *reward valuation*
- *default mode network*
- *defensive aggression*
- *Reception of non-facial communication*
- *Semantic network*
- *lethargy*

BROWSE ALL 695 CONCEPTS

Recently updated experimental

### TASKS

- *Children's Psychiatric Rating Scale*
- *Aberrant Behavior Checklist - Community*
- *Differential Ability Scales*
- *Beery-Buktenica Developmental Test of Visual-Motor Integration*
- *Conners 3rd Edition*
- *Social Communication Questionnaire*
- *Kaufman Brief Intelligence Test*

BROWSE ALL 512 TASKS

Recently updated **DISORDERS**

- *central sleep apnea*
- *depersonalization disorder*
- *seasonal affective disorder*
- *amnesic disorder*
- *specific developmental disorder*
- *agoraphobia*
- *amusia*
- *avoidant personality disorder*

BROWSE ALL 214 DISORDERS

Recently updated **COLLECTIONS**

- *Computerized Neurocognitive Battery ((CNB), Penn Neuropsych Battery)*
- *NIH Toolbox Sensation and Pain Battery*
- *NIH Toolbox Motor Battery*
- *Baddeley's model of working memory*
- *RDoC Working Memory Matrix*
- *RDoc Social Processes Matrix*

BROWSE ALL 20 COLLECTIONS

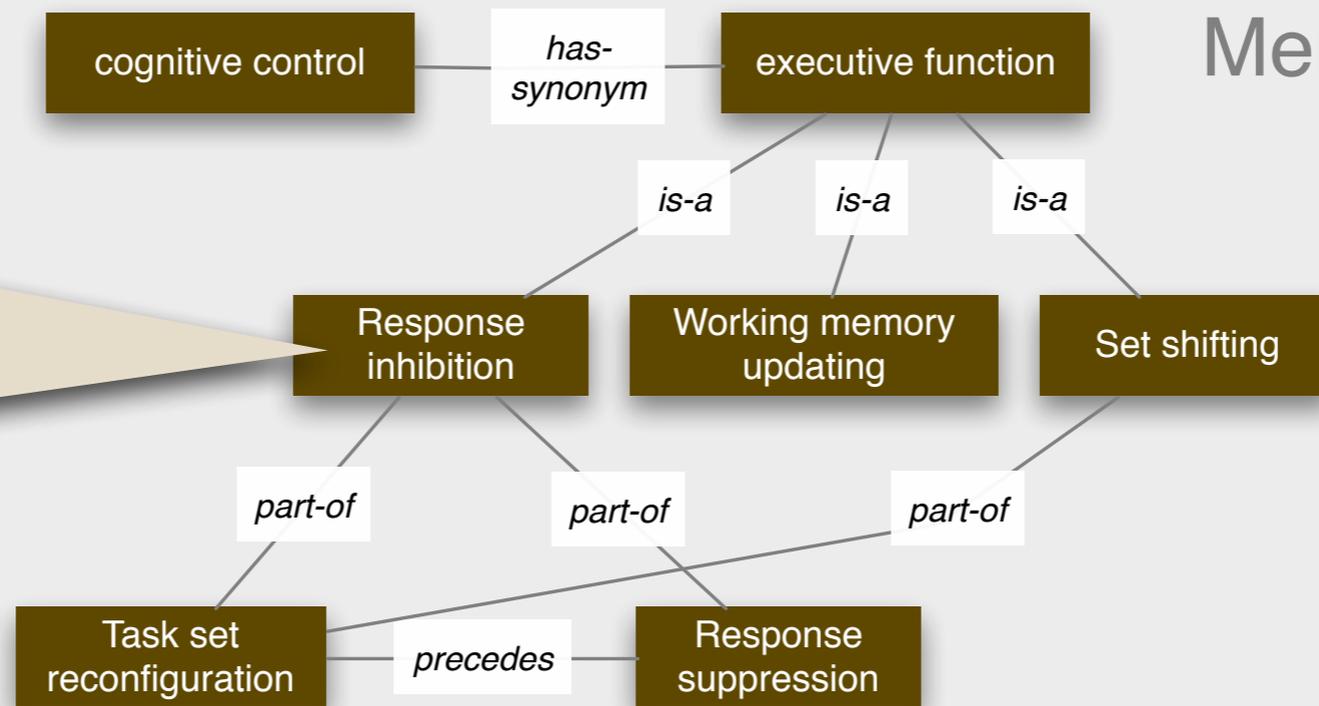
# Mental Concepts

## Response inhibition

Suppression of actions that are inappropriate in a given context and that interfere with goal-driven behavior.

### Bibliography

Logan, G. D. & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295-327.



*is-measured-by*

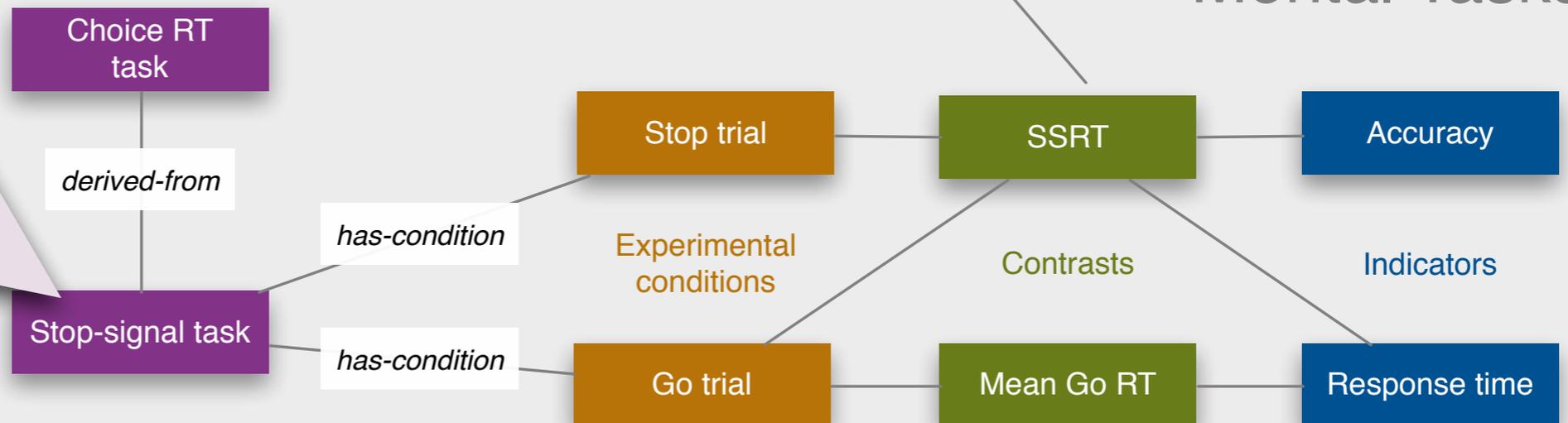
# Mental Tasks

## Stop-signal task

A task in which an external stimulus signals the participant to interrupt an already-initiated motor response.

### Bibliography

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, 12, 418-424.





a collaborative knowledge base characterizing the state of current thought in Cognitive Science.

CONCEPTS / 695

TASKS / 512

DISORDERS / 214

COLLECTIONS / 20

ABOUT

BLOG

# working memory

CONCEPT

active maintenance and flexible updating of goal/task relevant information (items, goals, strategies, etc.) in a form that resists interference but has limited capacity. These representations may involve flexible binding of representations, may be characterized by the absence of external support for the internally maintained representations, and are frequently temporary due to ongoing interference

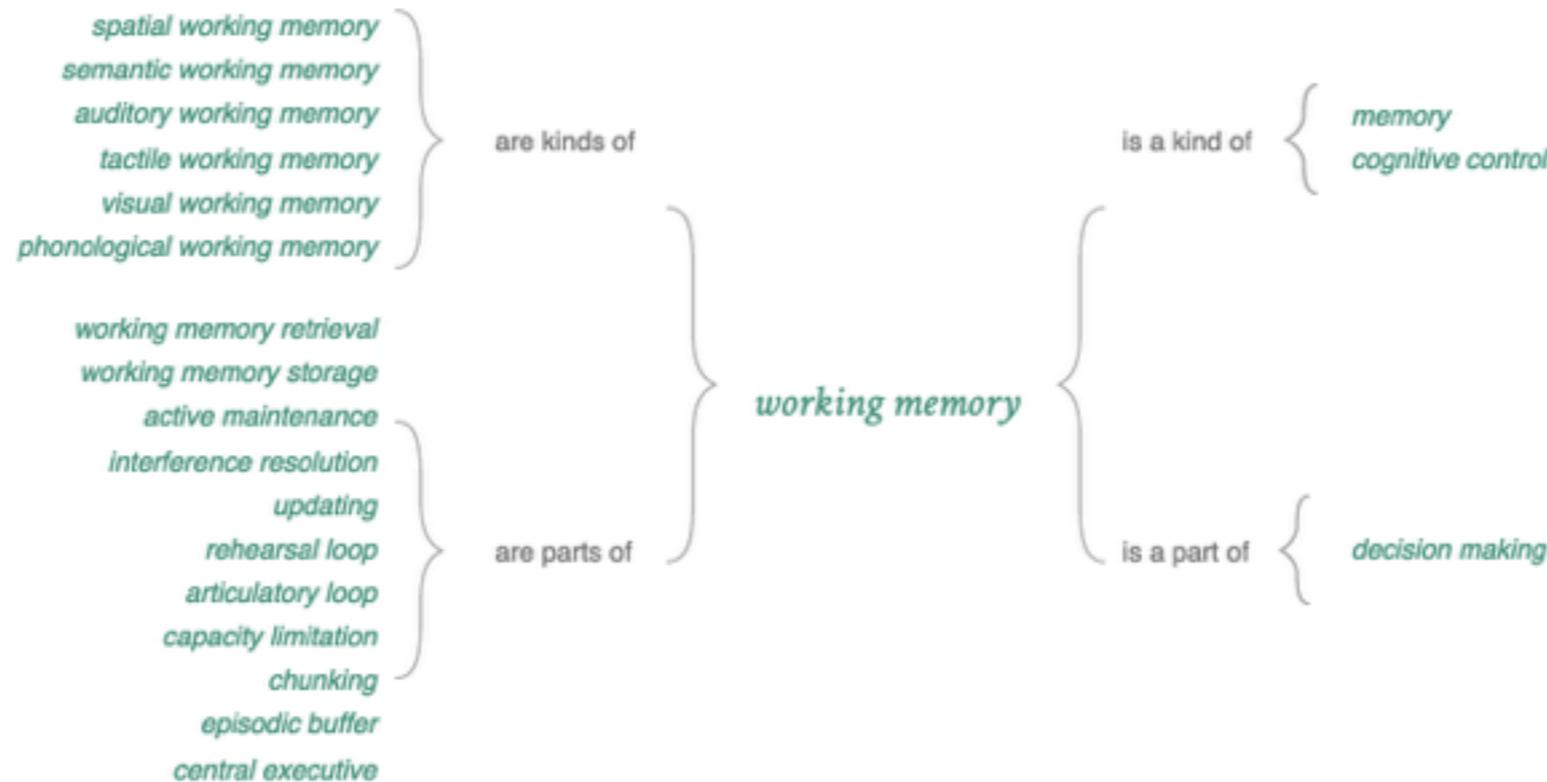
REFERENCED IN COLLECTIONS: [Baddeley's model of working memory](#) - [RDoC Working Memory Matrix](#)

Classified under [Executive/Cognitive Control](#)  
 Definition contributed by [RPoldrack](#) about three years ago.

## Asserted RELATIONSHIPS to other concepts

ATLAS VIEW

LIST VIEW



# Current state of the Cognitive Atlas

- 771 mental constructs
- 610 tasks
- 214 disorders (inherited from Disease Ontology)
- 22 collections
- Formal ontology (OWL) available via BioPortal

# Annotating data using the Cognitive Atlas

 **OpenfMRI**

[Home](#) [View Data Sets](#) [Add a Dataset](#) [FAQs](#) [Contact Us](#)

## User login

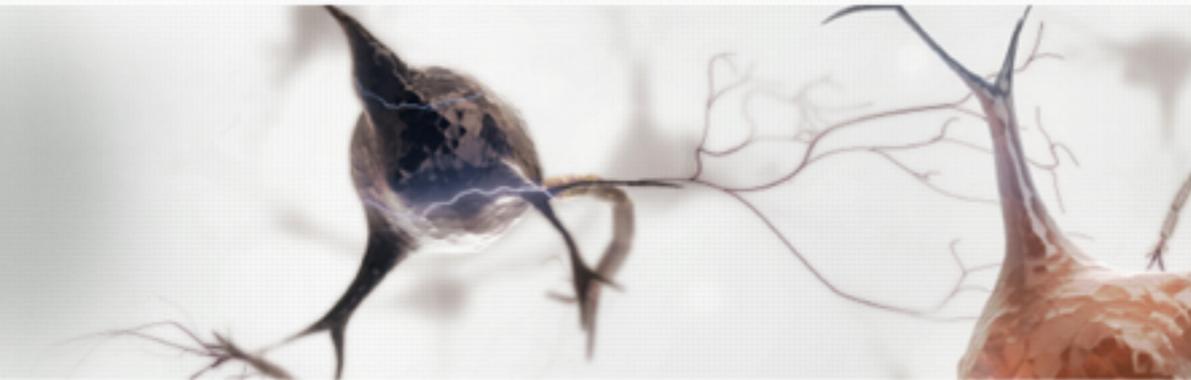
**LOG IN**

- [Create new account](#)
- [Request new password](#)

## Freedom to Share

OpenfMRI.org is a project dedicated to the free and open sharing of functional magnetic resonance imaging (fMRI) datasets, including raw data.

**Number of currently available datasets: 37**  
**Number of subjects across all datasets: 1411**



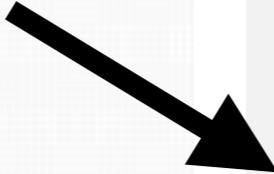
## Classification learning

Submitted by picchetti on Thu, 10/06/2011 - 11:36

Subjects performed a classification learning task with two different problems (across different runs), using a "weather prediction" task. In one (probabilistic) problem, the labels were probabilistically related to each set of cards. In another (deterministic) problem, the labels were deterministically related to each set of cards. After learning, subjects participated in an event-related block of judgment only (no feedback) in which they were presented with stimuli from both of the training problems.

### Tasks and Conditions:

- 001 [Probabilistic classification task](#)
  - 001 Probabilistic classification trials
  - 002 feedback
- 002 [deterministic classification](#)
  - 001 Deterministic classification trials
  - 002 feedback
- 003 [classification probe without feedback](#)
  - 001 Classification trials: Probabilistic
  - 002 Classification trials: Deterministic



## Probabilistic classification task TASK

Subjects are presented with a set of stimuli and must classify those stimuli into one of two categories. In a common version known as the "weather prediction task" the stimuli are cards with geometric shapes on them and the outcomes are rainy versus sunny weather. The feedback is probabilistic, and performance is measured by the proportion of statistically optimal responses.  
 Synonyms: *probabilistic classification learning task, weather prediction task*

Definition contributed by [RPoldrack](#) about two years ago  
 No relations have yet been associated.

**Probabilistic classification task** has been asserted to measure the following **CONCEPTS**

as measured by the contrast: **proportion of correct responses across all trials minus proportion correct on early trials**

**DISORDERS** associated with *Probabilistic classification task*  
 No associations have been added.

**IMPLEMENTATIONS** of *Probabilistic classification task*  
 No implementations have been added.

**EXTERNAL DATASETS** for *Probabilistic classification task*

- [Dataset #1](#) 🔗 Classification learning
- [Dataset #2](#) 🔗 Classification learning and stop-signal (1 year test-retest)

**CONDITIONS**

*Probabilistic classification trials (Trials on which the subject classifies items) feedback (presentation of feedback following response)*

*Probabilistic classification trials: Positive feedback (Trials on which subject classifies item and receives positive feedback)*

*Probabilistic classification trials: Negative feedback (Trials on which subject classifies item and receives negative feedback)*

**CONTRASTS**

*Probabilistic classification trials proportion of correct responses across all trials minus proportion correct on early trials*

*accuracy of participant minus average accuracy of controls*

In the Cognitive Atlas, we define a contrast as any function over experimental conditions. The simplest contrast is the indicator value for a specific condition; more complex contrasts include linear or nonlinear functions of the indicator across different experimental conditions.

**INDICATORS**

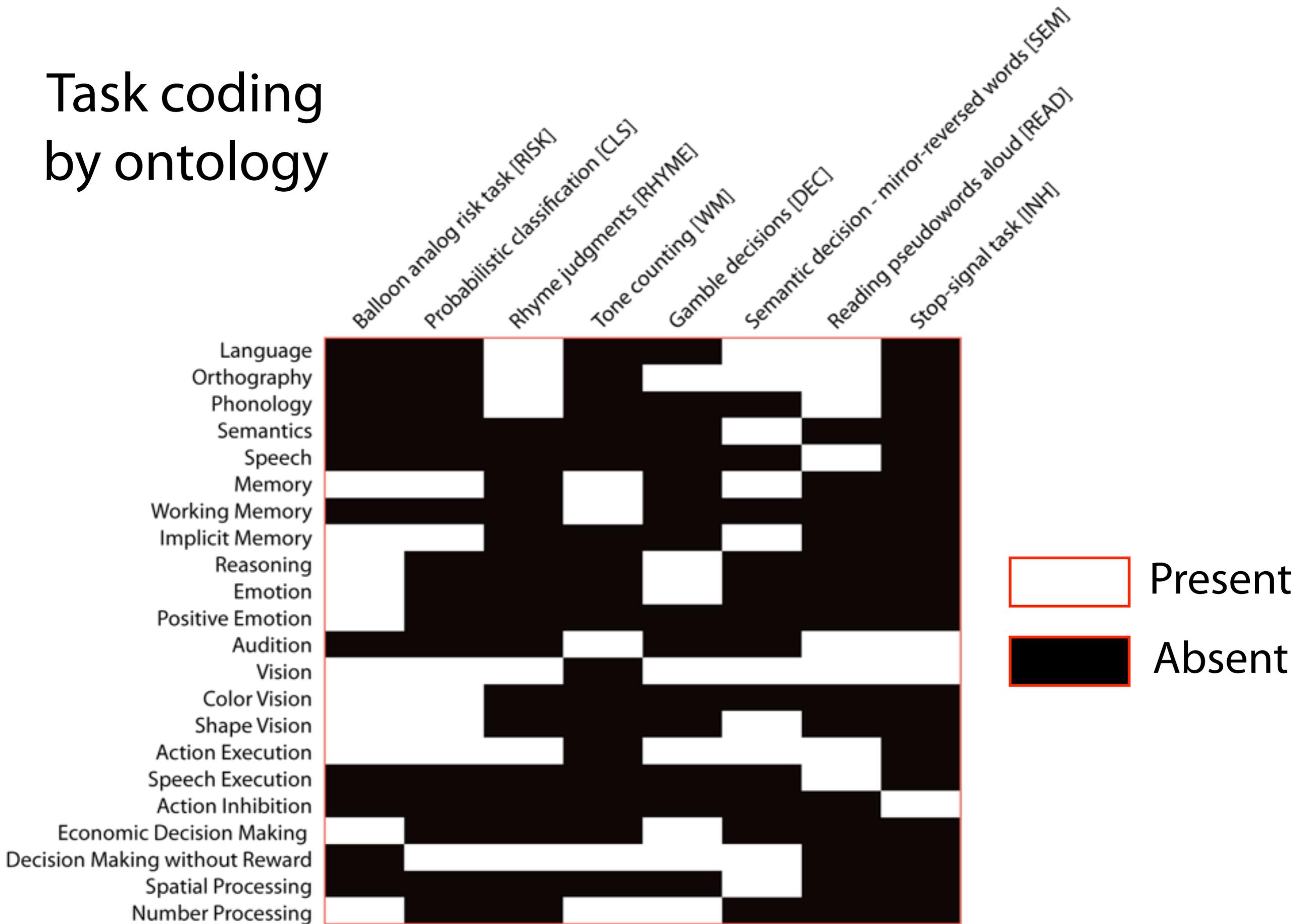
*Proportion optimal responses response time*

An indicator is a specific quantitative or qualitative variable that is recorded for analysis. These may include behavioral variables (such as response time, accuracy, or other measures of performance) or physiological variables (including genetics, psychophysiology, or brain imaging data).

# An initial proof of concept

- Obtain brain imaging data from a broad range of mental tasks
- In this case, 130 people doing one of 8 different tasks
- Code the the tasks using a preliminary cognitive ontology
- Map the brain systems onto the ontology

# Task coding by ontology

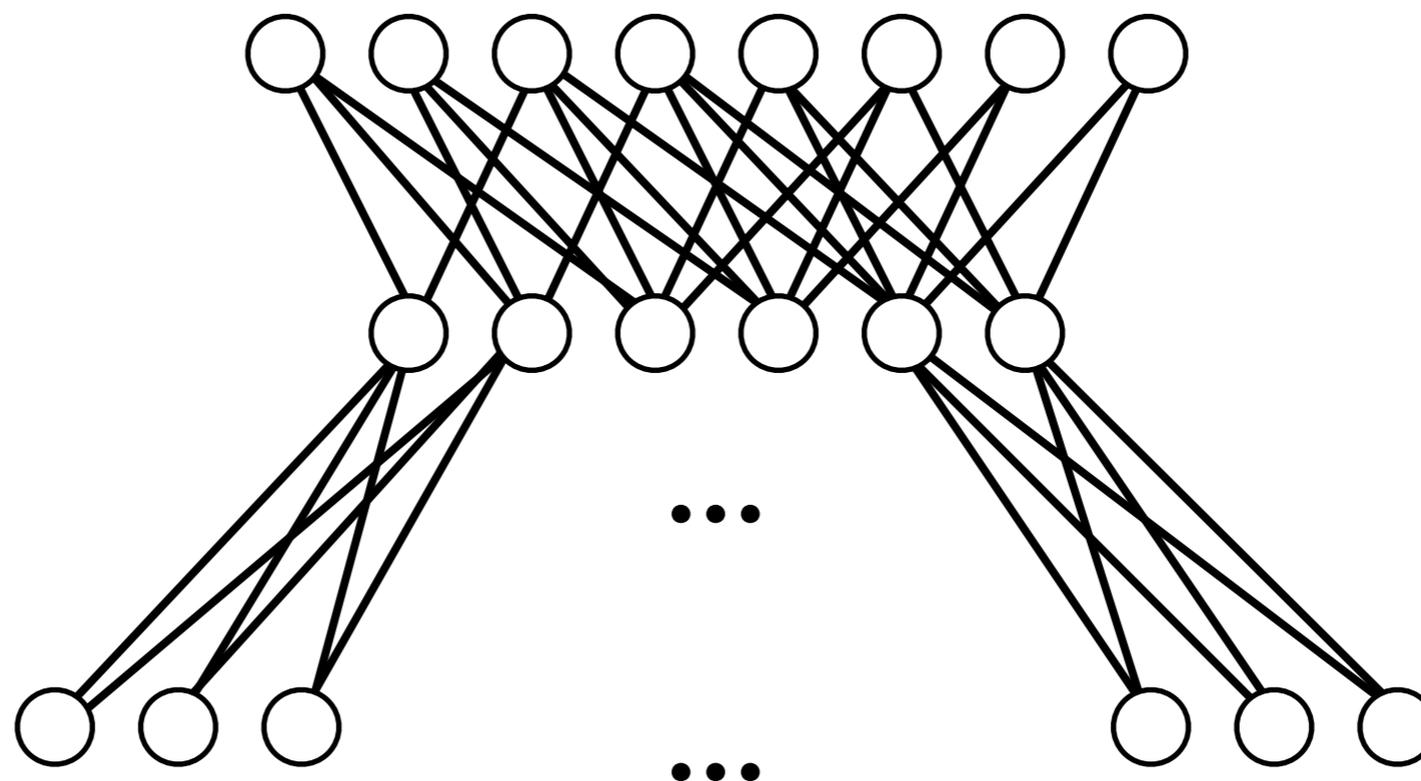


# Artificial neural network classifier

Output: Which of the  
8 tasks was the  
person performing?

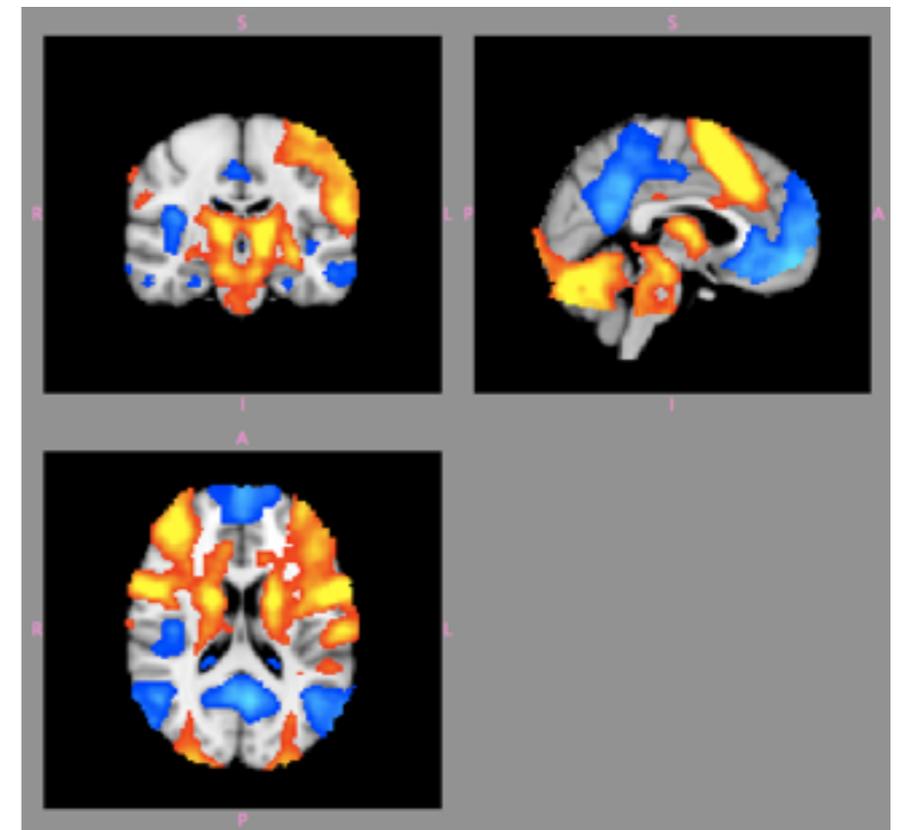
6 hidden units

Input: brain  
activity at ~2000  
locations



Use hidden unit patterns  
as low-dimensional  
representation  
for each subject

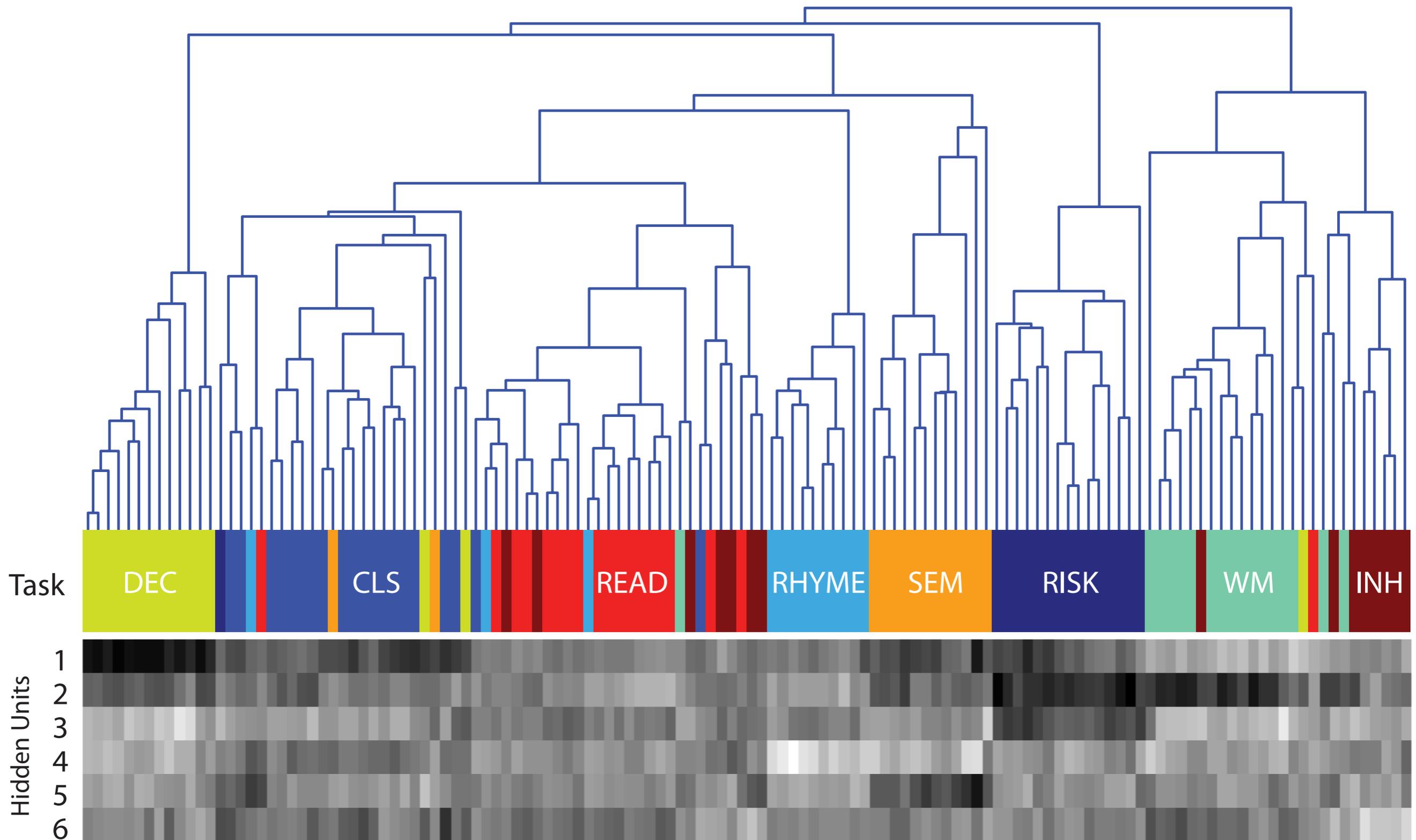
Analysis	Crossvalidated accuracy	# of voxels included
Union of all in-mask voxels across subjects (one-vs-one)	74%	417,231
Intersection of in-mask voxels across subjects (one-vs-many)	80.8%	214,940
Positively activated voxels only (across all 130 subjects, $t > 3$ , $p < .002$ ) (one-vs-many)	74.6%	83,825
Deactivated voxels only ( $t < -3$ , $p < .002$ ) (one-vs-many)	50.8%	23,736



Accuracy above 18.5% is significant at  $p < .05$  by randomization

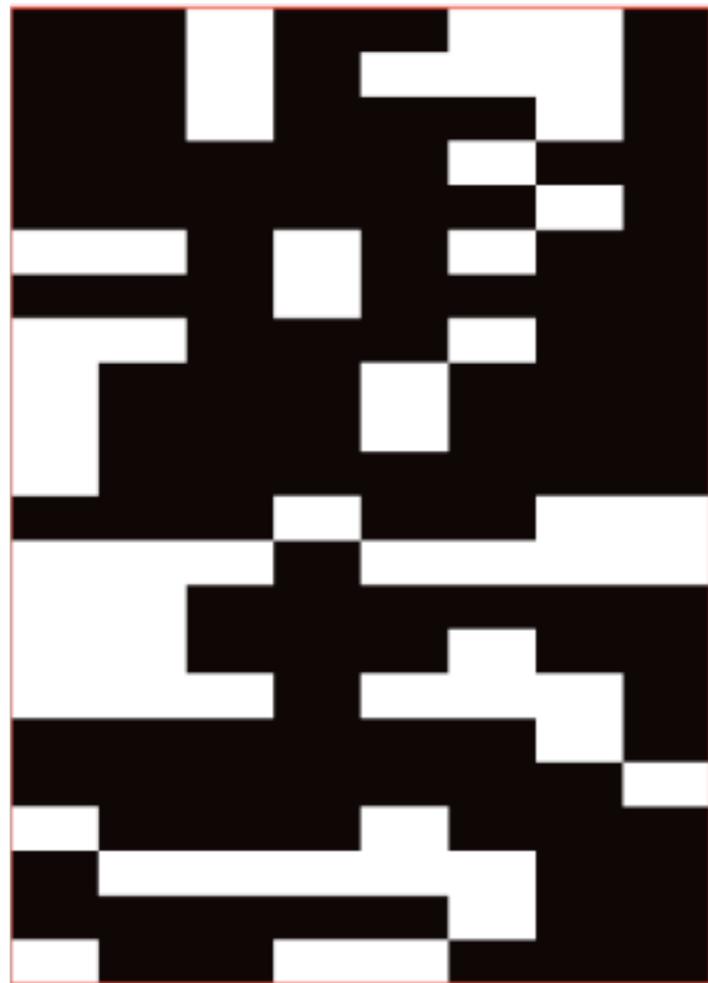
### Task chosen by classifier

	<i>Task 1</i>	<i>Task 2</i>	<i>Task 3</i>	<i>Task 4</i>	<i>Task 5</i>	<i>Task 6</i>	<i>Task 7</i>	<i>Task 8</i>
<i>Task 1</i>	<b>87.5</b>	6.0	0.0	0.0	6.0	0.0	0.0	0.0
<i>Task 2</i>	0.0	<b>90.0</b>	0.0	0.0	0.0	0.0	5.0	5.0
<i>Task 3</i>	8.0	23.0	<b>61.5</b>	0.0	0.0	8.0	0.0	0.0
<i>Task 4</i>	0.0	0.0	0.0	<b>82.4</b>	0.0	0.0	0.0	18.0
<i>Task 5</i>	0.0	38.0	0.0	0.0	<b>43.8</b>	18.2	0.0	0.0
<i>Task 6</i>	0.0	28.0	0.0	0.0	0.0	<b>71.4</b>	0.0	0.0
<i>Task 7</i>	0.0	11.0	0.0	0.0	0.0	0.0	<b>84.0</b>	5.0
<i>Task 8</i>	0.0	0.0	7.0	0.0	0.0	0.0	27.0	<b>63.0</b>



Hierarchical clustering on hidden unit values

# Mapping neural data into a cognitive ontology



Concept X task

**X**



Task X dimension

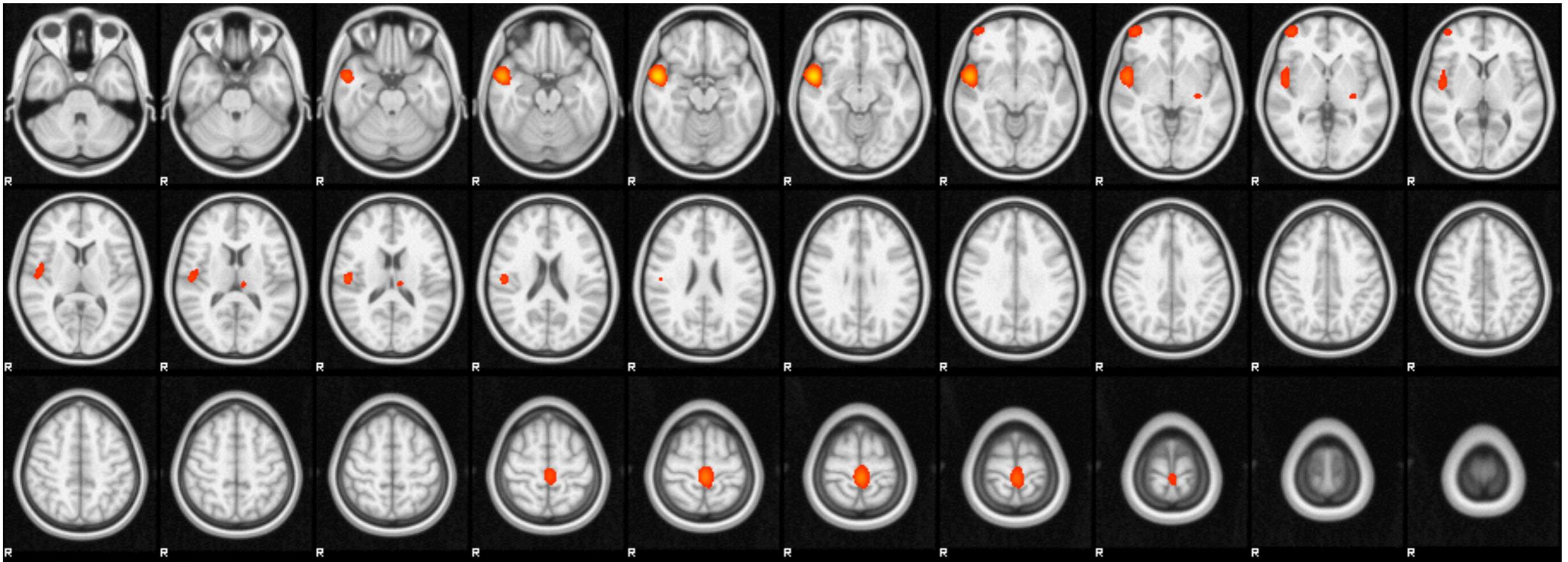
**=**



Concept X dimension

Poldrack, Halchenko, & Hanson, 2009

actionexecution **audition** colorvision decisionmaking emotion implicitmemory  
 language memory numericprocessing orthography **phonology**  
 positiveemotion reasoning **responseinhibition**  
 reward semantics shapevision spatialprocessing **speech** vision  
**workingmemory**

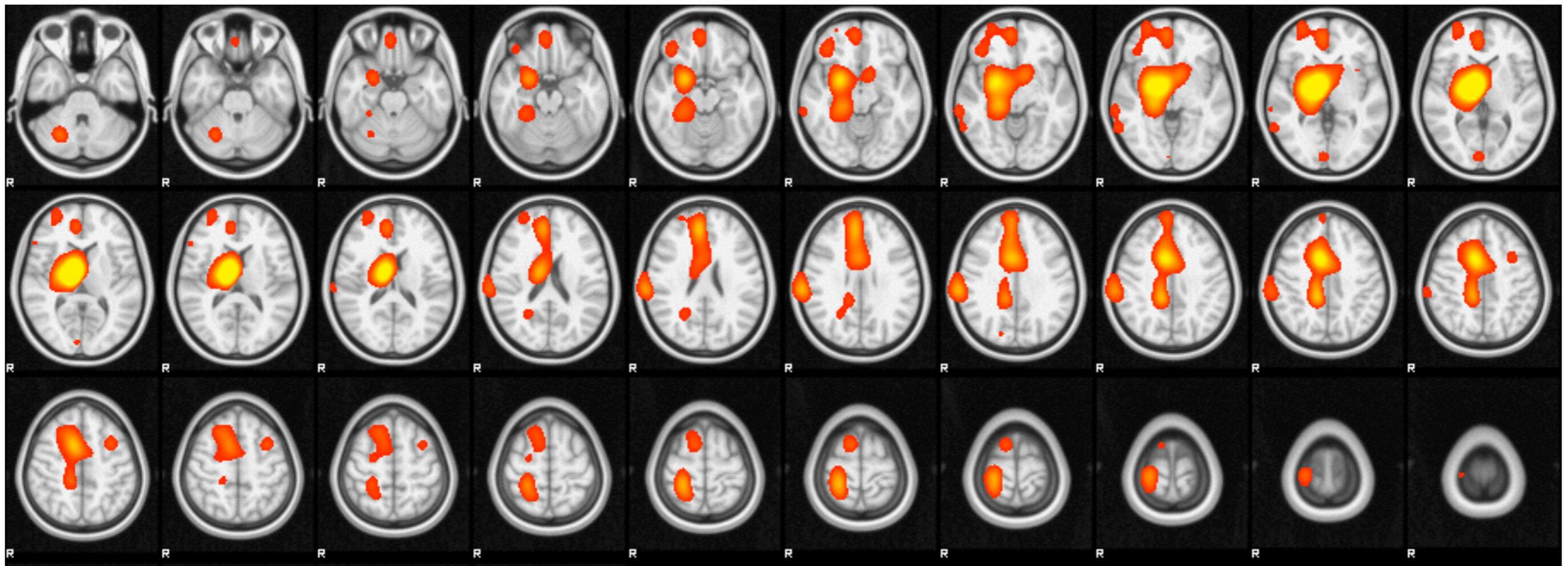


# decisionmaking

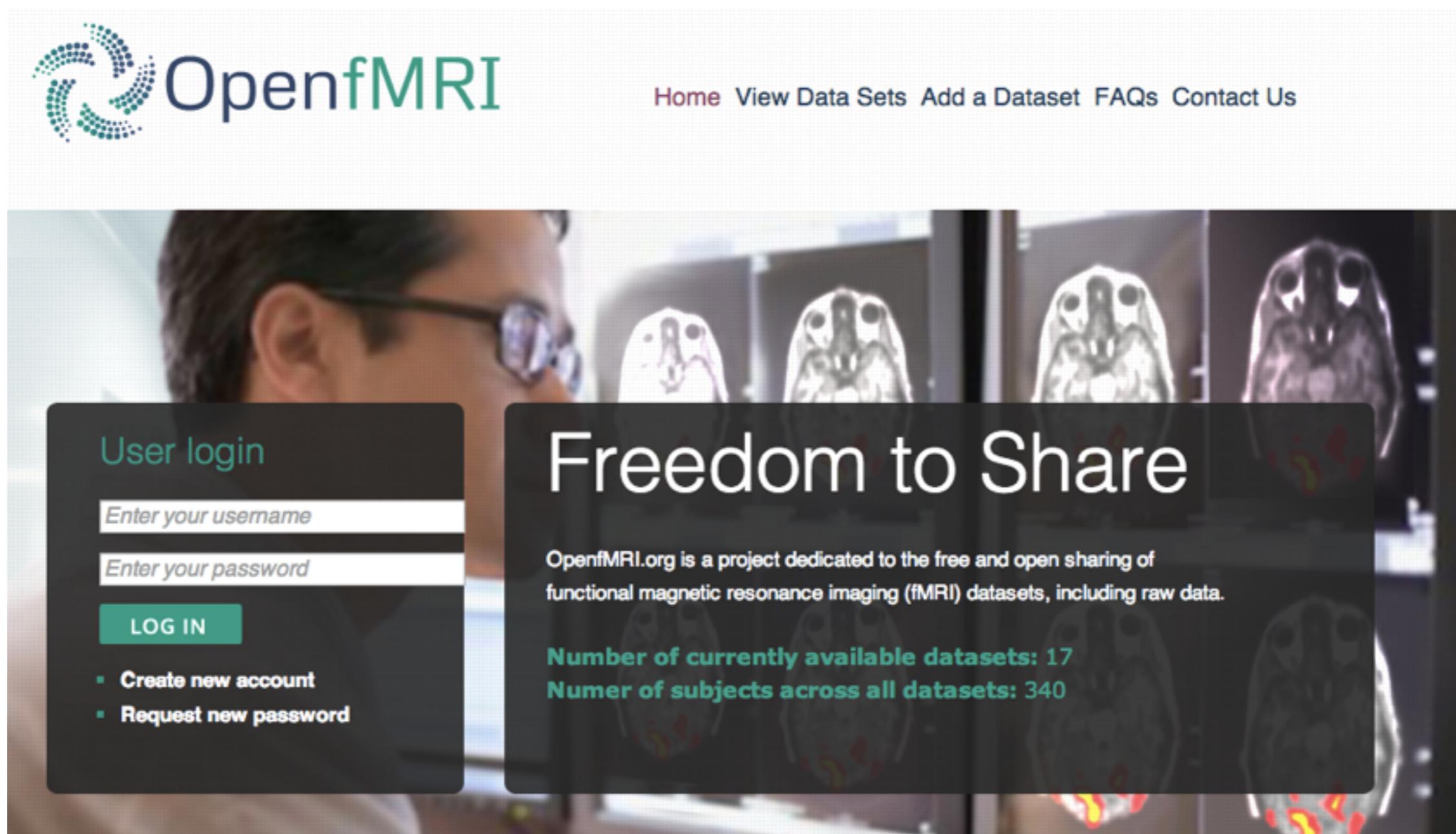
actionexecution audition colorvision

emotion implicitmemory language **memory** numericprocessing orthography phonology positiveemotion

reasoning responseinhibition reward semantics shapevision spatialprocessing **speech vision** workingmemory



# Larger-scale decoding

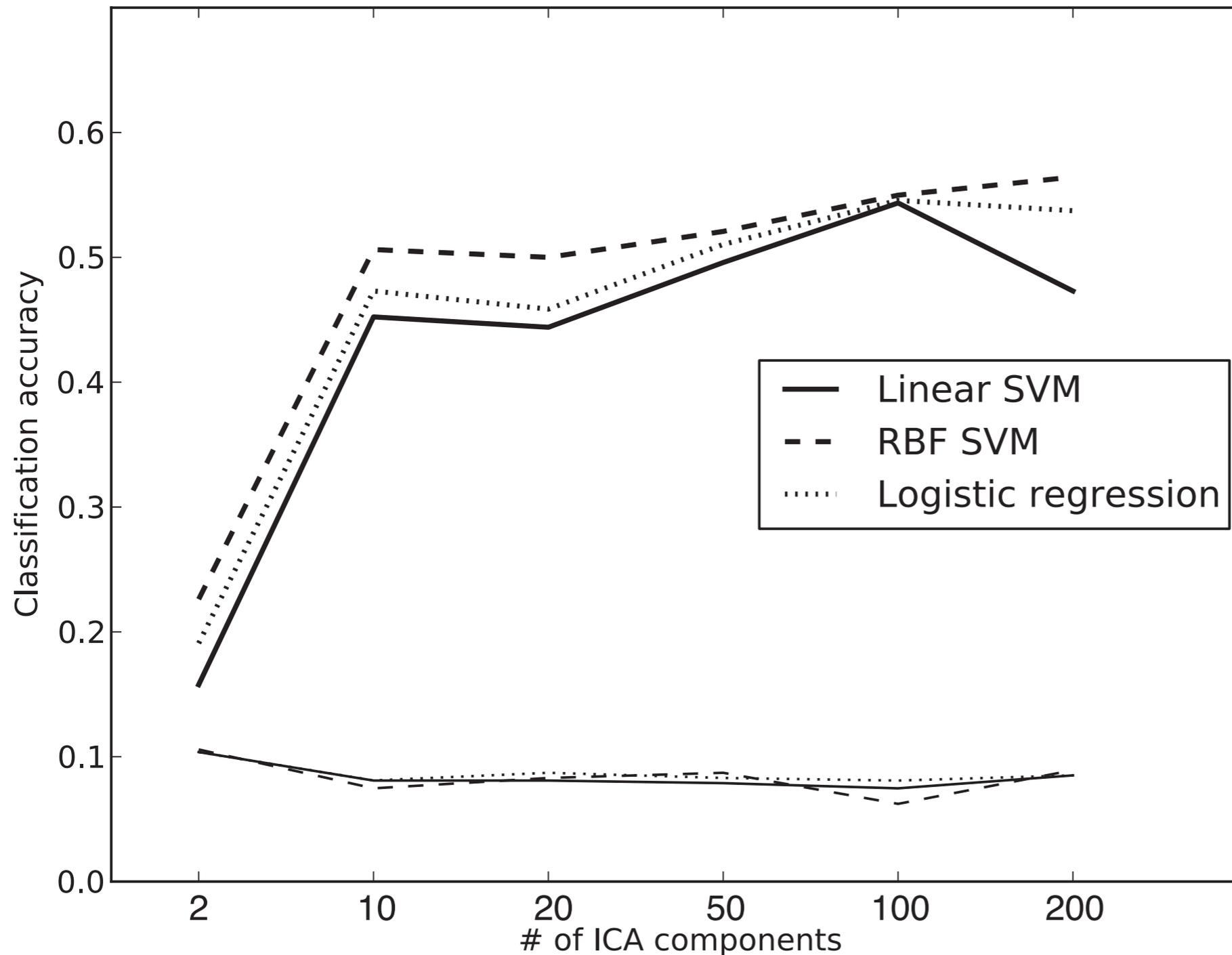


The screenshot shows the OpenfMRI website interface. At the top left is the OpenfMRI logo, and to its right is a navigation menu with links for Home, View Data Sets, Add a Dataset, FAQs, and Contact Us. Below the navigation is a large banner image of a person looking at MRI brain scans. Overlaid on the left is a 'User login' form with fields for 'Enter your username' and 'Enter your password', a 'LOG IN' button, and links for 'Create new account' and 'Request new password'. Overlaid on the right is a 'Freedom to Share' banner with the text: 'OpenfMRI.org is a project dedicated to the free and open sharing of functional magnetic resonance imaging (fMRI) datasets, including raw data.' Below this text are two statistics: 'Number of currently available datasets: 17' and 'Number of subjects across all datasets: 340'.

26 tasks, 482 images from 338 subjects

# Larger-scale decoding

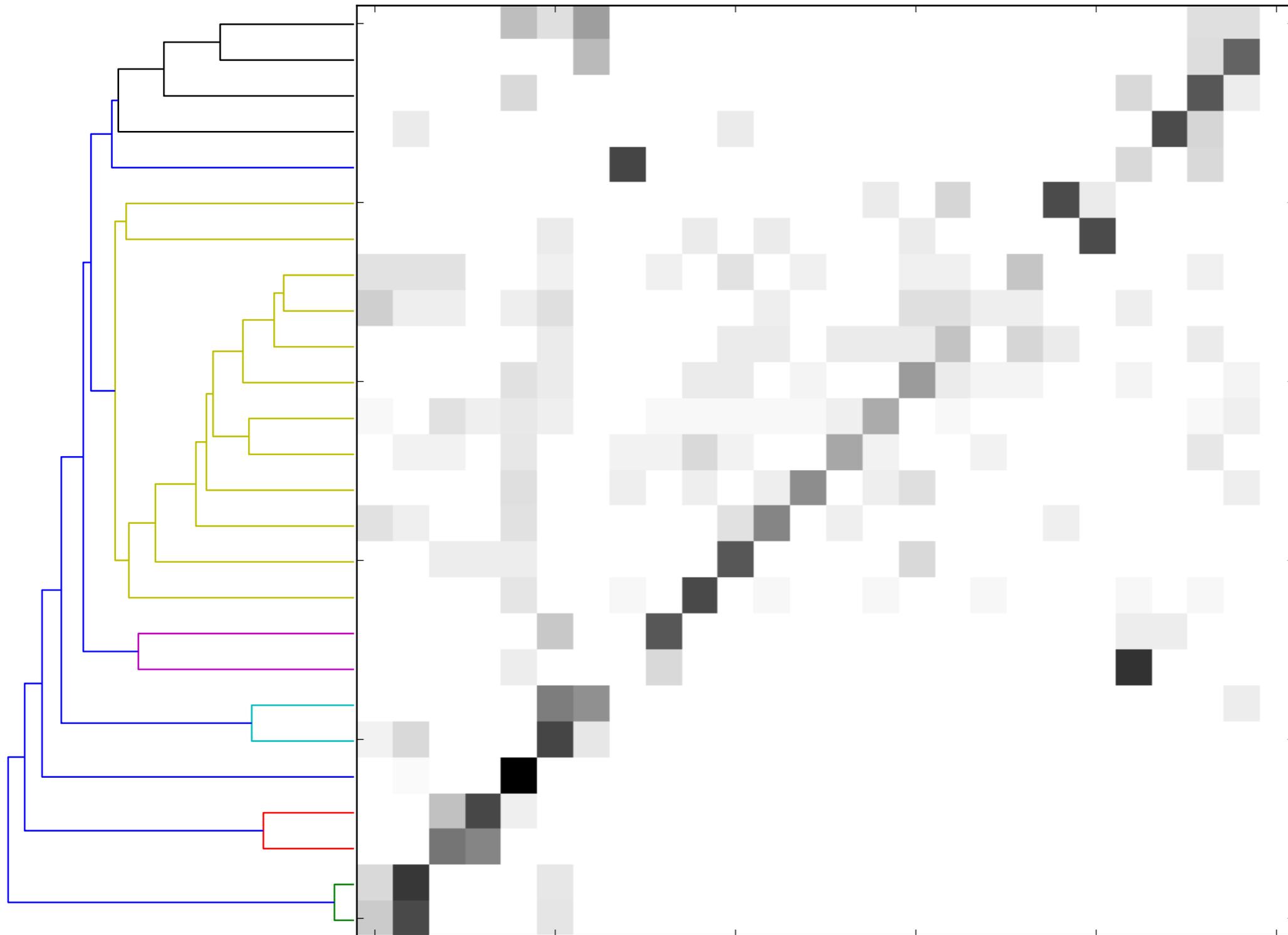
26 tasks, 482 images from 338 subjects



**Whole-brain:  
47% accuracy**

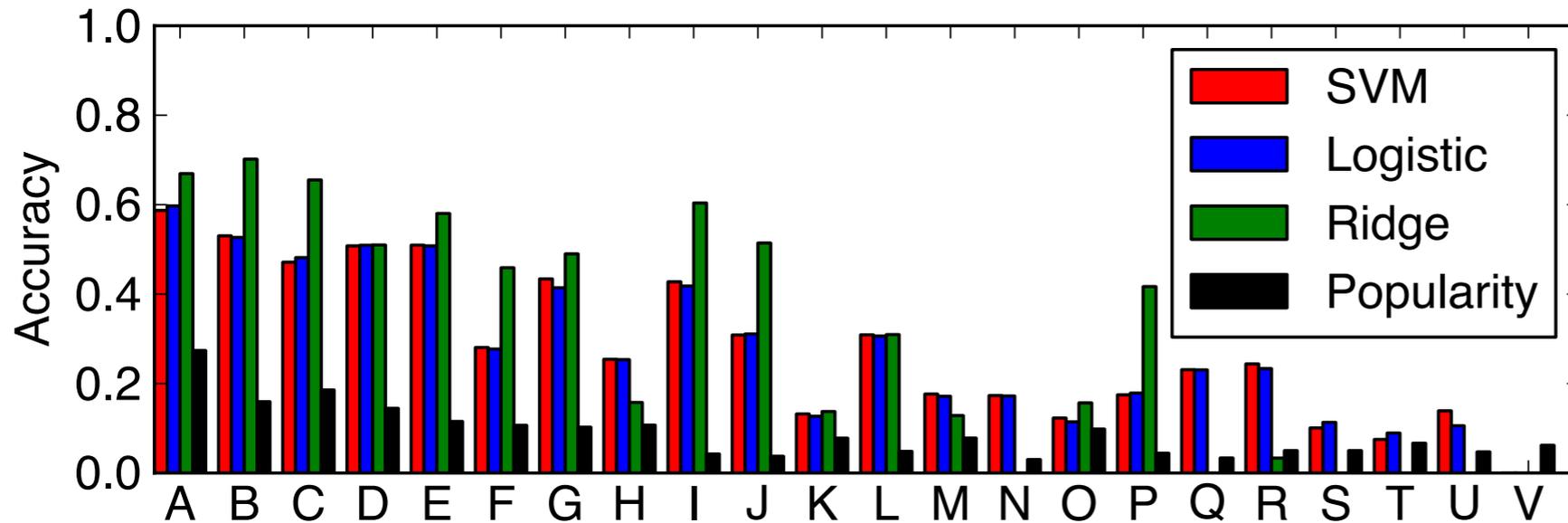
Poldrack et al., submitted

# Larger-scale decoding: Clustering



- ds017A (2): Conditional stop signal: go
- ds008 (1): Stop signal: successful stop
- ds011 (1): Tone counting
- ds003 (1): Rhyme judgment
- ds011 (3): Classification: dual-task
- ds052 (2): Classification: negative feedback
- ds052 (1): Classification: positive feedback
- ds110 (1): Memory encoding: subsequent
- ds005 (1): Gamble decisions: parameter
- ds051 (1): Abstract/concrete decisions:
- ds102 (1): Flanker task: incongruent vs
- ds108 (1): Emotion regulation: Regulation
- ds101 (1): Simon task: incorrect vs correct
- ds001 (1): BART: pumps vs. control (de
- ds002 (2): Classification: feedback
- ds006A (1): Mirror-reading: mirror vs. p
- ds109 (1): False belief task: false belief
- ds011 (4): Classification decision (no fe
- ds011 (2): Classification : single-task
- ds008 (2): Conditional stop signal task:
- ds007 (1): Stop signal task: go
- ds107 (1): One-back: objects vs scrambled
- ds002 (3): Classification decision (no fe
- ds002 (1): Classification: single-task
- ds007 (3): Stop signal task: pseudowor
- ds007 (2): Stop signal task: letter namin

# Decoding cognitive functions across subjects



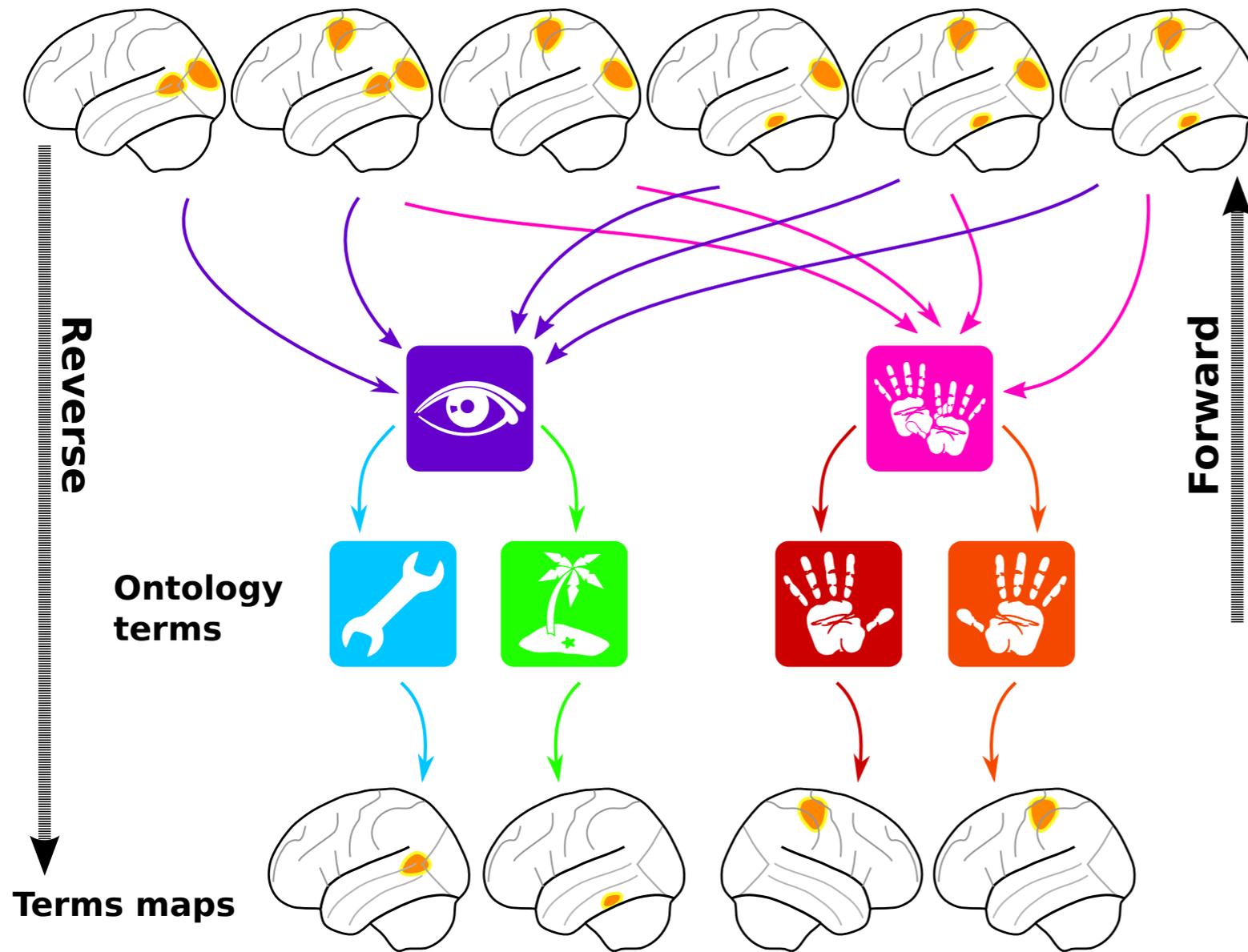
- A Vision
- B Action Execution
- C Decision Making
- D Orthography
- E Shape Vision
- F Audition
- G Phonology
- H Conflict
- I Semantics
- J Reinforcement Learning
- K Working Memory
- L Feedback
- M Response Inhibition
- N Reward
- O Stimulus-driven Attention
- P Speech
- Q Emotion Regulation
- R Mentalizing
- S Punishment
- T Error Processing
- U Memory Encoding
- V Spatial Attention

- Multilabel classifier trained using OpenfMRI data and Cognitive Atlas labels
- 26 task contrasts, 482 images from 338 subjects
- annotated with Cognitive Atlas

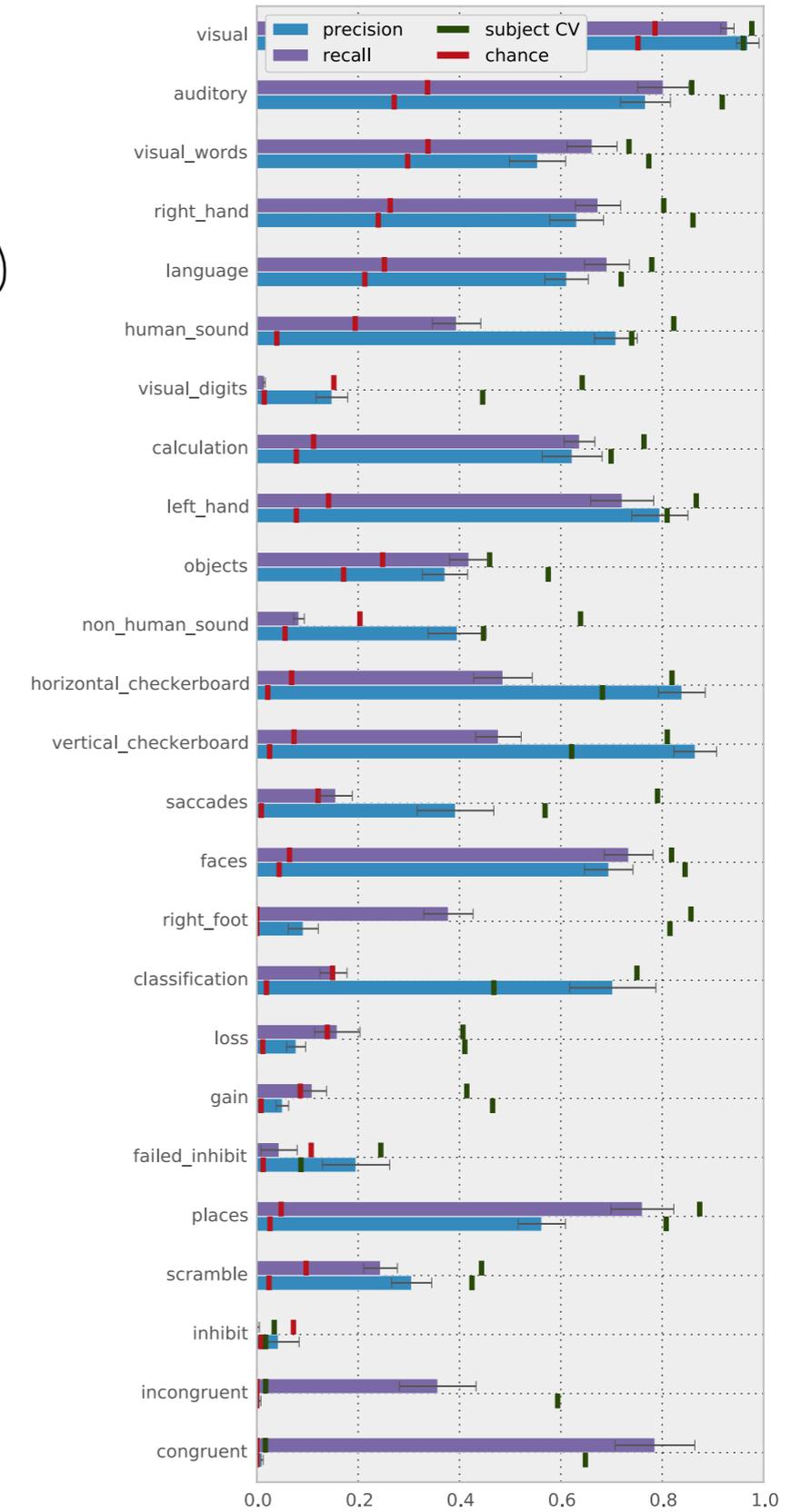
Koyejo & Poldrack, 2013

# Ontology-based decoding

## Experimental conditions



Schwartz et al., in prep



# Mining text using topic modeling

Terms

“decision”, “value”,  
“choice”, “risk”

“activation”, “scan”,  
“TR”, “EPI”

“nucleus accumbens”,  
“striatum”, “dopamine”

Topics

decision making

fMRI

basal ganglia

Documents

## Neural computations underlying action-based decision making in the human brain

Klaus Wunderlich<sup>1,2</sup>, Antonio Rangel<sup>1,3</sup>, and John P. O’Doherty<sup>1,4,5</sup>

<sup>1</sup>Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA, <sup>2</sup>Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, and <sup>3</sup>Institute of Neuroscience and School of Psychology, Trinity College, Dublin, Ireland

Edited by Raffaele Rumi, Universidad Nacional Autónoma de México, México, D.F., México, and approved August 6, 2009 (received for review February 4, 2009)

Action-based decision making involves choices between different physical actions to obtain rewards. To make such decisions the brain needs to assign a value to each action and then compare them to make a choice. Using fMRI in human subjects, we found evidence for action-value signals in supplementary motor cortex. Separate brain regions, most prominently ventromedial prefrontal cortex, were involved in encoding the expected value of the action that was ultimately taken. These findings differentiate two main forms of value signals in the human brain: those relating to the value of each available action, likely reflecting signals that are a precursor of choice, and those corresponding to the expected value of the action that is subsequently chosen, and therefore reflecting the consequence of the decision process. Furthermore, we also found signals in the dorsomedial frontal cortex that resemble the output of a decision comparator, which implicates this region in the computation of the decision itself.

ACC | action value | reinforcement learning | area | single

Consider a goalkeeper trying to stop a soccer ball during a crucial kick. Within a brief amount of time he needs to choose between jumping to the left or right goal posts. Repeated play against the same opponents allows him to learn about their scoring tendencies, which can be used to compute the values of a left and a right jump before making a decision. It is a long-established view in economics, psychology, and computational neuroscience that the brain makes choices among actions by first computing a value for each possible action, and then selecting one of them on the basis of those values (1–3). This raises two fundamental questions in decision neuroscience: (1) where in the brain are the values of different types of actions encoded? and (2) how and where does the brain compare those values to generate a choice?

An emerging view is that the brain compares values in a manner that organizes them to make a number of value-related computations to make even simple choices (4). Consider the case of actions that are exemplified by the goalkeeper’s problem. First, he needs to assign a value to each action under consideration. These signals, known as action values, encode the value of each action before choice and regardless of whether it is subsequently chosen or not, which allows them to serve as inputs into the decision-making process (5–7). Second, these action values are compared to generate a choice. Third, the value of the option that is selected, known as the chosen value, is tracked to be able to do reinforcement learning. In particular, by comparing the value of the outcome generated by the decision to the chosen value, the organism can compute a prediction-error signal that can be used to update the action value of the chosen option. Note that while the action values are computed before the decision is made, the chosen value and outcome of the comparator process signals are computed afterward.

Although a rapidly growing number of studies have found neural responses that are correlated with some form of value signals, little is known about how the brain encodes action values or about how it compares them. This central to understanding how the brain makes action-based choices. For example, a number of choice value signals have been found in the orbitofrontal and prefrontal cortex

(8, 9) and amygdala (10, 11). Note that these signals are quite distinct from action values, and are not precursors to choice, because they reflect the value of the actions that were selected in the decision. For similar reasons, the value signals that have been found in lateral intraparietal cortex (LIP) during saccadic action-based choices (12, 13) are also not pure action values since they are strongly modulated by whether an action is subsequently taken. This suggests that instead of serving as inputs to the comparison process, they reflect its output. Several studies found orbitofrontal cortex to encode the value of different goals (14–16). Although these signals are precursors of choice, they are not instances of action values since they are stimulus-based and independent of the action required to obtain them. To date, only three monkey electrophysiology studies have found evidence for the presence of action-value signals for hand and eye movements in the striatum during simple decision-making tasks (5–7). This study extends their findings in three directions. First, as of yet no evidence has been presented for the existence of action-value signals in the human brain. Second, using fMRI we are able to look for action-value signals in the entire brain, whereas the previous electrophysiology studies have limited their attention to the striatum. As a result, no previous study has looked for action-value signals in the cortex. This is important because, as discussed below, there are a priori reasons to believe that action value signals might be found in the motor and supplementary motor cortices. Finally, we investigate how such signals might be compared to actually compute the decision itself and where neuronal correlates of the output of this decision process are represented, an issue about which very little is known.

We studied these questions using fMRI in humans while subjects performed a variant of a two-armed bandit task to obtain probabilistically delivered monetary rewards (Fig. 1A). A critical feature of the task was that they had to select a motor response in one of two distinct response modalities in every trial, they could choose to make either a saccade to the right of a fixation cross, or to press a button with the right hand. This design allowed us to exploit the fact that different regions of the cortex are involved in the planning of eye and hand movements (17). We hypothesized that value representations for the two actions would be separable within these cortical areas at the spatial resolution available to fMRI. The probability of being rewarded on each of the two actions drifted randomly over time and was independent of the probability of being rewarded on the other (Fig. 1B). This characteristic ensured that value estimates for eye and hand movements were uncorrelated, which gave us maximum sensitivity with which to dissociate the neural representations of the two action values.

Author contributions: K.W., A.R., and J.P.O. designed research; K.W. performed research; K.W. analyzed data; K.W., A.R., and J.P.O. wrote the paper.

The authors declare no conflict of interest.

This article is a U.S. Government work and, as such, is in the public domain in the United States of America.

Reprints and permissions: [www.sagepub.com/journalsPermissions.nav](http://www.sagepub.com/journalsPermissions.nav)

DOI: 10.1073/pnas.0907177106

PNAS | October 6, 2009 | vol. 106 | no. 40 | 17399–17404

## LETTERS

### Cortical substrates for exploratory decisions in humans

Nathaniel D. Daw<sup>1,2</sup>, John P. O’Doherty<sup>1,3</sup>, Peter Dayan<sup>1</sup>, Ben Seymour<sup>4</sup> & Raymond J. Dolan<sup>1</sup>

Decision making in an uncertain environment poses a conflict between the opposing demands of gathering and exploiting information. In a classic illustration of this ‘exploration–exploitation’ dilemma, a gambler choosing between multiple slot machines balances the desire to select what seems, on the basis of accumulated experience, the richest option, against the desire to choose a less familiar option that might turn out more advantageous (and thereby provide information for improving future decisions). Far from representing idle curiosity, such exploration is often critical for organisms to discover how best to harvest resources such as food and water. An appetitive choice, substantial experimental evidence, underpinned by computational reinforcement learning (RL) theory, indicates that a dopaminergic, ‘striatal’ and medial prefrontal network mediates learning to exploit. In contrast, although exploration has been well studied from both theoretical and ethological perspectives, its neural substrates are much less clear. Here we show, in a gambling task, that human subjects’ choices can be characterized by a computationally well-regarded strategy for addressing the exploration/exploitation dilemma. Furthermore, using this characterization to classify decisions as exploratory or exploitative, we employ functional magnetic resonance imaging to show that the frontopolar cortex and intraparietal sulcus are preferentially active during exploratory decisions. In contrast, regions of striatum and ventromedial prefrontal cortex exhibit activity characteristic of an involvement in value-based exploratory decision making. The results suggest a model of action selection under uncertainty that involves switching between exploratory and exploitative behavioural modes, and provide a computationally precise characterization of the contribution of key decision-related brain systems to each of these functions.

Exploration is a computationally refined capacity, demanding careful regulation. Two possibilities for this regulation arise. On the one hand, we might expect the involvement of cognitive, prefrontal control systems that can supersede ‘over simple’ dopamine-mediated habitual mechanisms. On the other hand, theoretical work on optimal exploration<sup>1</sup> indicates a more unified architecture, according to which actions can be associated with the use of a metric that integrates both primary reward and the informational value of exploration, even in simple, habitual decision systems. We studied patterns of behaviour and brain activity in 14 healthy subjects while they performed a ‘four-armed bandit’ task involving repeated choices between four slot machines (Fig. 1). See Supplementary Methods. The slow paid-off points (to be exchanged for money) mostly around four different means. Unlike standard slots, the mean payoffs changed randomly and independently from trial to trial, with subjects finding information about the current worth of a slot only through sampling it actively. This feature of the experimental design, together with a model-based analysis, allowed us to study exploratory and exploitative decisions under uniform conditions, in the context of a single task.

We asked subjects to post-task interviews to describe their choice strategies. The majority (11 of 14) reported occasionally trying the different slots to work out which currently had the highest payoffs (exploring) while at other times choosing the slot they thought had the highest payoffs (exploiting). To investigate this behaviour quantitatively, we considered RL (ref. 2) strategies for exploration. These strategies come in three flavours, differing in how exploratory actions are directed. The simplest method, known as ‘ $\epsilon$ -greedy’, is straightforward: it chooses the ‘greedy’ option (the one believed to be best) most of the time, but occasionally (with probability  $\epsilon$ ) substitutes a random action. A more sophisticated approach is to guide exploration by expected value, as in the ‘softmax’ rule. With softmax, the decision to explore and the choice of which suboptimal action to take are determined probabilistically on the basis of the actions’ relative expected values. Last, exploration can additionally be directed by avoiding bonuses in this latter decision towards actions whose consequences are uncertain, specifically to those for which exploration will be most informative. The optimal strategy for a restricted class of simple bandit tasks has this characteristic, as do standard heuristics<sup>3,4</sup> for exploration in more complicated RL tasks such as ours, for which the optimal solution is computationally intractable.

We compared the fit of three distinct RL models, embodying the aforementioned strategies, to our subjects’ behavioural choices. All the models learned the values of actions with the use of a Kalman filter (see Supplementary Methods), an error-driven prediction algorithm that generalizes the temporal-difference learning algorithm (used in most RL theories of dopamine)<sup>5</sup> by also tracking uncertainty about the value of each action. The models differed only in their choice rules. We compared models by using the likelihood of the subjects’ choices over their experience, optimized over free parameters. This comparison (Supplementary Tables 1 and 2) revealed strong evidence for value-sensitive (softmax) over undirected ( $\epsilon$ -greedy) exploration. There was no evidence to justify the introduction of an extra parameter that allowed exploration to be directed towards uncertainly (softmax with an uncertainty bonus), or optimal fit, the bonus was negligible, making the model equivalent to the simpler softmax. We conducted additional model fits (see Supplementary Information) to verify that these findings were not an artefact of our assumptions about the timing of free parameters between subjects.

Having characterized subjects’ behaviour computationally, we used the best-fitting softmax model to generate regressions containing value predictions, prediction errors and choice probabilities for each subject on each trial. We used statistical parametric mapping to

## Nucleus Accumbens D2/3 Receptors Predict Trait Impulsivity and Cocaine Reinforcement

Jeffrey W. Dalley<sup>1,2\*</sup>, Tim D. Fryer<sup>1,3</sup>, Laurent Riches<sup>1,4</sup>, Emma S. J. Robinson<sup>1,2</sup>, David E. N. Theobald<sup>1,2</sup>, Kristian Lüske<sup>1,2</sup>, Yolanda Peña<sup>1,2</sup>, Emily K. Murphy<sup>1,2</sup>, Yasmine Shah<sup>1,2</sup>, Karim Probst<sup>1,2</sup>, Ulag Akhramov<sup>1,2</sup>, Franklin J. Aghajanian<sup>1,2</sup>, Hugh K. Richards<sup>1,2</sup>, Young Hoang<sup>1,2</sup>, Jean-Claude Barrot<sup>1,2</sup>, Barry J. Everitt<sup>1,2</sup>, Trevor W. Robbins<sup>1,2</sup>

Stimulant addiction is often linked to excessive risk taking, sensation seeking, and impulsivity, but in ways that are poorly understood. We report here that a form of impulsivity in rats predicts high rates of intravenous cocaine self-administration and is associated with changes in dopamine (DA) function before drug exposure. Using positron emission tomography, we demonstrated that D2/3 receptor availability is significantly reduced in the nucleus accumbens of impulsive rats that were never exposed to cocaine and that such effects are independent of DA release. These data demonstrate that trait impulsivity predicts cocaine reinforcement and that D2 receptor dysfunction in abstinent cocaine addicts may, in part, be determined by premitochondrial influences.

Accumulating evidence suggests that certain personality traits, including sensation seeking and impulsivity, and antisocial conduct disorder, may predispose humans to drug abuse and addiction (1–4). However, from studies of human drug addicts alone, it is difficult to determine whether compulsive impulsivity and cognitive dysfunction (5, 6) pre-date the onset of drug use or emerge as a consequence of chronic drug use. Current hypotheses suggest that long-term drug use impairs inhibitory control functions mediated by the prefrontal cortex and the associated limbic brain circuitry, leading to a loss of inhibition or to impulsivity (7, 8). However, there is little evidence to date that cocaine, exposure to cocaine and other psychostimulant drugs leads to long-term increases in impulsive behaviour in animals (9, 10).

The view that individual differences in drug abuse reflect distinct behavioural and physiological traits is widely supported by studies in nonhuman primates in which cocaine is more readily self-administered by individuals with high novelty-seeking behaviour (11–13). Rats that are selected for high novelty-seeking behaviour activity more readily acquire instrumental responses to drug-paired cues (14, 15) and show a higher preference for drug-paired cues (16, 17). The view that individual differences in drug abuse reflect distinct behavioural and physiological traits is widely supported by studies in nonhuman primates in which cocaine is more readily self-administered by individuals with high novelty-seeking behaviour (11–13). Rats that are selected for high novelty-seeking behaviour activity more readily acquire instrumental responses to drug-paired cues (14, 15) and show a higher preference for drug-paired cues (16, 17).

\*To whom correspondence should be addressed. E-mail: [jeff.dalley@york.ac.uk](mailto:jeff.dalley@york.ac.uk)

Present address: <sup>1</sup>Department of Psychology, University of York, Heslington, York YO10 5DD, UK; <sup>2</sup>Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK; <sup>3</sup>Wolfson Brain Imaging Centre, Department of Clinical Neurosciences, University of Cambridge, Addenbrooke’s Hospital, Cambridge CB2 2RU, UK; <sup>4</sup>Centre for Cognitive and Behavioral Neuroscience, University of Cambridge, Addenbrooke’s Hospital, Cambridge CB2 2RU, UK

Received July 1, 2009; revised August 11, 2009; accepted August 11, 2009. This article is a U.S. Government work and, as such, is in the public domain in the United States of America.

Reprints and permissions: [www.sagepub.com/journalsPermissions.nav](http://www.sagepub.com/journalsPermissions.nav)

DOI: 10.1073/pnas.0907177106

PNAS | October 6, 2009 | vol. 106 | no. 40 | 17399–17404

REPORTS

involve the brain dopamine (DA) systems, in particular the mesolimbic and mesocortical DA pathways innervating the nucleus accumbens and prefrontal cortex (18–21). Positron emission tomography (PET) studies in nonhuman primates have indicated a role for DA D2 receptors in determining individual differences in intravenous cocaine self-administration (19, 20). Specifically, low D2 receptor availability in the nucleus accumbens predicts subsequent levels of intravenous cocaine self-administration in these nonmonkeys (20), a result apparently similar to that seen in studies of human cocaine abusers (23).

However, it is not clear how individual differences in D2 receptor availability relate to a specific behavioural endophenotype or behavioural process that confers vulnerability to drug addiction. In addition, there have been few, if any, studies where DA release *in vivo* has been combined with PET estimates of D2 receptor availability. This is important because D2 receptor availability is influenced by both receptor density and competing DA release (24, 25). Thus, there is a need to conduct analogous PET studies in animals to investigate the predictive relationship between D2 receptor availability and trait behavioural markers of drug-abuse vulnerability.

We investigated the relevance of a spontaneously occurring form of impulsivity in outbred Lister hooded (LH) rats to intravenous cocaine self-administration and to underlying changes in striatal DA function, as measured by microPET and *in vivo* microdialysis (26). We defined impulsivity as high levels of anticipatory responses made before the presentation of a

**Fig. 1.** Behavioral attributes of trait impulsivity on the 5-CSF task. (A) Impulsive rats exhibit high levels of premature responding on days 1–3 after trial initiation (day 0) as compared with nonimpulsive rats. (B) Premature responding (mean  $\pm$  SEM) on days 1, 2, 4, and 5 of 7 after trial initiation (day 0) as compared with nonimpulsive rats. (C) Mean  $\pm$  SEM variance (ANOVA) of premature responses revealed a significant main effect of day [F(2, 60) = 34.9,  $P < 0.001$ ] and a significant main effect of group [F(1, 30) = 26.1,  $P < 0.001$ ]. However, there were no significant effects on other measures of task performance, including (D) error measures of task performance, including (E) latency to collect food reward [F < 1, not significant (n.s.)], (F) omissions [F < 1, n.s.], (G) latency to respond correctly [F(1, 120) = 3.0,  $P = 0.081$ ], and (H) the time required to complete both standard and challenge long ITI sessions [F < 1, n.s.]. Black circles, high-impulsive rats; white circles, non-impulsive rats.

Author contributions: J.W.D., T.D.F., L.R., E.S.J.R., D.E.N.T., K.P., Y.H., J.C.B., B.J.E., and T.W.R. designed research; J.W.D., T.D.F., L.R., E.S.J.R., D.E.N.T., K.P., Y.H., J.C.B., B.J.E., and T.W.R. performed research; J.W.D., T.D.F., L.R., E.S.J.R., D.E.N.T., K.P., Y.H., J.C.B., B.J.E., and T.W.R. analyzed data; J.W.D., T.D.F., L.R., E.S.J.R., D.E.N.T., K.P., Y.H., J.C.B., B.J.E., and T.W.R. wrote the paper.

The authors declare no conflict of interest.

This article is a U.S. Government work and, as such, is in the public domain in the United States of America.

Reprints and permissions: [www.sagepub.com/journalsPermissions.nav](http://www.sagepub.com/journalsPermissions.nav)

DOI: 10.1073/pnas.0907177106

PNAS | October 6, 2009 | vol. 106 | no. 40 | 17399–17404

EMBARGOED UNTIL 2PM U.S. EASTERN TIME ON THE THURSDAY BEFORE THIS DATE:

[www.pnas.org](http://www.pnas.org) SCIENCE VOL 315 2 MARCH 2007

17399

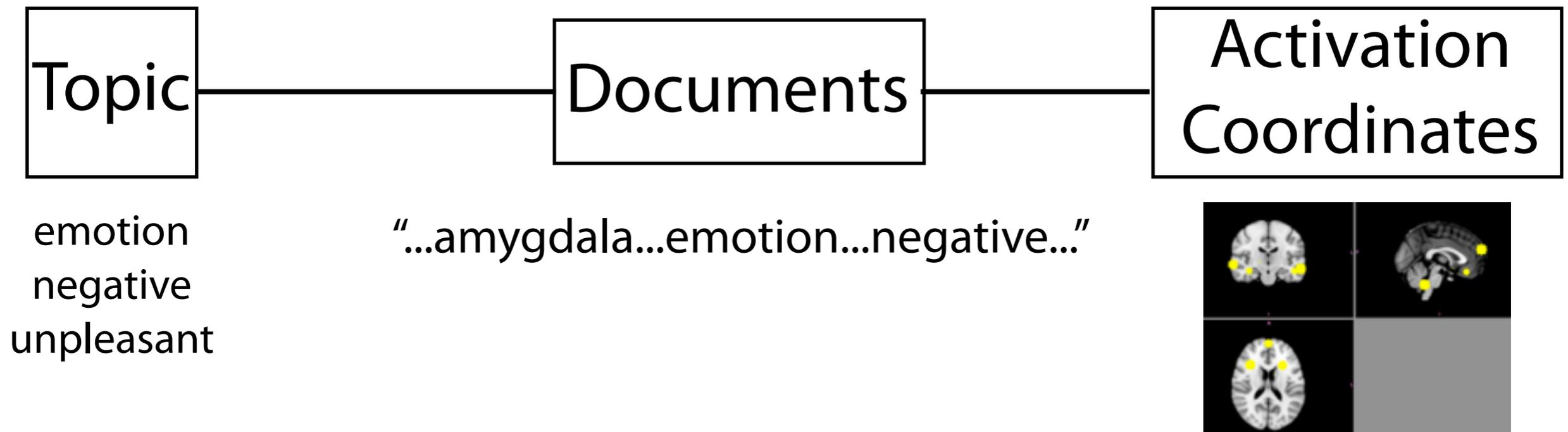
© 2009 Nature Publishing Group

576

polracklab.org

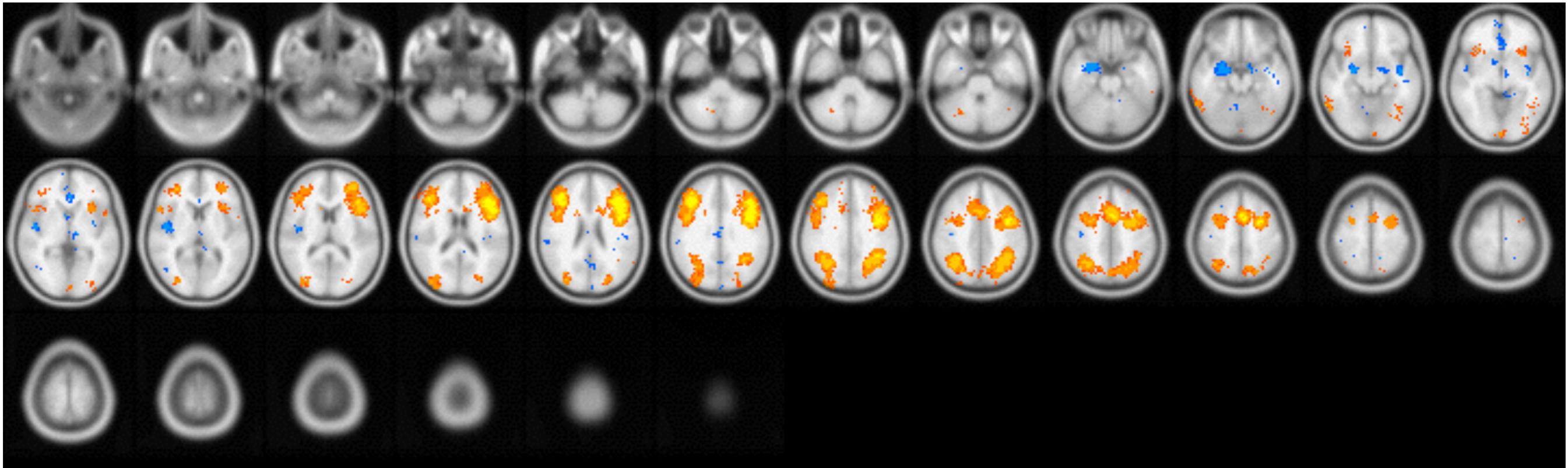
# Topic Mapping

- Perform topic modeling using latent Dirichlet allocation with Cognitive Atlas terms
- Each document has a loading on each topic
  - On average, each document loads on ~6.5 topics
- Extract activation coordinates for 5,809 papers in NeuroSynth
- Perform voxelwise chi-square test with FDR correction to examine association between topics and activation



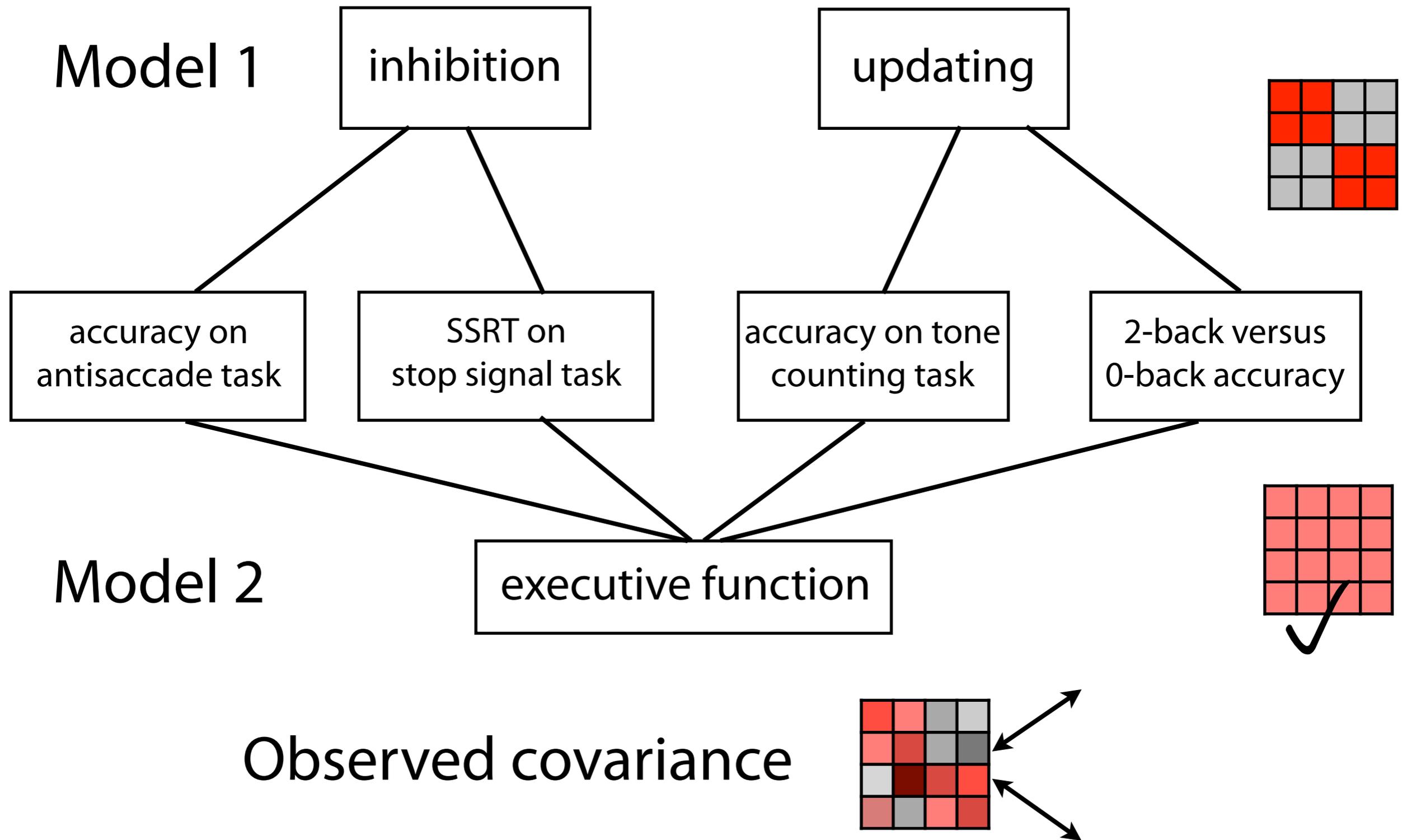
Poldrack et al., 2012, *PLoS Comp Biology*

# Meta-analysis using Cognitive Atlas terms



Topic 61 (442 docs): memory working\_memory  
maintenance visual\_working\_memory  
spatial\_working\_memory manipulation episodic\_buffer  
retention rehearsal retrieval

# Towards meta-analytic testing of cognitive frameworks



# Conclusion

- Cognitive ontologies can provide a more formal definition of cognitive functions
- Ontologies plus meta-analysis may provide the means to test between different conceptual frameworks

# Acknowledgments

Jeanette Mumford  
Sarah Helfinstein  
Tom Schonberg  
Craig Fox  
Koji Jimura  
Sanmi Koyejo  
Chris Gorgolewski  
Tyler Davis  
Tal Yarkoni  
Jessica Cohen  
Robert Bilder  
Eliza Congdon  
Eydie London  
Tyrone Cannon  
Nelson Freimer  
Stephen J. Hanson  
Yaroslav Halchenko  
Yannick Schwartz  
Gael Varoquaux  
Bertrand Thirion



The Poldrack Lab @ Stanford

Visit:

[www.openfmri.org](http://www.openfmri.org)

[www.cognitiveatlas.org](http://www.cognitiveatlas.org)



James S. McDonnell  
Foundation



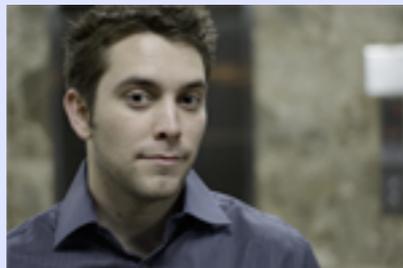


# Acknowledgments

## UCLA



Robert Bilder



Don Kalar



Fred Sabb

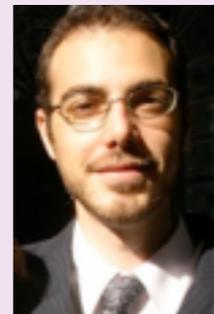


D. Stott Parker

## UT Austin



Jeanette Mumford



Tal Yarkoni

## Stanford



Sanmi Koyejo

## Carnegie-Mellon



Niki Kittur

Get involved!  
[www.cognitiveatlas.org](http://www.cognitiveatlas.org)



James S. McDonnell  
Foundation