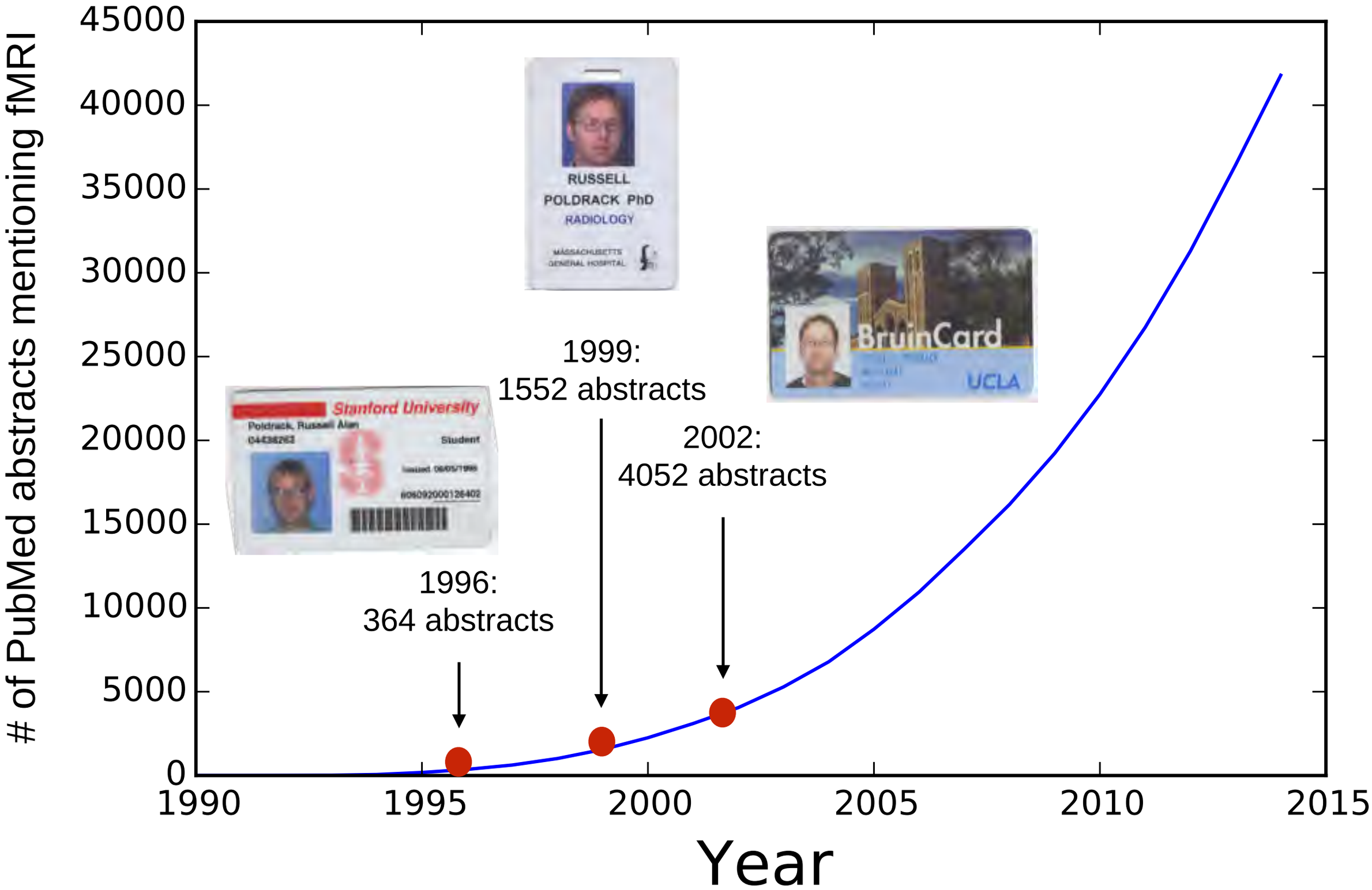
The future of fMRI in cognitive neuroscience

Russell Poldrack

Department of Psychology
Stanford University


# of PubMed abstracts mentioning fMRI
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
1990
1995
2000
2005
2010
2015
Year
RUSSELL
POLDRACK PhD
RADIOLOGY
BruinCard
UCLA
Stanford University
1999:
1552 abstracts
2002:
4052 abstracts
1996:
364 abstracts

Do we really know an order of magnitude more now than we did in 2002?

Will 40,000 more of the same kind of fMRI papers tell us twice what we already know?

The future of fMRI in cognitive neuroscience

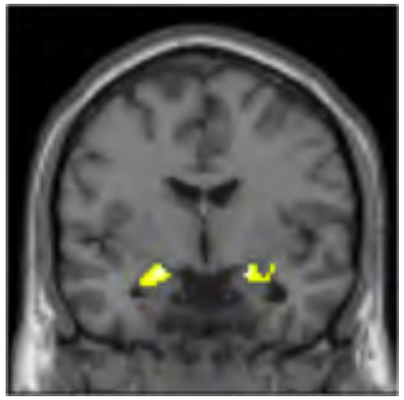
- Selective inference
- Predictive fMRI
- Individual variability
- Reproducible research practices

What does the anterior cingulate cortex do?

- "anterior cingulate" and fMRI
  - 3683 abstracts in PubMed

The standard imaging paradigm

Manipulate mental process → Observe activation in region X

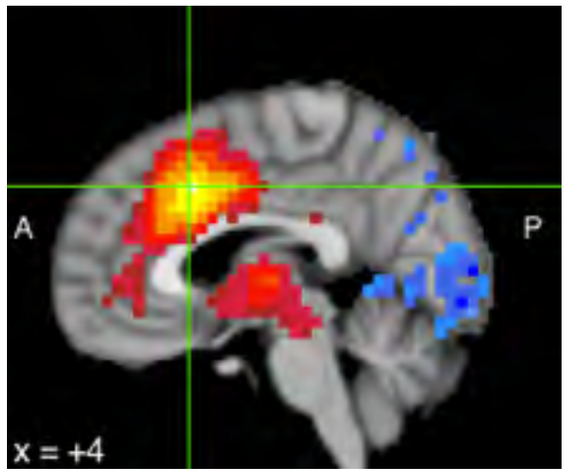


Forward Inference: Amygdala is involved in fear

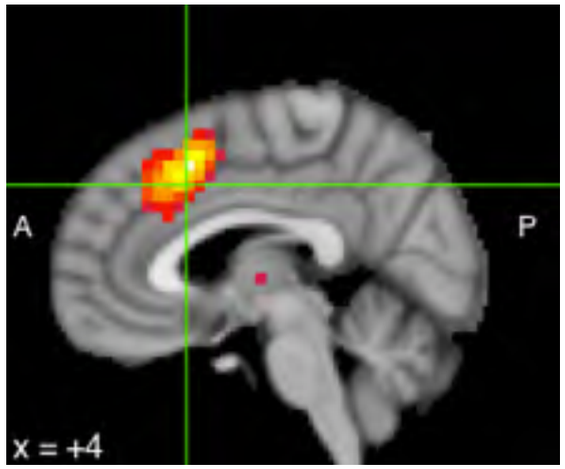
What does the ACC do?

maintenance

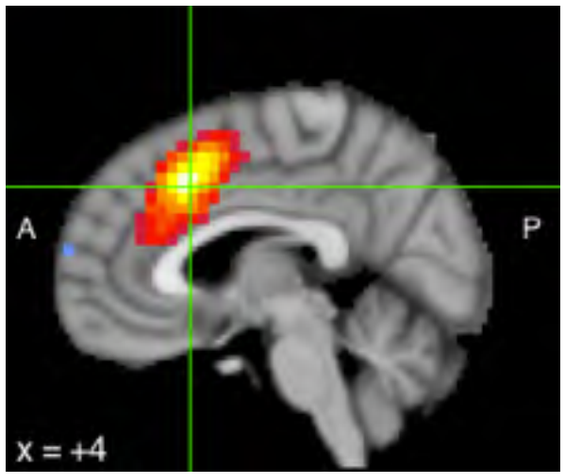
pain



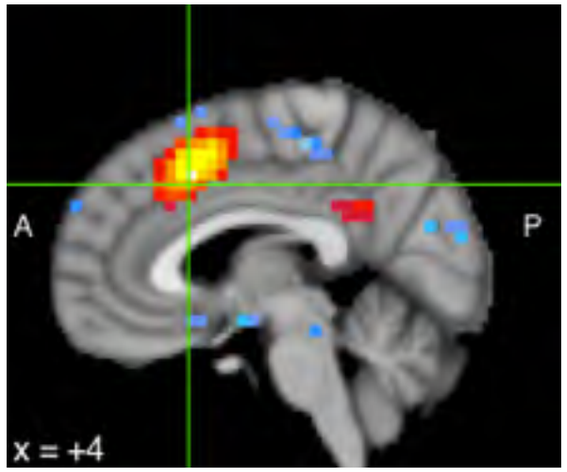
phonology



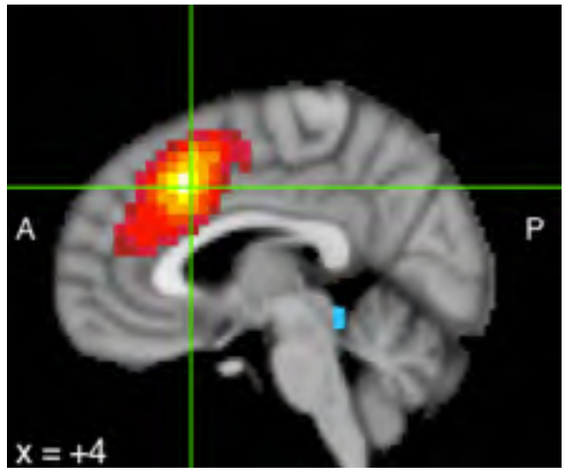
interference



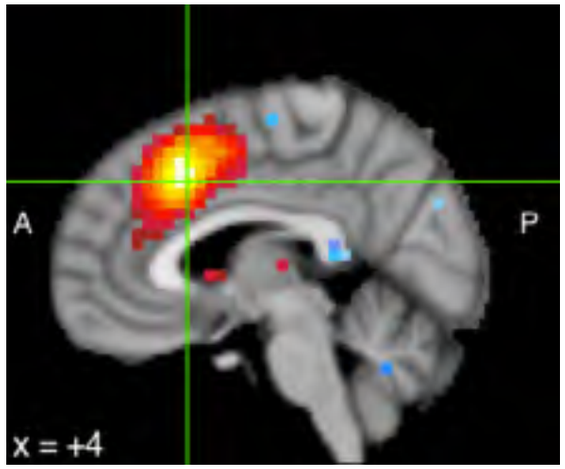
difficulty



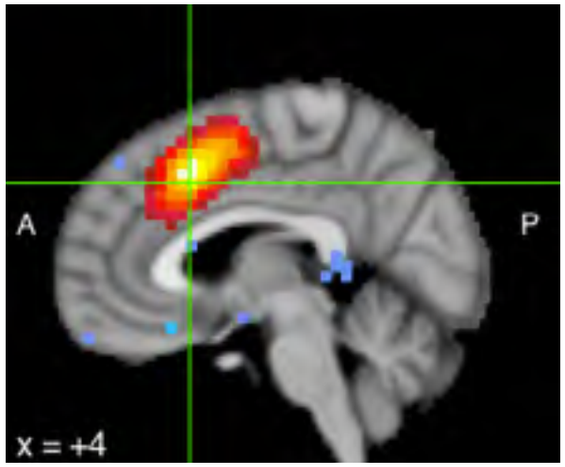
conflict



errors



attention



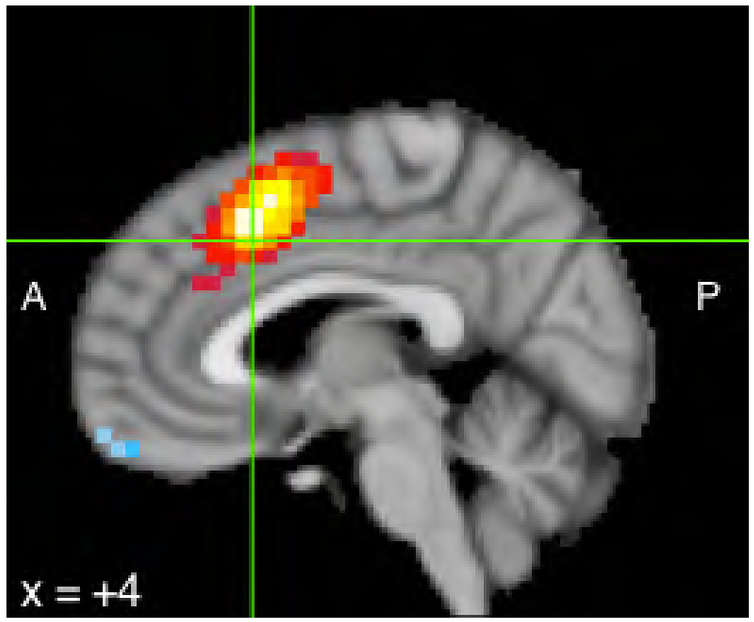
forward inference Z estimated using neurosynth.org

Neural correlates

Manipulate some mental process

Observe associated brain activation

working memory maintenance



working memory is sufficient to activate ACC

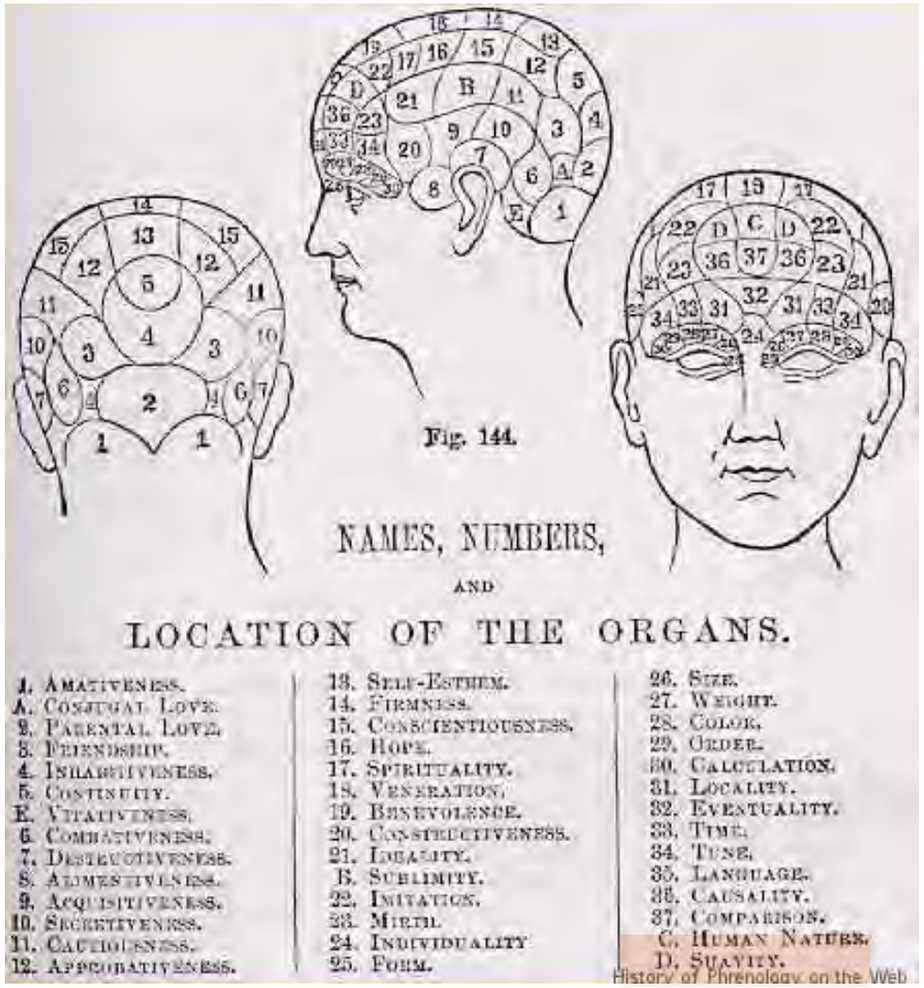
~~working memory is necessary to activate ACC~~

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

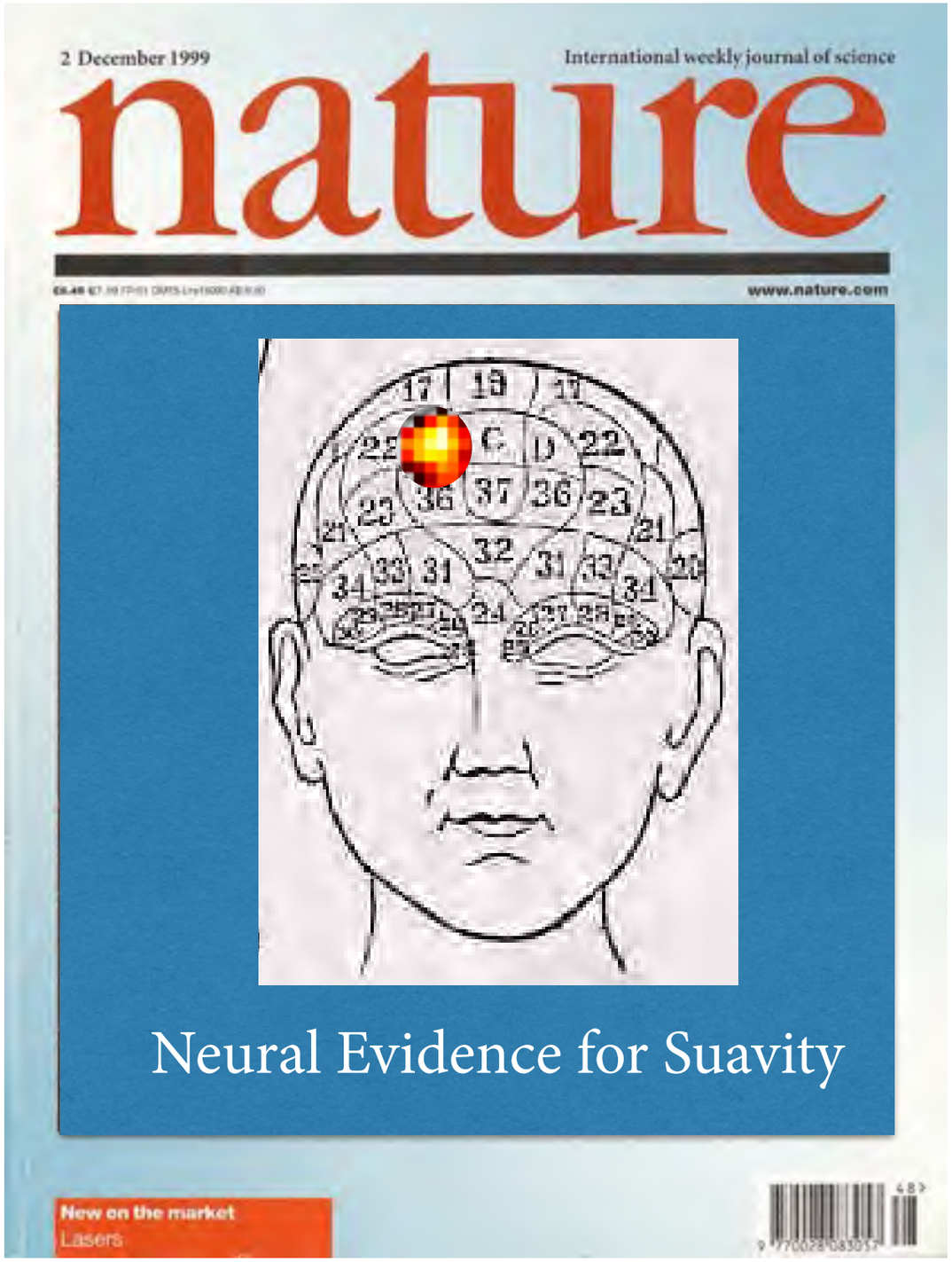
Some alternatives

- 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?

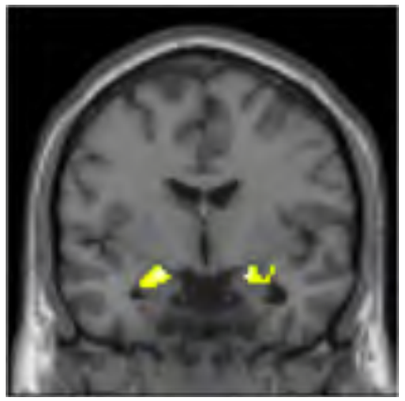


-



The reverse inference paradigm

Observe activation in region X → Predict the mental process engaged by the subject



→

Reverse Inference: The person is experiencing fear

Do you really love your iPhone?

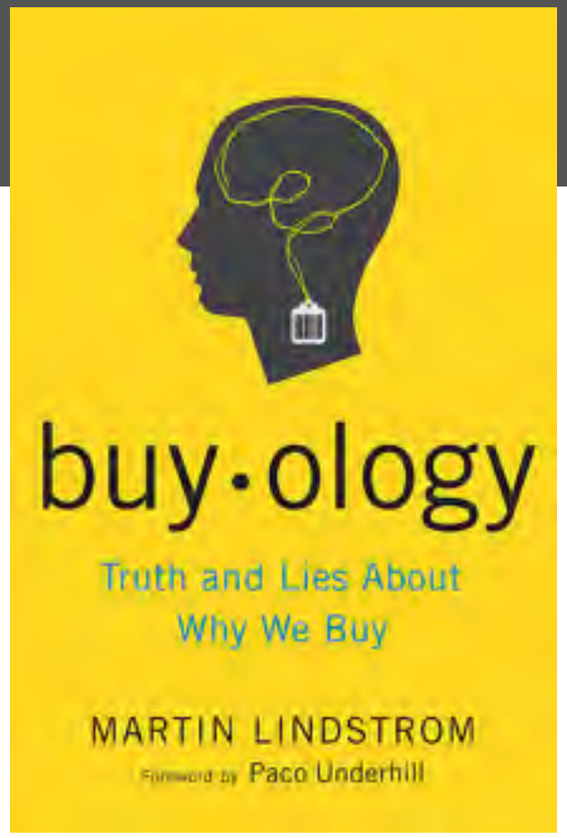
The New York Times

The Opinion Pages

OP-ED CONTRIBUTOR

You Love Your iPhone. Literally.

By MARTIN LINDSTROM
Published: September 30, 2011



- "Earlier this year, I carried out an fMRI experiment to find out whether iPhones were really, truly addictive, no less so than alcohol, cocaine, shopping or video games. In conjunction with the San Diego-based firm MindSign Neuromarketing, I enlisted eight men and eight women between the ages of 18 and 25. Our 16 subjects were exposed separately to audio and to video of a ringing and vibrating iPhone...most striking of all was the flurry of activation in the insular cortex of the brain, which is associated with feelings of love and compassion. The subjects' brains responded to the sound of their phones as they would respond to the presence or proximity of a girlfriend, boyfriend or family member. In short, the subjects didn't demonstrate the classic brain-based signs of addiction. Instead, they loved their iPhones.

To the Editor:

"You Love Your iPhone. Literally," by Martin Lindstrom (Op-Ed, Oct. 1), purports to show, using brain imaging, that our attachment to digital devices reflects not addiction but instead the same kind of emotion that we feel for human loved ones.

However, the evidence the writer presents does not show this.

The brain region that he points to as being "associated with feelings of love and compassion" (the insular cortex) is active in as many as one-third of all brain imaging studies.

Further, in studies of decision making the insular cortex is more often associated with negative than positive emotions.

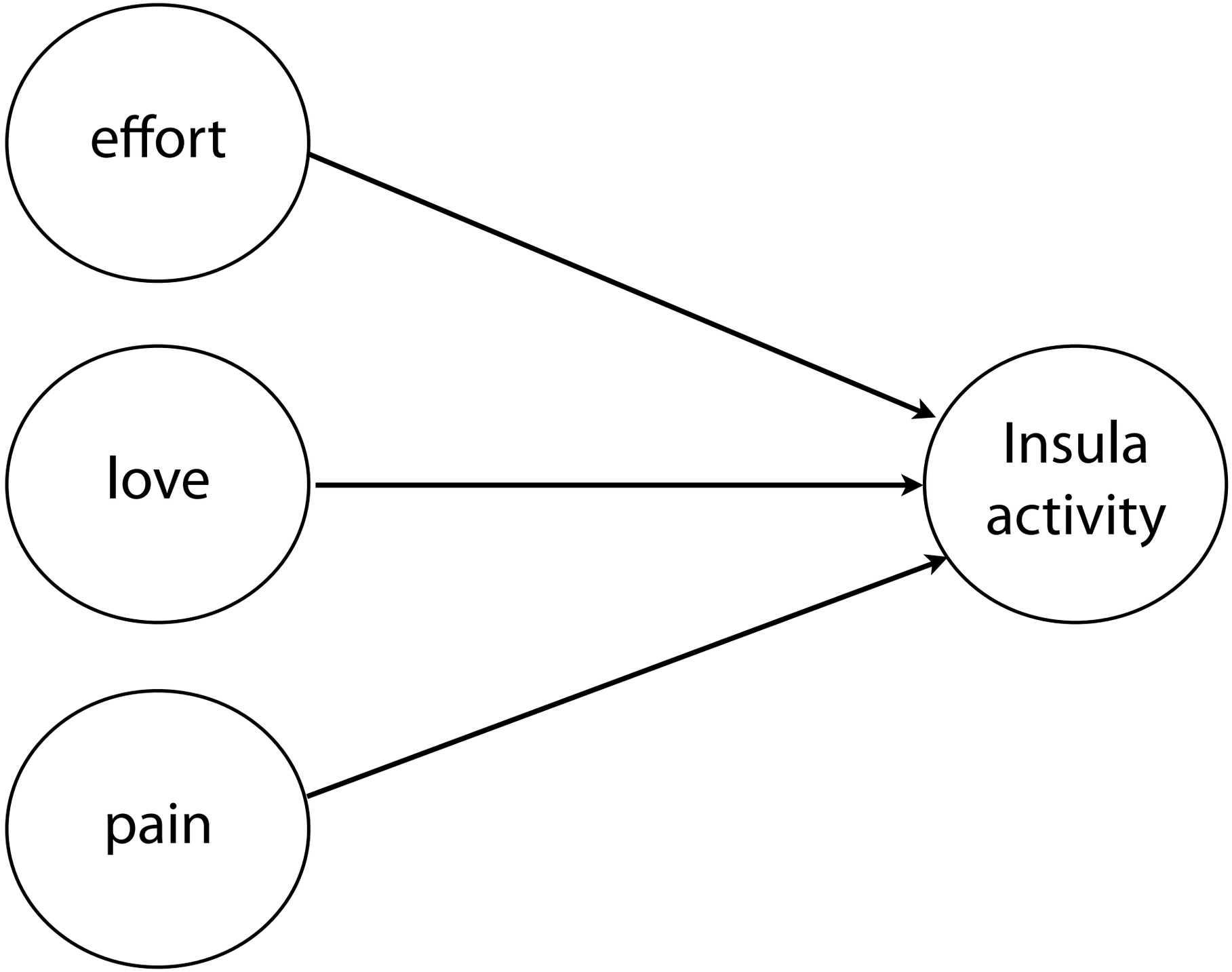
The kind of reasoning that Mr. Lindstrom uses is well known to be flawed, because there is rarely a one-to-one mapping between any brain region and a single mental state; insular cortex activity could reflect one or more of several psychological processes.

We find it surprising that The Times would publish claims like this that lack scientific validity.

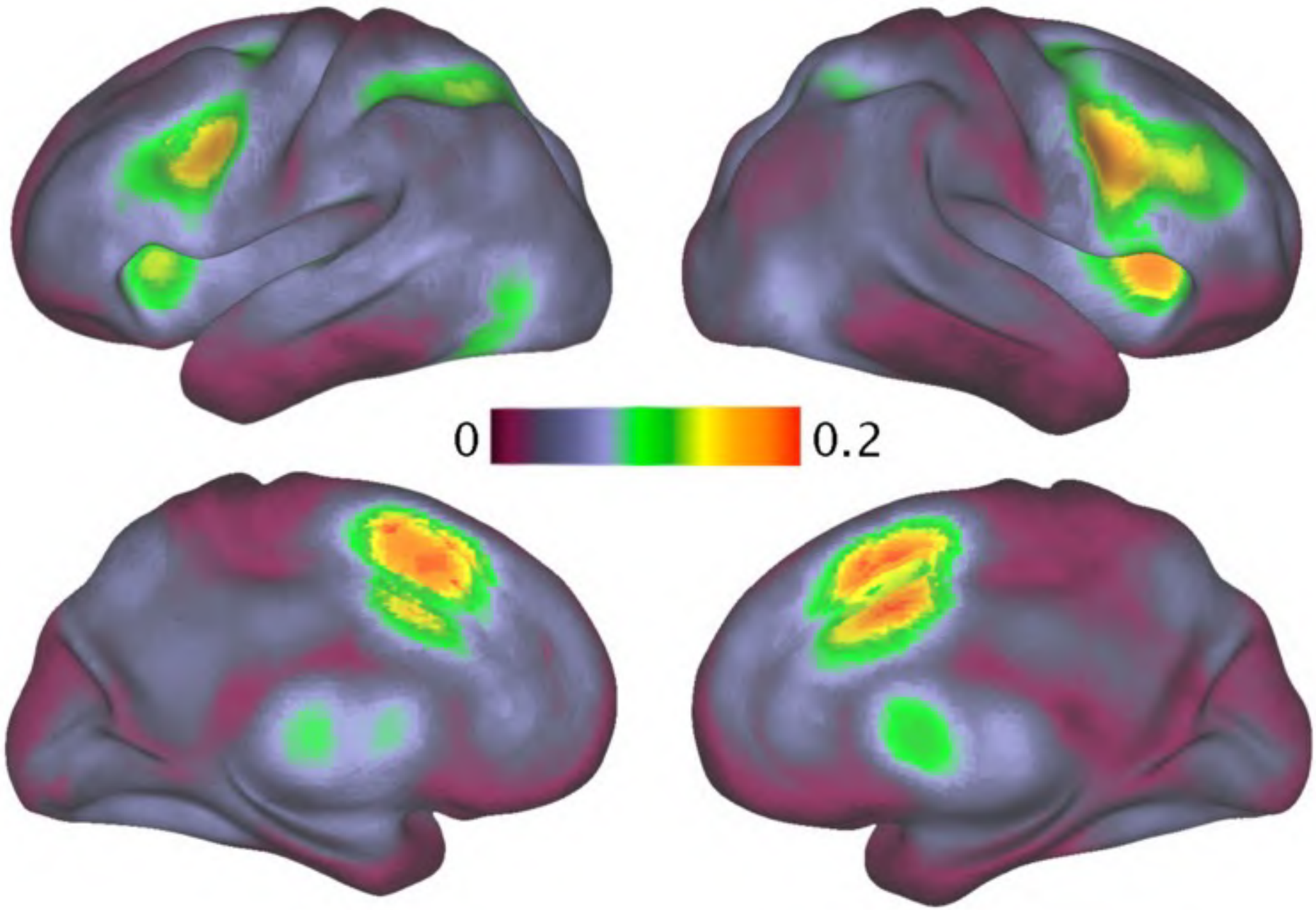
RUSSELL POLDRACK
Austin, Tex., Oct. 3, 2011

The writer is a professor of psychology and neurobiology at the University of Texas at Austin. His letter was signed by 44 other neuroscientists.

Does reverse inference work?



Insula activation is weakly selective



Some voxels active in more than 20% of studies

Yarkoni et al., 2011

Reverse inference

- Informal reverse inference provides relatively weak evidence

Opinion

Can cognitive processes be inferred from neuroimaging data?

Russell A. Poldrack

Departments of Psychology and Brain Research Institute, UCLA, Los Angeles, CA 90095-1563, USA

There is much interest currently in using functional neuroimaging techniques to understand better the nature of cognition. One particular practice that has become common is 'reverse inference', by which the engagement of a particular cognitive process is inferred from the activation of a particular brain region. Such inferences are not deductively valid, but can still provide some information. Using a Bayesian analysis of the BrainMap neuroimaging database, I characterize the amount of additional evidence in favor of the engagement of a cognitive process that can be offered by a reverse inference. Its usefulness is particularly limited by the selectivity of activation in the region of interest. I argue that cognitive neuroscientists should be circumspect in the use of reverse inference, particularly when selectivity of the region in question cannot be established or is known to be weak.

Introduction

Functional neuroimaging techniques such as functional magnetic resonance imaging (fMRI) provide a measure of local brain activity in response to cognitive tasks undertaken during scanning. These data allow the cognitive neuroscientist to infer something about the role of particular brain regions in cognitive function. However, there is increasing use of neuroimaging data to make the opposite inference; that is, to infer the engagement of particular cognitive functions based on activation in particular brain regions. My goal here is to analyze this practice, known as 'reverse inference', and to characterize some limitations on the effectiveness of this strategy. The companion paper in this issue by Henson [1] discusses a complementary strategy for using neuroimaging to distinguish competing cognitive theories.

The goal of cognitive psychology is to understand the underlying mental architecture that supports cognitive functions. To this end, cognitive psychologists examine the effects of task manipulations on behavioral variables, such as response time or accuracy, and use these data to test models of cognitive function. However, it is often not possible to determine on the basis of behavioral variables alone whether a particular cognitive process is engaged, or whether a particular theory of cognitive architecture is correct; for example, there are well-known examples of theoretical indeterminacy based on behavioral data [2]. If neuroimaging were able to provide information regarding what cognitive processes were engaged in performance of a particular task, cognitive psychologists would have gained a powerful new tool. Researchers outside cognitive psychology are also sometimes interested in using neuroimaging to determine the engagement of particular cognitive processes. For example, philosophers might wish to know the degree to which emotion versus deliberative reasoning plays a role in moral judgments [3].

Inference in neuroimaging

The usual kind of inference that is drawn from neuroimaging data is of the form '*if cognitive process X is engaged then brain area Z is active*'. Perusal of the discussion sections of a few fMRI articles will quickly reveal, however, an epidemic of reasoning taking the following form:

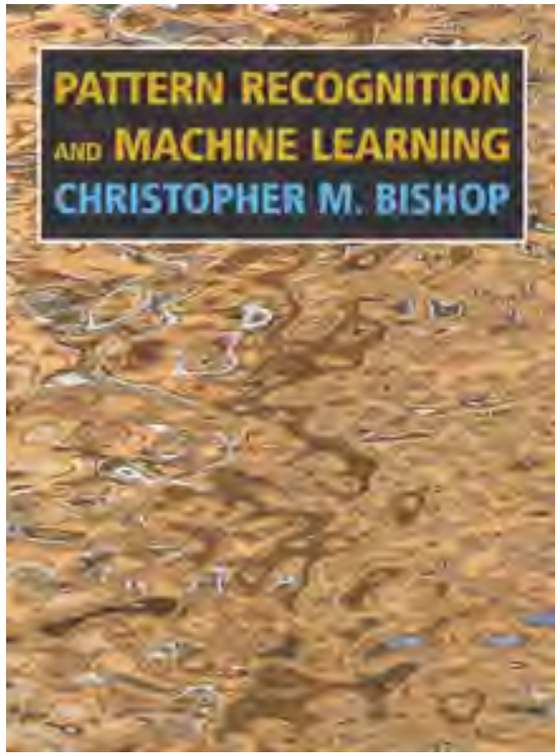
(1) In the present study, when task comparison A was presented, brain area Z was active.
(2) In other studies, when cognitive process X was putatively engaged, then brain area Z was active.
(3) Thus, the activity of area Z in the present study demonstrates engagement of cognitive process X by task comparison A.

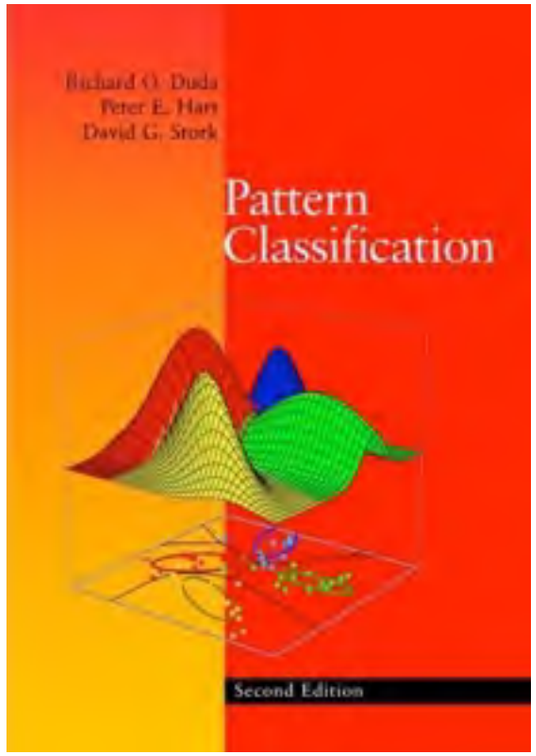
This is a 'reverse inference', in that it reasons backwards from the presence of brain activation to the engagement of a particular cognitive function.

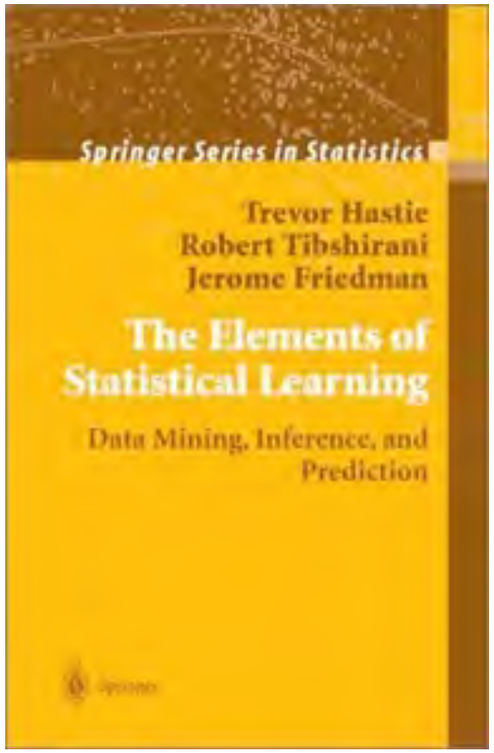
In many cases the use of reverse inference is informal; the presence of unexpected activation in a particular region is explained by reference to other studies that found activation in the same region. However, in some studies the reverse inference is a central feature. In one study [4], subjects were scanned using PET while they performed an economic exchange task in which they had the chance to punish those who defected. Activation was observed in the dorsal striatum when participants subjected defectors to effective punishment; this activation was inferred to reflect the rewarding properties of altruistic punishment. Similarly, a study using fMRI in rats [5] compared activity during pup suckling versus cocaine administration. Greater activity in the dorsal and ventral striatum during suckling compared with cocaine administration led the authors to conclude that 'pup suckling is more rewarding than cocaine' (p. 149). In each of these studies, a cognitive process ('reward') was inferred from activation in a particular brain system (the striatum). Nearly every

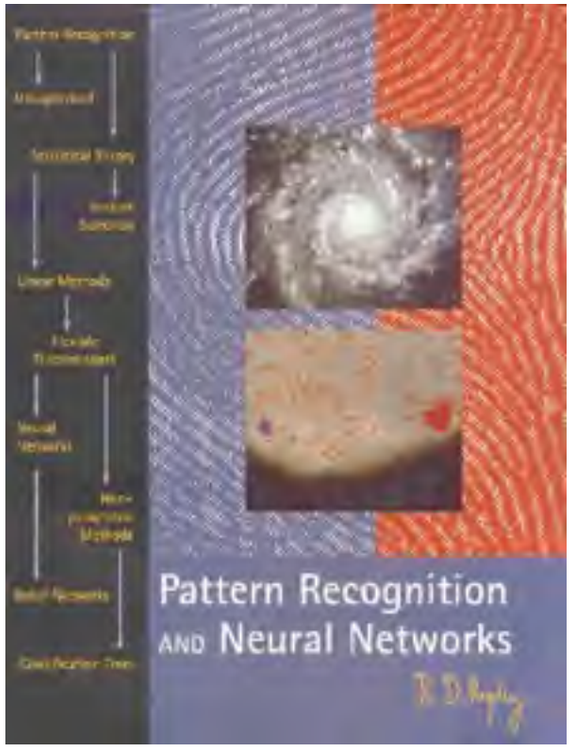
TICS, 2006

How to make better predictions

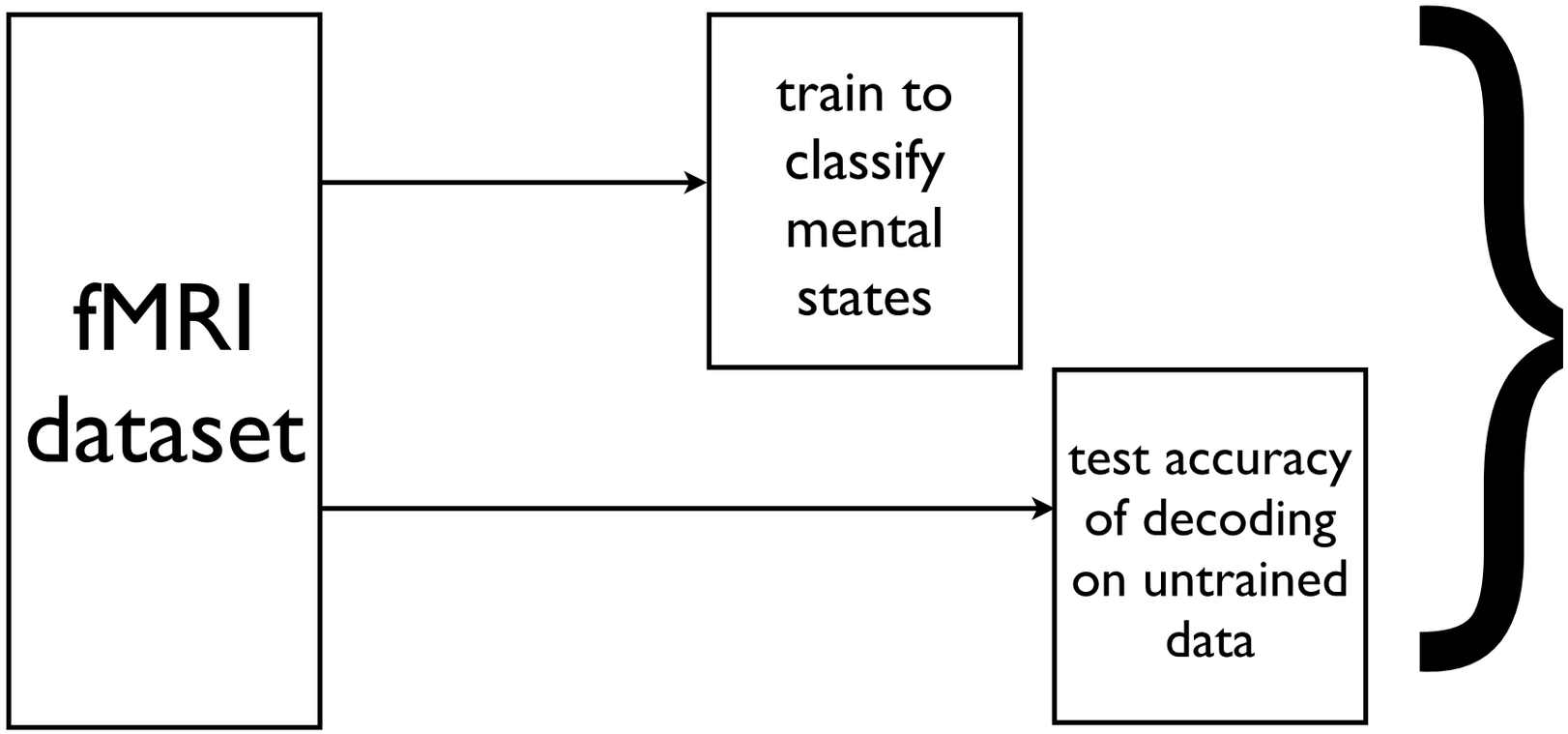








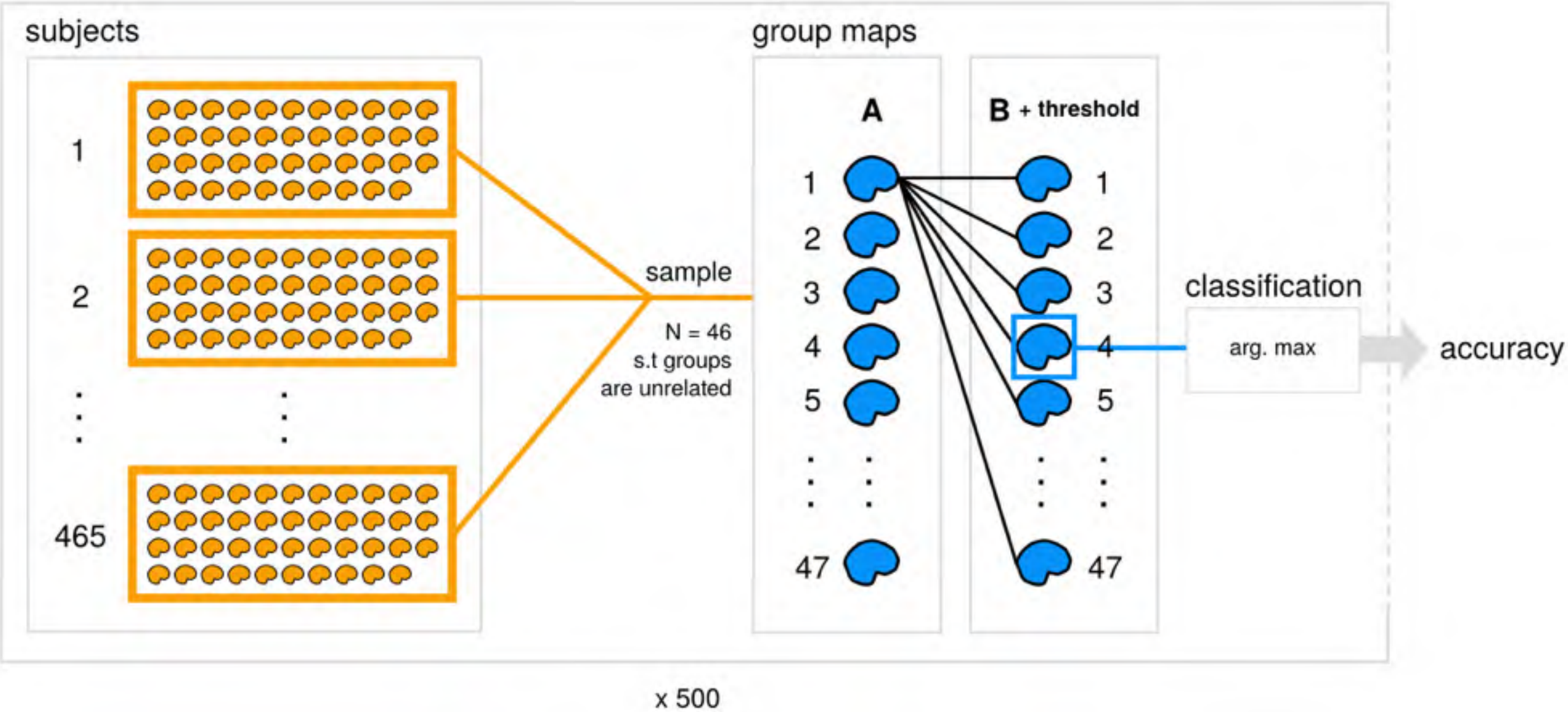
Decoding mental states using machine learning



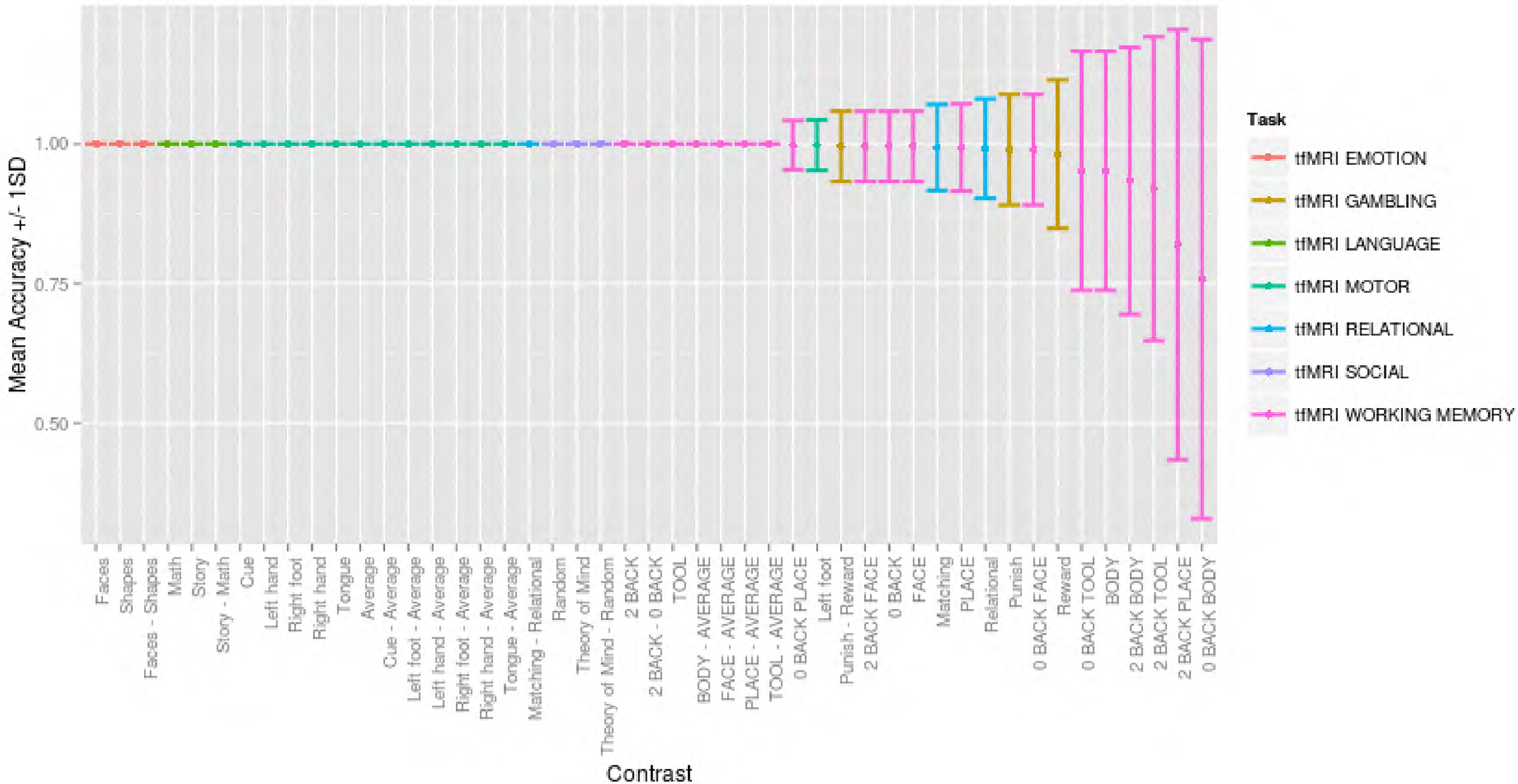
Cross-validation:

- Train for each split of size k
- Compute average predictive accuracy on left-out data

An example: Classifying the HCP data

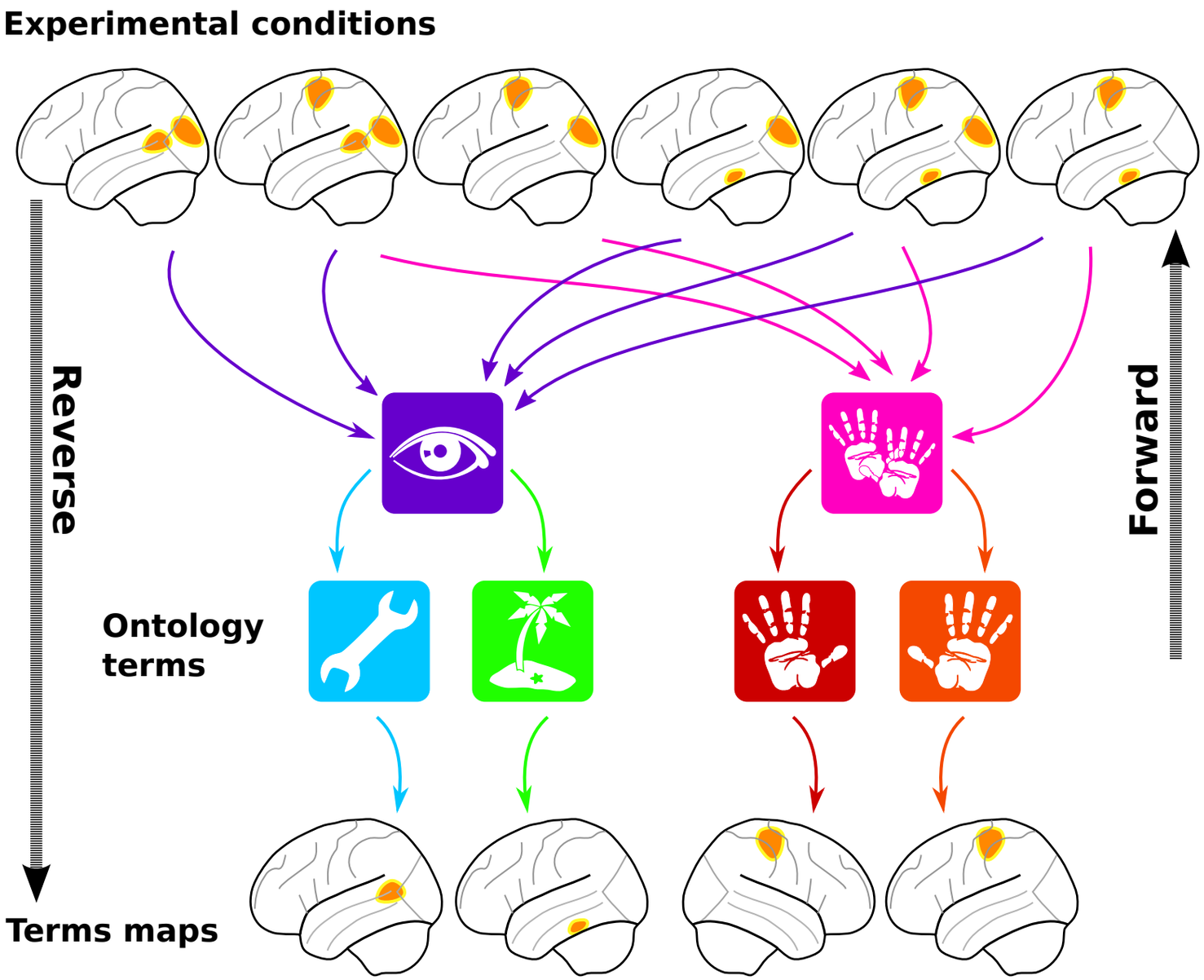


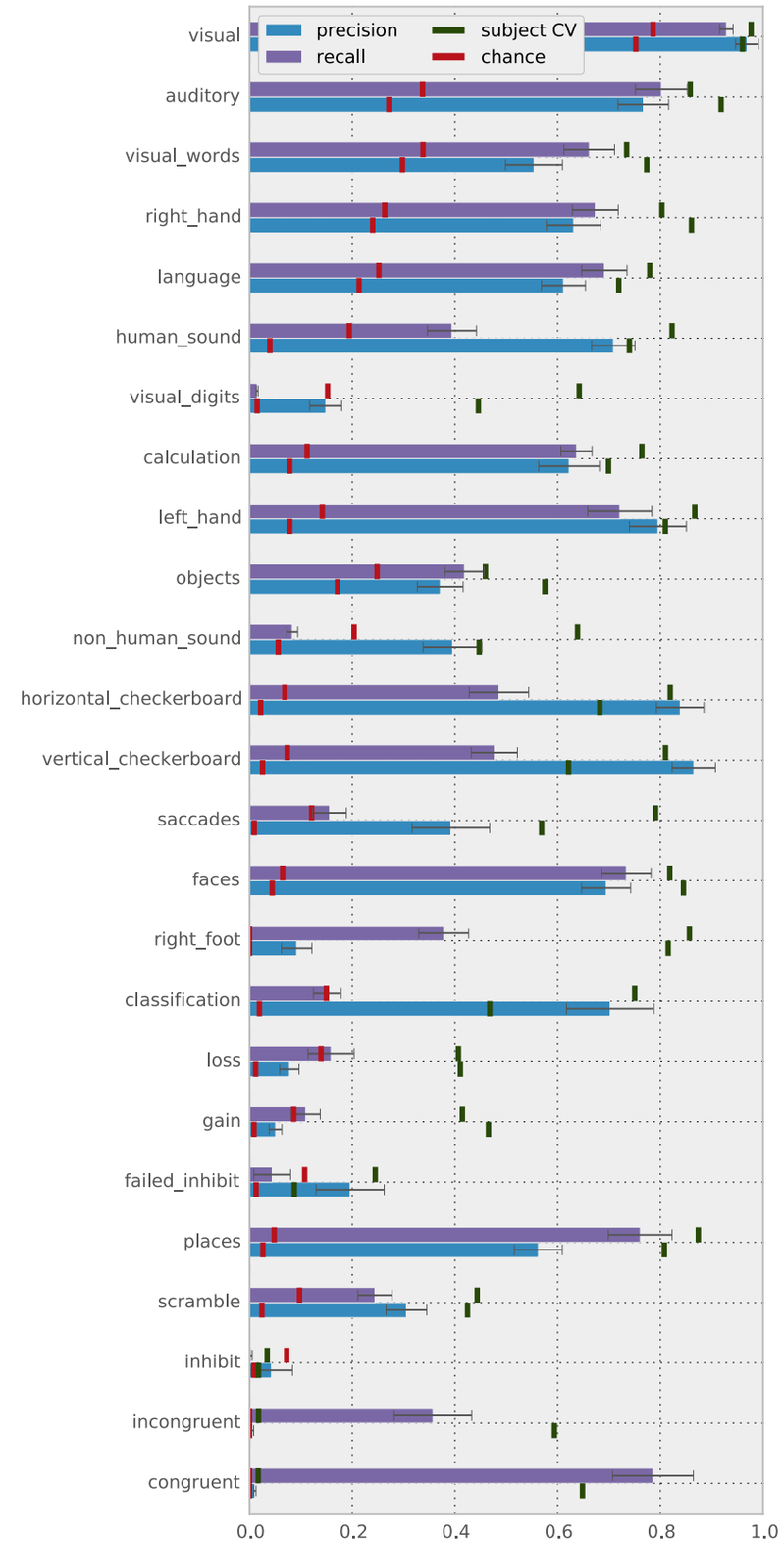
Task contrasts are (mostly) easily classified



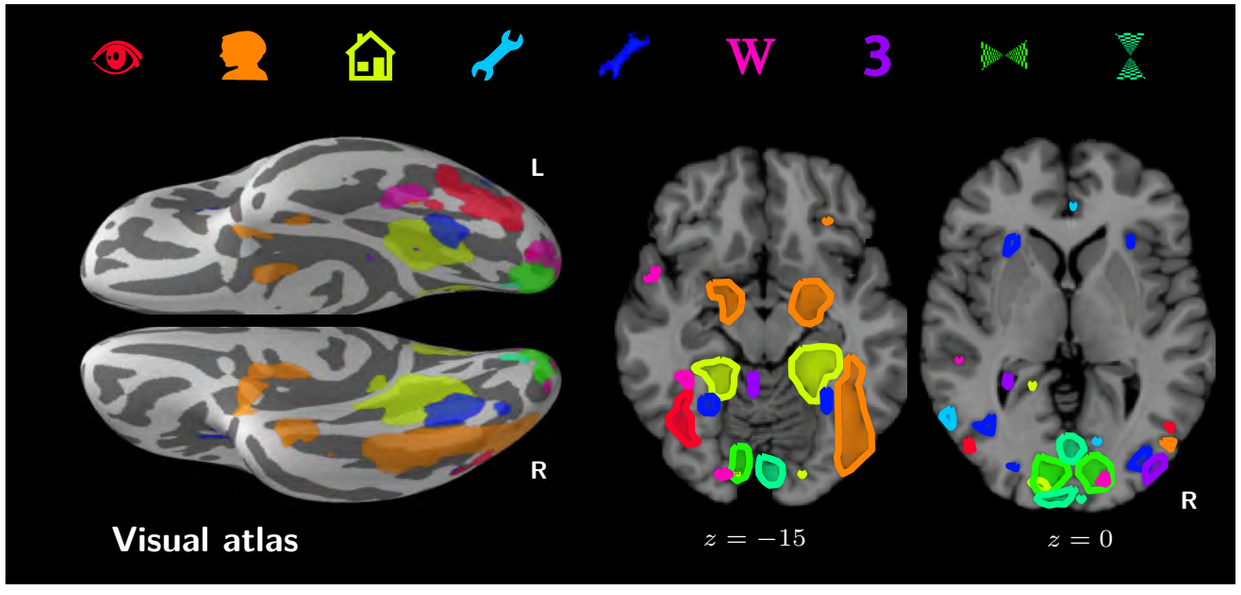
- What we really care about is decoding psychological functions, not tasks

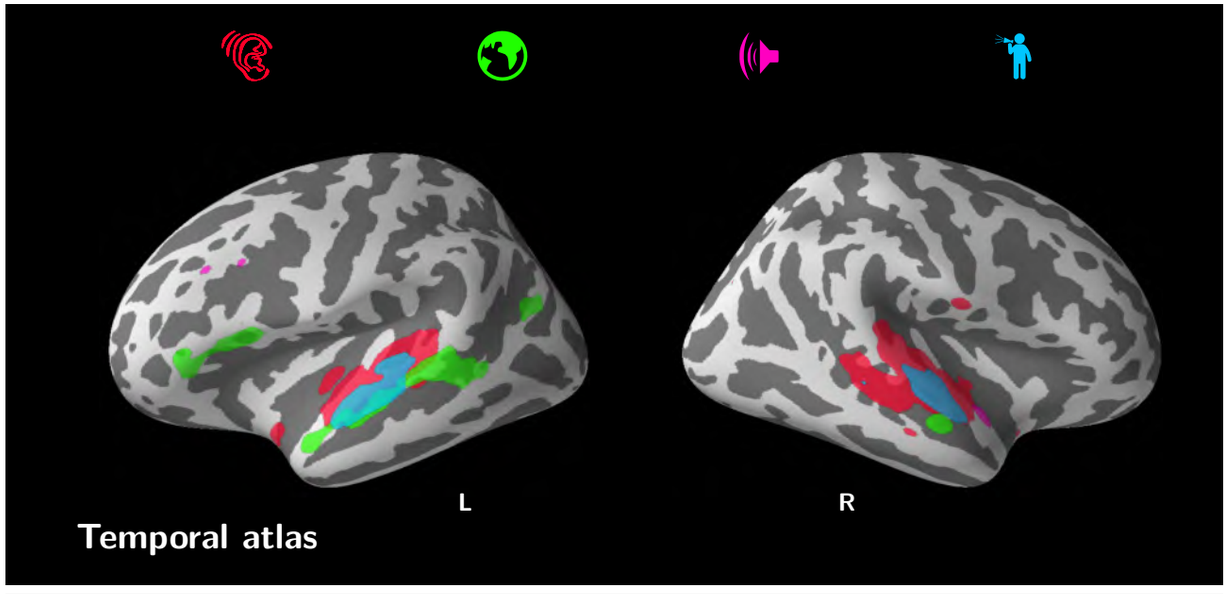
Ontology-based decoding

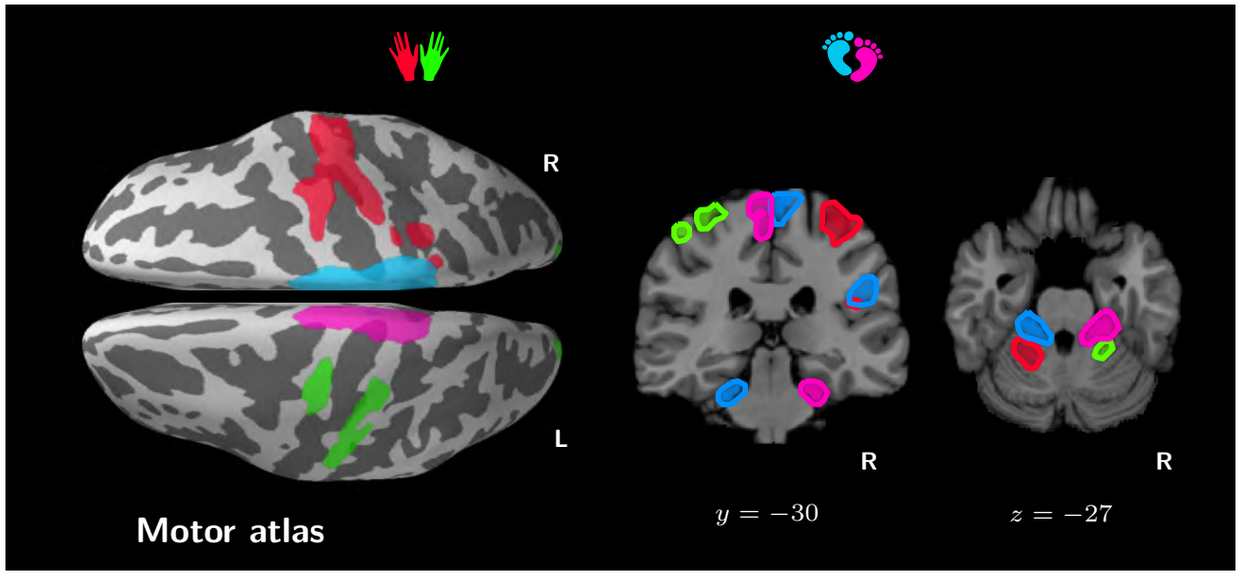


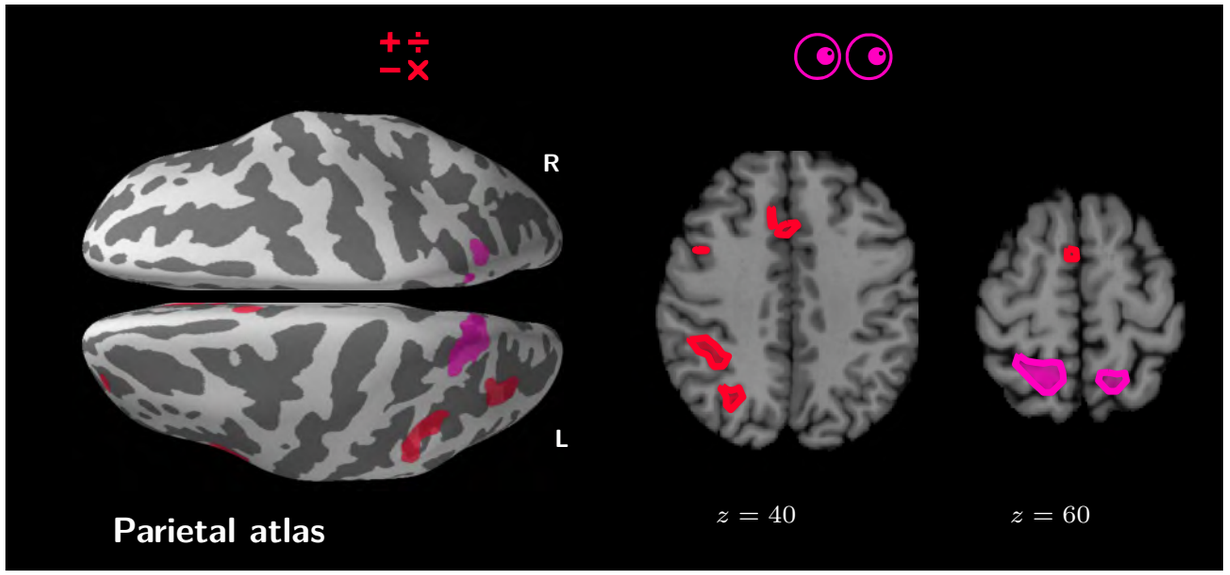


Schwartz et al., in prep

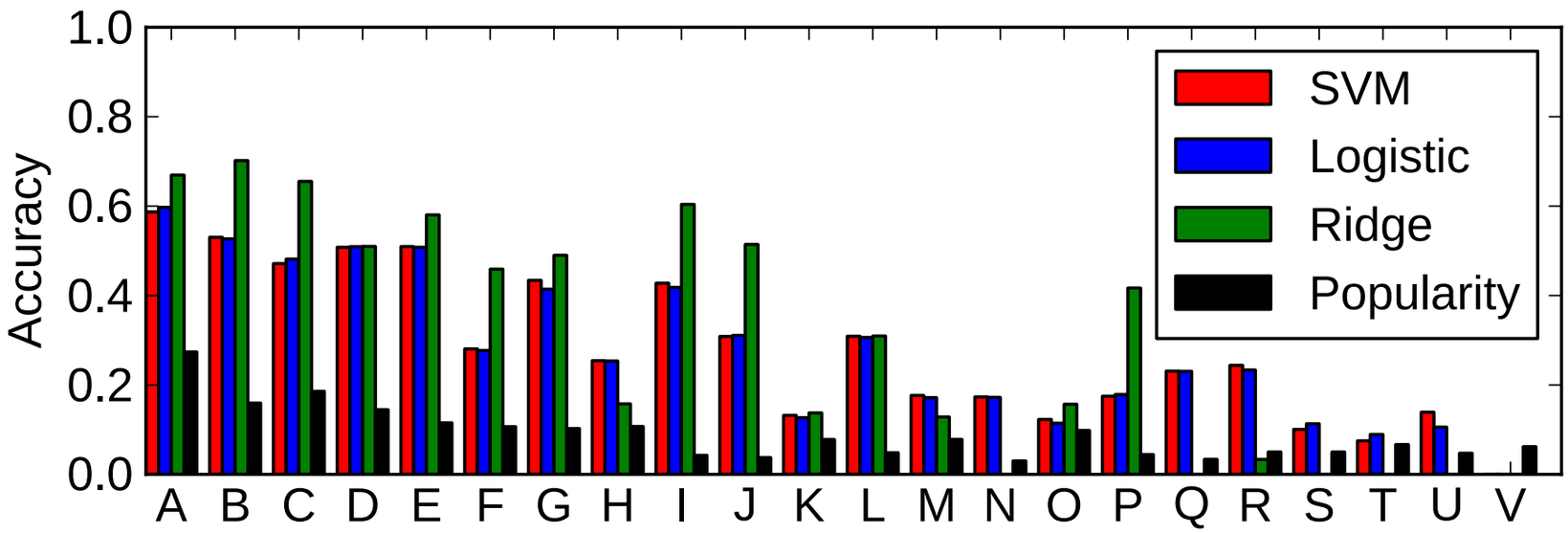

L
R
Visual atlas
$z = -15$
$z = 0$
R


L
R
Temporal atlas


R
L
Motor atlas
R
$y = -30$
R
$z = -27$


R
L
Parietal atlas
$z = 40$
$z = 60$

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

Koyejo & Poldrack, 2013

How might predictive neuroimaging be useful?

Will I get Alzheimer's disease?

Which depression drug is best for this patient?

Will this person commit another crime?

Does this person have terrorist intent?

What do we mean by "prediction"?

- Pubmed search: "predict"[TI] AND "fMRI"
  - 96 results

Presurgery resting-state local graph-theory measures predict neurocognitive outcomes after brain surgery in temporal lobe epilepsy.

Recollection-related increases in functional connectivity predict individual differences in memory accuracy.

Structural Connectivity Fingerprints Predict Cortical Selectivity for Multiple Visual Categories across Cortex.

Individual differences in symptom severity and behavior predict neural activation during face processing in adolescents with autism.

Resting-state networks predict individual differences in common and specific aspects of executive function.

Dissociated signals in human dentate gyrus and CA3 predict different facets of recognition memory.

Network measures predict neuropsychological outcome after brain injury.

First steps in using machine learning on fMRI data to predict intrusive memories of traumatic film footage.

Childhood maltreatment and combat posttraumatic stress differentially predict fear-related fronto-subcortical connectivity.

Global genetic variations predict brain response to faces.

Balancing reward and work: anticipatory brain activation in NAcc and VTA predict effort differentially.

Buffering social influence: neural correlates of response inhibition predict driving safety in the presence of a peer.

Pre-existing brain states predict risky choices.

Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions.

Individual differences in intrinsic brain connectivity predict decision strategy.

Prediction as correlation

The Journal of Neuroscience, January 28, 2015 • 35(4):1763–1772 • 1763

Behavioral/Cognitive

Recollection-Related Increases in Functional Connectivity Predict Individual Differences in Memory Accuracy

Danielle R. King, Marianne de Chastelaine, Rachael L. Elward, Tracy H. Wang, and Michael D. Rugg
Center for Vital Longevity and School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, Texas 75235

"in all three experiments the magnitude of connectivity increases correlated across individuals with recollection accuracy in areas diffusely distributed throughout the brain"

Prediction as prospective correlation

Presurgery resting-state local graph-theory measures predict neurocognitive outcomes after brain surgery in temporal lobe epilepsy

*Gaelle E. Doucet, *Robert Rider, *Nathan Taylor, *Christopher Skidmore, †Ashwini Sharan, *Michael Sperling, and *‡Joseph I. Tracy

Epilepsia, **(*):1–10, 2015
doi: 10.1111/epi.12936

Objective: This study determined the ability of resting-state functional connectivity (rsFC) graph-theory measures to predict neurocognitive status postsurgery in patients with temporal lobe epilepsy (TLE) who underwent anterior temporal lobectomy (ATL).

Linear regression analyses were computed to predict the change in each neurocognitive domain.

Observed correlation ≉ predictive accuracy

J. R. Statist. Soc. B (1983),
45, *No.* 3, *pp.* 311–354

Regression, Prediction and Shrinkage

By J. B. COPAS

University of Birmingham, UK

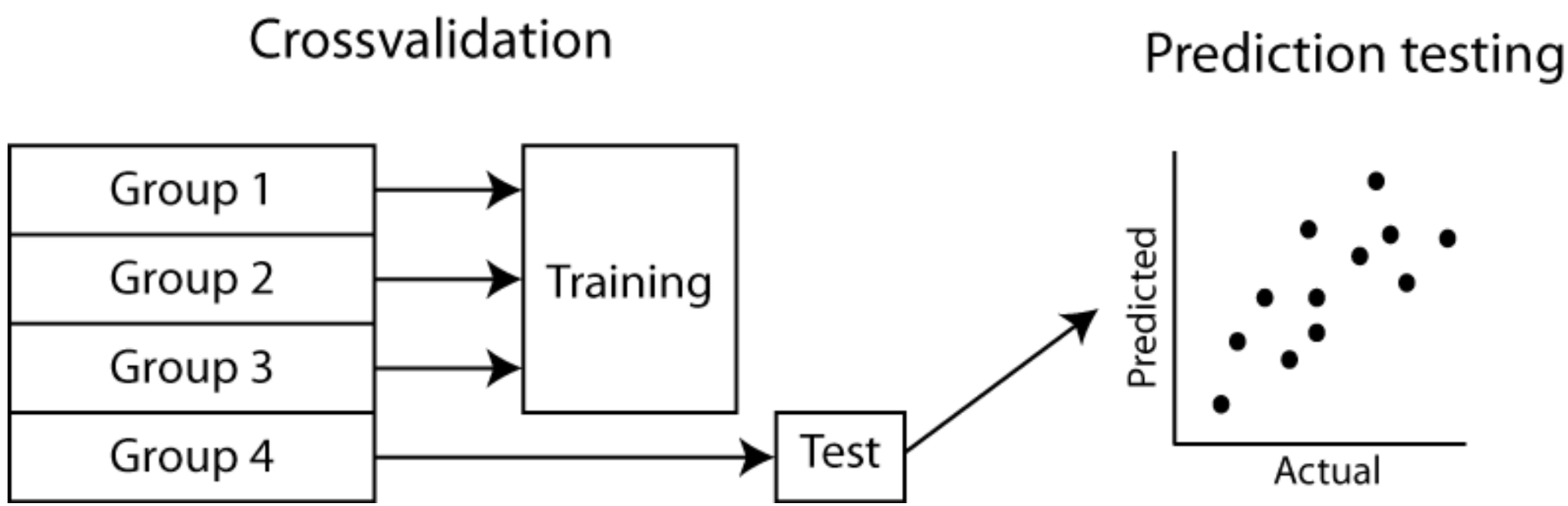
[*Read before the* Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday, January 12th, 1983,* Professor R. N. Curnow *in the Chair*]

SUMMARY

The fit of a regression predictor to new data is nearly always worse than its fit to the original data. Anticipating this shrinkage leads to Stein-type predictors which, under certain assumptions, give a uniformly lower prediction mean squared error than least squares. Shrinkage can be particularly marked when stepwise fitting is used: the shrinkage is then closer to that expected of the full regression rather than of the subset regression actually fitted. Preshrunk predictors for selected subsets are proposed and tested on a number of practical examples. Both multiple and binary (logistic) regression models are considered.

“Since any assessment of retrospective fit "uses the data twice", it is obvious that it gives too optimistic a picture of the validation fit likely to be obtained on new data.”

Testing prediction using cross-validation



Proceedings of the National Academy of Sciences of the United States of America

PNAS

Neuroprediction of future rearrest

Eyal Aharoni[a,b,1,2], Gina M. Vincent[c], Carla L. Harenski[a], Vince D. Calhoun[a,d], Walter Sinnott-Armstrong[e], Michael S. Gazzaniga[f], and Kent A. Kiehl[a,b,2]

Author Affiliations

Edited by Robert Desimone, Massachusetts Institute of Technology, Cambridge, MA, and approved February 27, 2013 (received for review November 7, 2012)

Published online before print March 27, 2013, doi: 10.1073/pnas.1219302110
PNAS **March 27, 2013**

"The present analysis shows that hemodynamic activity within the brain prospectively predicted rearrest in an offender sample."

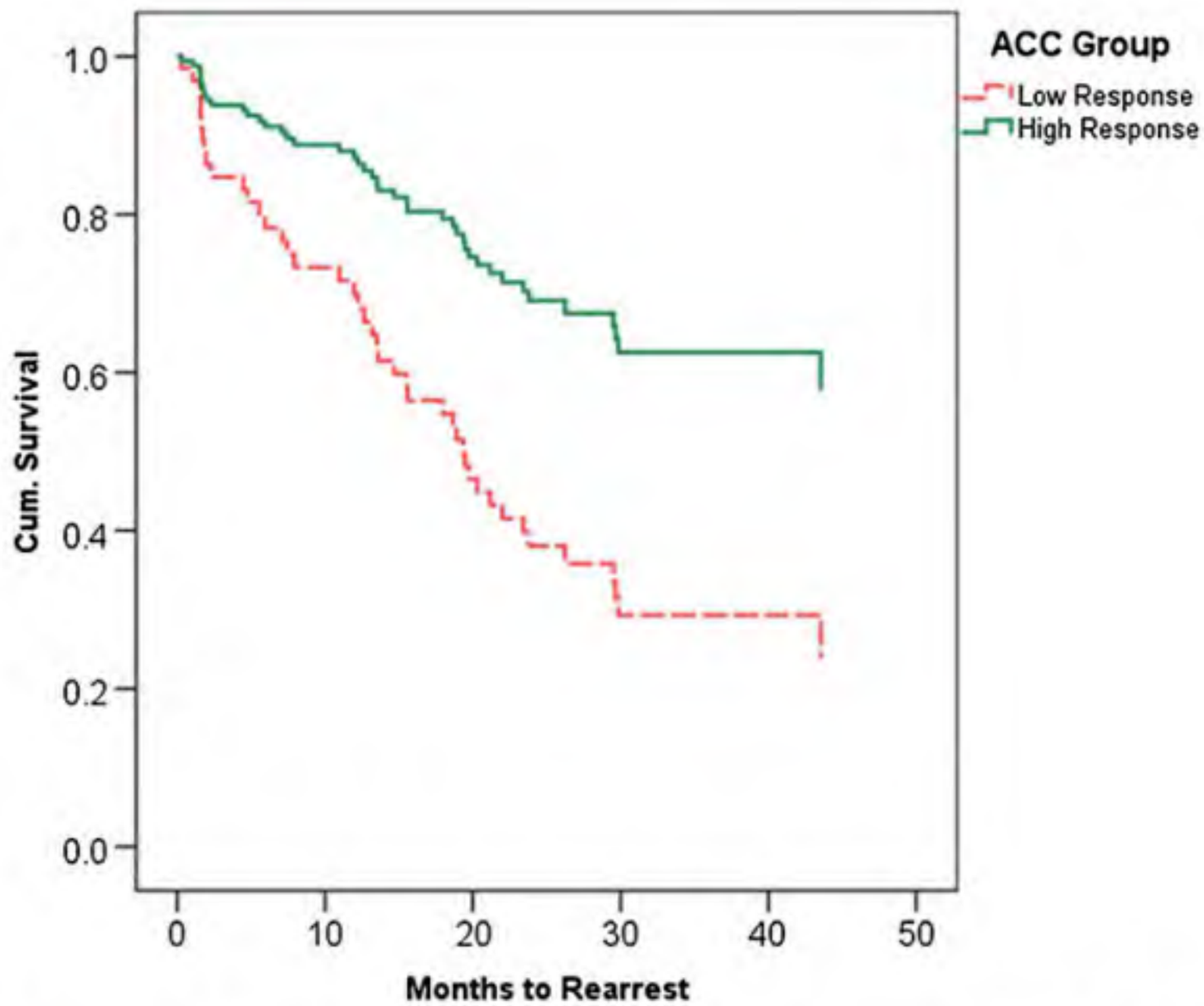
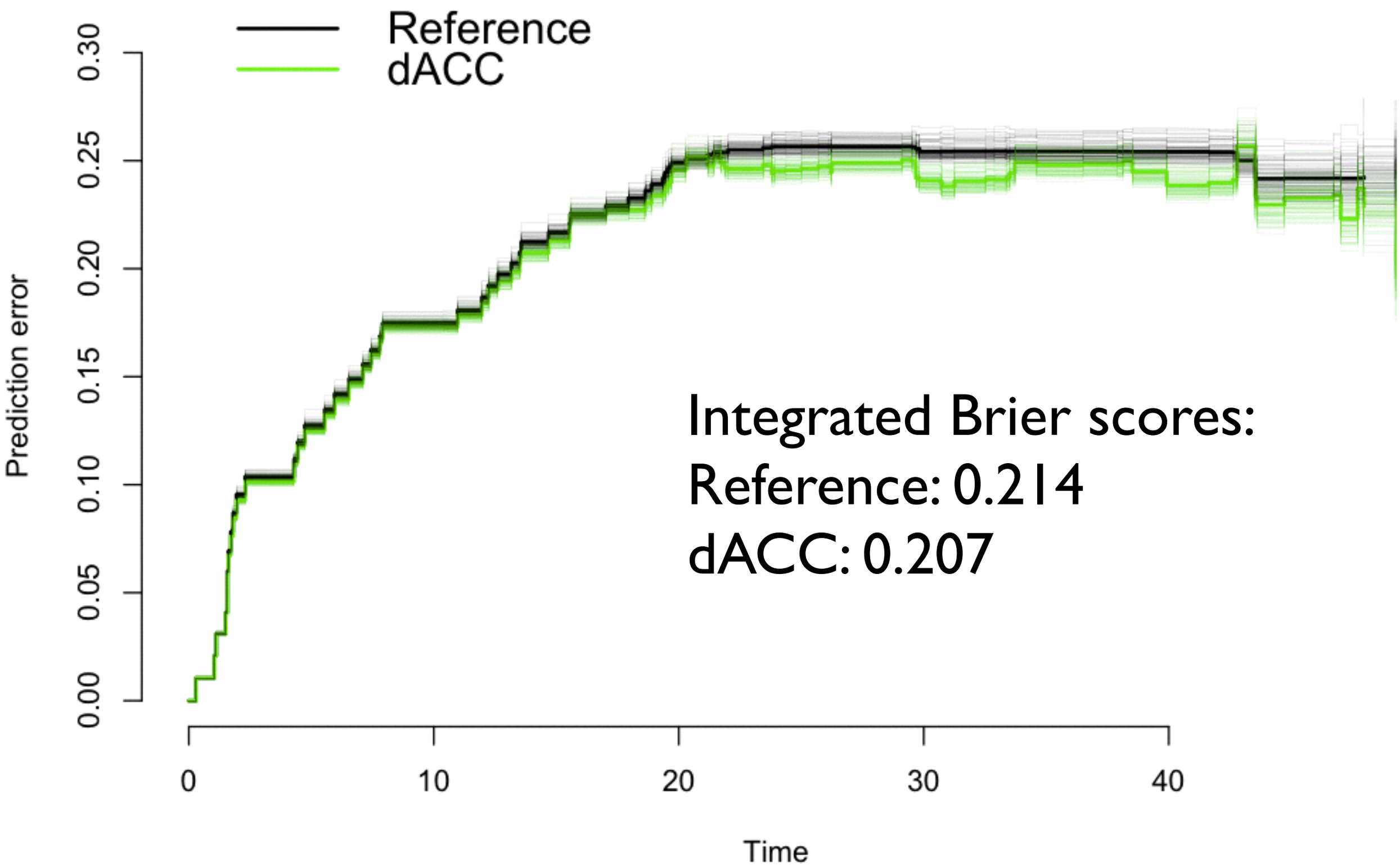


Fig. 1. Cox survival function showing proportional rearrest survival rates of high (solid green) vs. low (dashed red) ACC response groups for any crime over a 4-y period. Results of this median split analysis were equivalent to that of the parametric model: bootstrapped $B = 0.96$; SE = 0.40; $P < 0.01$; 95% CI, 0.29–1.84. The mean survival times to rearrest for the low and high ACC activity groups were 25.27 (2.80) mo and 32.42 (2.73) mo, respectively. The overall probabilities of rearrest were 60% for the low ACC group and 46% for the high ACC group.

Kiehl et al., 2013

Prediction error using crossvalidation



Integrated Brier scores:
Reference: 0.214
dACC: 0.207

http://www.russpoldrack.org/2013/04/how-well-can-we-predict-future-criminal.html

https://github.com/poldrack/criminalprediction

"Diagnosing" psychiatric disorders

IOP Publishing
Journal of Neural Engineering
J. Neural Eng. **7** (2010) 016011 (7pp)
doi:10.1088/1741-2560/7/1/016011

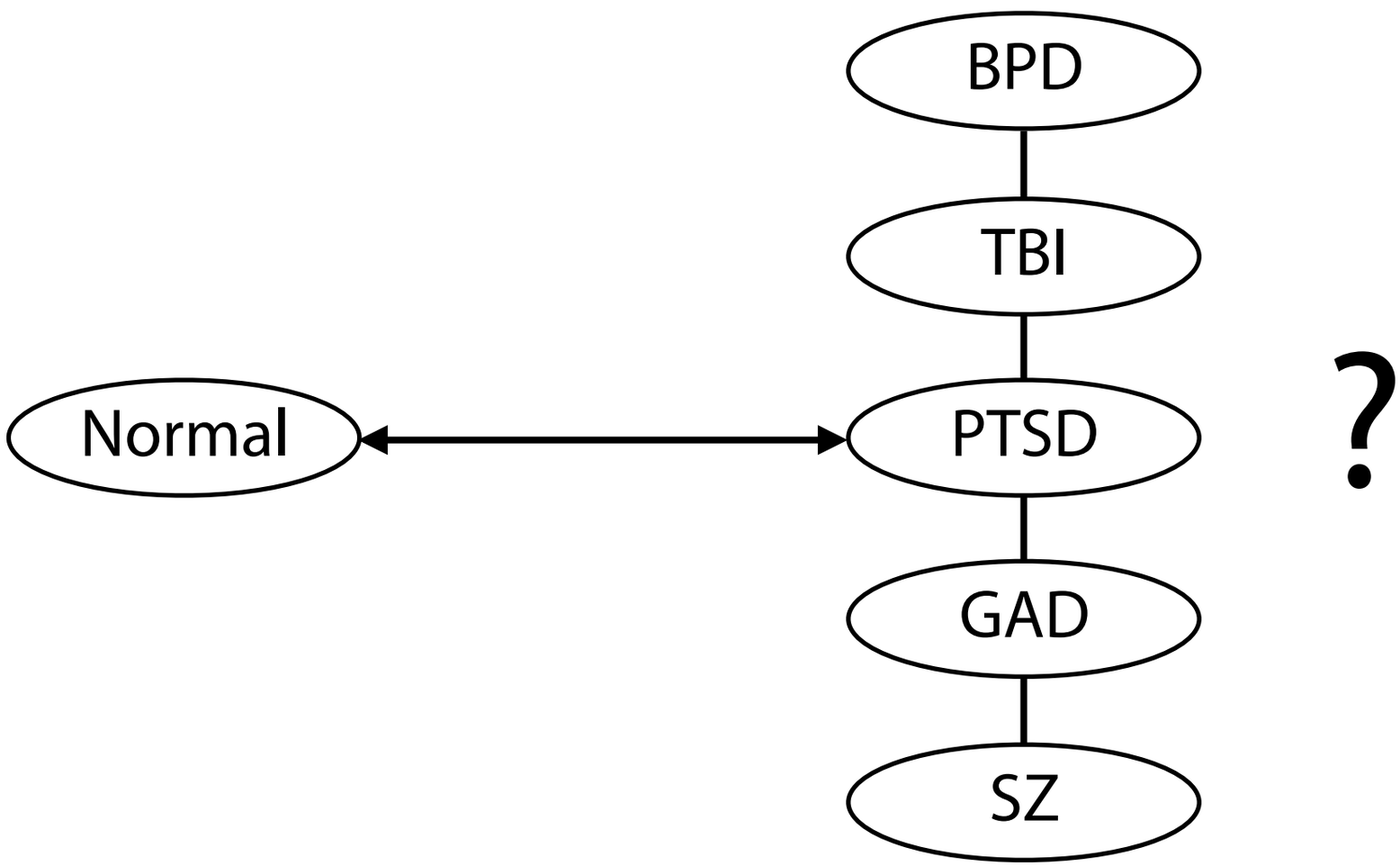
The synchronous neural interactions test as a functional neuromarker for post-traumatic stress disorder (PTSD): a robust classification method based on the bootstrap

A P Georgopoulos[1,2,4,5,6,9], H-R M Tan[1,2,8], S M Lewis[1,3], A C Leuthold[1,2], A M Winskowski[1,6], J K Lynch[1,2] and B Engdahl[6,7]

"Here we show that the synchronous neural interactions (SNI) test which assesses the functional interactions among neural populations derived from magnetoencephalographic (MEG) recordings can successfully *differentiate PTSD patients from healthy control subjects*. ... Altogether, these findings document robust differences in brain function between the PTSD and control groups that *can be used for differential diagnosis*..." (my italics)

Detection versus differential diagnosis

- The ability of a biomarker to discriminate between patient and normal does not demonstrate its ability to diagnose the disorder versus other disorders
  - "brain schmutz" (R. Bilder)



Disease "diagnosis": A cautionary tale

frontiers in
SYSTEMS NEUROSCIENCE

OPINION ARTICLE
published: 05 September 2012
doi: 10.3389/fnsys.2012.00062

The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience

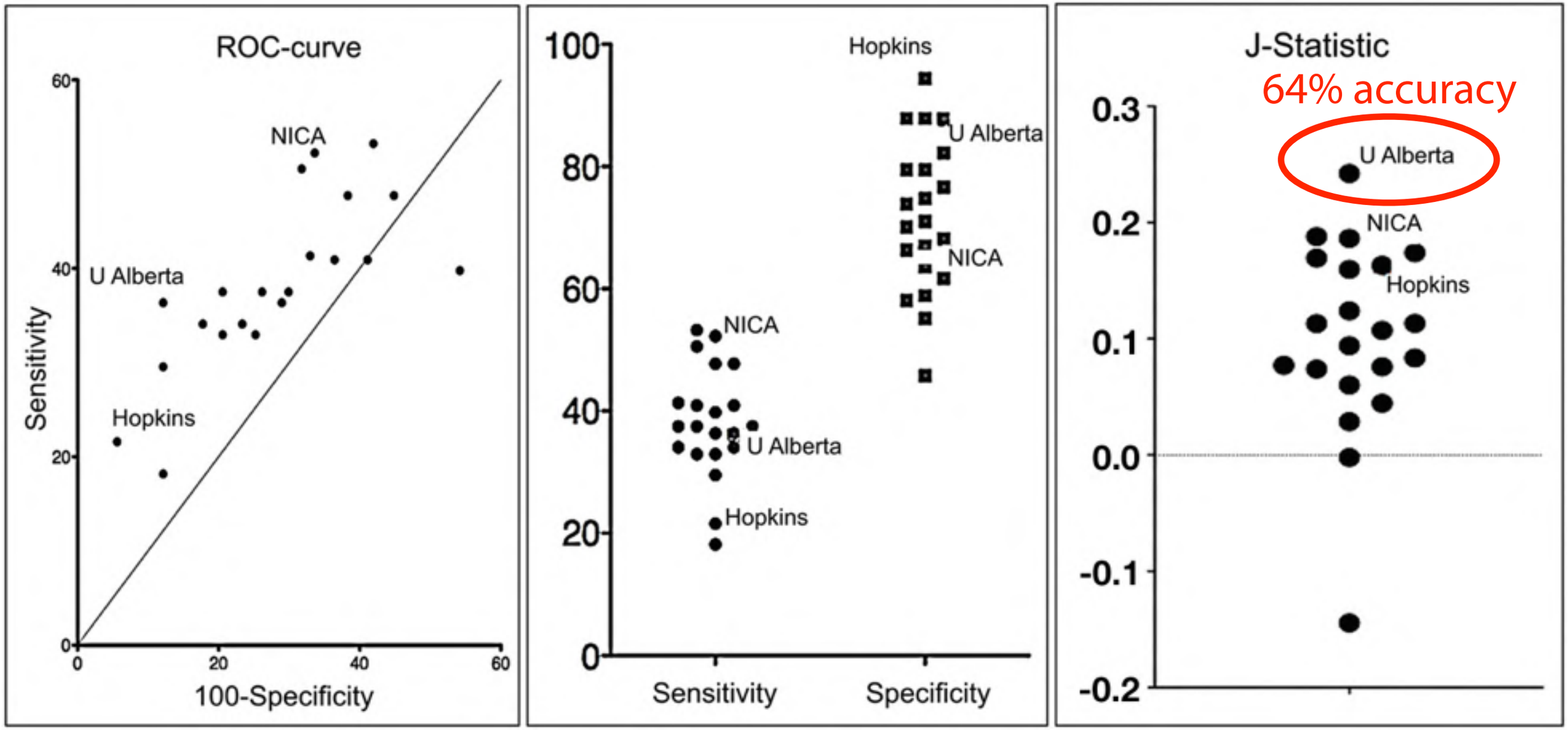
*The ADHD-200 Consortium**

Challenge: Diagnose typical controls from ADHD, and diagnose ADHD subtypes (attentive vs combined)

Training data (released 3/1/2011):
- 491 typical controls
- 285 ADHD

Test data (released 6/1/2011):
- 197 unlabeled datasets

ADHD-200 results



Sensitivity: ability to identify true ADHD cases
Specificity: ability to correctly identify healthy cases
J-statistic: sensitivity + specificity -1

How did Alberta win?

frontiers in
SYSTEMS NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 28 September 2012
doi: 10.3389/fnsys.2012.00069

ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements

Matthew R. G. Brown[1]*, Gagan S. Sidhu[2,3], Russell Greiner[2,3], Nasimeh Asgarian[2,3], Meysam Bastani[2,3], Peter H. Silverstone[1], Andrew J. Greenshaw[1] and Serdar M. Dursun[1]

[1] *Department of Psychiatry, University of Alberta, Edmonton, AB, Canada*
[2] *Department of Computing Science, University of Alberta, Edmonton, AB, Canada*
[3] *Alberta Innovates Centre for Machine Learning, Edmonton, AB, Canada*

- Used only age, sex, handedness, and IQ
- Capitalized on differences in gender and IQ between patient and control groups
- Brings results of other groups into question: where they also capitalizing on these confounds?

Why decoding DSM diagnoses may be a fool's errand

Original Investigation | META-ANALYSIS

Identification of a Common Neurobiological Substrate for Mental Illness

Madeleine Goodkind, PhD; Simon B. Eickhoff, DrMed; Desmond J. Oathes, PhD; Ying Jiang, MD; Andrew Chang, BS; Laura B. Jones-Hagata, MA; Brissa N. Ortega, BS; Yevgeniya V. Zaiko, BA; Erika L. Roach, BA; Mayuresh S. Korgaonkar, PhD; Stuart M. Grieve, DPhil; Isaac Galatzer-Levy, PhD; Peter T. Fox, MD; Amit Etkin, MD, PhD

Our findings suggest that a general mapping exists between a broad range of symptoms and the integrity of an anterior insula/dACC-based network across a wide variety of neuropsychiatric illnesses. These results do not imply that phenotypic differences between diagnoses are negligible. Rather, they provide an organizing model that emphasizes the import of shared endophenotypes across psychopathology, which is not currently an explicit component of psychiatric nosology. This

Lancet. 2013 April 20; 381(9875): 1371–1379. doi:10.1016/S0140-6736(12)62129-1.

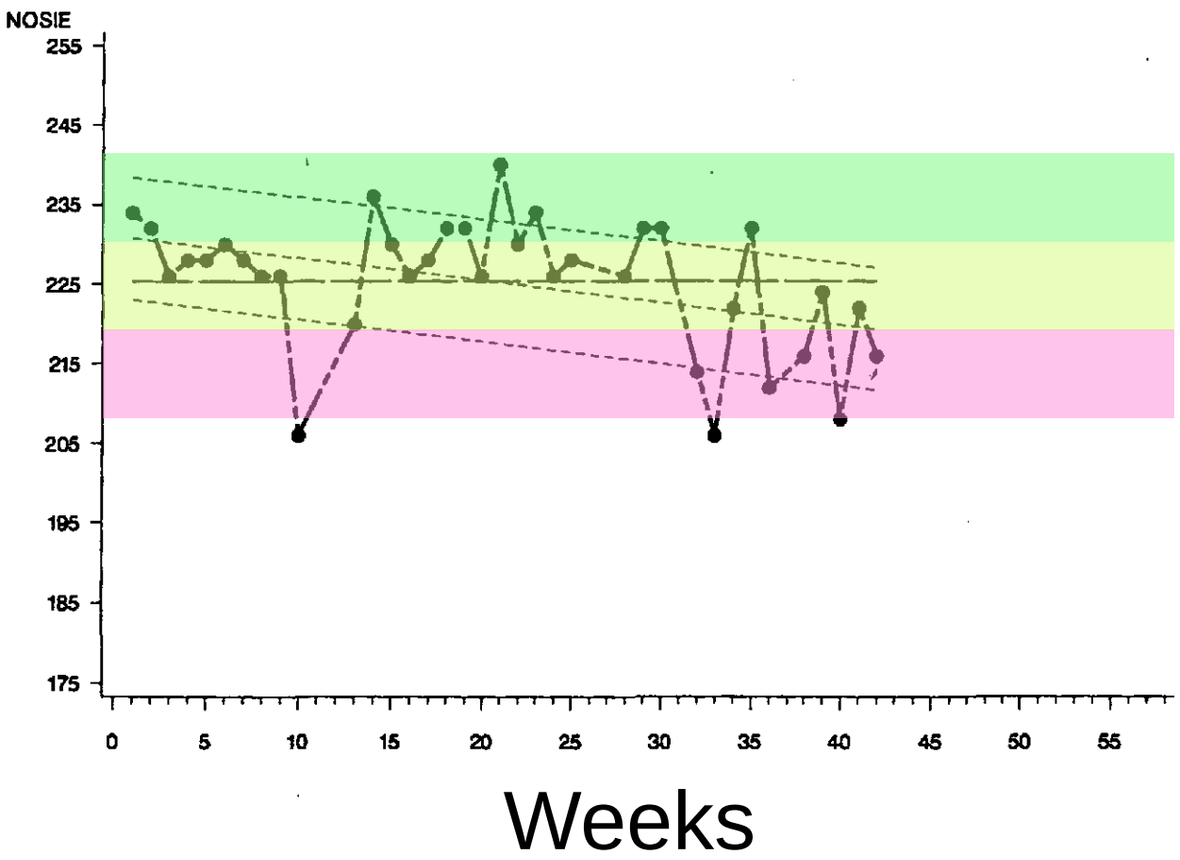
Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis

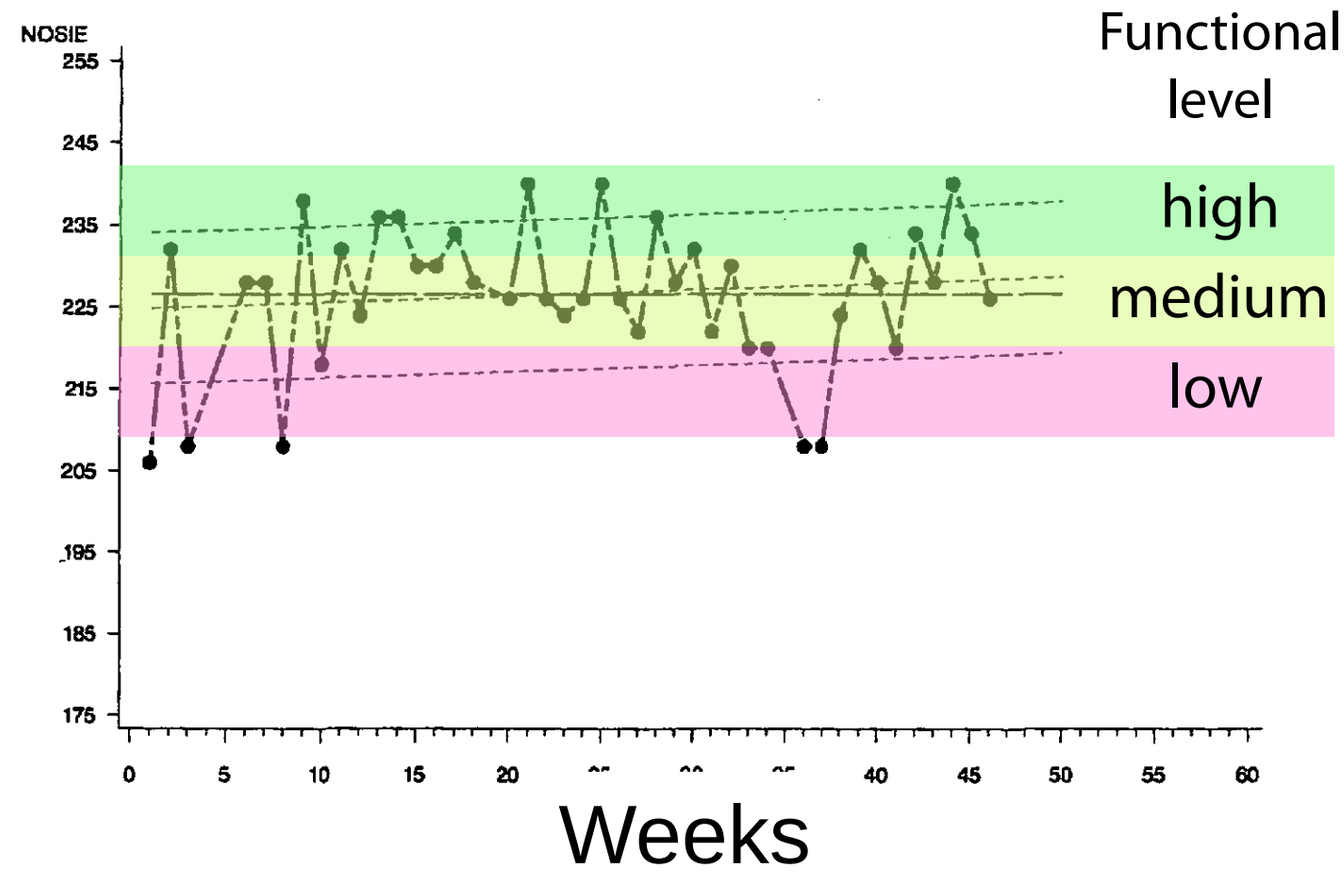
Cross-Disorder Group of the Psychiatric Genomics Consortium*

Our findings show that specific SNPs are associated with a range of psychiatric disorders of childhood onset or adult onset. In particular, variation in calcium-channel activity genes seems to have pleiotropic effects on psychopathology. These results provide evidence relevant to the goal of moving beyond descriptive syndromes in psychiatry, and towards a nosology informed by disease cause.

Psychiatric disease is not a steady state

Fluctuations in daily life function of schizophrenic patients



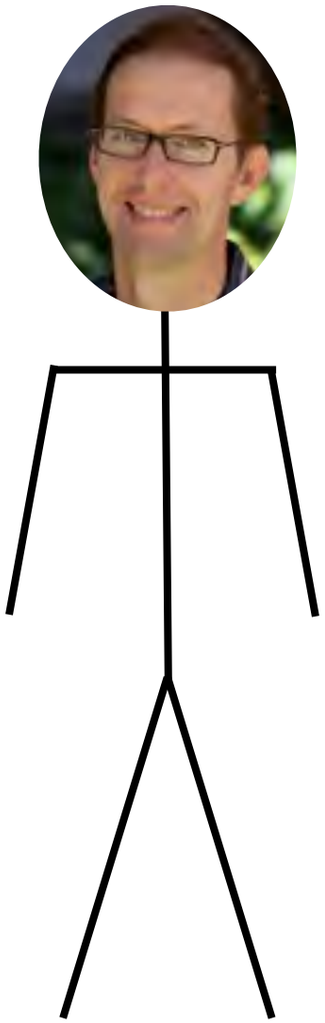


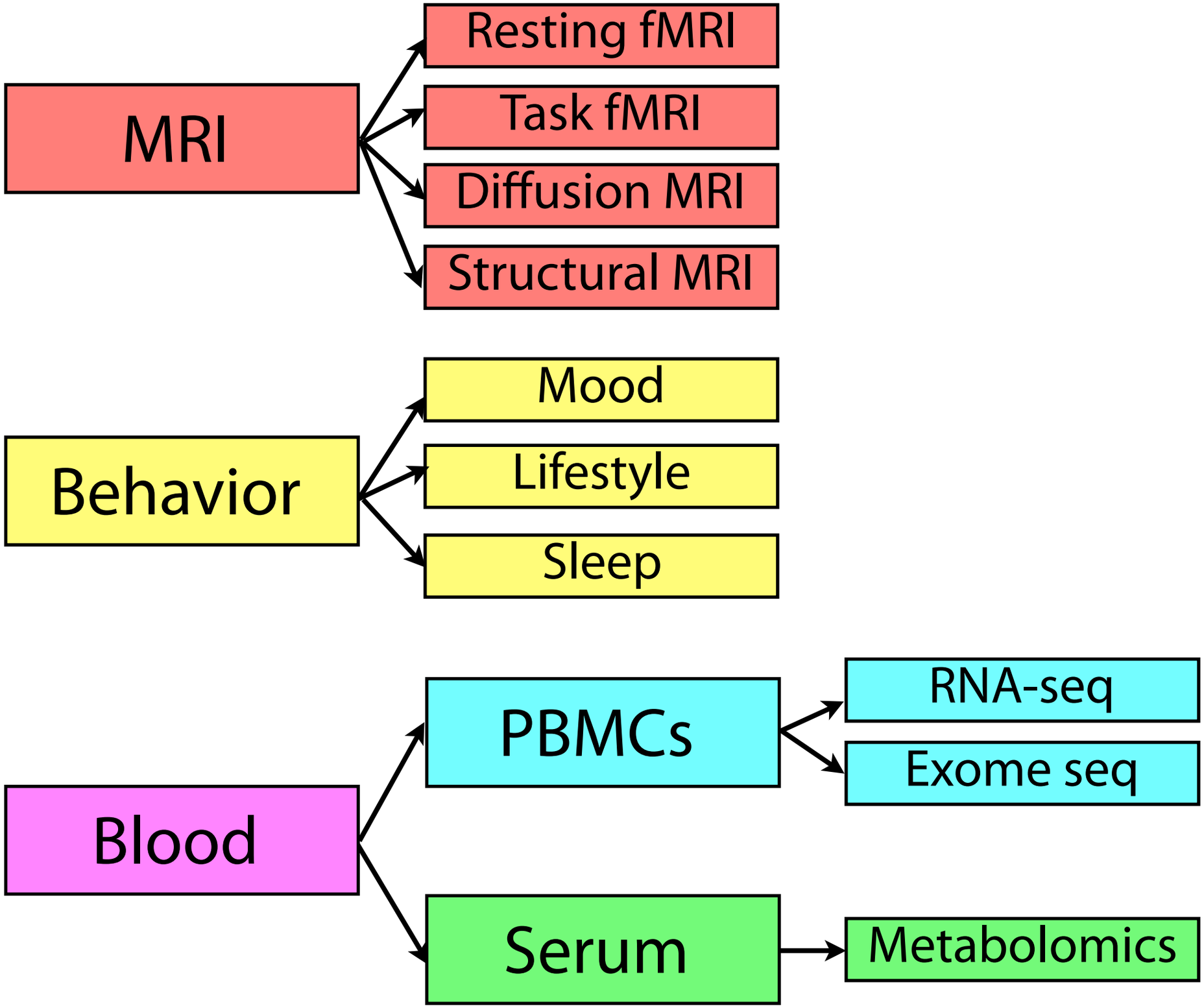
Kupper & Hoffman, 2000, Schiz. Bull.

How does the healthy brain fluctuate over time?

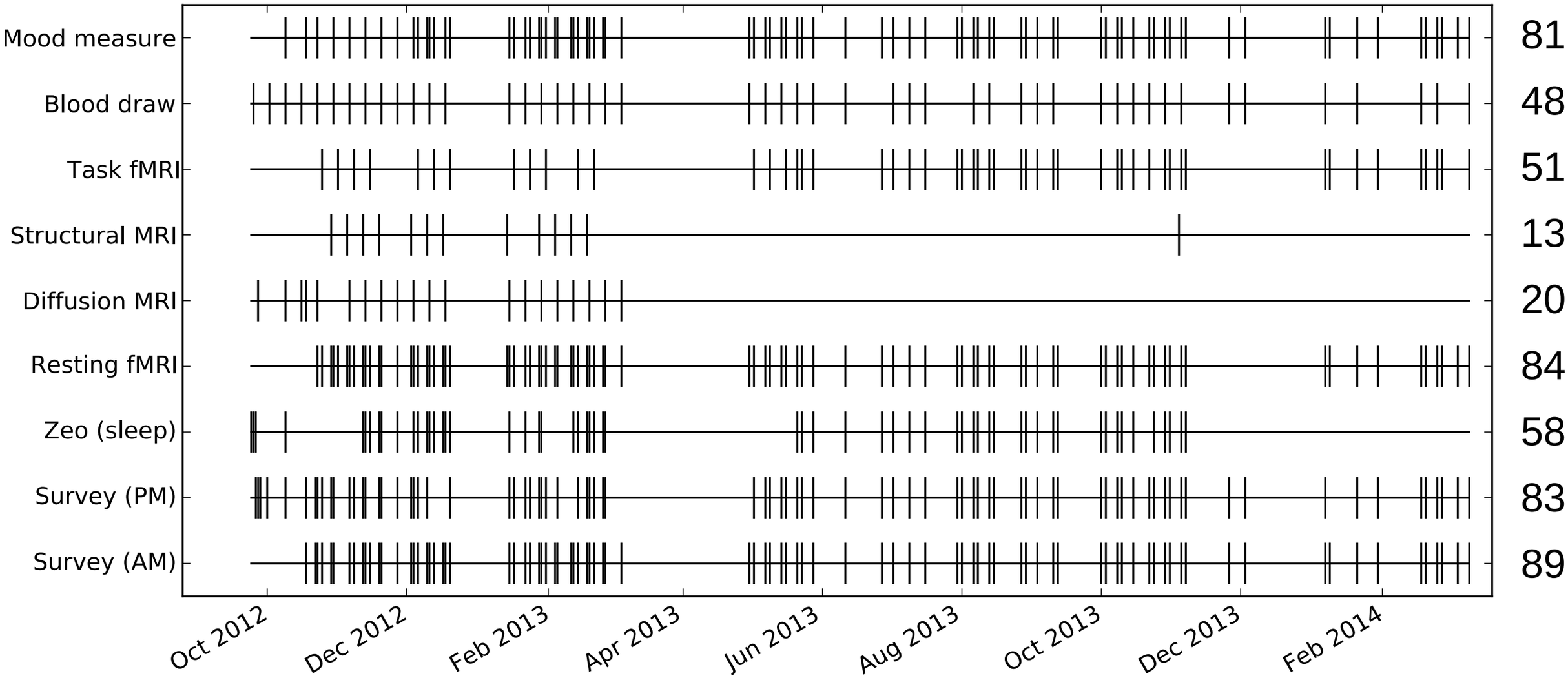
?

The MyConnectome Project



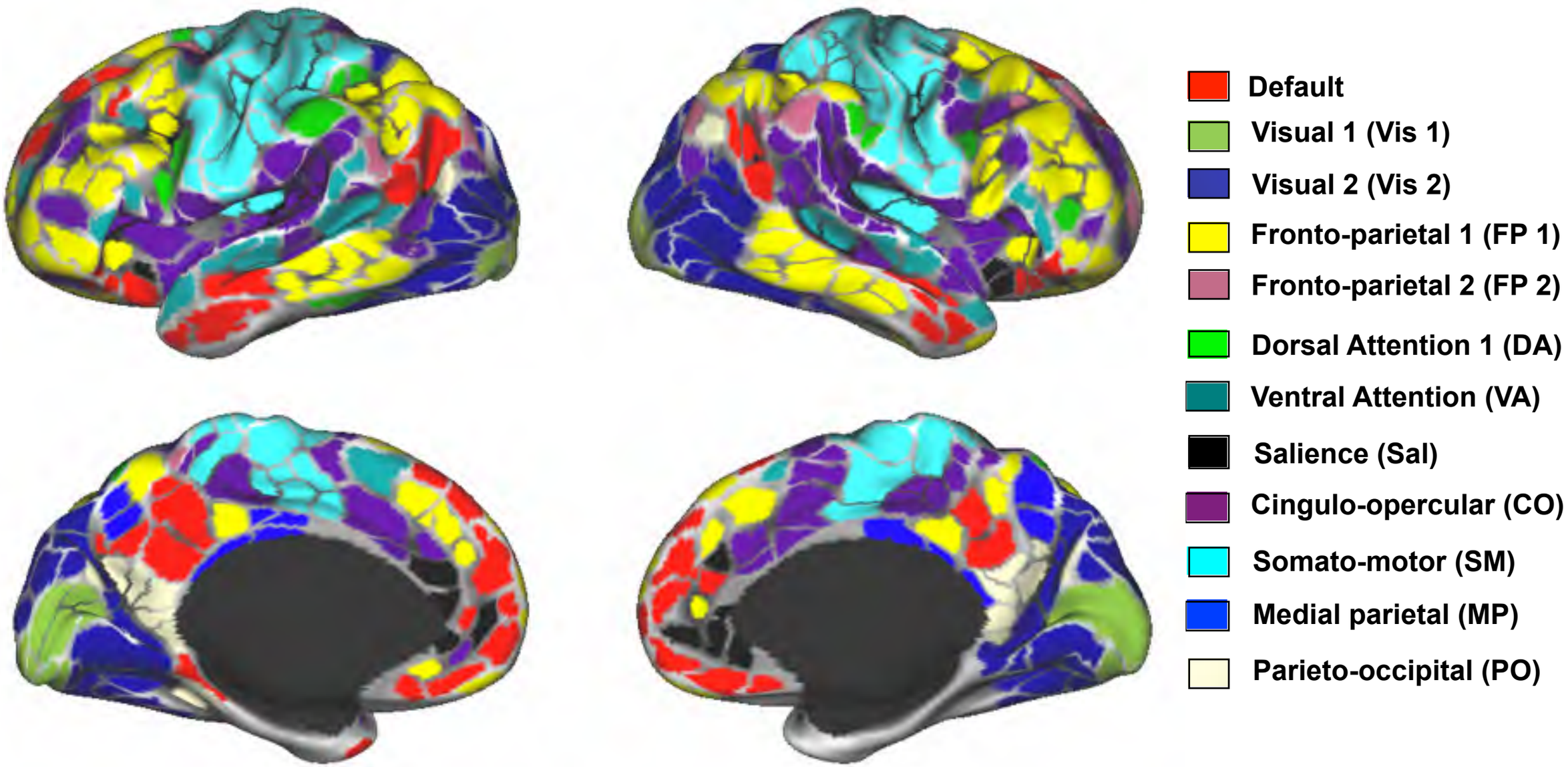


Timeline and measurements



Resting state fMRI: Surface parcellation

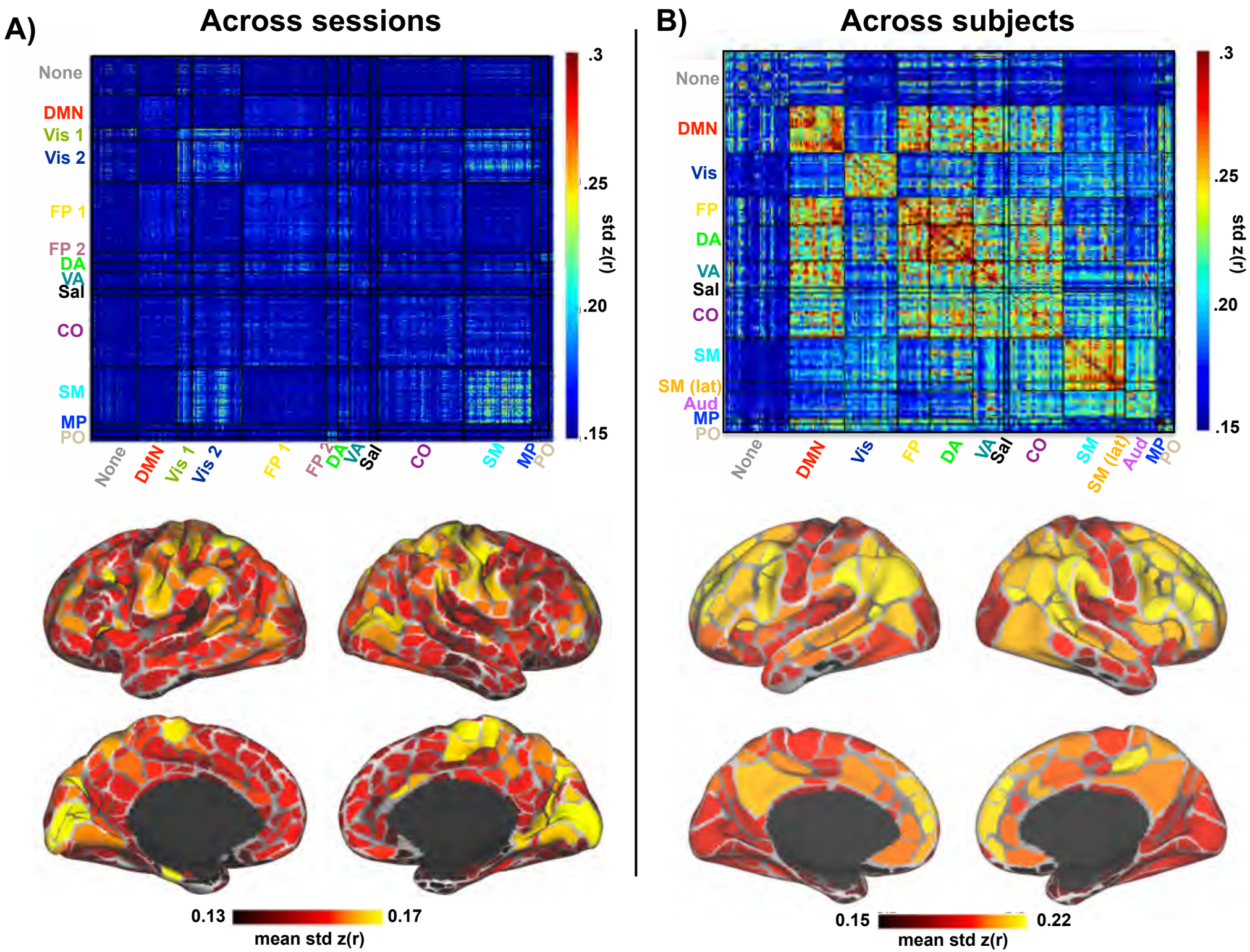
616 cortical parcels + 14 subcortical regions



Similar community structure to previous group studies
(Power et al., 2011)

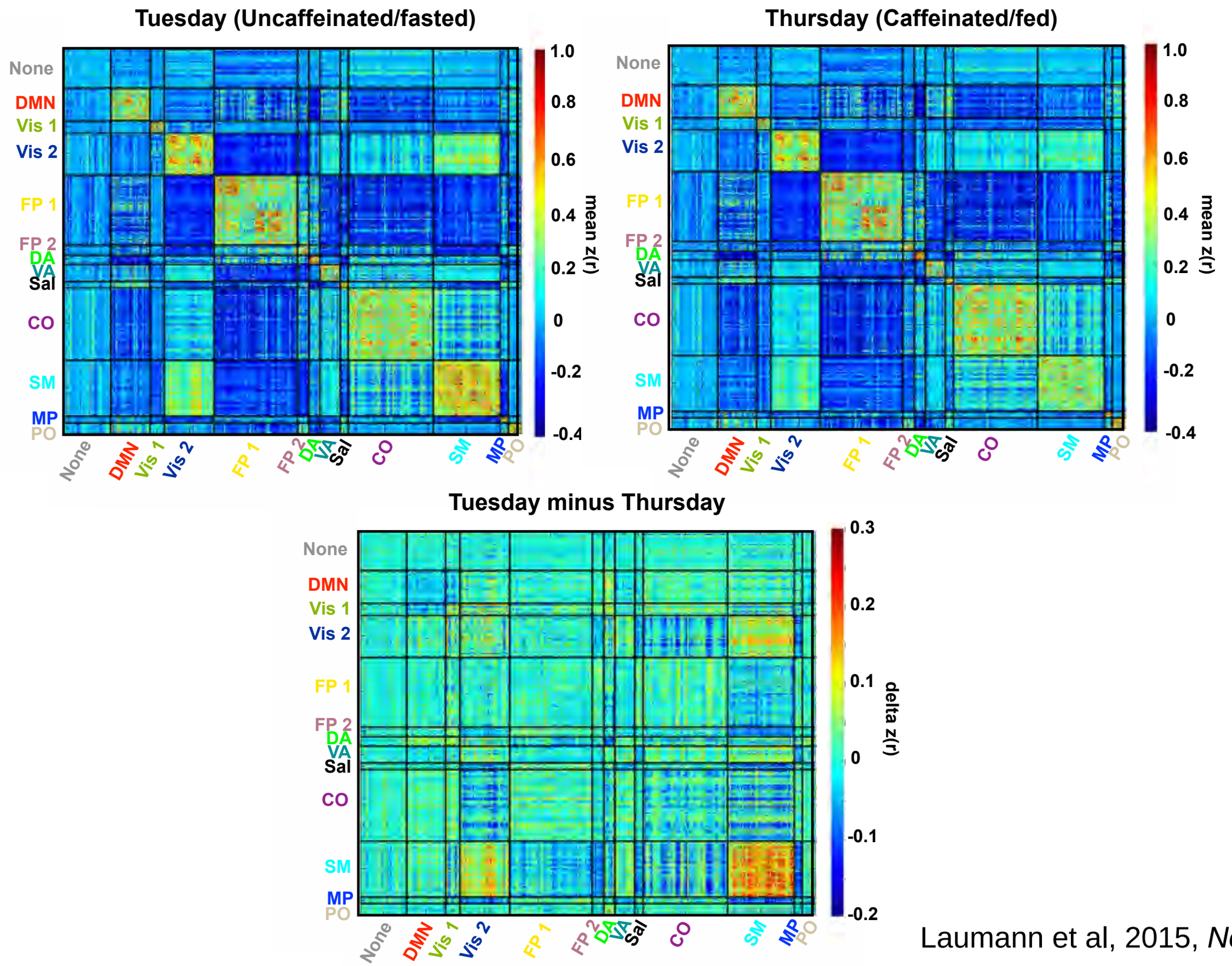
Laumann et al, 2015, *Neuron*

Within-subject and between subject variability differ drastically



Laumann et al, 2015, *Neuron*

Food/caffeine drastically affects brain connectivity

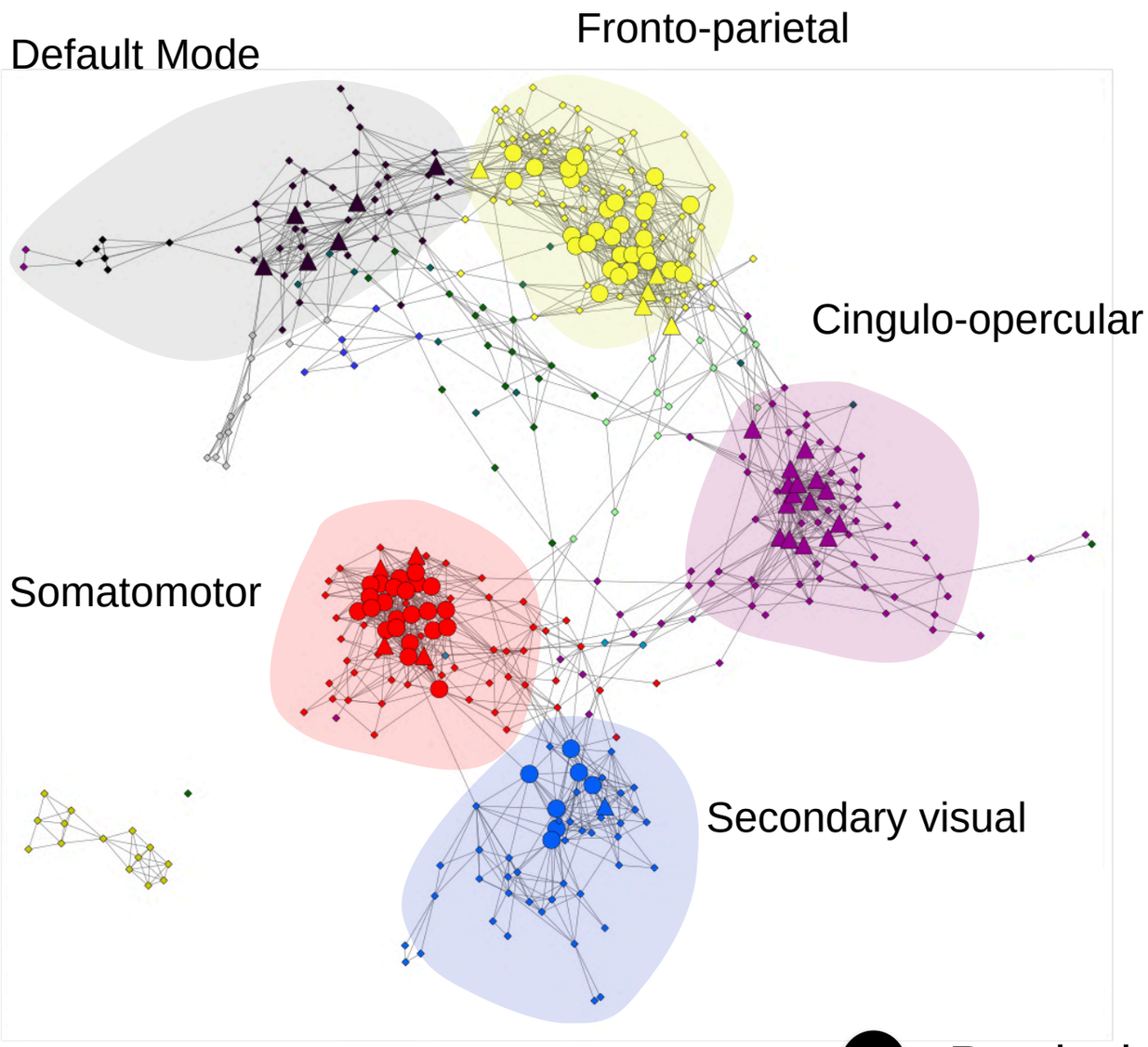


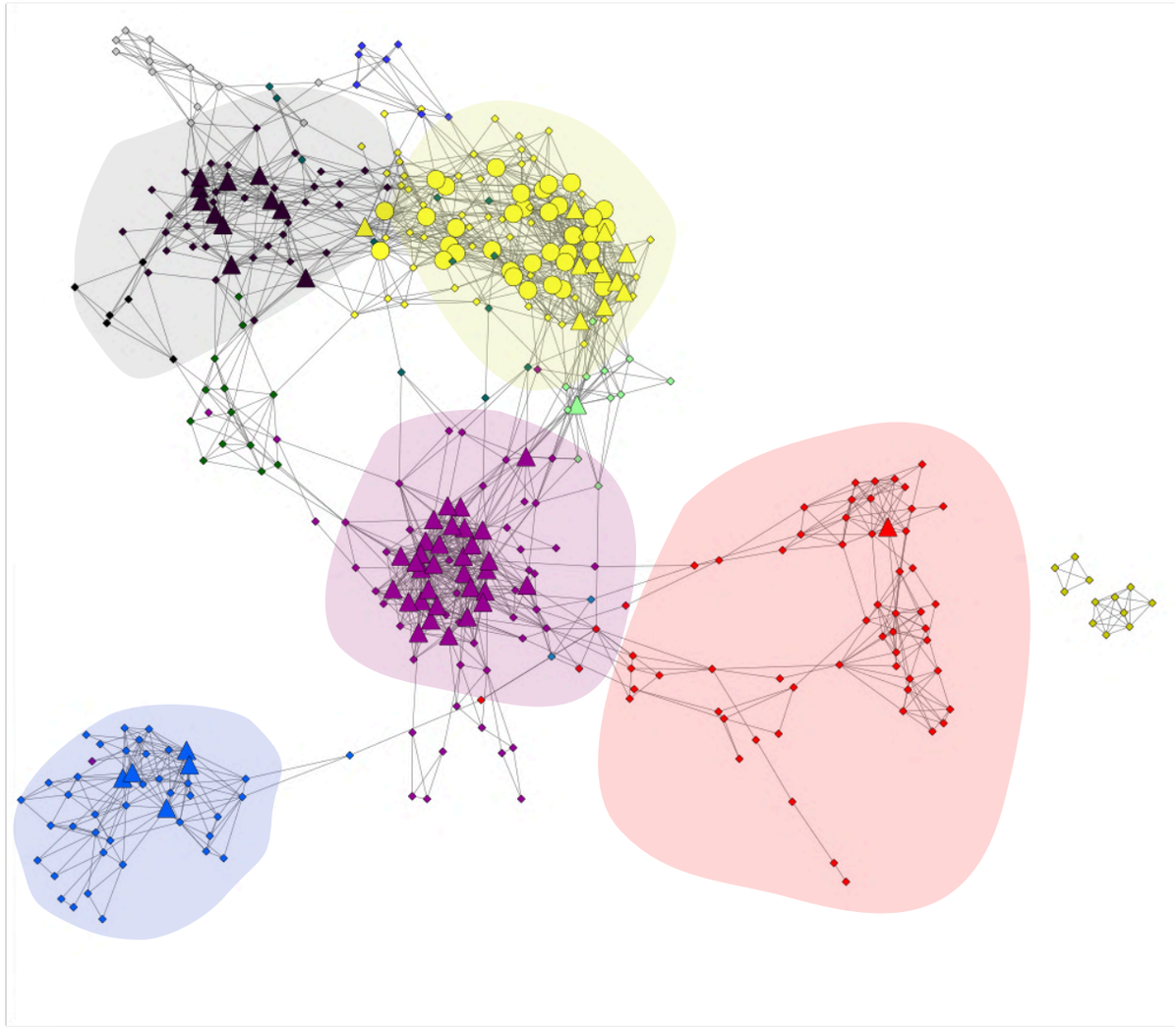
Laumann et al, 2015, *Neuron*

Caffeine/food affects large-scale network structure

Tuesday (fasted/no caffeine)

Thursday (fed/caffeinated)





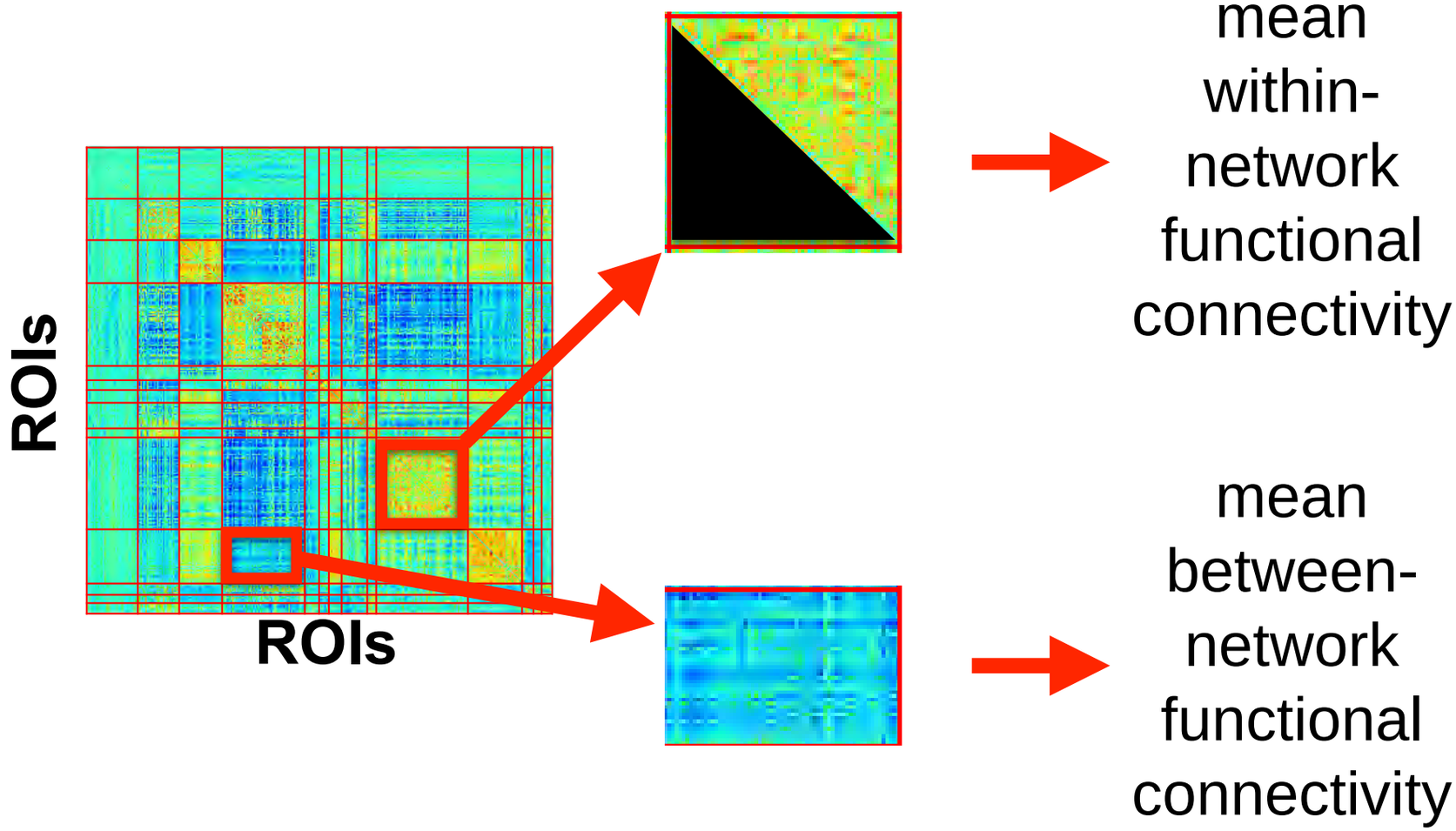
Provincial hub

Connector hub

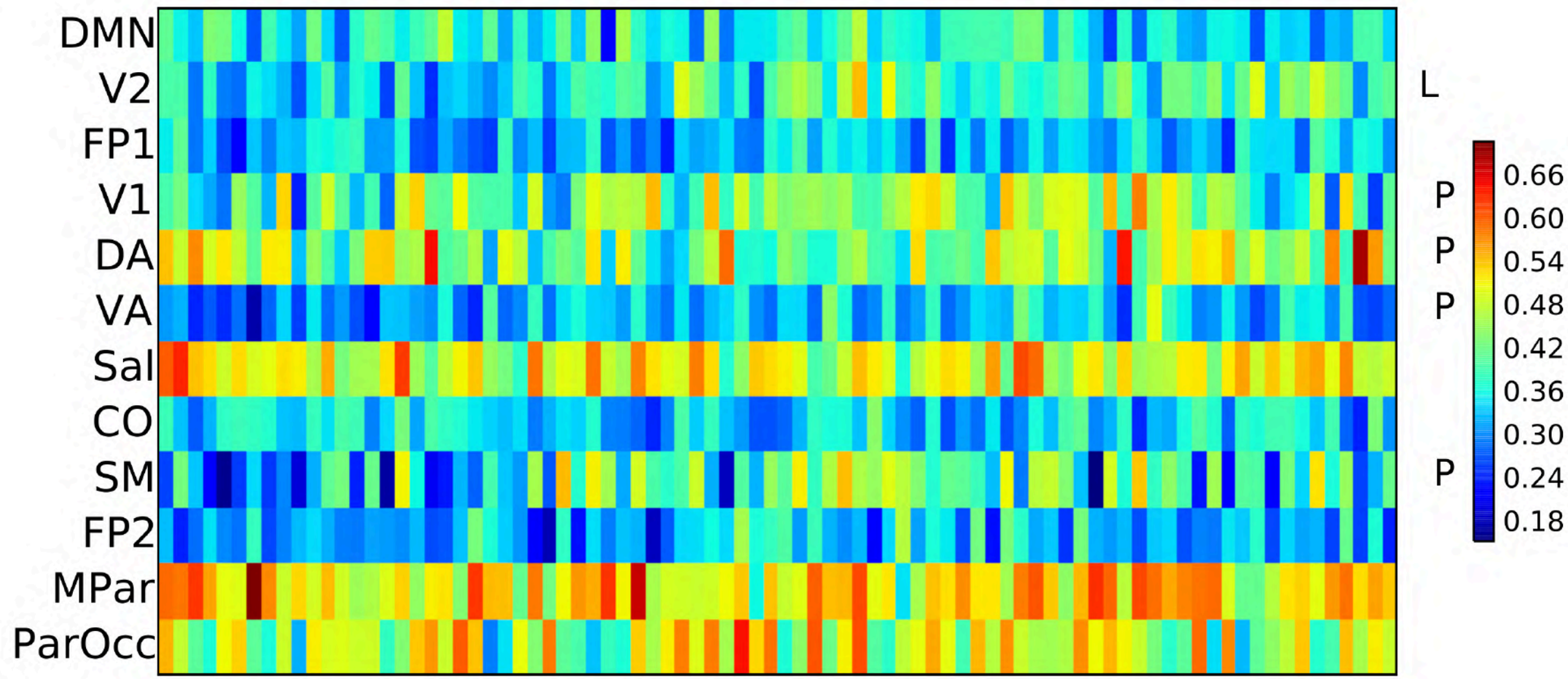
Poldrack et al, submitted

Reduced rsfMRI dimensionality: within-module connectivity

Compute mean connectivity within each of the 12 resting state networks

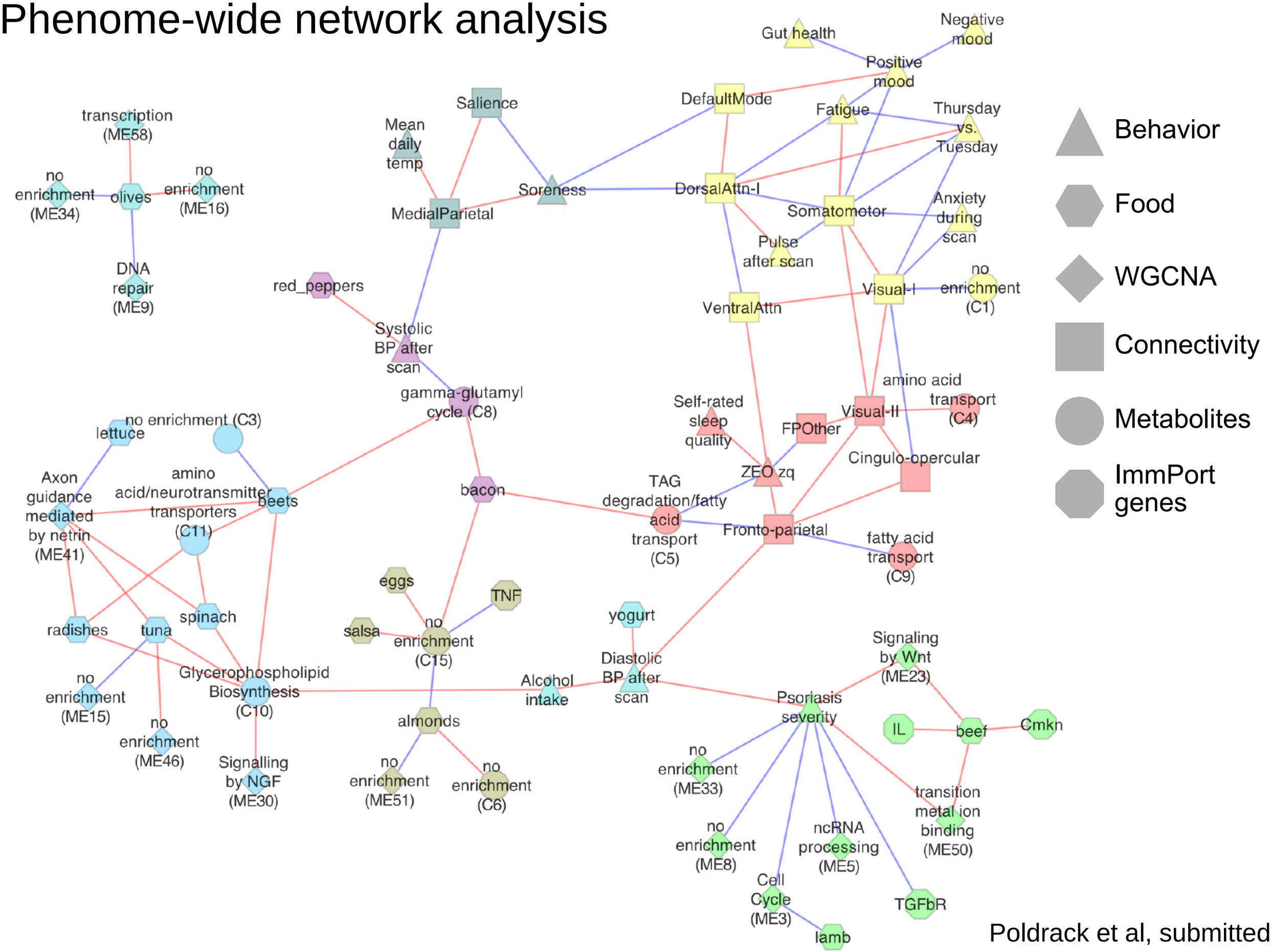


Within-module connectivity



Poldrack et al, submitted

Phenome-wide network analysis



Poldrack et al, submitted

Google Cal | https://www | https://twitt | https://app | Nature | Shared with | Handling M | poldrack/m | Russ

GitHub, Inc. [US] https://github.com/poldrack/myconnectome-vm

Apps | Google Calendar - W | SERA | Axess | PubMed | sklearn | nipype

This repository Search | Pull requests | Issues | Gist

poldrack / myconnectome-vm

Unwatch 2 | Star 0 | Fork 1

Virtual machine setup for MyConnectome data analysis — Edit

86 commits | 1 branch | 0 releases | 3 contributors

Branch: master | myconnectome-vm / +

Merge pull request #14 from vsoch/master ...

poldrack authored 23 days ago | latest commit 3b45da4ddb

LICENSE	Initial commit	2 months ago
README.md	Update README.md	24 days ago
Vagrantfile	removing supervisor controller from application - will be run with st...	23 days ago

README.md

MyConnectome-VM: A virtual machine to implement MyConnectome analyses.

The MyConnectome project is a project meant to investigate the relations between mind, brain, and body across an extended period of time in a single individual. One of the major goals of the project is to serve as a testbed for reproducible analysis practices. For this reason, we have released the data and as much code as possible for the processing and analyses.

Code
Issues 0
Pull requests 0
Wiki
Pulse
Graphs
Settings

SSH clone URL
git@github.com:poldr

You can clone with HTTPS, SSH, or Subversion.

Clone in Desktop
Download ZIP

nihms-470697.pdf | yoi140096.pdf | Show All

Home Log Data Explorer Download

MYCONNECTOME

MyConnectome Analyses

100% Complete

Analyses Complete

Explore Data

Learn more about the MyConnectome Virtual Machine

Learn more about the MyConnectome Project

Timeseries analyses

Timeseries analysis results

Timeseries plots

Table of top timeseries results

Timeseries longitudinal heatmaps

Listing of all files

RNA-seq analyses

RNA-seq data preparation

RNA-seq QA results

RNA-seq WGCNA analysis

Snyderome vs. MyConnectome analysis

Listing of all files

Metabolomic analyses

Metabolomics data preparation

Listing of all files

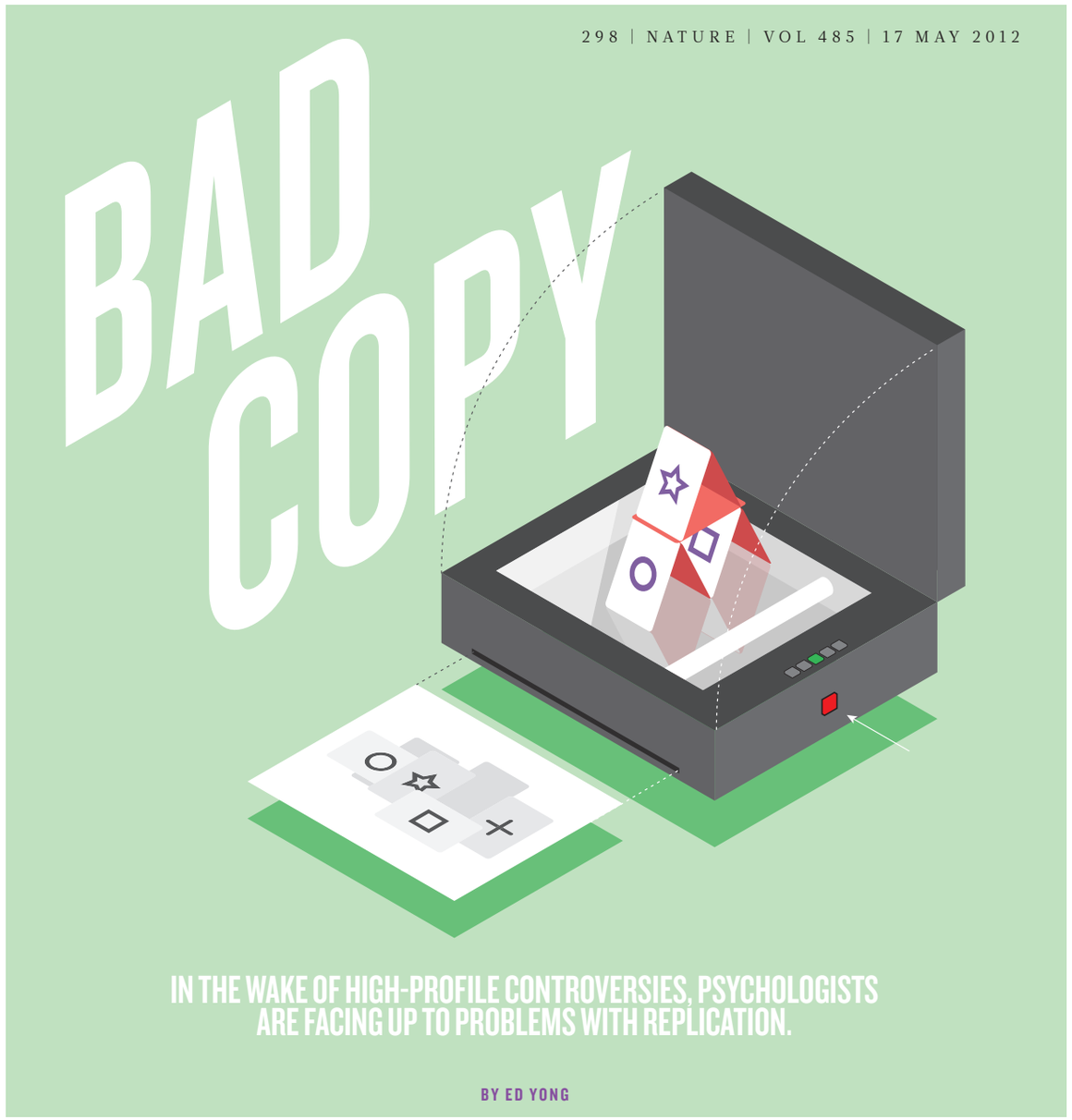
Resting state fMRI analyses

Resting fMRI QA results

Listing of all files

why bother?

Science in crisis (?)



Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

PLoS Medicine | www.plosmedicine.org 0696 August 2005 | Volume 2 | Issue 8 | e124

PERSPECTIVES ON PSYCHOLOGICAL SCIENCE

Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition[1]

Edward Vul,[1] Christine Harris,[2] Piotr Winkielman,[2] & Harold Pashler[2]

[1]Massachussetts Institute of Technology and [2]University of California, San Diego

Raise standards for preclinical cancer research

C. Glenn Begley and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

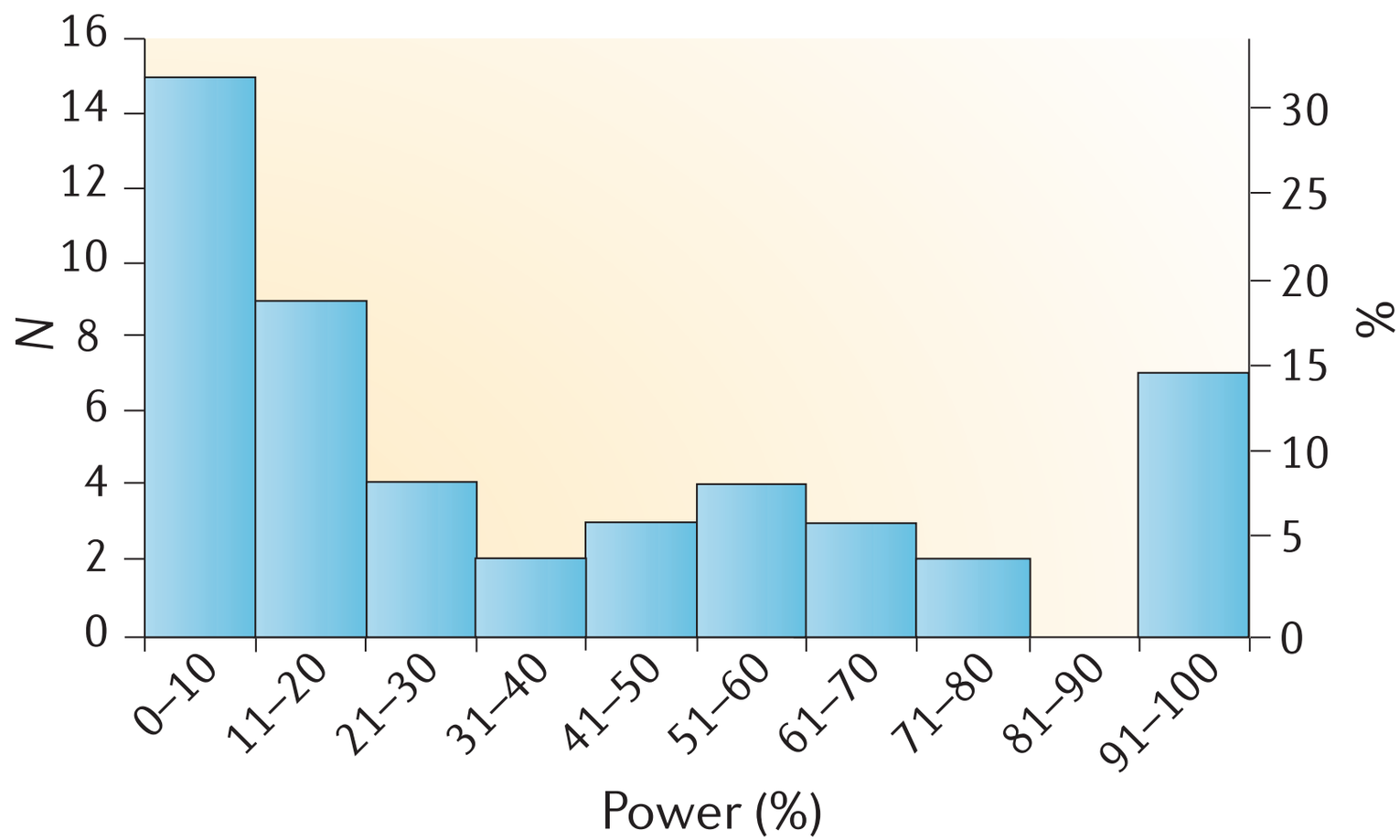
29 MARCH 2012 | VOL 483 | NATURE | 531

Threats to reproducibility: Low power

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button[1,2], John P. A. Ioannidis[3], Claire Mokrysz[1], Brian A. Nosek[4], Jonathan Flint[5], Emma S. J. Robinson[6] and Marcus R. Munafò[1]

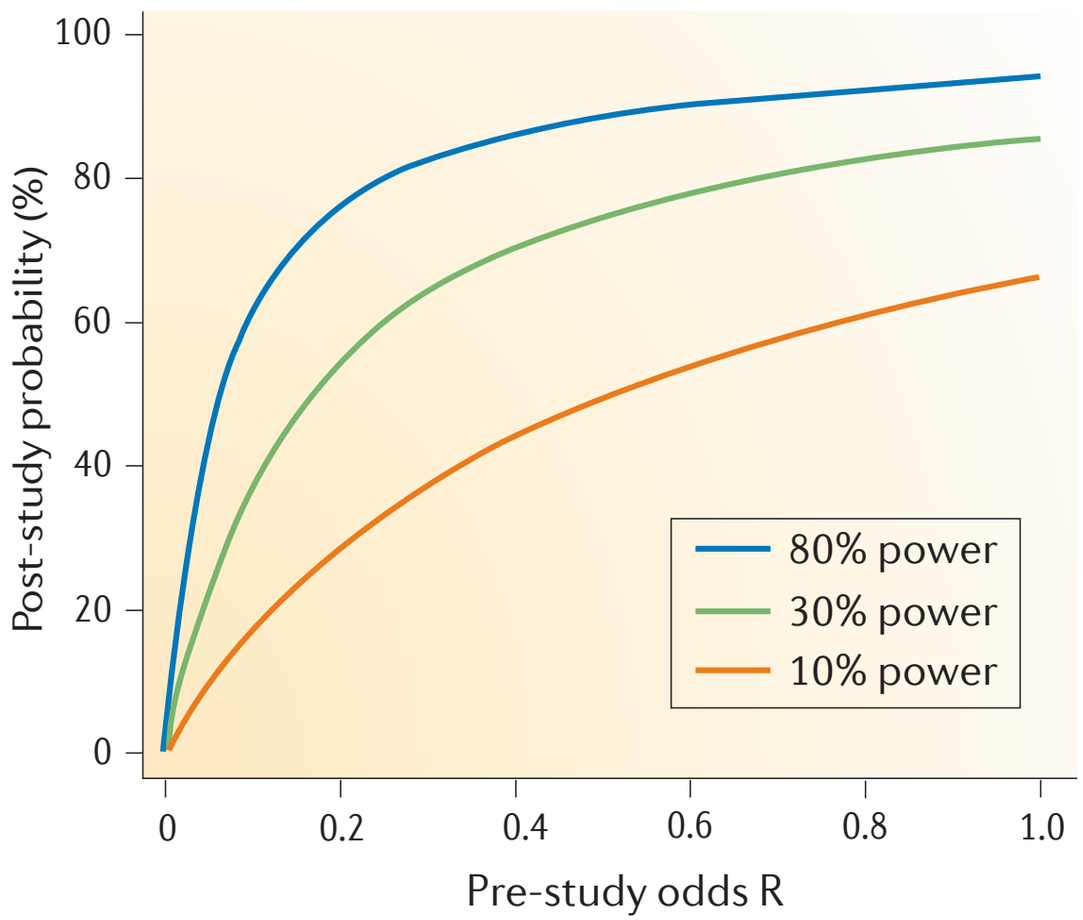
NATURE REVIEWS | NEUROSCIENCE VOLUME 14 | MAY 2013 | 365



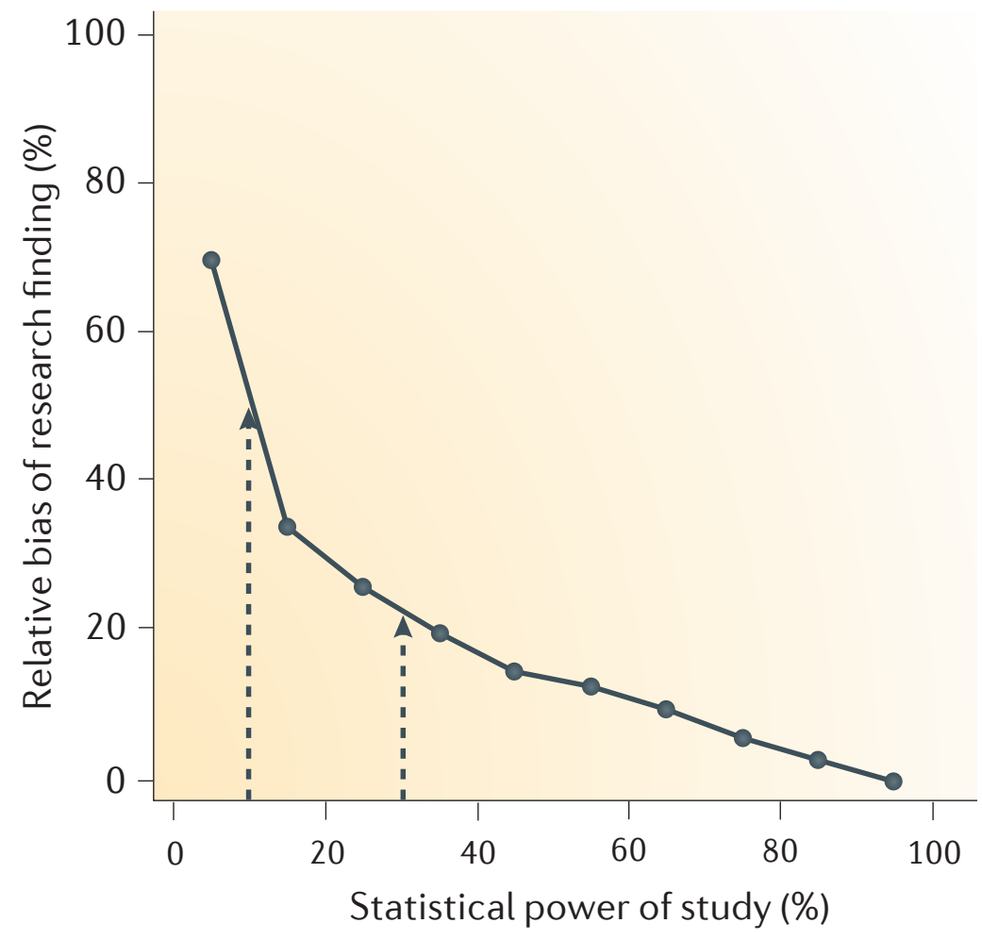
Low power -> unreliable science

Positive Predictive Value (PPV): The probability that a positive result is true

$$PPV = ([1 - \beta] \times R) / ([1 - \beta] \times R + \alpha)$$



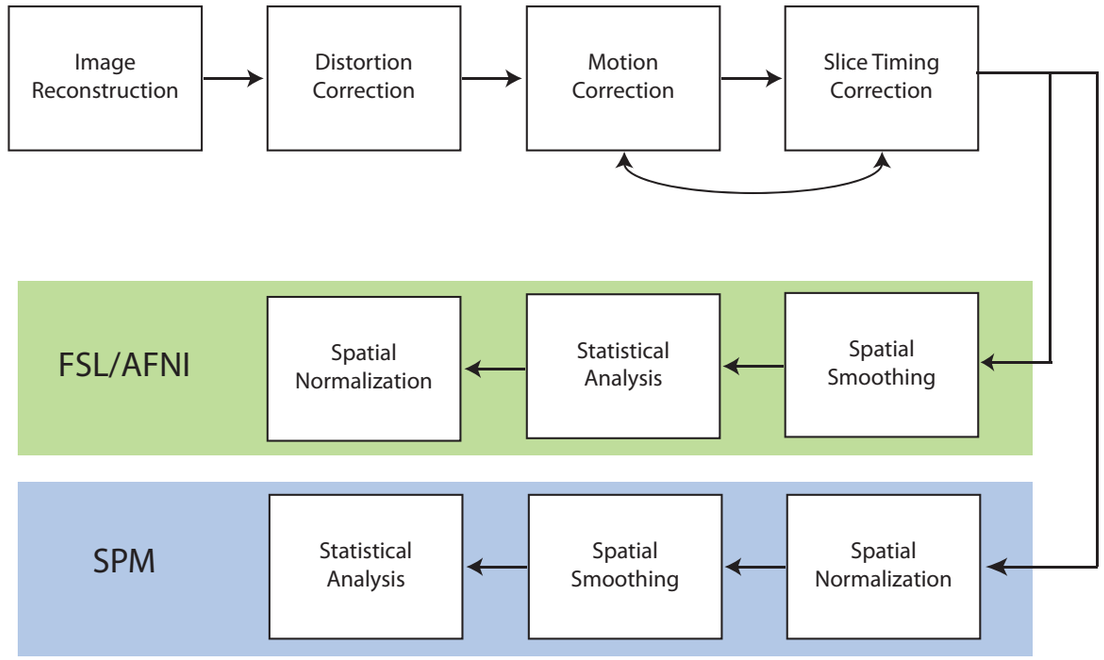
Winner's Curse: overestimation of effect sizes for significant results



Button et al., 2013

Threats to reproducibility: Analytic flexibility

- Using standard FSL analysis options
  - 5,376 possible analysis workflows



Poldrack et al., 2011

Threats to reproducibility: Analytic flexibility

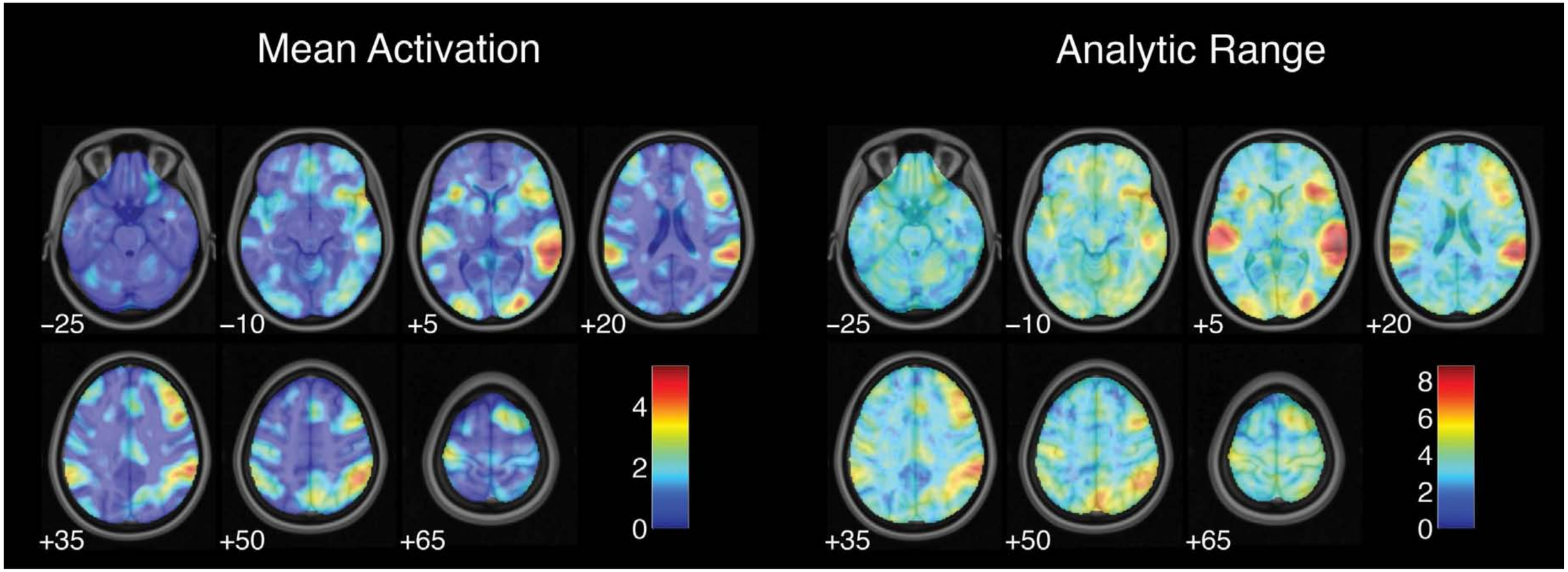
frontiers in NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 11 October 2012
doi: 10.3389/fnins.2012.00149

On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments

*Joshua Carp**

6,912 pipelines



P-hacking: Anything can become significant

Study 2: musical contrast and chronological rejuvenation

…we asked 20 University of Pennsylvania undergraduates to listen to either "When I'm Sixty-Four" by The Beatles or "Kalimba." Then, in an ostensibly unrelated task, they indicated their birth date (mm/dd/ yyyy) and their father's age. We used father's age to control for variation in baseline age across participants.

An ANCOVA revealed the predicted effect: According to their birth dates, people were nearly a year-and-a-half younger after listening to "When I'm Sixty-Four" (adjusted $M = 20.1$ years) rather than to "Kalimba" (adjusted $M = 21.5$ years), $F(1, 17) = 4.92$, $p = .040$.

-Simmons et al., 2011, Psychological Science

P-hacking: Anything can become significant

Table 1. Likelihood of Obtaining a False-Positive Result

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

-Simmons et al., 2011, Psychological Science

- “My result isn’t significant, so I need to add more subjects…”

Sample size flexibility

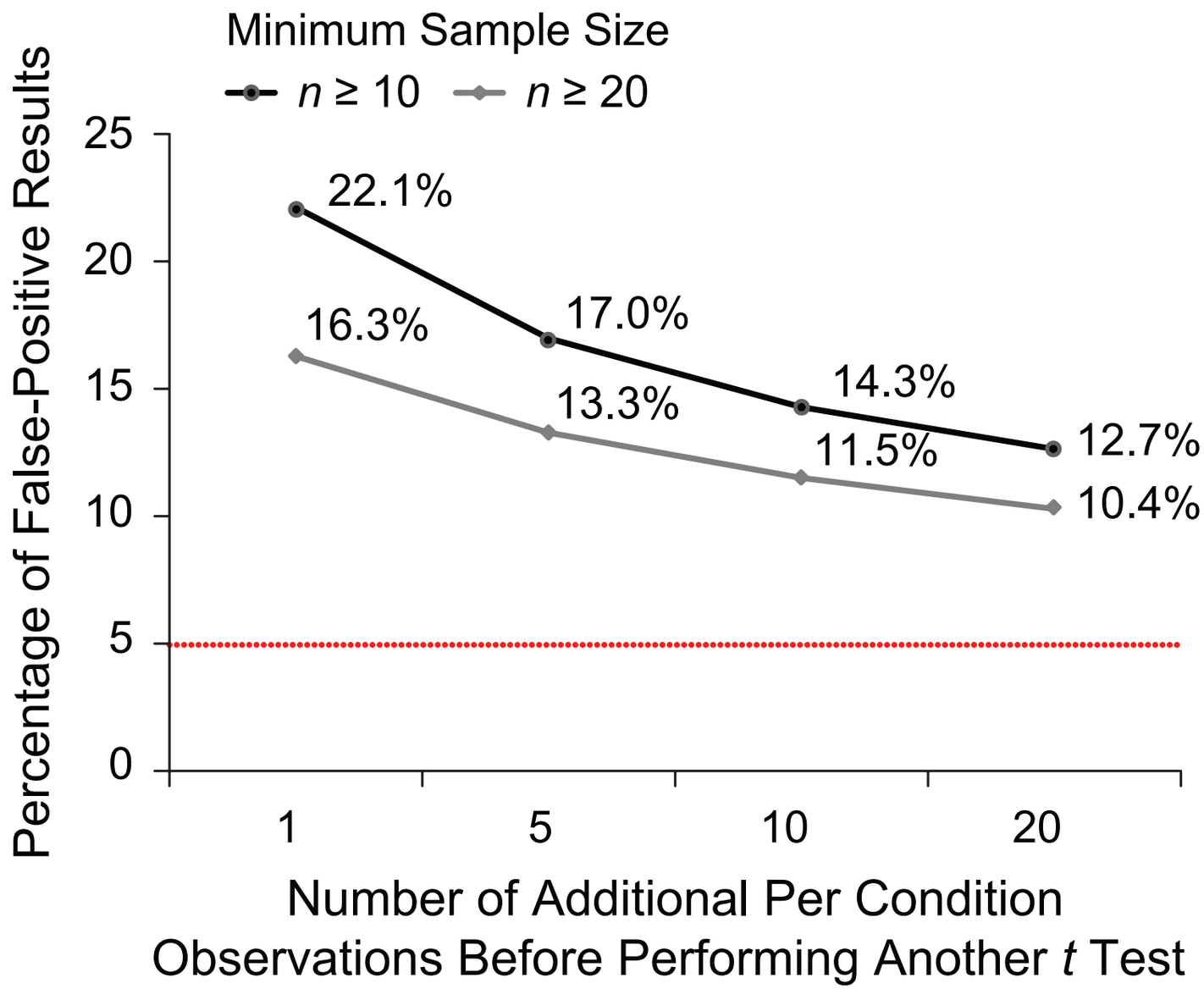
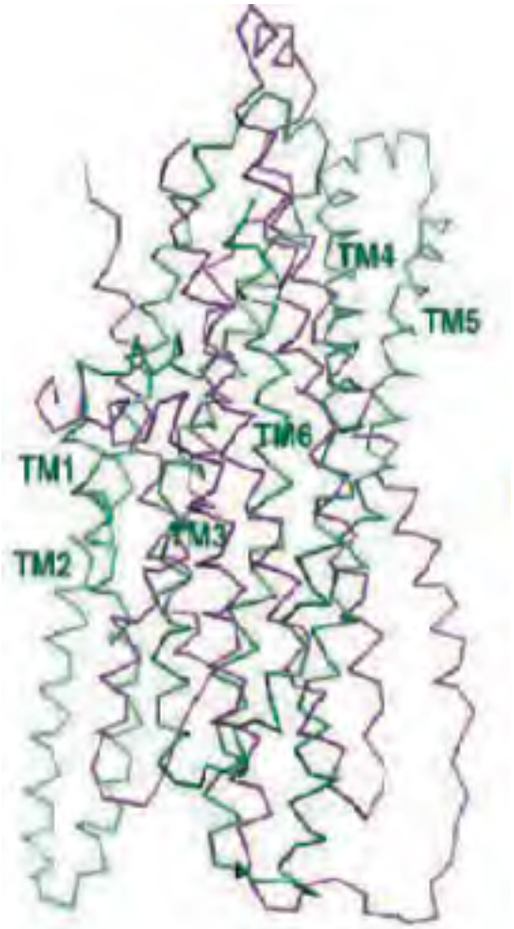


Fig. 1. Likelihood of obtaining a false-positive result when data collection ends upon obtaining significance ($p \leq .05$, highlighted by the dotted line). The figure depicts likelihoods for two minimum sample sizes, as a function of the frequency with which significance tests are performed.

-Simmons et al., 2011, Psychological Science

Threats to reproducibilty: software errors

Geoffrey Chang



Structure of MsbA from *E. coli*: A Homolog of the Multidrug Resistance ATP Binding Cassette (ABC) Transporters

Geoffrey Chang* and Christopher B. Roth

Multidrug resistance (MDR) is a serious medical problem and presents a major challenge to the treatment of disease and the development of novel therapeutics. ABC transporters that are associated with multidrug resistance (MDR-ABC transporters) translocate hydrophobic drugs and lipids from the inner to the outer leaflet of the cell membrane. To better elucidate the structural basis for the "flip-flop" mechanism of substrate movement across the lipid bilayer, we have determined the structure of the lipid flippase MsbA from *Escherichia coli* by x-ray crystallography to a resolution of 4.5 angstroms. MsbA is organized as a homodimer with each subunit containing six transmembrane α-helices and a nucleotide-binding domain. The asymmetric distribution of charged residues lining a central chamber suggests a general mechanism for the translocation of substrate by MsbA and other MDR-ABC transporters. The structure of MsbA can serve as a model for the MDR-ABC transporters that confer multidrug resistance to cancer cells and infectious microorganisms.

www.sciencemag.org SCIENCE VOL 293 7 SEPTEMBER 2001

Structure of the ABC Transporter MsbA in Complex with ADP·Vanadate and Lipopolysaccharide

Christopher L. Reyes and Geoffrey Chang*

Select members of the adenosine triphosphate (ATP)–binding cassette (ABC) transporter family couple ATP binding and hydrolysis to substrate efflux and confer multidrug resistance. We have determined the x-ray structure of MsbA in complex with magnesium, adenosine diphosphate, and inorganic vanadate ($Mg{\cdot}ADP{\cdot}V_i$) and the rough-chemotype lipopolysaccharide, Ra LPS. The structure supports a model involving a rigid-body torque of the two transmembrane domains during ATP hydrolysis and suggests a mechanism by which the nucleotide-binding domain communicates with the transmembrane domain. We propose a lipid "flip-flop" mechanism in which the sugar groups are sequestered in the chamber while the hydrophobic tails are dragged through the lipid bilayer.

13 MAY 2005 VOL 308 SCIENCE www.sciencemag.org

X-ray Structure of the EmrE Multidrug Transporter in Complex with a Substrate

Owen Pornillos, Yen-Ju Chen, Andy P. Chen, Geoffrey Chang*

EmrE is a prototype of the Small Multidrug Resistance family of efflux transporters and actively expels positively charged hydrophobic drugs across the inner membrane of *Escherichia coli*. Here, we report the x-ray crystal structure, at 3.7 angstrom resolution, of one conformational state of the EmrE transporter in complex with a translocation substrate, tetraphenylphosphonium. Two EmrE polypeptides form a homodimeric transporter that binds substrate at the dimerization interface. The two subunits have opposite orientations in the membrane and adopt slightly different folds, forming an asymmetric antiparallel dimer. This unusual architecture likely confers unidirectionality to transport by creating an asymmetric substrate translocation pathway. On the basis of available structural data, we propose a model for the proton-dependent drug efflux mechanism of EmrE.

23 DECEMBER 2005 VOL 310 SCIENCE www.sciencemag.org

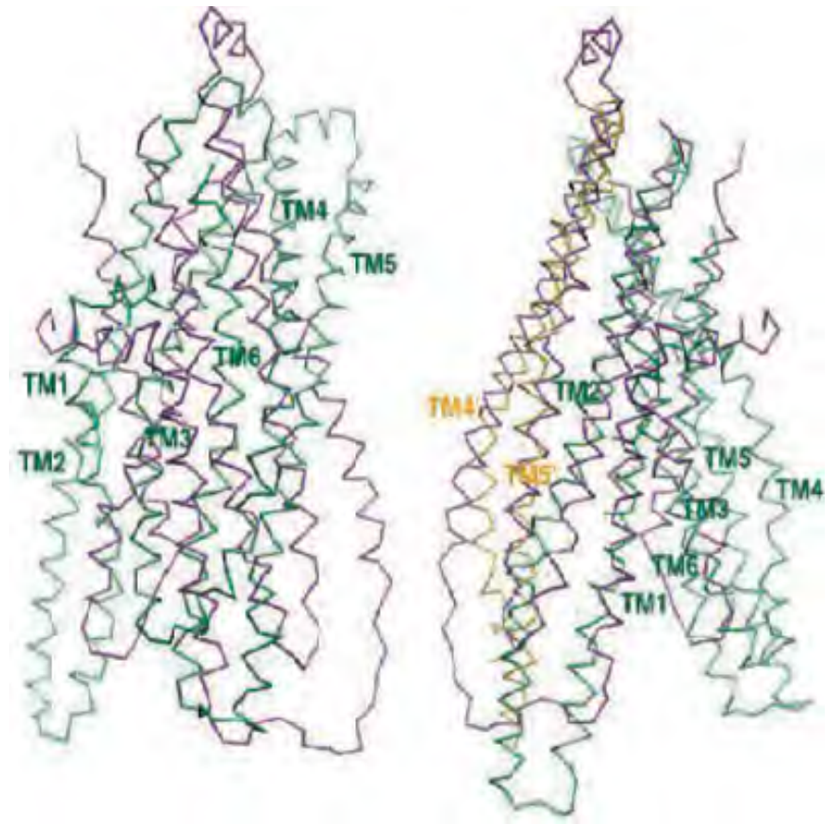
Threats to reproducibilty: software errors

Retraction

WE WISH TO RETRACT OUR RESEARCH ARTICLE "STRUCTURE OF MsbA from *E. coli*: A homolog of the multidrug resistance ATP binding cassette (ABC) transporters" and both of our Reports "Structure of the ABC transporter MsbA in complex with ADP•vanadate and lipopolysaccharide" and "X-ray structure of the EmrE multidrug transporter in complex with a substrate" (*1*–*3*).

The recently reported structure of Sav1866 (*4*) indicated that our MsbA structures (*1*, *2*, *5*) were incorrect in both the hand of the structure and the topology. Thus, our biological interpretations based on these inverted models for MsbA are invalid.

An in-house data reduction program introduced a change in sign for anomalous differences. This program, which was not part of a conventional data processing package, converted the anomalous pairs (I+ and I–) to (F– and F+), thereby introducing a sign change. As the diffraction data collected for each set of MsbA crystals and for the EmrE crystals were processed with the same program, the structures reported in (*1*–*3*, *5*, *6*) had the wrong hand.



Small errors can have big effects

```
# 23-class classification problem

skf=StratifiedKFold(labels,8)

if trainsvm:
    pred=N.zeros(len(labels))
    for train,test in skf:
        clf=LinearSVC()
        clf.fit(data[train],labels[train])
        pred[test]=clf.predict(data[test])
```

data[:,train]

data[:,test]

Results:
93% accuracy

Results:
53% accuracy

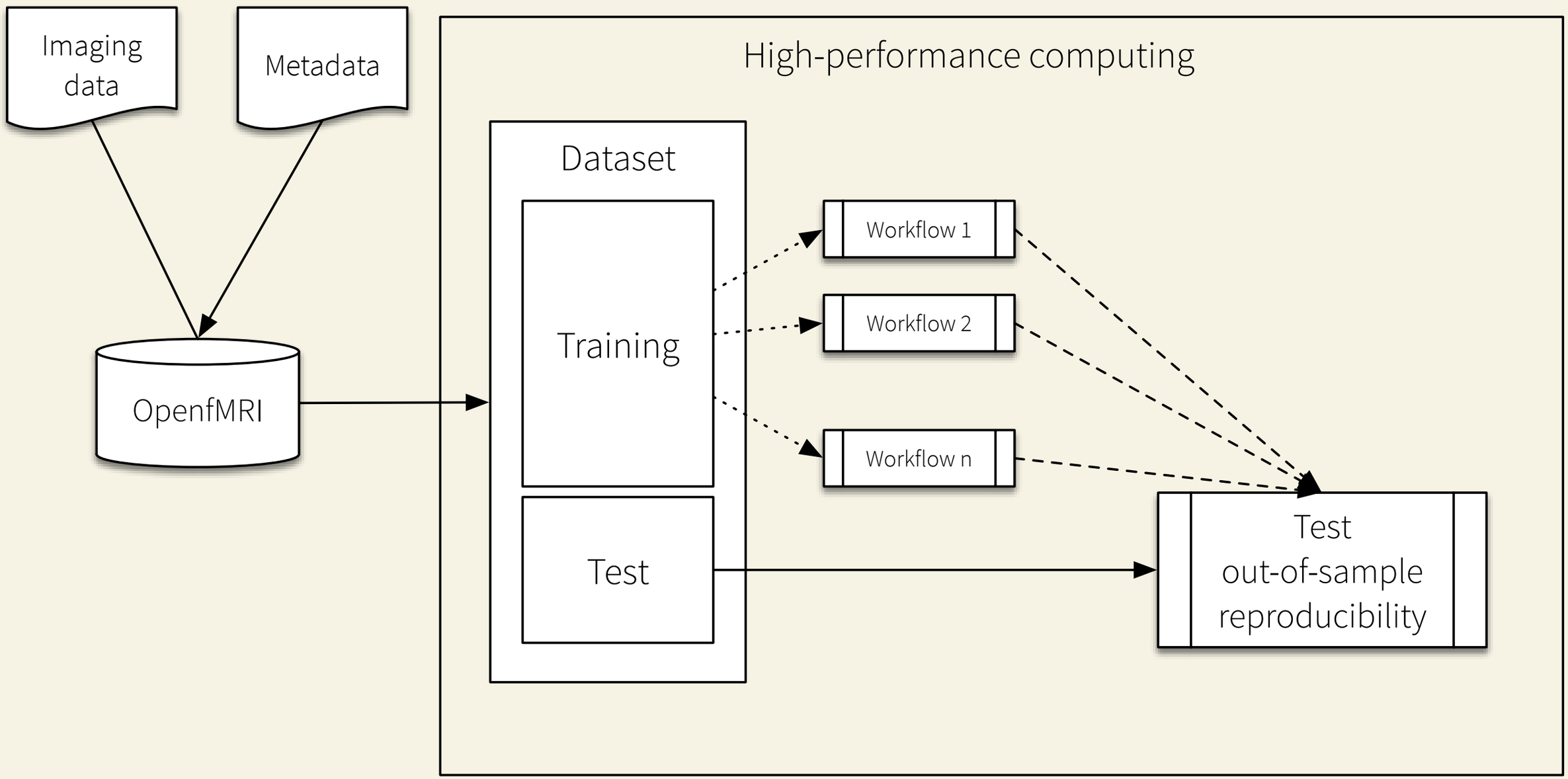
http://www.russpoldrack.org/2013/02/anatomy-of-coding-error.html

Solutions for reproducible research

- Statistical power
- Openness and transparency
- Planning and documentation
- Publication of positive and negative results
- Focus on generalization rather than statistical significance

Stanford | Center for Reproducible Neuroscience

Analyzing for reproducibility

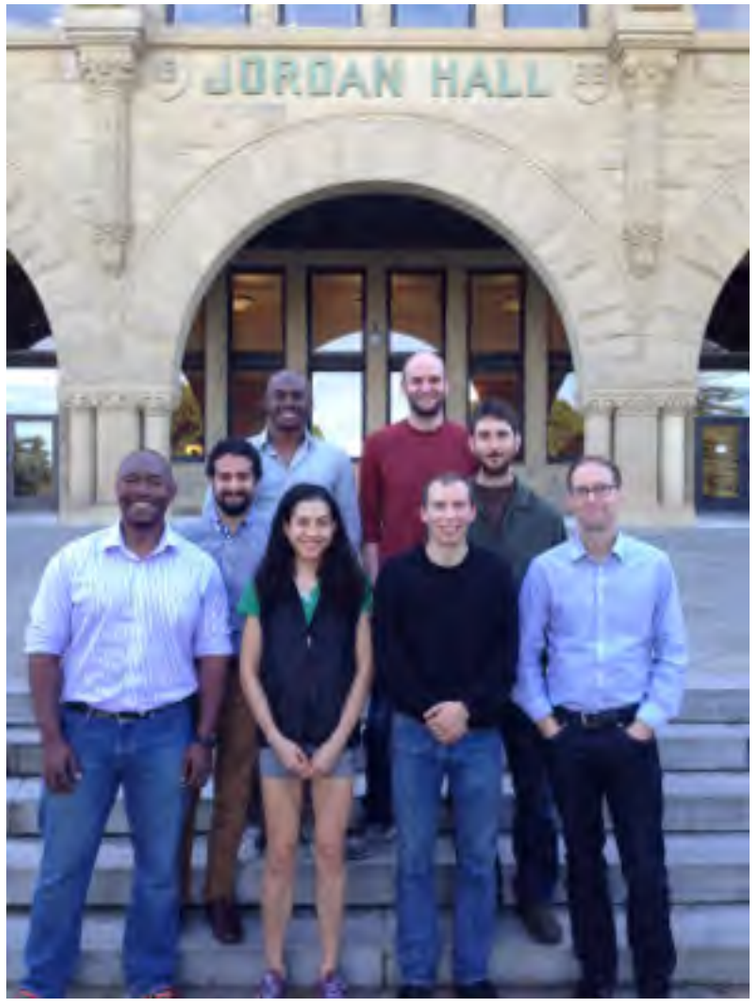


Conclusions

- fMRI has come a long way in 20 years, but we have a long way to go
- There are many threats to research quality and reproducibility that arise from daily research practices in neuroimaging
- “The first principle is that you must not fool yourself and you are the easiest person to fool”
  - R. Feynman

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



The Poldrack Lab @ Stanford

http://reproducibility.stanford.edu

NIMH National Institute of Mental Health

NIH National Institute on Drug Abuse

ljaf laura and john arnold foundation®

James S. McDonnell Foundation

NSF

Data sets and code will be made available at www.openfmri.org