Open Science Controversies in Neuroimaging

Adam Thomas
Data Science and Sharing Team, NIMH IRP
Slides: https://github.com/agt24

The crisis of reproducibility in neuroimaging (small N vs big N)

Adam Thomas & Peter Bandettini
Data Science and Sharing Team, FMRIF, NIMH

FMRIF

CMN

NIH National Institute of Mental Health

The crisis of reproducibility in neuroimaging (small N vs big N)

Adam Thomas & Peter Bandettini
Data Science and Sharing Team, FMRIF, NIMH

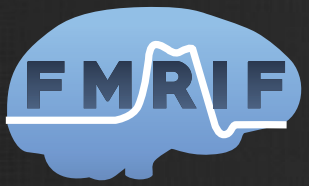
FMRIF

CMN

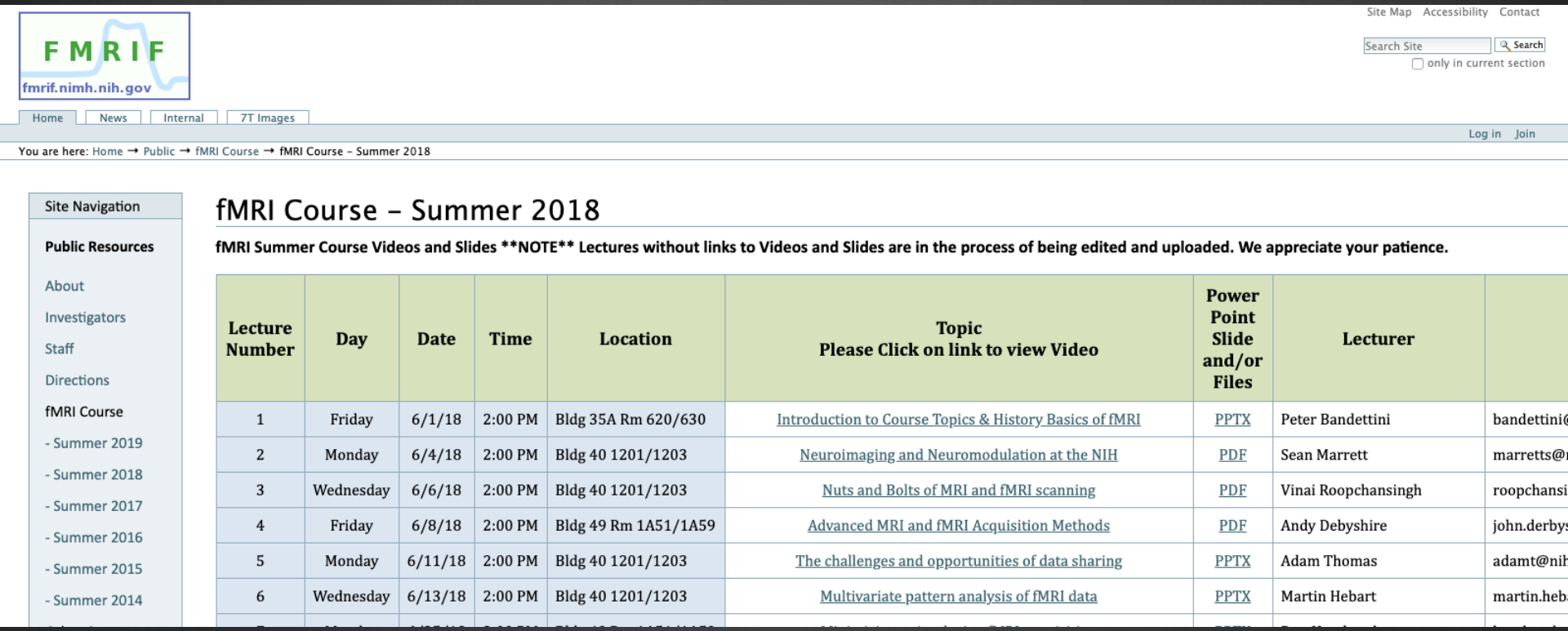
NIH National Institute of Mental Health

Open Science in Neuroimaging

Adam Thomas
Data Science and Sharing Team, FMRIF, NIMH



Open Science Talk



FMRIF
fmrif.nimh.gov

Site Map Accessibility Contact

Search Site Search
only in current section

Home News Internal 7T Images

Log in Join

You are here: Home → Public → fMRI Course → fMRI Course – Summer 2018

Site Navigation

Public Resources

About
Investigators
Staff
Directions
fMRI Course
- Summer 2019
- Summer 2018
- Summer 2017
- Summer 2016
- Summer 2015
- Summer 2014

fMRI Course – Summer 2018

fMRI Summer Course Videos and Slides **NOTE Lectures without links to Videos and Slides are in the process of being edited and uploaded. We appreciate your patience.**

Lecture Number	Day	Date	Time	Location	Topic Please Click on link to view Video	Power Point Slide and/or Files	Lecturer	
1	Friday	6/1/18	2:00 PM	Bldg 35A Rm 620/630	Introduction to Course Topics & History Basics of fMRI	PPTX	Peter Bandettini	bandettini@
2	Monday	6/4/18	2:00 PM	Bldg 40 1201/1203	Neuroimaging and Neuromodulation at the NIH	PDF	Sean Marrett	marretts@
3	Wednesday	6/6/18	2:00 PM	Bldg 40 1201/1203	Nuts and Bolts of MRI and fMRI scanning	PDF	Vinai Roopchansingh	roopchansi
4	Friday	6/8/18	2:00 PM	Bldg 49 Rm 1A51/1A59	Advanced MRI and fMRI Acquisition Methods	PDF	Andy Debyshire	john.derby
5	Monday	6/11/18	2:00 PM	Bldg 40 1201/1203	The challenges and opportunities of data sharing	PPTX	Adam Thomas	adamt@nih
6	Wednesday	6/13/18	2:00 PM	Bldg 40 1201/1203	Multivariate pattern analysis of fMRI data	PPTX	Martin Hebart	martin.heb

Outline of my typical Standard Open Science Talk

- Why do we need Open Science?
- What is Open Science?
- How do I do Open Science?

Take homes of my other Open Science talk

- Science has a problem and it is changing (for the better) in both scope (big) and culture (open) to address future challenges
- Open science strives to maximize reproducibility and transparency of data, code, and papers
- Adopting Open Science practices yields benefits in productivity, impact, and reach
- You don't have to do it all at once, and you don't have to do it alone


Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

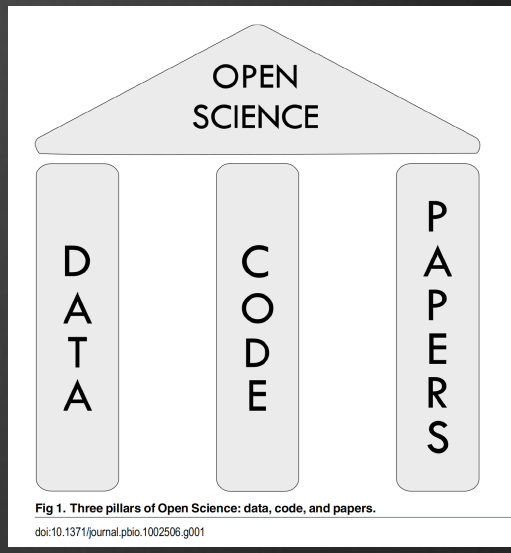


Fig 1. Three pillars of Open Science: data, code, and papers.

doi:10.1371/journal.pbio.1002506.g001

Current
Data Science & Sharing Team

Adam Thomas

Dylan Nielson

Travis Riddle

- NIMHDSST@mail.nih.gov
- http://cmn.nimh.nih.gov/dsst
- Building 10, Room 3D39
- Open door policy
- Co-located with the Machine Learning Team

Future Data Science & Sharing Team

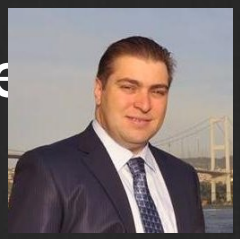
Adam Thomas

Dustin Moraczewski
July 15th

Travis Riddle

IRTA: Nikhil Goyal
Sept 2019

- NIMHDSST@mail.nih.gov
- http://cmn.nimh.nih.gov/dsst
- Building 10, Room 3D39
- Open door policy
- Co-located with the Machine Learning Team & the Scientific Information Officer, Asim Okur



Take homes of my other Open Science talk

- Science has a problem and it is changing (for the better) in both scope (big) and culture (open) to address future challenges
- Open science strives to maximize reproducibility and transparency of data, code, and papers
- Adopting Open Science practices yields benefits in productivity, impact, and reach
- You don't have to do it all at once, and you don't have to do it alone


Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

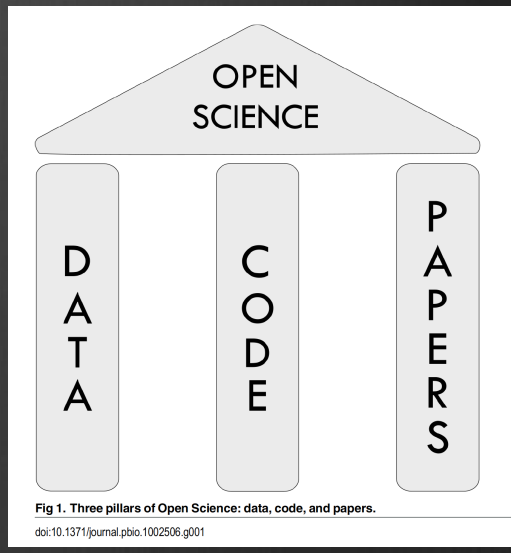
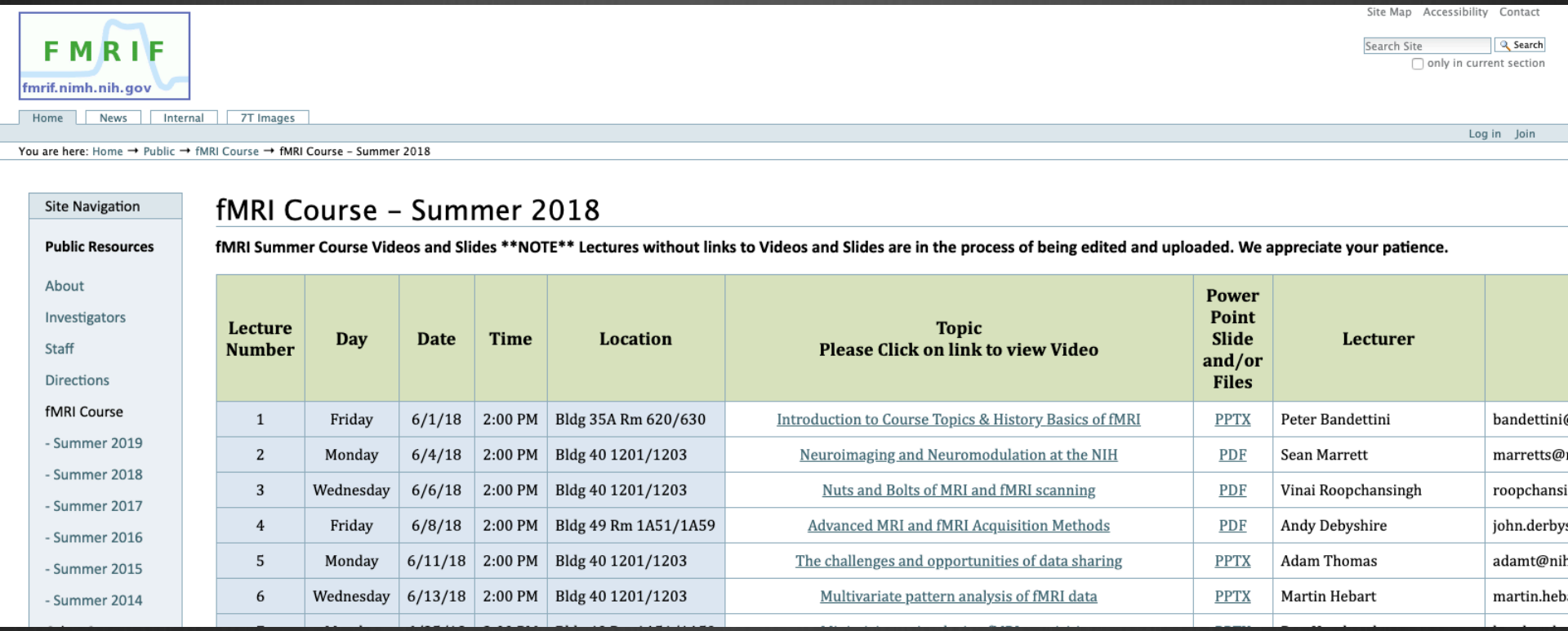


Fig 1. Three pillars of Open Science: data, code, and papers.

doi:10.1371/journal.pbio.1002506.g001

Typical Open Science Talk



FMRIF
fmrif.nimh.nih.gov

Site Map Accessibility Contact

Search Site Search
only in current section

Home News Internal 7T Images

Log in Join

You are here: Home → Public → fMRI Course → fMRI Course – Summer 2018

Site Navigation

Public Resources

About
Investigators
Staff
Directions
fMRI Course
- Summer 2019
- Summer 2018
- Summer 2017
- Summer 2016
- Summer 2015
- Summer 2014

fMRI Course – Summer 2018

fMRI Summer Course Videos and Slides **NOTE Lectures without links to Videos and Slides are in the process of being edited and uploaded. We appreciate your patience.**

Lecture Number	Day	Date	Time	Location	Topic Please Click on link to view Video	Power Point Slide and/or Files	Lecturer	
1	Friday	6/1/18	2:00 PM	Bldg 35A Rm 620/630	Introduction to Course Topics & History Basics of fMRI	PPTX	Peter Bandettini	bandettini@
2	Monday	6/4/18	2:00 PM	Bldg 40 1201/1203	Neuroimaging and Neuromodulation at the NIH	PDF	Sean Marrett	marretts@
3	Wednesday	6/6/18	2:00 PM	Bldg 40 1201/1203	Nuts and Bolts of MRI and fMRI scanning	PDF	Vinai Roopchansingh	roopchansi
4	Friday	6/8/18	2:00 PM	Bldg 49 Rm 1A51/1A59	Advanced MRI and fMRI Acquisition Methods	PDF	Andy Debyshire	john.derby
5	Monday	6/11/18	2:00 PM	Bldg 40 1201/1203	The challenges and opportunities of data sharing	PPTX	Adam Thomas	adamt@nih
6	Wednesday	6/13/18	2:00 PM	Bldg 40 1201/1203	Multivariate pattern analysis of fMRI data	PPTX	Martin Hebart	martin.heb

The contraversies of reproducibility in neuroimaging (small N vs big N)

Adam Thomas & Peter Bandettini
Data Science and Sharing Team, FMRIF, NIMH

FMRIF

CMN

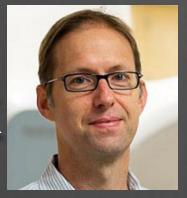
NIH National Institute of Mental Health

Outline

- How can one do reproducible science with limited resources?
  - Perils of the pilot study
  - Other options
- Talking about reproducibility – is Science broken?

Credits

Material borrowed, adapted, and/or stolen from:

- Russ Poldrack 
- Kathleen Hall Jamieson 

Neuron

NeuroView

The Costs of Reproducibility

Russell A. Poldrack[1,*]
[1]Department of Psychology, Stanford University, Stanford, CA, USA
*Correspondence: poldrack@stanford.edu
https://doi.org/10.1016/j.neuron.2018.11.030

PNAS

Crisis or self-correction: Rethinking media narratives about the well-being of science

Kathleen Hall Jamieson[a,b,1]

[a]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104; and [b]Annenberg Public Policy Center, University of Pennsylvania, Philadelphia, PA 19104

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved September 25, 2017 (received for review June 16, 2017)

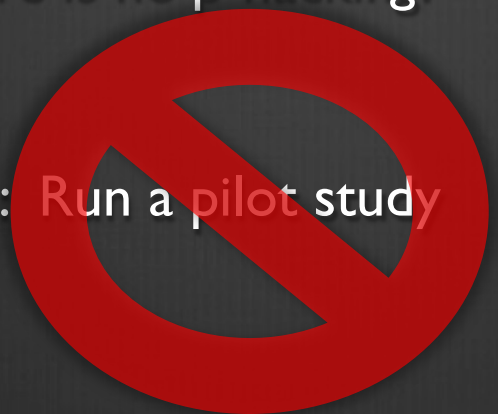
After documenting the existence and exploring some implications of oldest, hardiest, and most popular of all literary genres" (11).

Outline

- How can one do reproducible science with limited resources?
- The cost of Reproducibility
- Talking about reproducibility – is Science broken?

How can one do reproducible science with limited resources?

- Imagine you are a graduate student who desperately wants to do fMRI research, but your mentor doesn't have a large grant to support your study. You cobble together funds to collect a dataset of 20 subjects performing your new cognitive task, and you wish to identify the whole-brain activity pattern associated with the task.
- Then you happen to read "Scanning the Horizon" which argues that the likelihood of such a study identifying a true positive result if it exists is very low, and the likelihood of any positive results being false is high (as outlined by Button et al, 2013), even if the study was fully pre-registered and there is no p-hacking.
- What are you to do?
- Most common answer: Run a pilot study



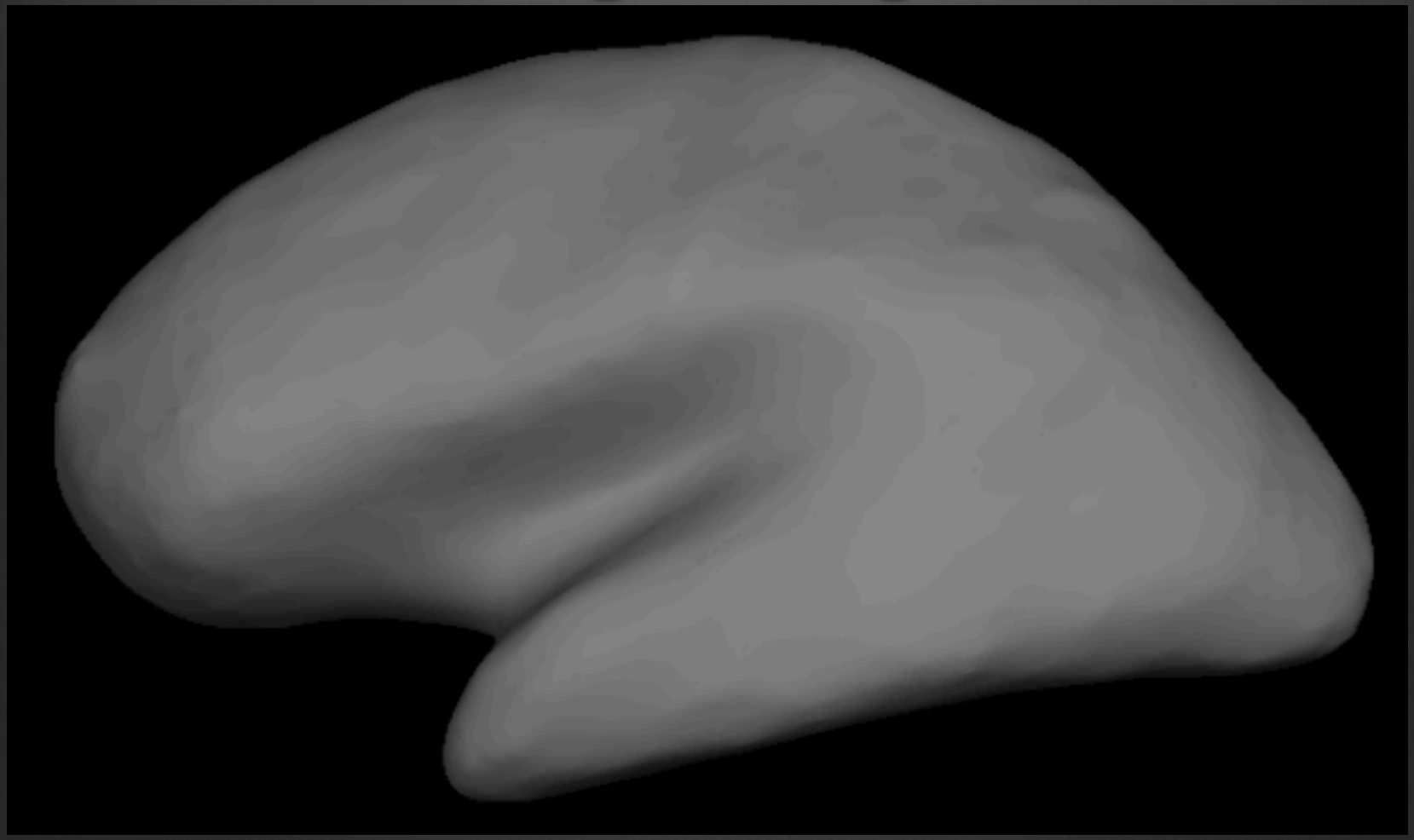
10.1038/nrn.2016.167
10.1038/nrn3475

Can't I just focus on big effects?

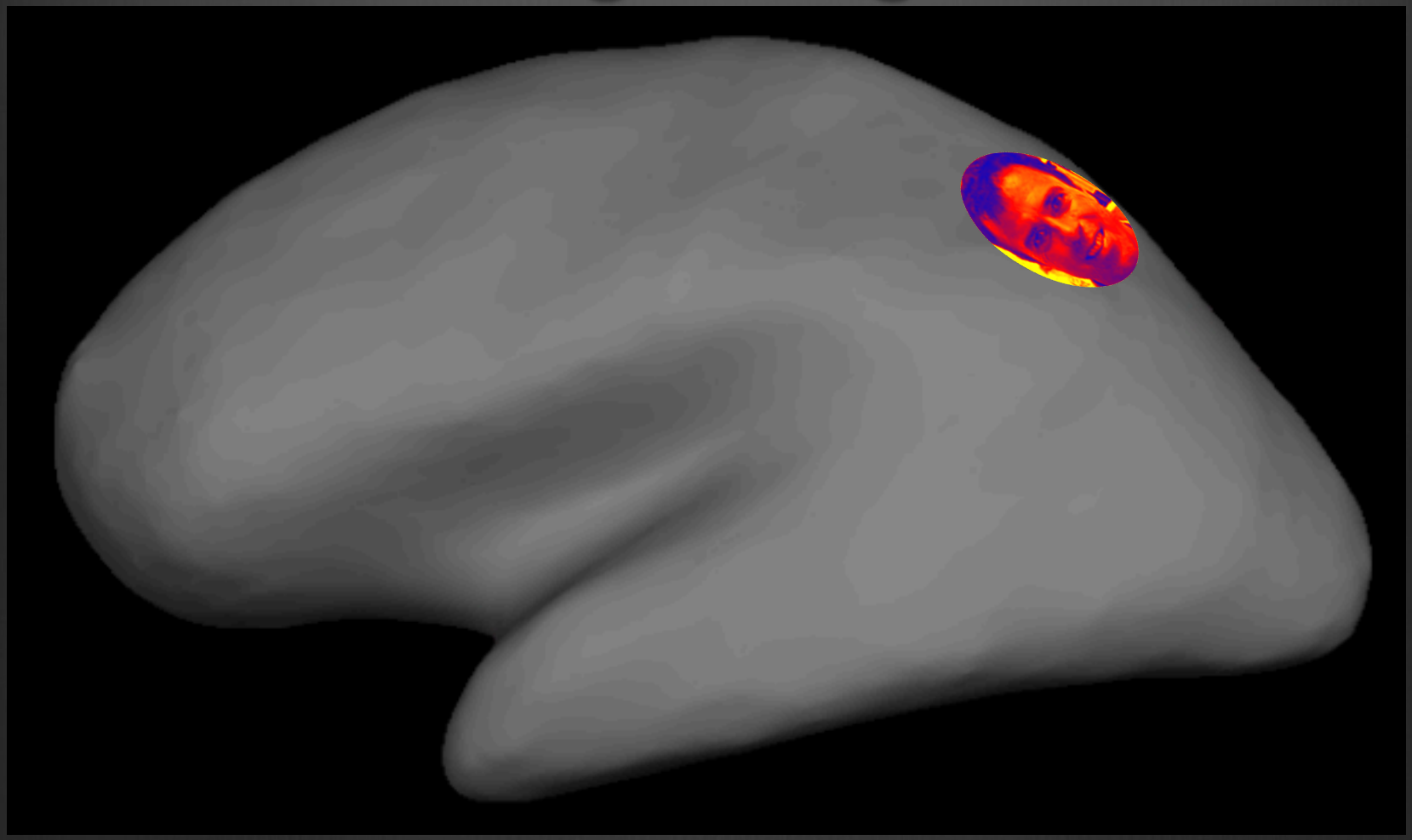
- All experiments need power
  - Threshold (e.g. $P < 0.001$)
  - Sample size (N)
  - Effect Size
- "Big data is for small effects. I'm only interested in big effects"

Alan Koretsky, NINDS Investigator & former Scientific Director

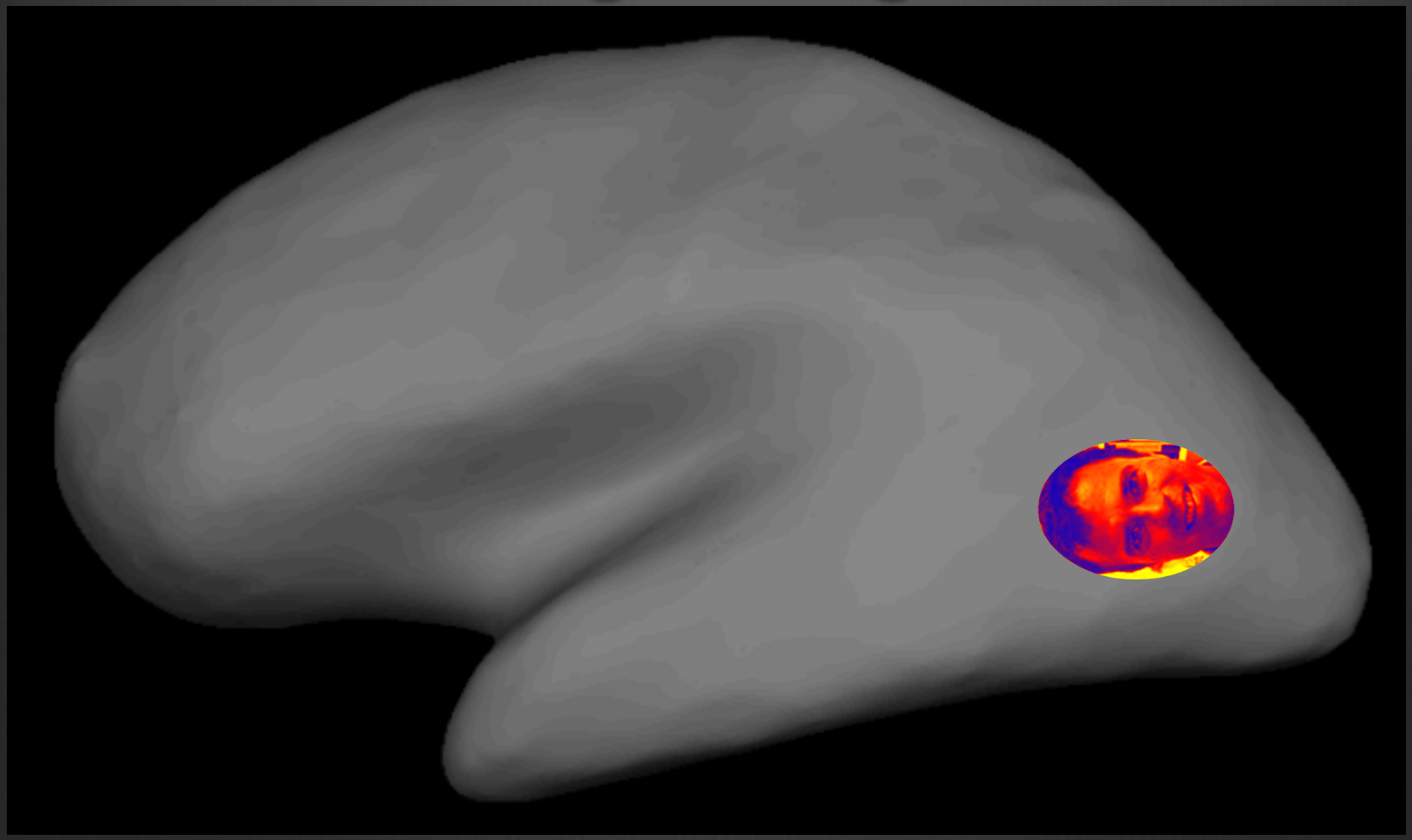
A parable about a pilot study: Searching for big effects



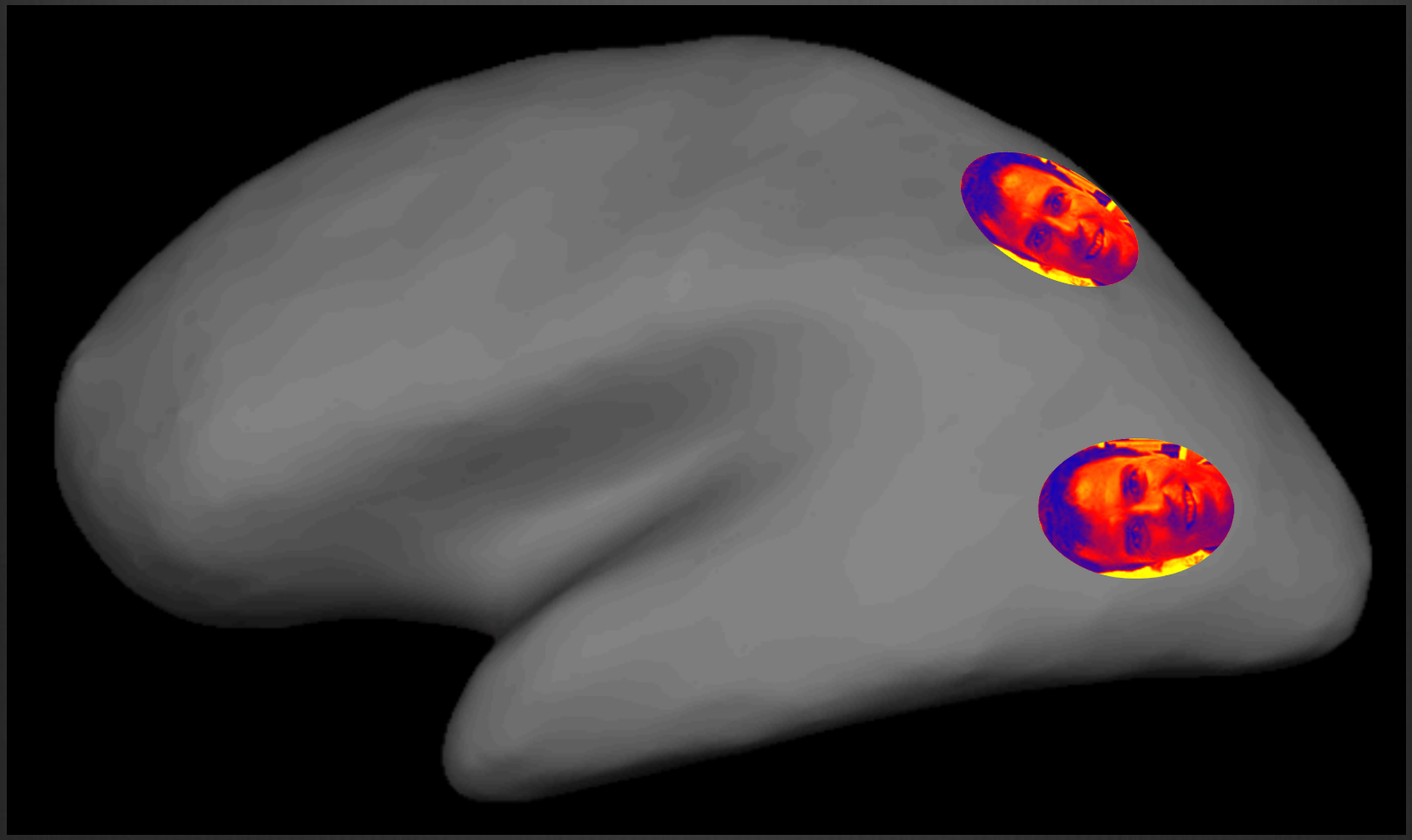
A parable about a pilot study: Searching for big effects



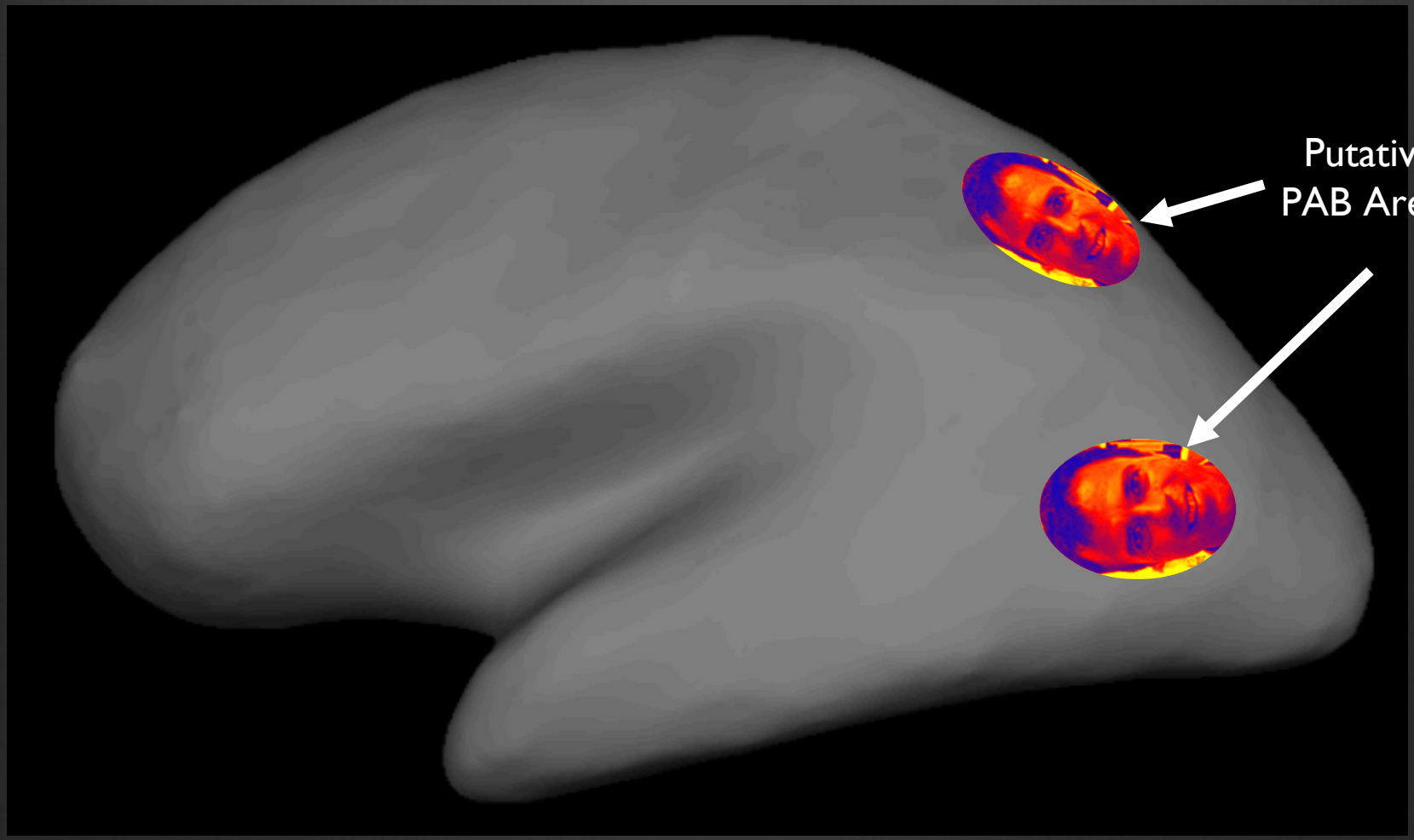
A parable about a pilot study: Searching for big effects



A parable about a pilot study: Searching for big effects

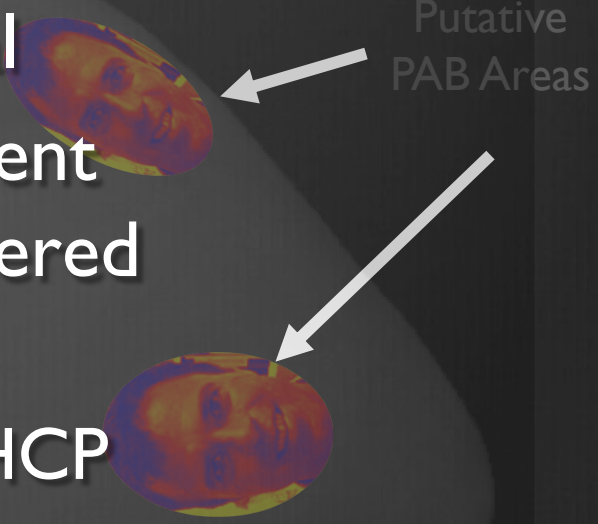


A parable about a pilot study: Searching for big effects



A parable about a pilot study: Searching for big effects

- Options:
  - Publish it in a high profile journal
  - Try to validate it in an independent dataset with a strong, pre-registered hypothesis
    - Can you find PAB's face in the HCP data?



Can't I just focus on big effects?

- All experiments need power
  - Threshold (e.g. $P < 0.001$)
  - Sample size (N)
  - Effect Size
- It's very hard to estimate your effect size in advance
- It is easy and common to overestimate effect size
  DOI: 10.1038/nrn3475
- Better to plan your study for more modest effect size

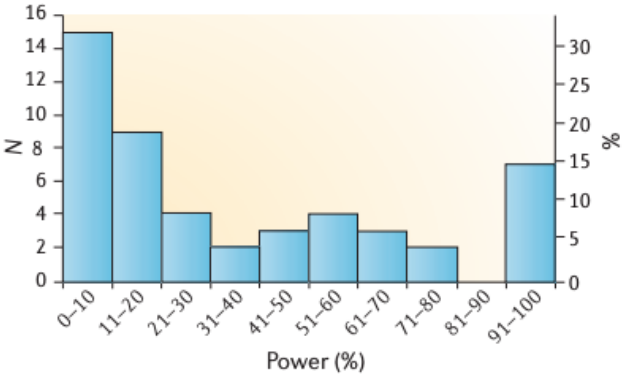


Figure 3 | **Median power of studies included in neuroscience meta-analyses.** The figure shows a

Outline

- How can one do reproducible science with limited resources?
  - Perils of the pilot study
  - Other options
- Talking about reproducibility – is Science broken?

How can one do reproducible science with limited resources?

- Pivot

http://www.russpoldrack.org/2018/04/how-can-one-do-reproducible-science.html

Options for doing reproducible science: Pivot

The Quest for the FFA and Where It Led

JNeurosci THE JOURNAL OF NEUROSCIENCE

Nancy Kanwisher
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

- "I had never worked on face perception because I considered it to be a special case, less important than the general case of object perception. But I needed to stop messing around and discover something, so I cultivated an interest in faces. To paraphrase Stephen Stills, if you can't answer the question you love, love the question you can."

- Asking a more narrow, tractable question is a great option. But be sure to publish it, even if it's a negative result
- Use shared dataset to form testable hypotheses

http://www.russpoldrack.org/2018/04/how-can-one-do-reproducible-science.html

DOI:10.1523/JNEUROSCI.1706-16.2016

How can one do reproducible science with limited resources?

- Pivot
- Collaborate
- Use Shared Data
- Embrace Theory
- Think like a Visual Neuroscientist

http://www.russpoldrack.org/2018/04/how-can-one-do-reproducible-science.html

How can one do reproducible science with limited resources?

- Think like a Visual Neuroscientist
  - Big N doesn't have to mean lots of subjects

THEORETICAL REVIEW

Small is beautiful: In defense of the small-*N* design

Philip L. Smith[1] · Daniel R. Little[1]

Psychonomic Bulletin & Review

Published online: 19 March 2018

PB&R

http://www.russpoldrack.org/2018/04/how-can-one-do-reproducible-science.html

How can one do reproducible science with limited resources?

- Pivot
- Collaborate
- Use Shared Data
- Embrace Theory
- Think like a Visual Neuroscientist

- The Incentives Are Not Yet Aligned
- Making your efforts known
  - CV section, social media
- Ladder pulling

http://www.russpoldrack.org/2018/04/how-can-one-do-reproducible-science.html

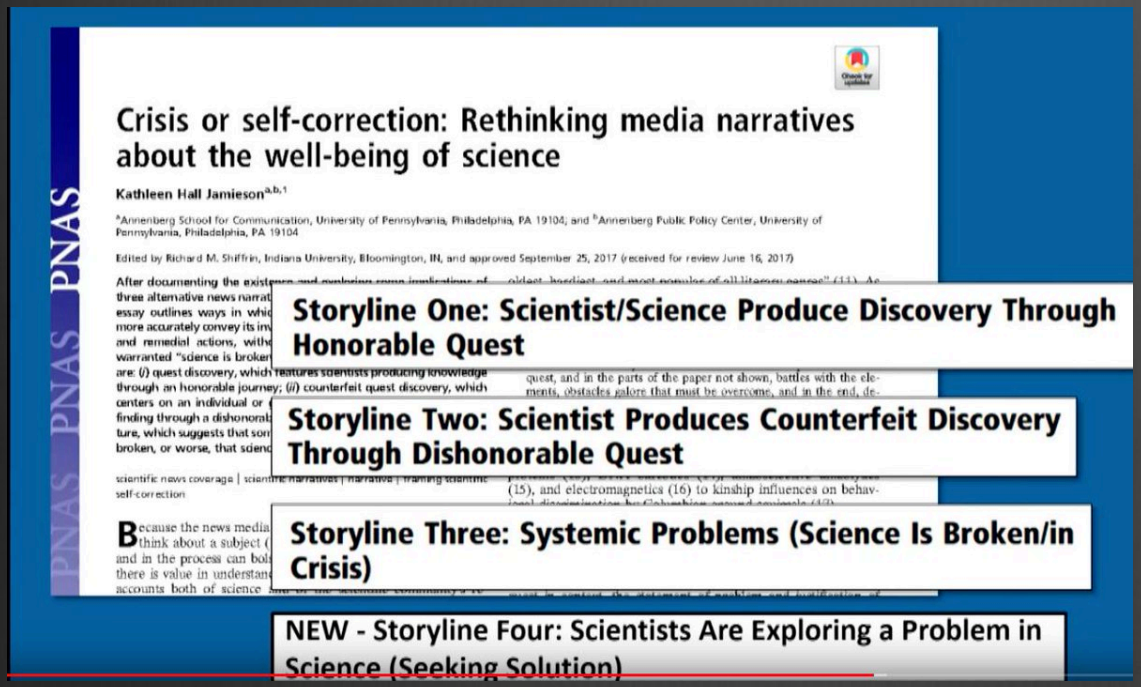
Outline

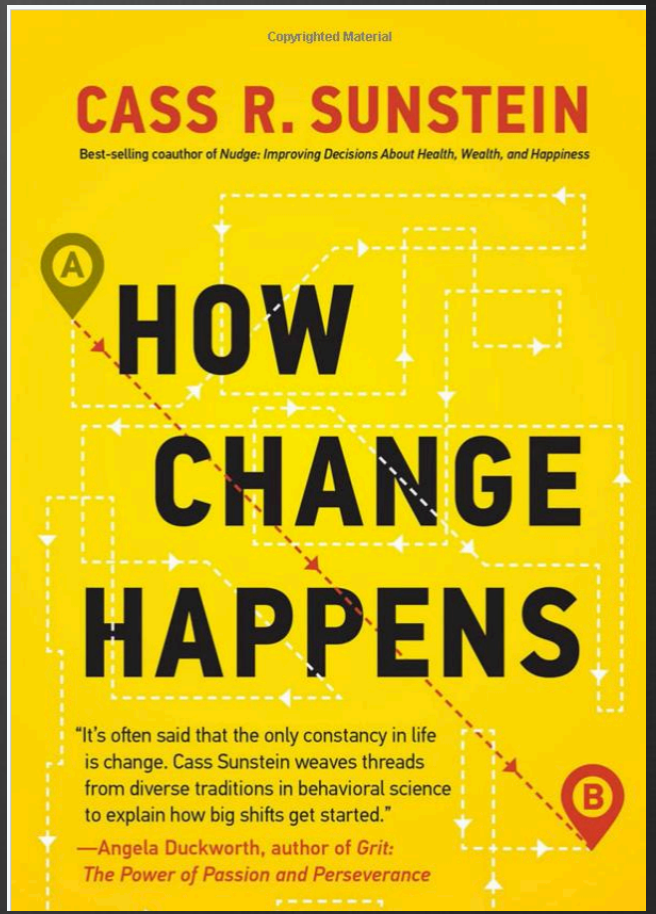
- How can one do reproducible science with limited resources?
  - Perils of the pilot study
  - Other options
- Talking about reproducibility – Is Science broken?

How should we talk about the problem?

http://www.nasonline.org/about-nas/events/annual-meeting/nas156/symposium.html

How should we talk about the reproducibility problem?





View

False Stories Spread Fast. So Do Some True Ones.

The falsehood problem is measurably alarming. Beware of ready solutions.

By Cass R. Sunstein
March 13, 2018, 9:00 AM EDT

http://www.nasonline.org/about-nas/events/annual-meeting/nas156/symposium.html

How should we talk about the problem?

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

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for

factors that influence this problem and some corollaries thereof.

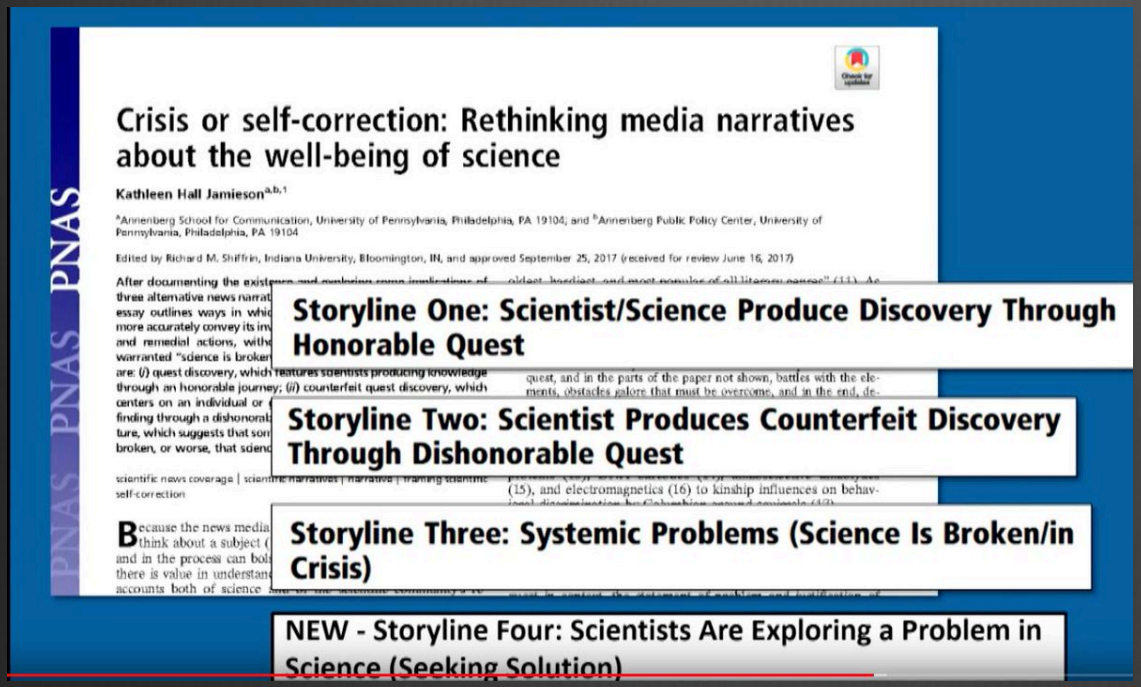
Modeling the Framework for False Positive Findings

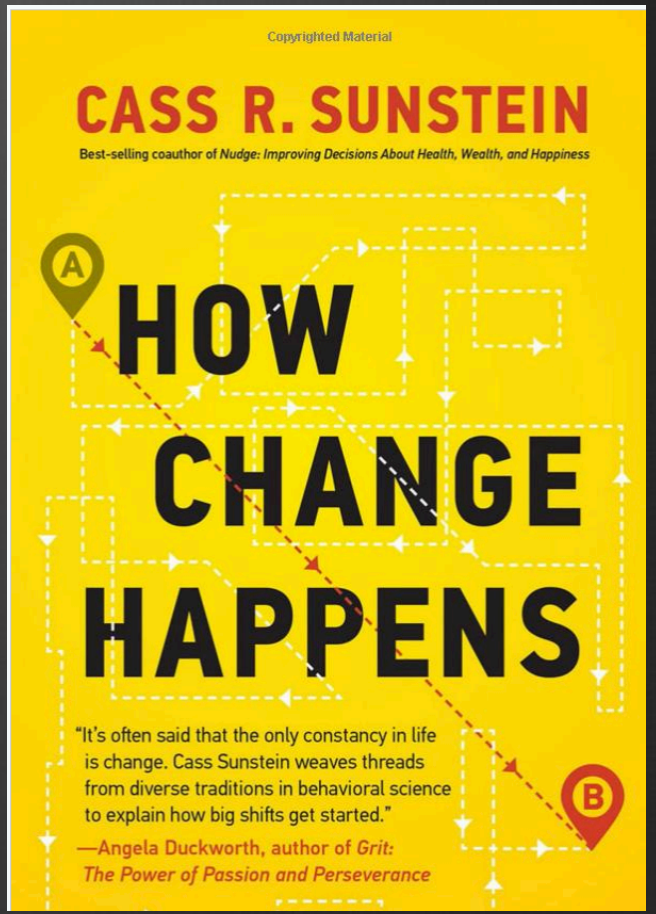
Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R + 1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that c relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research

How should we talk about the reproducibility problem?





View

False Stories Spread Fast. So Do Some True Ones.

The falsehood problem is measurably alarming. Beware of ready solutions.

By Cass R. Sunstein
March 13, 2018, 9:00 AM EDT

http://www.nasonline.org/about-nas/events/annual-meeting/nas156/symposium.html

How should we talk about the reproducibility problem?

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Anders Eklund, Thomas E. Nichols, and Ha

PNAS July 12, 2016 113 (28) 7900-7905; first publishe

Edited by Emery N. Brown, Massachusetts General H
review February 12, 2016)

This article has a correction and Letters. Pl

Correction for Eklund et al., Cluster failur
inflated false-positive rates

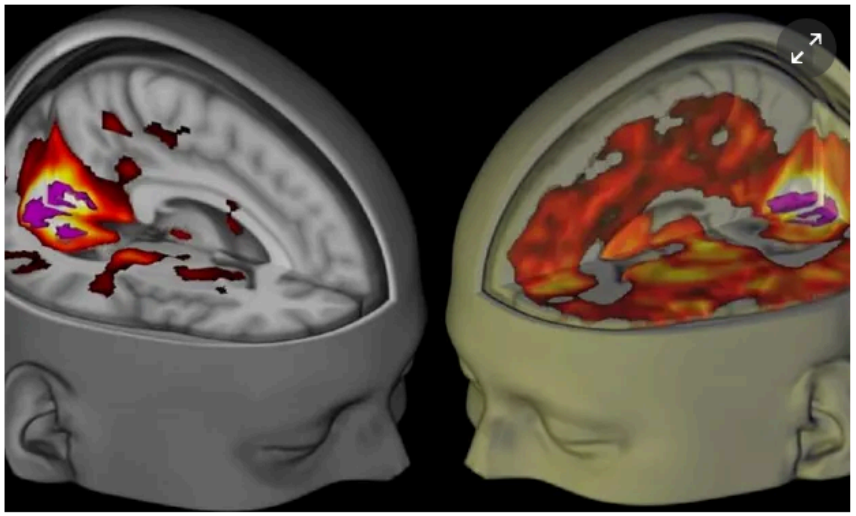
fMRI clustering and false-positive rates

Controversy in statistical analysis of func

Reevaluating "cluster failure" in fMRI usir

Has a software bug really called decades of brain imaging research into question?

Over the summer, some headlines suggested that a study highlighting issues in the way we analyse fMRI data renders the technique irretrievably flawed. But the reality is much more nuanced



▲ Is functional MRI research in a sorry state, or is the science just going through a normal process of refinement and improvement? Photograph: Handout/Reuters

Take homes

- Running underpowered studies is frustrating and futile
- It's possible to design reproducible neuroimaging studies with limited resources using shared data, targeted questions, and subject-centered designs
- When talking/tweeting/publishing about problems in science, carefully consider the potential downsides of your message.

- Thanks!
- Questions?
- Slides and older talks at http:/fmrif.nimh.nih.gov

Thanks!
Data Science & Sharing Team

Adam Thomas

Dustin Moraczewski
July 15th

Travis Riddle

IRTA: Nikhil Goyal
Sept 2019

- NIMHDSST@mail.nih.gov
- http://cmn.nimh.nih.gov/dsst
- Building 10, Room 3D39
- Open door policy
- Co-located with the Machine Learning Team & the Scientific Information Officer, Asim Okur