The parts of an fMRI study from design to interpretation

Daniel Handwerker

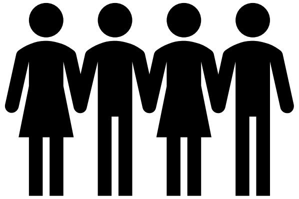
Section on Functional Imaging Methods
National Institute of Mental Health

July 7, 2026

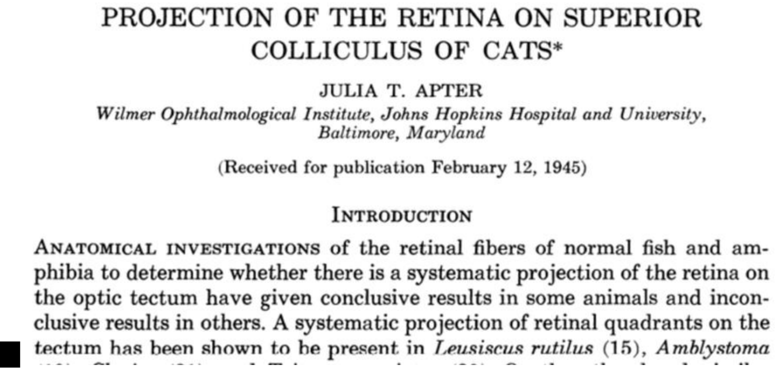
National Institute
of Mental Health

The basic fMRI study

+



=



PROJECTION OF THE RETINA ON SUPERIOR COLLICULUS OF CATS*

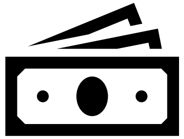
JULIA T. APTER

Wilmer Ophthalmological Institute, Johns Hopkins Hospital and University, Baltimore, Maryland

(Received for publication February 12, 1945)

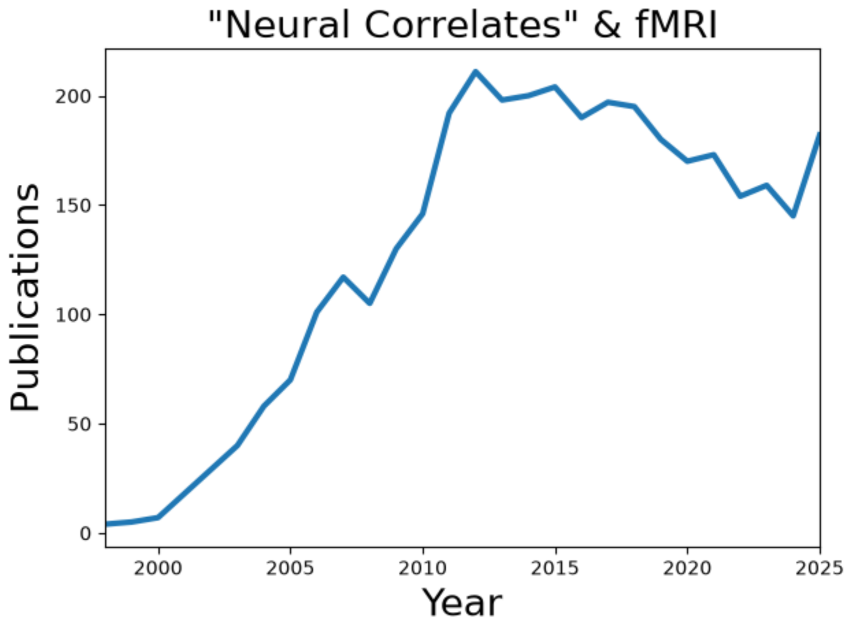
INTRODUCTION

ANATOMICAL INVESTIGATIONS of the retinal fibers of normal fish and amphibia to determine whether there is a systematic projection of the retina on the optic tectum have given conclusive results in some animals and inconclusive results in others. A systematic projection of retinal quadrants on the tectum has been shown to be present in *Leusiscus rutilus* (15), *Amblystoma*



“Neural correlates” publications in 2001

- Neural correlates of driving.
- Neural correlates of conscious self-regulation of emotion.
- Neural correlates of change detection and change blindness.
- Neural correlates of attention and working memory deficits in HIV patients.
- Neural correlates of verbal memory encoding during semantic and structural processing tasks.
- Neural correlates of emotions in psychiatric patients in the light of functional neuroimaging findings.
- Neural correlates of traumatic memories in posttraumatic stress dis
- Neural correlates of dual task interference can be dissociated from study.
- Neural correlates of person familiarity. A functional magnetic resona implications.
- Neural correlates of response inhibition for behavioral regulation in magnetic resonance imaging.
- Neural correlates of formal thought disorder in schizophrenia: prelir magnetic resonance imaging study.



Goals of this talk

Learn key phases of fMRI study design and analysis

Understand the types of decisions to make in each phase

Ideal: Know when to ask for help

Not the goals of this talk

Learn best options for each phase of a study

Explain all fMRI study design and analysis methods

Cannot cover details of all fMRI studies

Task fMRI: Each voxel fit to a predefined model

Visual temporal frequency preference shows a distinct cortical architecture using fMRI

Yuhui Chai [a,*], Daniel A. Handwerker [a], Sean Marrett [b], Javier Gonzalez-Castillo [a], Elisha P. Merriam [c], Andrew Hall [a], Peter J. Molfese [a], Peter A. Bandettini [a,b]

Resting fMRI: Correlations between predefined regions

Theta-burst TMS to the posterior superior temporal sulcus decreases resting-state fMRI connectivity across the face processing network

In Special Collection: CogNet

Daniel A. Handwerker, Geena Ianni, Benjamin Gutierrez, Vinai Roopchansingh, Javier Gonzalez-Castillo, Gang Chen, Peter A. Bandettini, Leslie G. Ungerleider, David Pitcher

Task fMRI with <1mm^3 resolution (mesoscopic)

Topographical and laminar distribution of audiovisual processing within human planum temporale

Yuhui Chai [a,*], Tina T. Liu [b], Sean Marrett [c], Linqing Li [c], Arman Khojandi [a], Daniel A. Handwerker [a], Arjen Alink [d], Lars Muckli [e], Peter A. Bandettini [a,c]

Task fMRI with multi-voxel pattern analyses

Revealing representational content with pattern-information fMRI—an introductory guide

Marieke Mur,[1,2] Peter A. Bandettini,[1,3] and Nikolaus Kriegeskorte[1]

Task & Rest with whole brain connectivity measures

Manifold learning for fMRI time-varying functional connectivity

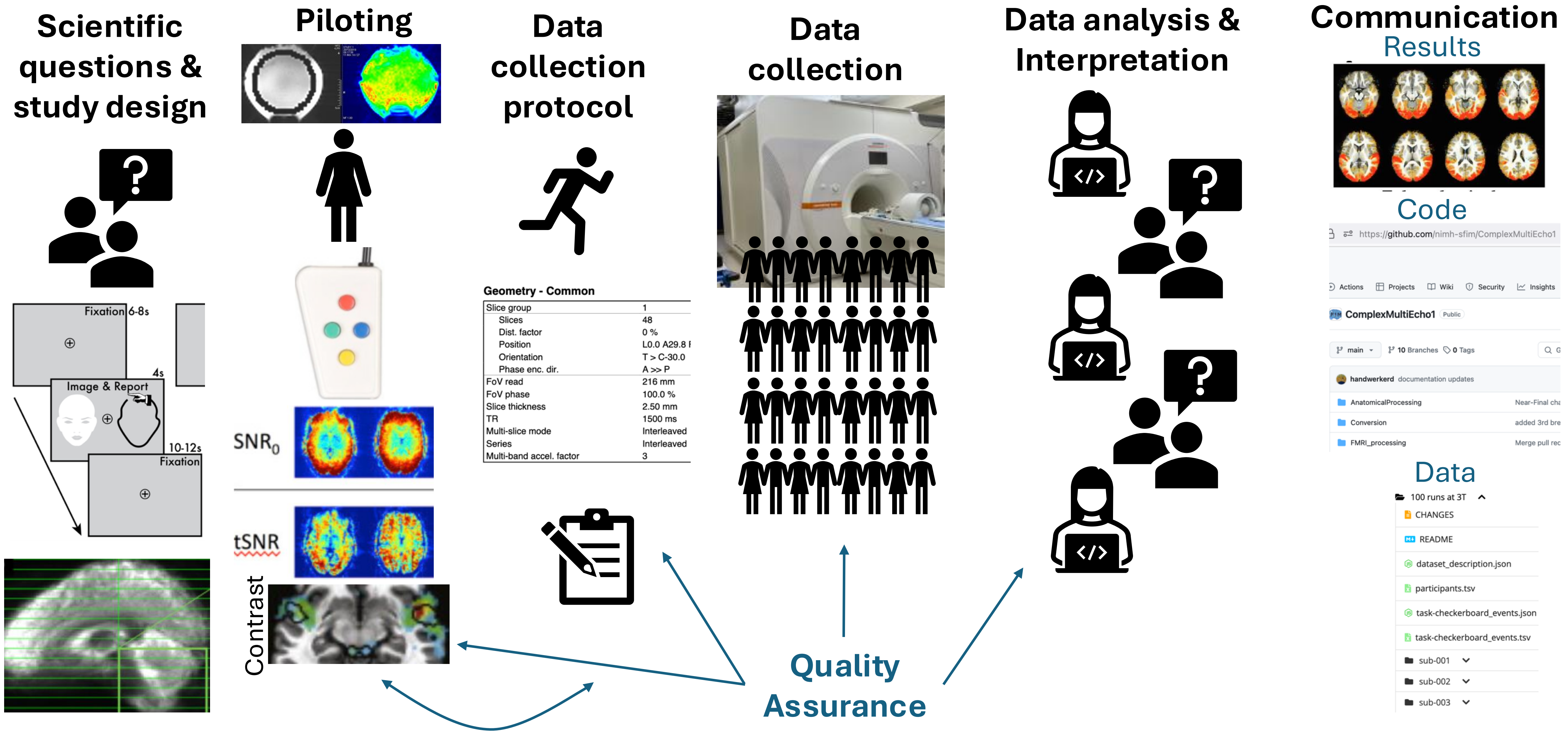
Javier Gonzalez-Castillo[1*], Isabel S. Fernandez[1], Ka Chun Lam[2], Daniel A. Handwerker[1], Francisco Pereira[2] and Peter A. Bandettini[1,3]

Movie viewing (naturalistic) with correlations across people

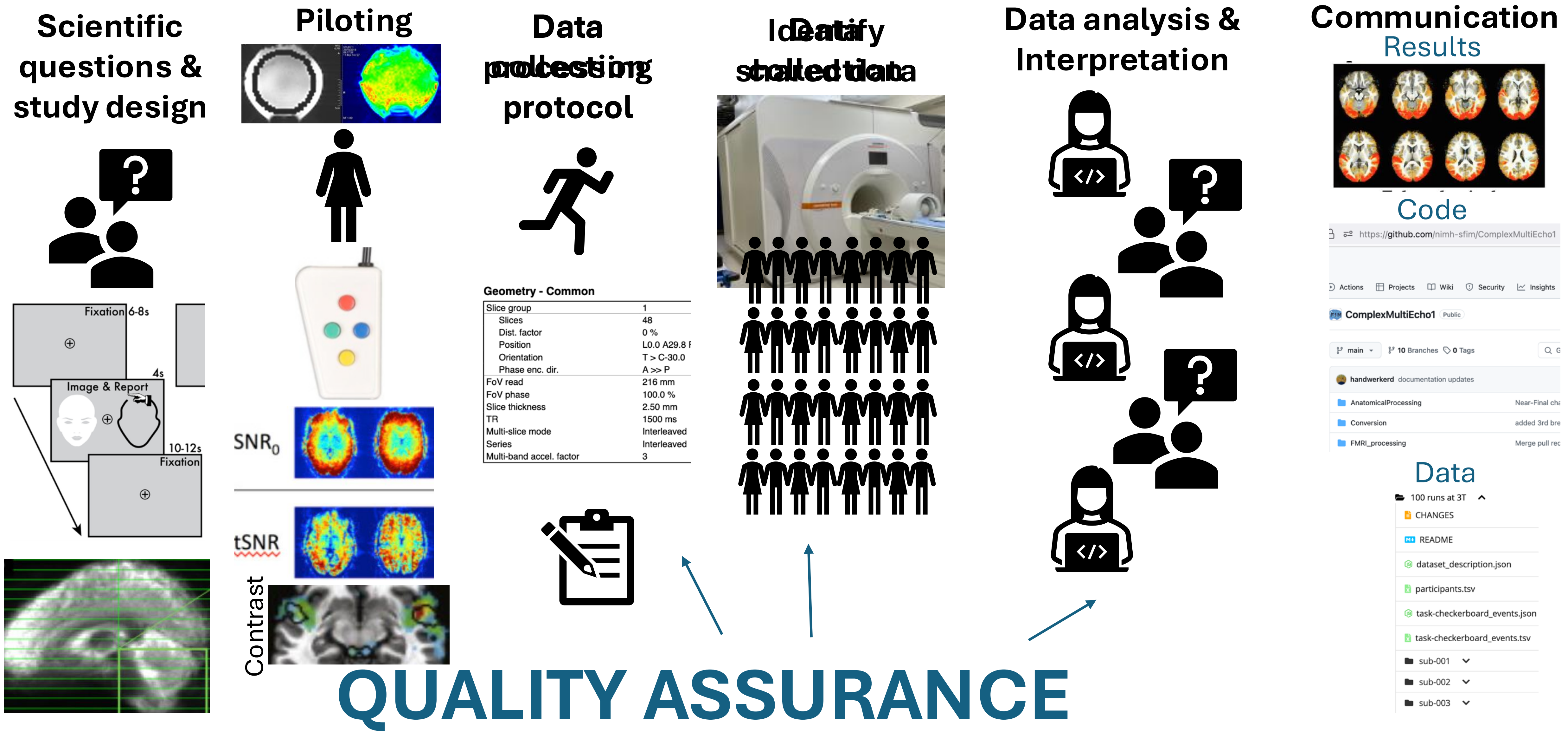
Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging

Emily S. Finn [a,*], Enrico Glerean [b], Arman Y. Khojandi [a], Dylan Nielson [c], Peter J. Molfese [a], Daniel A. Handwerker [a], Peter A. Bandettini [a]

The basic fMRI study



The basic fMRI study (existing data)

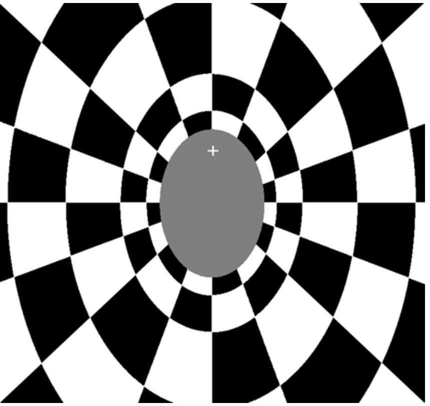


Scientific Questions & Study Design
What measures to collect?

- **What is your question?**
- **What types of data have the potential to accurately and effectively answer your scientific question?**
- Task? Movie? Rest?
  - Task design, which movies, what type of “rest”
- Structural MRIs? Calibration measures?
- Simultaneous acquisitions: Pulse, Respiration, EEG, eye tracking, participant responses, …

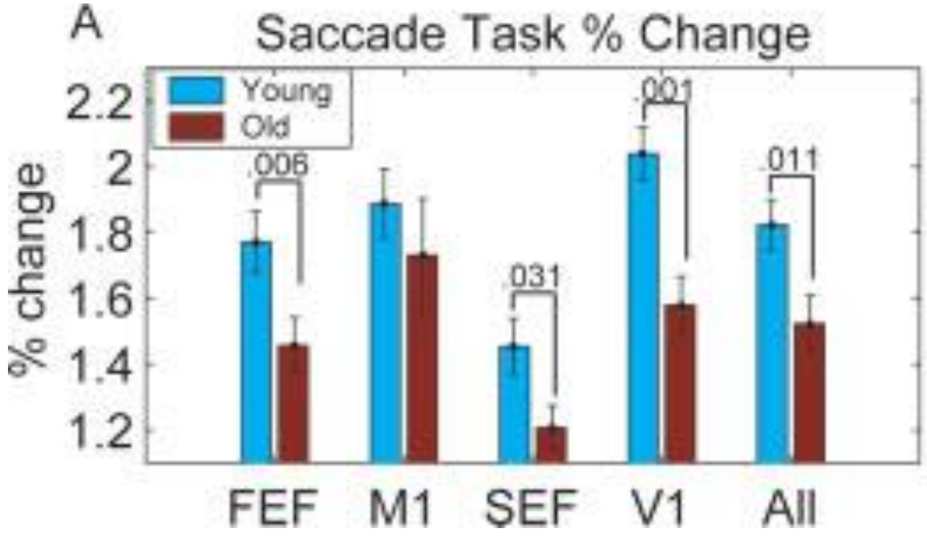
Don’t neglect peripheral measurements!

A breathless cautionary tale



Present a 200ms flickering checkerboard every 18-24s

Volunteers press a button and move their eyes



Handwerker, Gazzaley, et al 2007

The unpublished part

- Stimuli presented for 3s, 6s & 12s durations
- A non-trivial # of volunteers held their breath for the stimulus duration
  - Breath holding causes a robust, brain-wide BOLD response
- No respiration data →

  Visually appealing results →

  Misinterpreted results

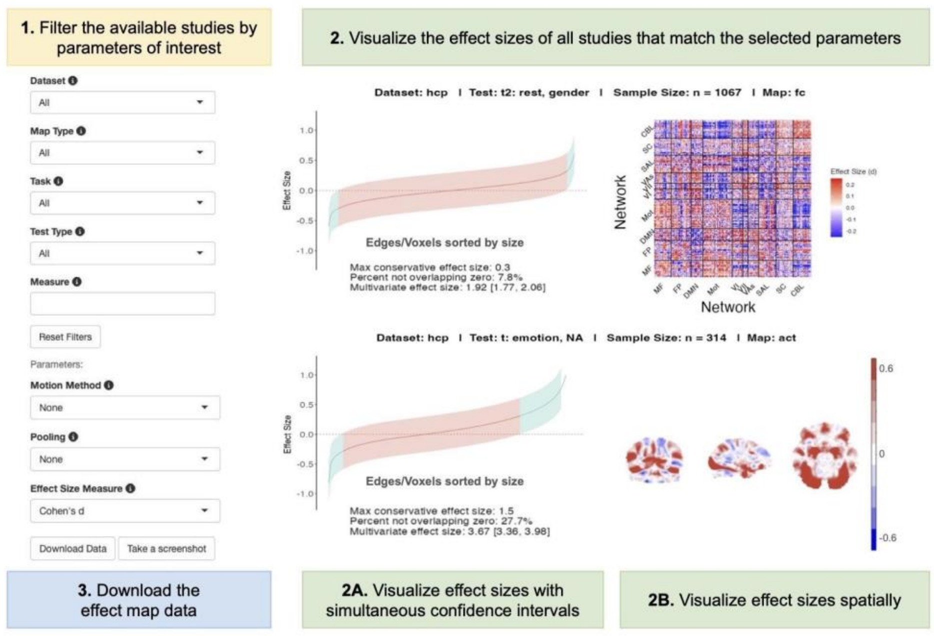
See also: Birn, Murphy, et al, NeuroImage 2009

Scientific Questions & Study Design
Study Level Questions

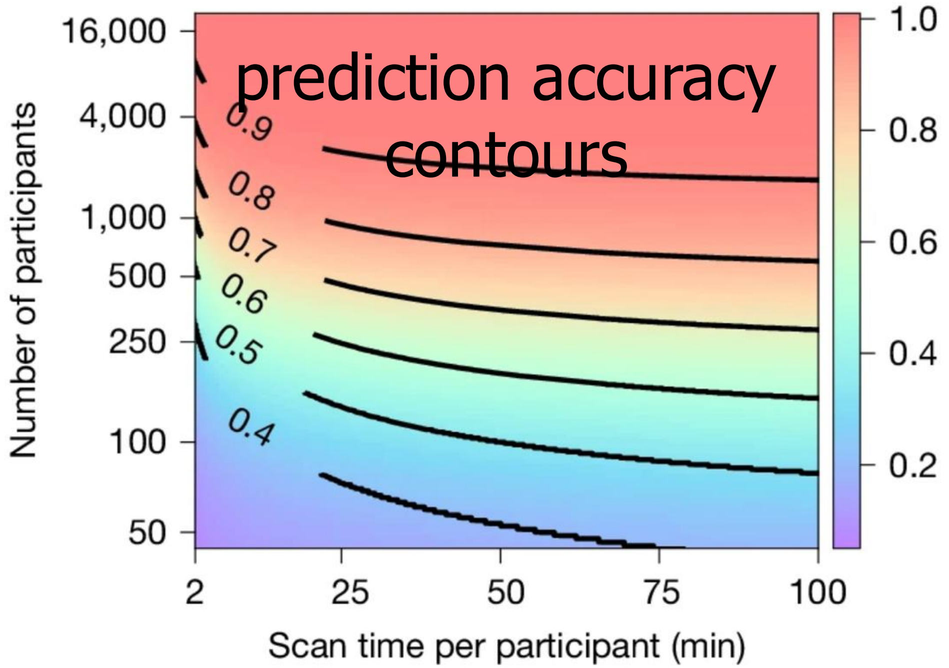
- Number of participants & amount of data / participants
  - Looking for individual-level effects?
  - Looking for consistent group effects?
  - Looking for behavioral/fMRI response variation across participants?
  - Looking for diagnostic or treatment relevant markers?

Scientific Questions & Study Design

Study Level power analyses are important



Shearer, et al. BrainEffeX: A web app for exploring fMRI effect sizes. *Aperture Neuro* 2025 https://neuroprismlab.shinyapps.io/BrainEffeX



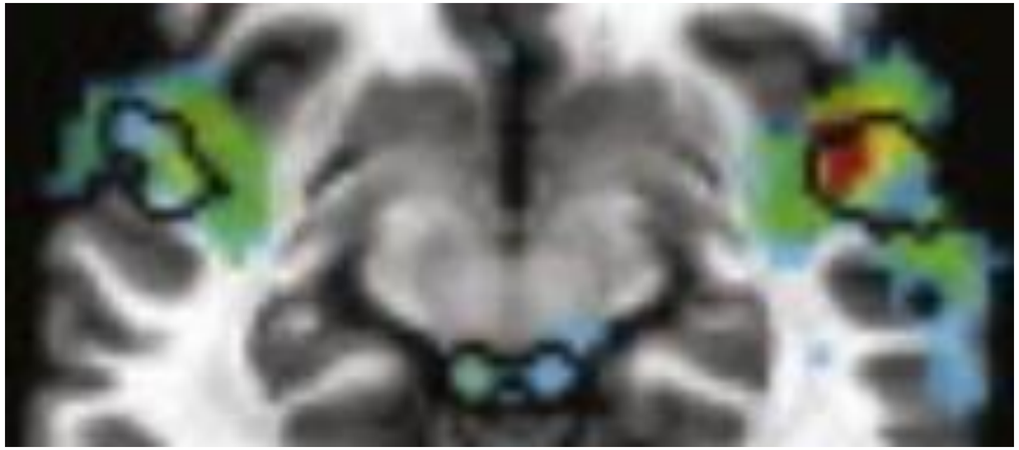
Ooi, Orban, et. al. *Nature* 2024 *Longer scans boost prediction and cut costs in brain-wide association studies*

Power analyses can be over-emphasized and over-simplified

Making sure each dataset is high quality will improve results more than increasing sample size or focusing on statistical thresholding methods.

Scientific Questions & Study Design
fMRI acquisition parameters

- No universally best parameters
- Never copy parameters from another study without understanding why they made their choices
- The question drives the study design. What matters?
  - Brain coverage
  - Dropout
  - Distortion
  - Temporal Signal-to-Noise Ratio (TSNR)
  - Contrast-to-Noise Ratio (CNR)
  - Head motion sensitivity



Scientific Questions & Study Design

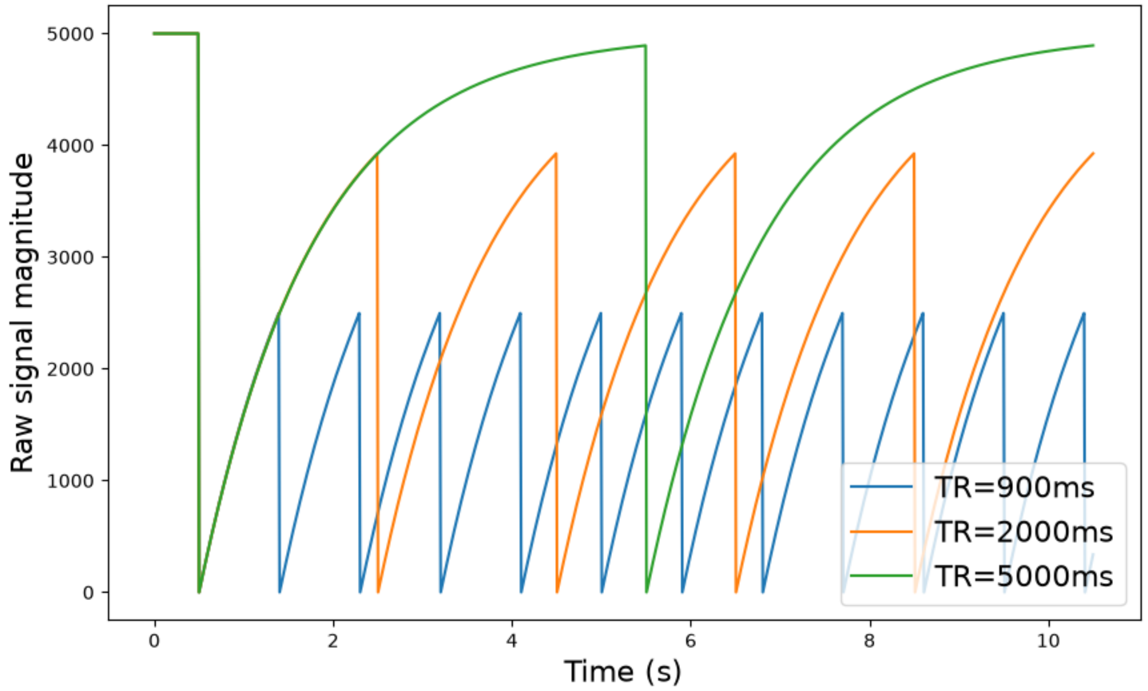
fMRI acquisition parameters

- TR (sampling rate)
- Spatial resolution
- Acceleration methods
- 2D vs 3D
- Multi-echo fMRI
- Non-BOLD contrasts
- Flip angle

Scientific Questions & Study Design
fMRI acquisition parameters

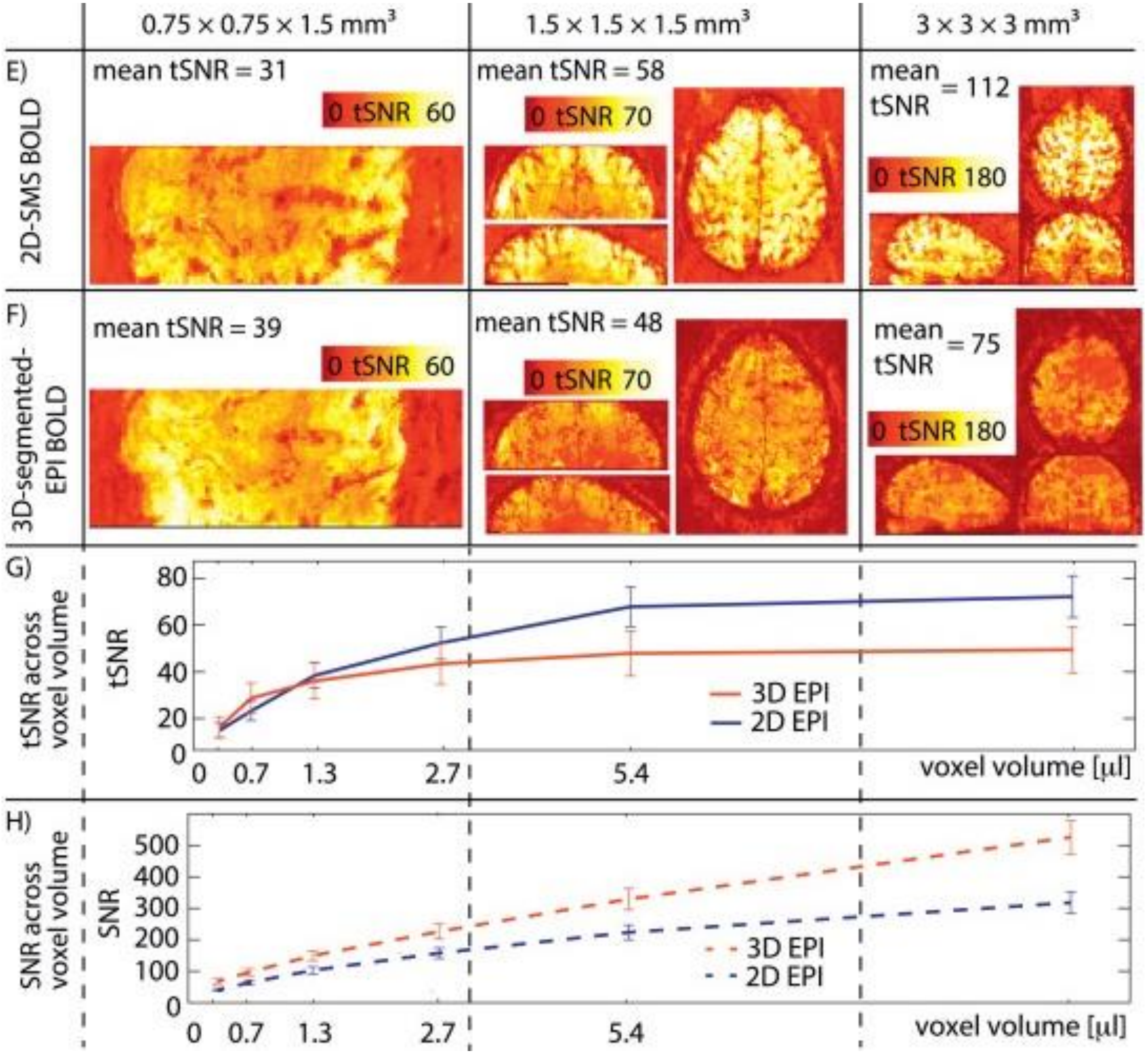
TR (sampling rate)

- Finer temporal resolution
- More volumes over the same time
- More noise per volume (acceleration & steady state diffs)
- Acceleration to get faster TRs potentially makes more artifacts
- Possibly less brain coverage and larger voxels



Scientific Questions & Study Design

fMRI acquisition parameters



- Smaller voxels can give more spatial precision
- More noise
- Slower

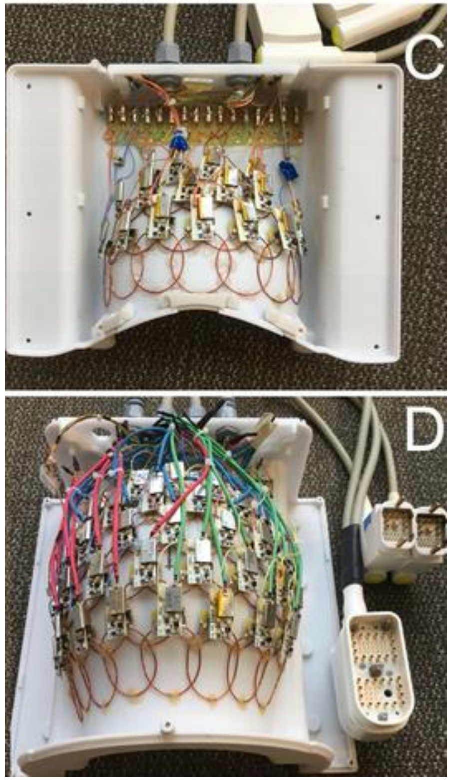
Huber, NeuroImage 2018

Piloting

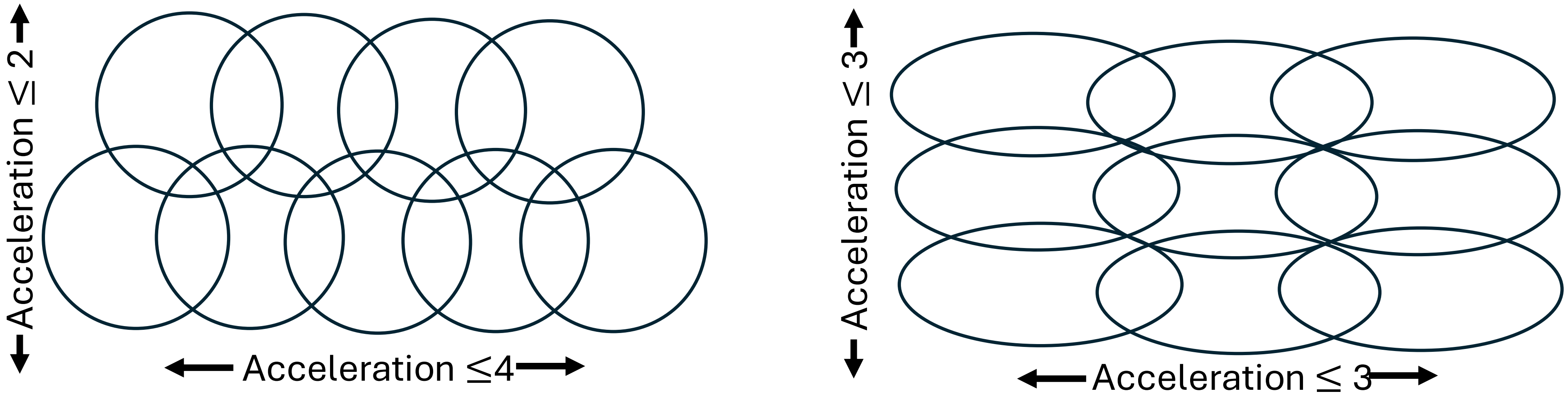
- Comparing options for potential study designs
- Collecting data to make sure design decisions are working as expected
- Making sure acquisition workflow works with real people
- Decide how to log key information, store, and process data
- Start to make a Quality Assurance (QA) workflow

Piloting

- Do not use scanning parameters just because they worked for someone else
- Do not use scanning parameters just because they worked for someone else
- If you’re **not an expert** in MRI physics, things you did not consider might affect data quality.



Alternate geometries for a 9-channel head coil

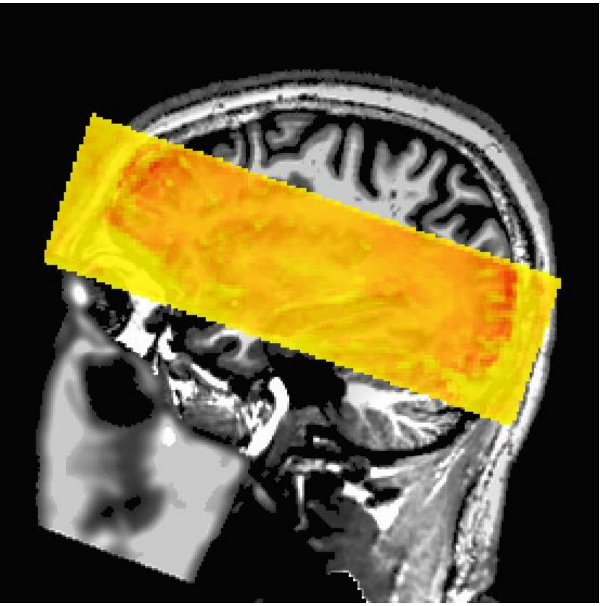


- If you **are an expert** in MRI physics, you already know you should collect pilot data

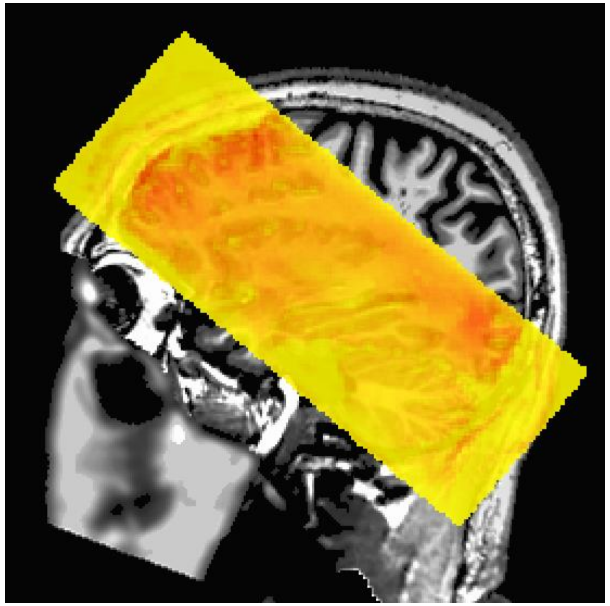
A piloting example: Object category memory task

TSNR

2D Acq

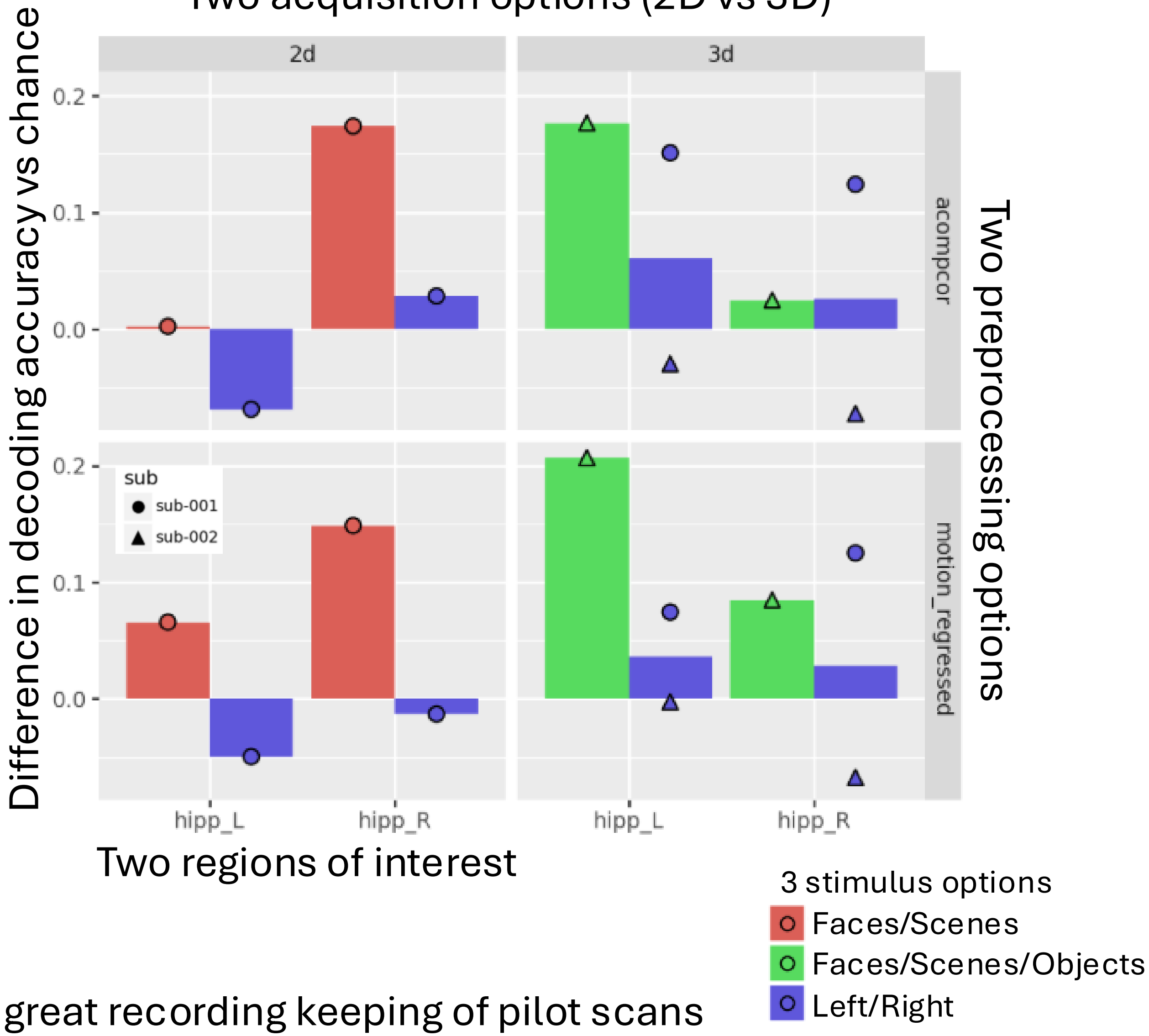


3D Acq



	TSNR			
Acquisition	Hipp L	Hipp R	DLPFC L	DLPFC R
2D	13.6	13.1	11.1	15.5
3D	19.5	19.4	24.8	39.4

Two acquisition options (2D vs 3D)



Thanks to Catherine Walsh for data & visualizations… and for great recording keeping of pilot scans

Data Collection Protocol

- Checklists and workflows for both regulations, safely, and experimental needs
- What to save?
  - MRI, task behavioral logs, respiration/cardiac, acquisition timing, anything atypical

Lots of hair on top of head. **1-2cm lower in the coil** than most other volunteers
There seems to be a dark area slicing through the occipital lobe. Looking at the anatomical, it might be that a slightly larger than normal sulcus is perfectly horizontal in the slice. No other obvious scanner-based causes for this issue. **Worth looking at more carefully during data processing.**

Scan 5 WNW run 4, 350 volumes. **Said volume was a bit low and hard to hear.** Everything looked reasonable from experimenter's side. Raised volume for subsequent runs.
Scan 6 WNW run 1 Better volume but asked to have it raised slightly more for next scan
Scan 7 WNW run 8
Scan 8 Movie with respiration pattern A, 304 volumes. **Noise cancelation out of range and reset a few times between volumes 170 & 220 (volume lowered slightly). Again at volume 250**
Scan 9 Respiration pattern A
Scan 10. Noise cancellation sounded off (heard bad audio artifact) so stopped scan & restarted optoacoustics system. Artifact was scanner noise volume increasing & decreasing in volume in a cycle as if noise cancellation was out of phase.
Scan 11 Movie with respiration pattern B
Scan 12. Same noise cancellation artifact. Turned off system & restarted

Data Collection Protocol

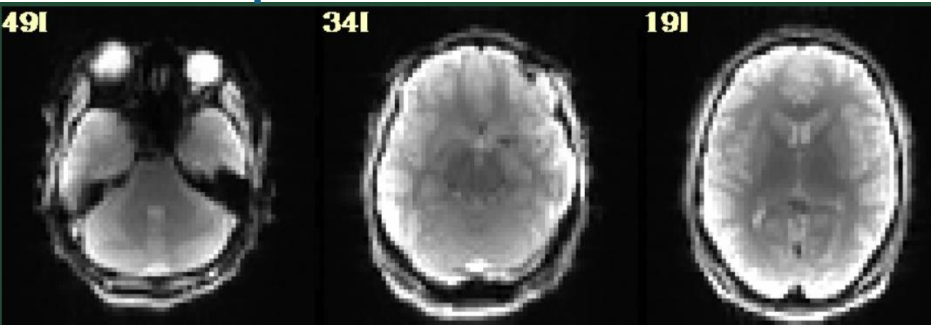
- Organize data in ways that facilitate future sharing
  - Consistent structure across individuals both for: data, information explaining the data, scanning notes, & results
  - Making analysis workflows that your future self will understand
- Quality Assurance steps to run **during and soon** after each acquisition
  - Behavioral tasks presented & performed as expected?
  - Data quality looks as expected
  - (Come back for July 28th presentation on Quality Assurance)
  - Teves et al "The art and science of using quality control to understand and improve fMRI data" Front. Neurosci. (2023)

Data Collection!

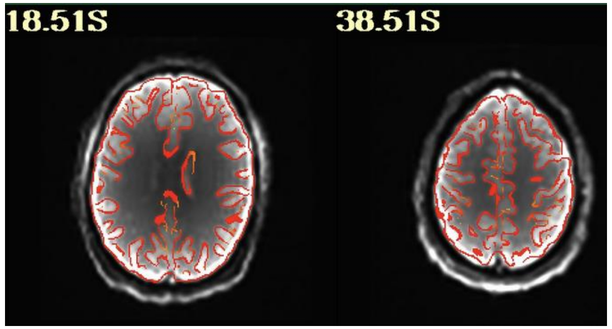
- If planning was done well, then data collection follows a script
- You will go off script. Log when you go off script Decide if the script should change
- QA soon after each acquisition
  - Identify fixable problems before a lot of data are collected
  - Identify data anomalies or unexpected variations that might skew or hide key results
  - Do acquisition parameters match the protocol?
- **All datasets have problems**
  - Not checking → Incorrect or misleading interpretations of results
  - Checking → **Fewer** unknown problems

AFNI automatic QC Report after processing data

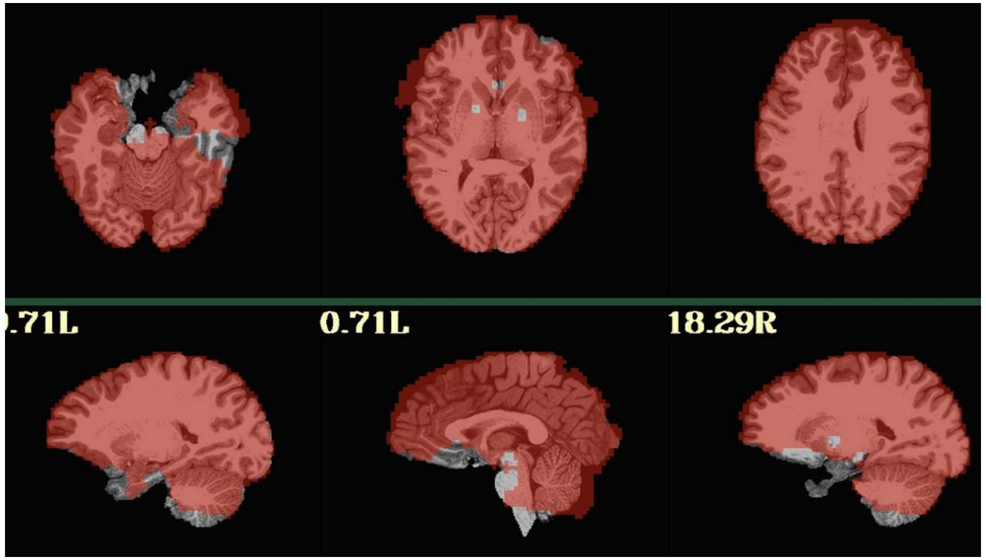
Unprocessed fMRI EPI



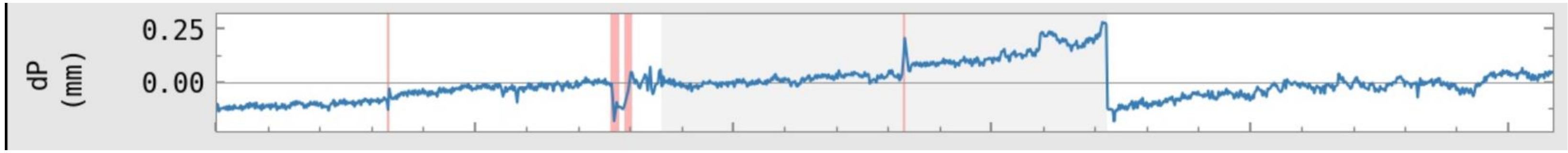
EPI to anatomical alignment



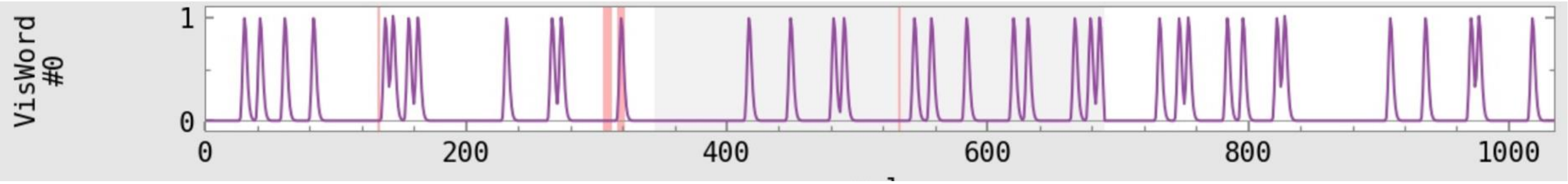
Brain regions without fMRI data



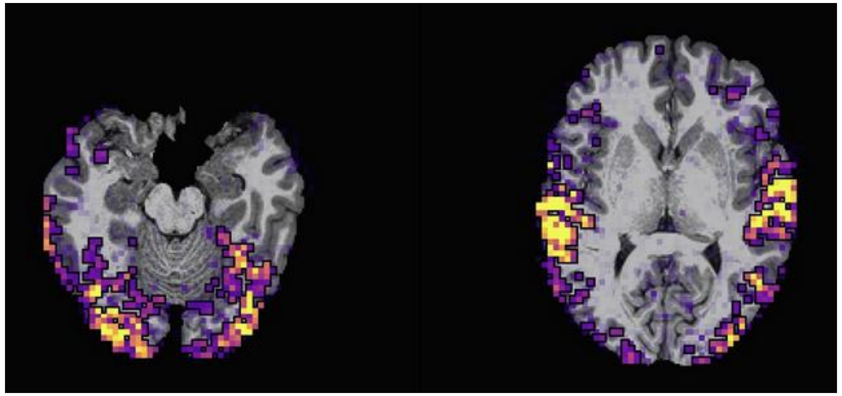
Head motion & censored volumes



Modeled task timing and censored volumes



Statistical Maps



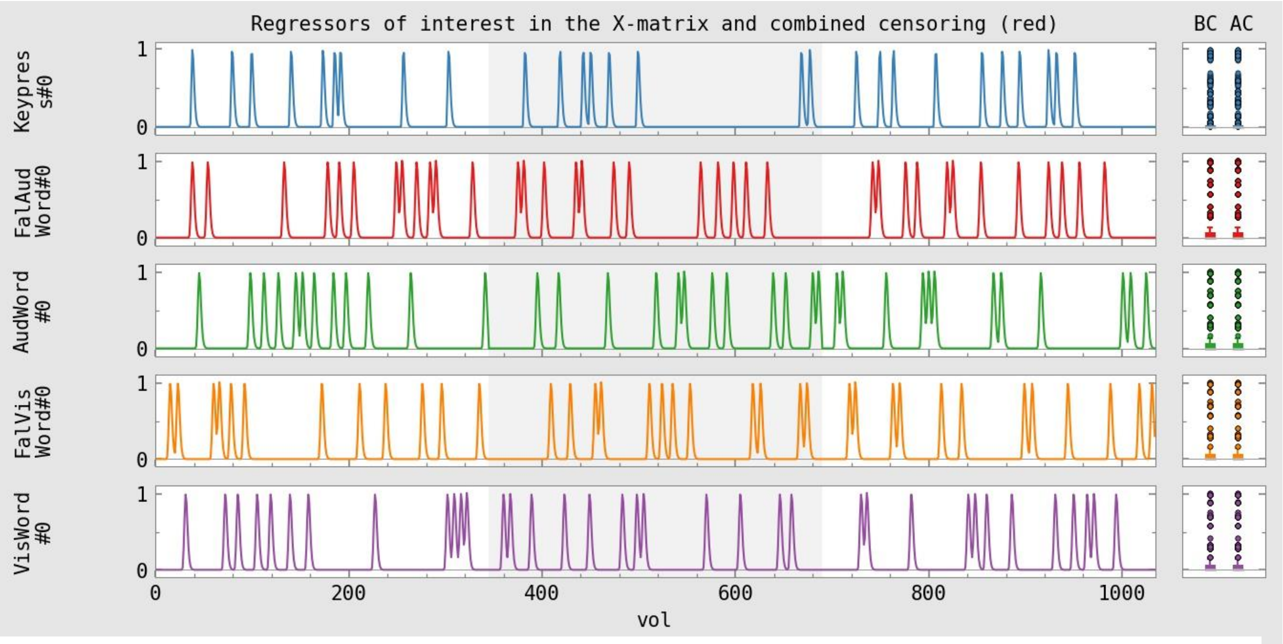
Analysis and interpretation
General advice

- Make sure analysis scripts fail in clear ways
- Save intermediate processing data
- Look at your data!
- Track provenance (how processed data were created)

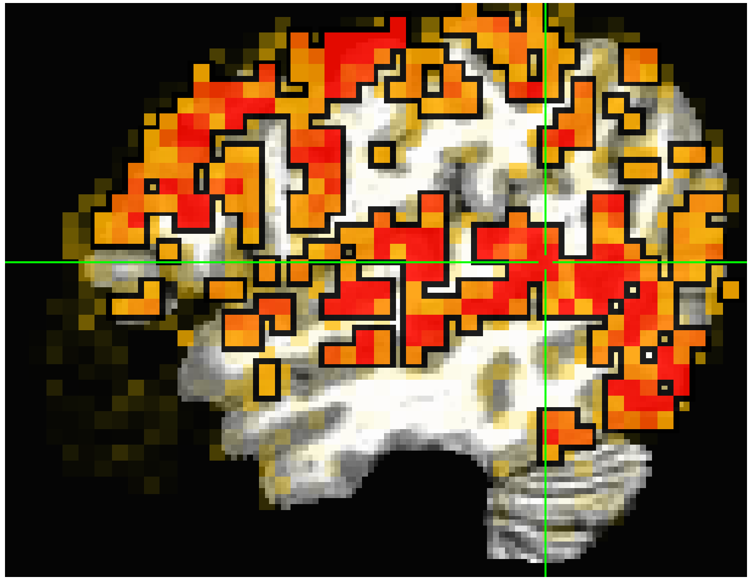
Analysis and interpretation

Task-based General Linear Model (GLM)

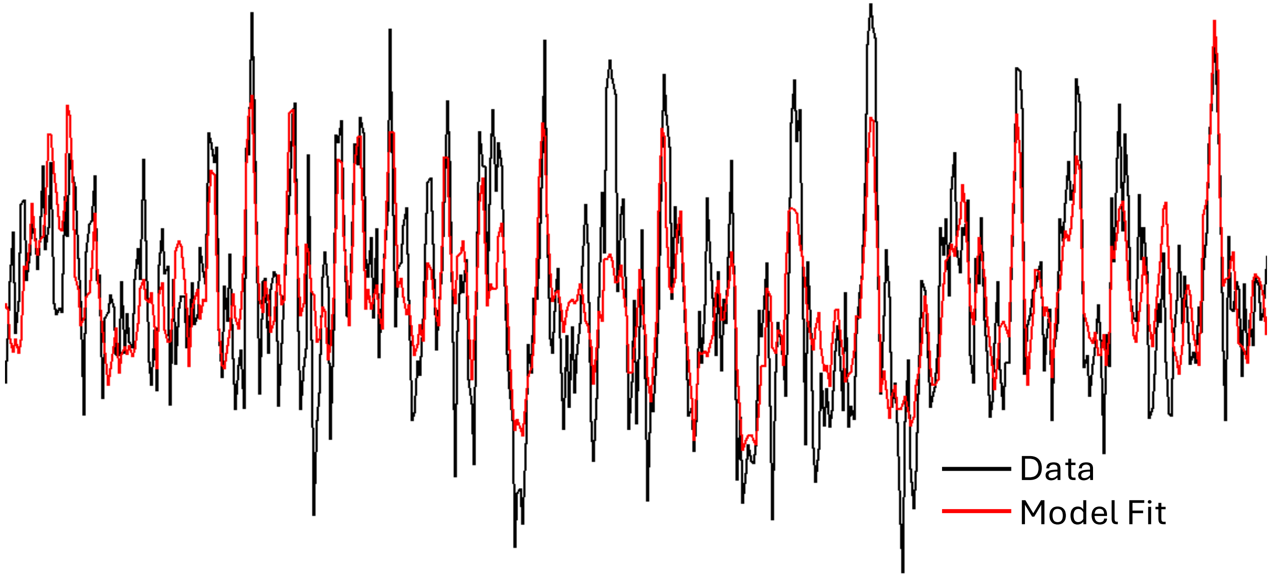
Task design



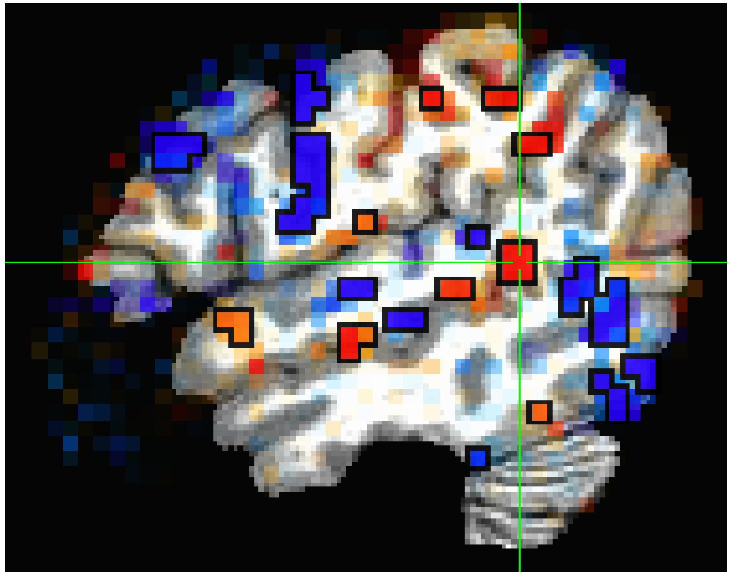
F statistic fit to full model



Single Voxel

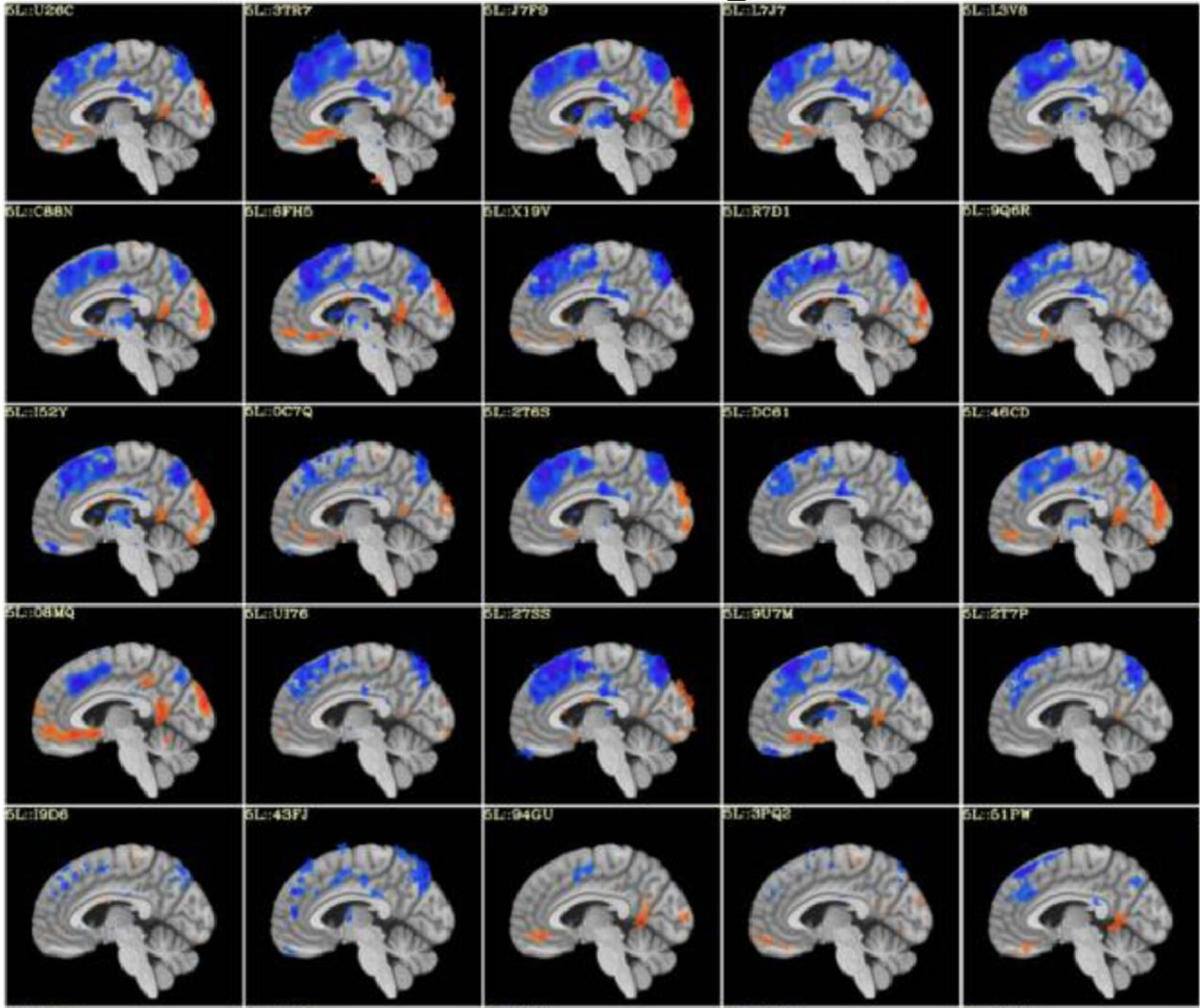


t statistic fit to Word – Nonword contrast



Interpretation & Communication

"Not significant" ≠ meaningless

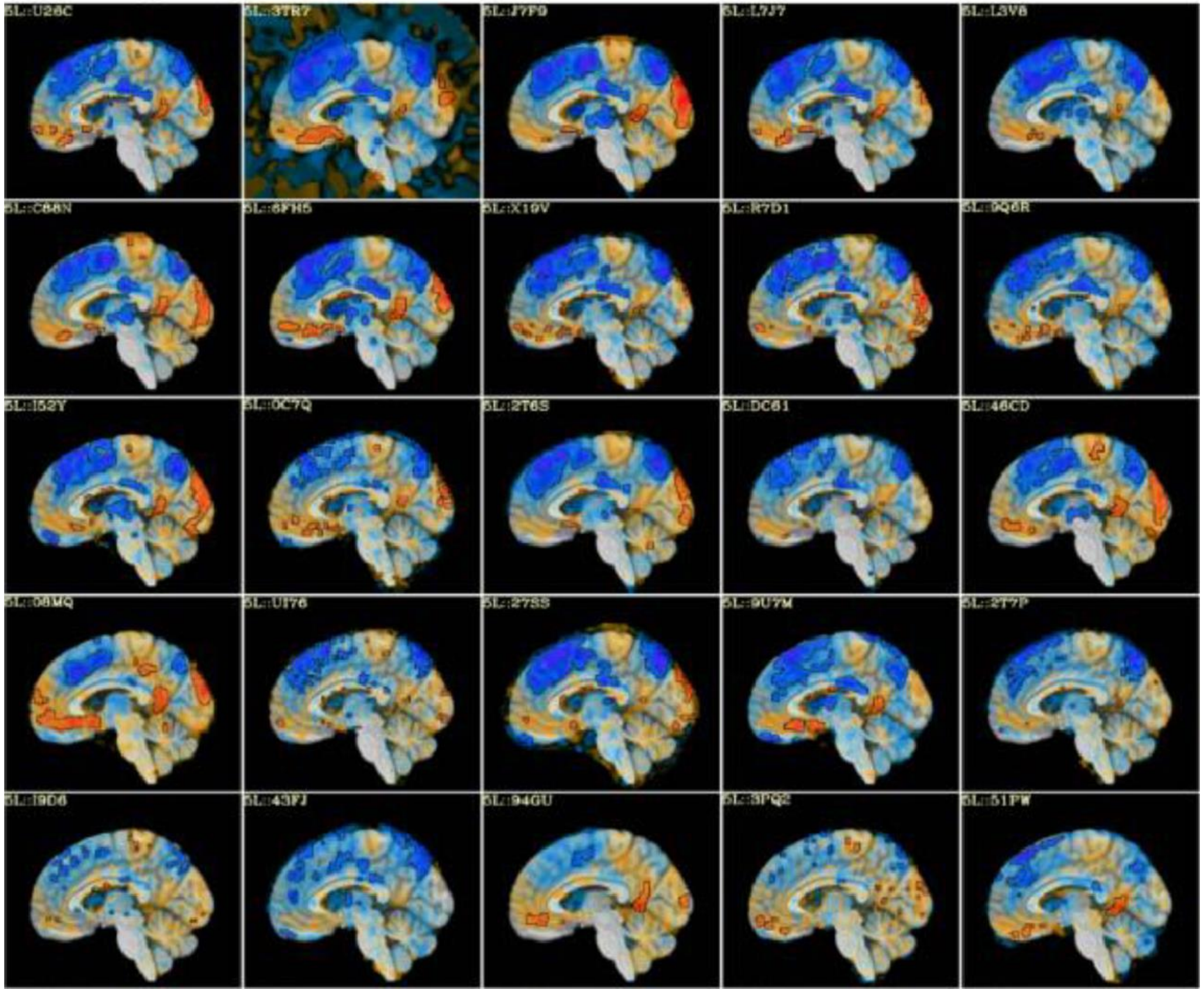


Variability in the analysis of a single neuroimaging dataset by many teams
"This flexibility resulted in sizeable variation in the results of hypothesis tests, even for teams whose statistical maps were highly correlated at intermediate stages of the analysis pipeline"
Botvinik-Nezer *et al.*. *Nature* 2020

Revisualization of data:
Taylor et al "Highlight results, don't hide them..."
NeuroImage 2023

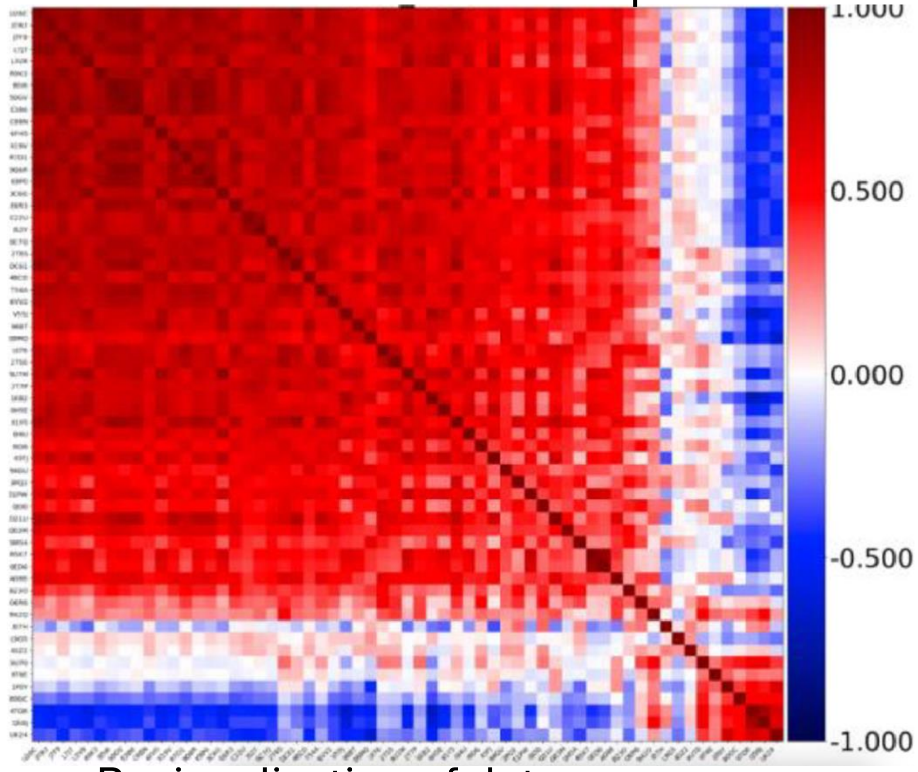
Interpretation & Communication

“Not significant” ≠ meaningless



Statistical significance thresholds are essential AND arbitrary

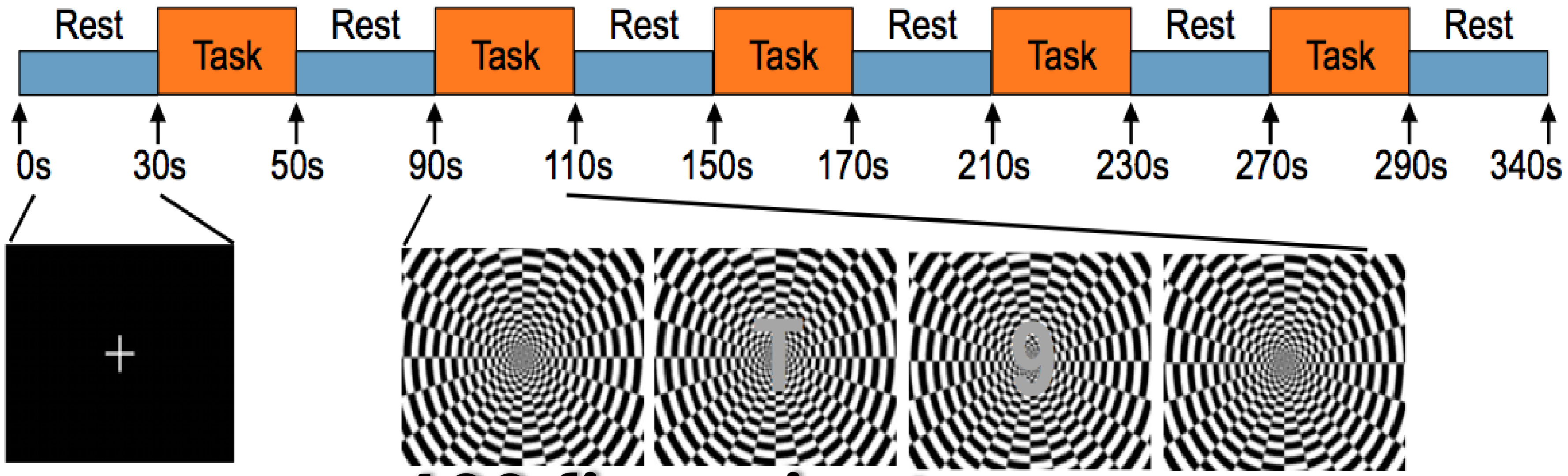
Correlation coefficients between unthresholded maps



Revisualization of data:
Taylor et al “Highlight results, don’t hide them…” *NeuroImage* 2023

Interpretation
What is significant?

3 volunteers

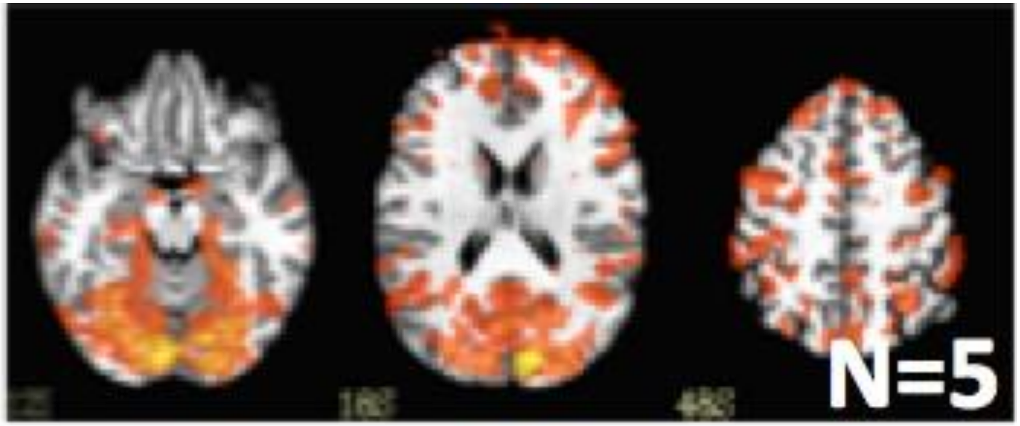


100 five minute runs
9 hours of functional data per volunteer

Gonzalez-Castillo … Handwerker … PNAS 2012

Interpretation

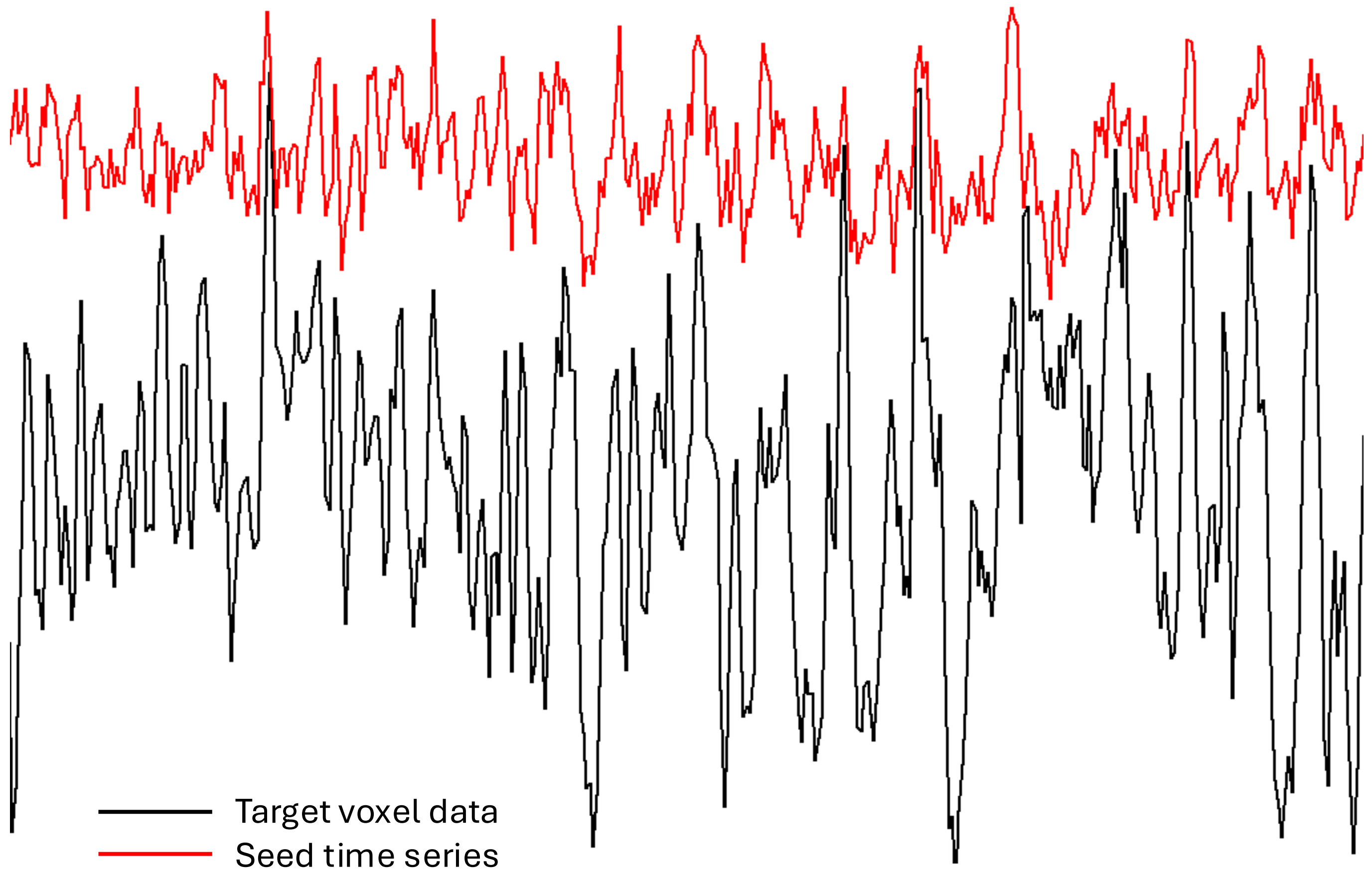
More data -> More significant voxels

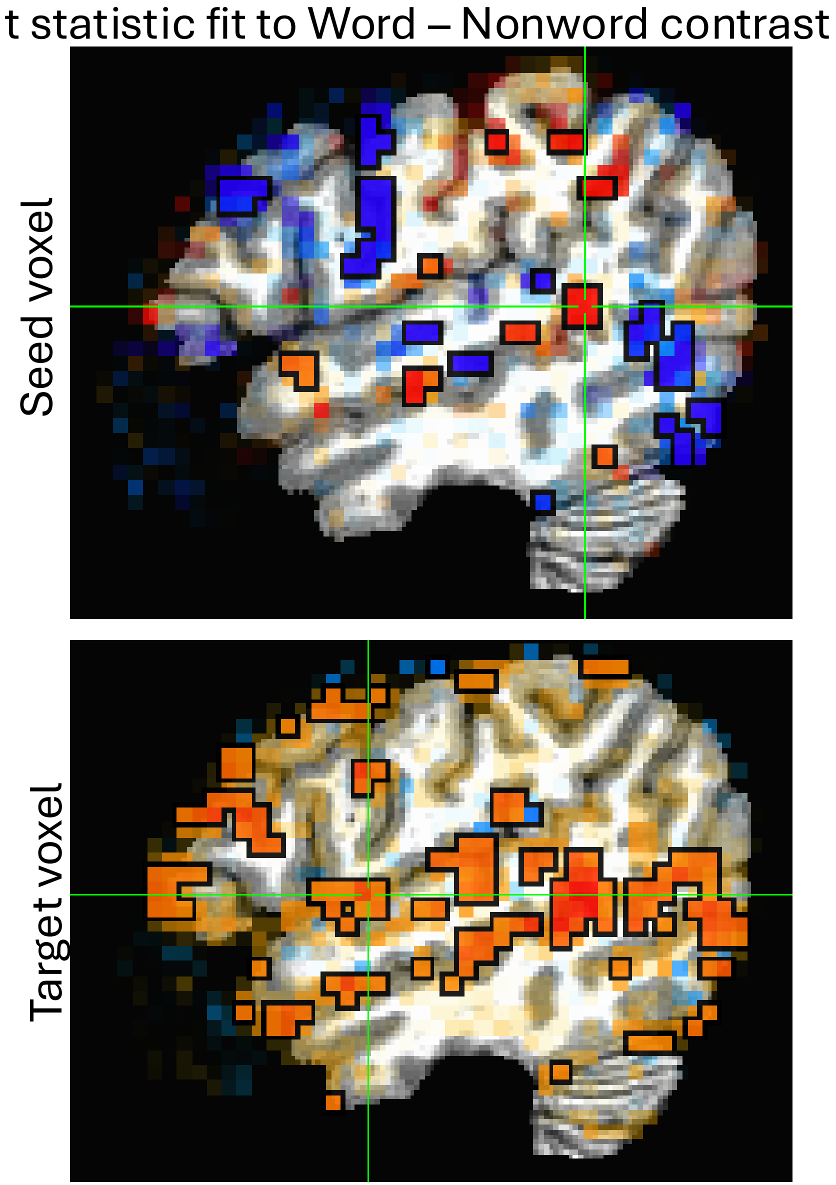


Gonzalez-Castillo ... Handwerker ... PNAS 2012

Analysis and interpretation

Functional Connectivity (correlation)

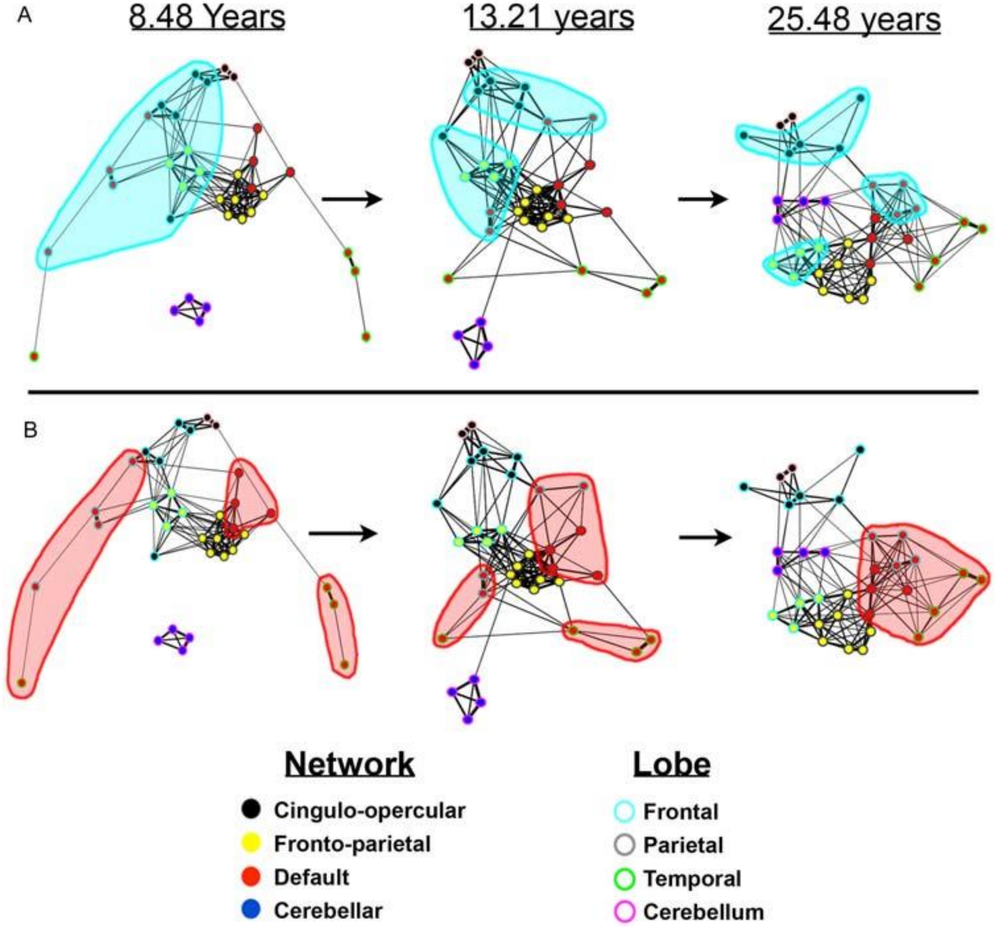




Interpretation limits

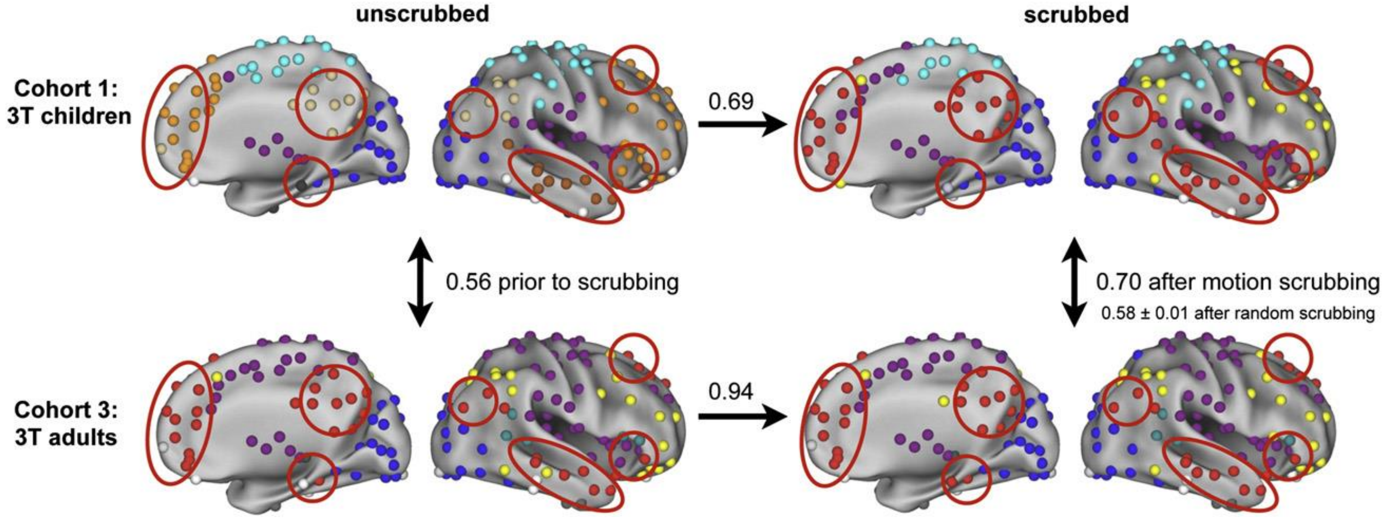
Elegant results ≠ Correct results

Functional Brain Networks Develop from a "Local to Distributed'' Organization



Fair... Petersen, PLoS Comp Bio 2009

Noise from head motion drove the network result



When they scrubbed data for areas of higher head motion (more common in children), the main network differences disappeared

Power, ... Petersen, Neuroimage 2012

"It really, really, really sucks. My favorite result of the last five years is an artifact," Steve Petersen

https://www.thetransmitter.org/spectrum/movement-during-brain-scans-may-lead-to-spurious-patterns

Anything that affects correlation will affect results:

Neural, cardiac/respiration, head motion, scanner artifacts, ...

Communication

- It is easier to share code and data if it’s set up with sharing in mind
  - Write code that you can share with your future self
- If you share data, share quality assessment measures!

Take home messages

- Make time to think through choices at every step of a study
- Quality Assessment is a critical part of every study
- No one is an expert in every step of every neuroimaging study
  - Behavioral psychology
  - Clinical Medicine
  - Neuroanatomy
  - MRI physics
  - Statistics
  - Software Development
  - Scientific Communication
- Learning what you don't know and who to ask is almost as important as what you do know.