

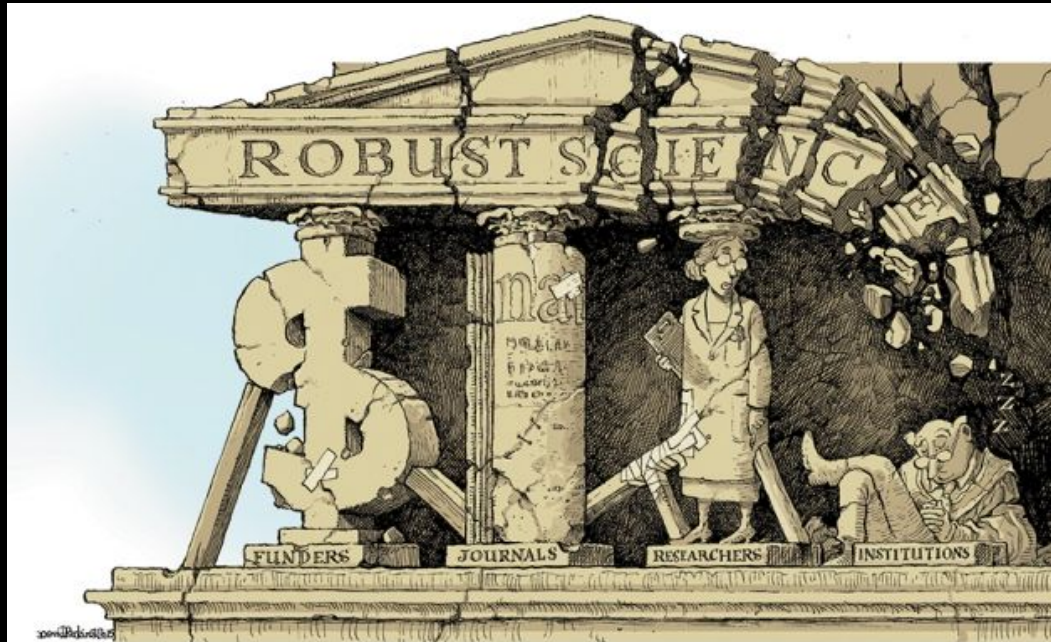
Progress Toward Openness, Transparency, and Reproducibility in Cognitive Neuroscience

Rick O. Gilmore^a, Michele T. Diaz^{a,b}, Brad A. Wyble^a, and
Tal Yarkoni^c

open science so far
and
the problems

Where are we at - the bad

many findings in psychological science may be difficult to reproduce



→ of 190 studies examined in 6 large-scale efforts combined, 47% were successfully replicated

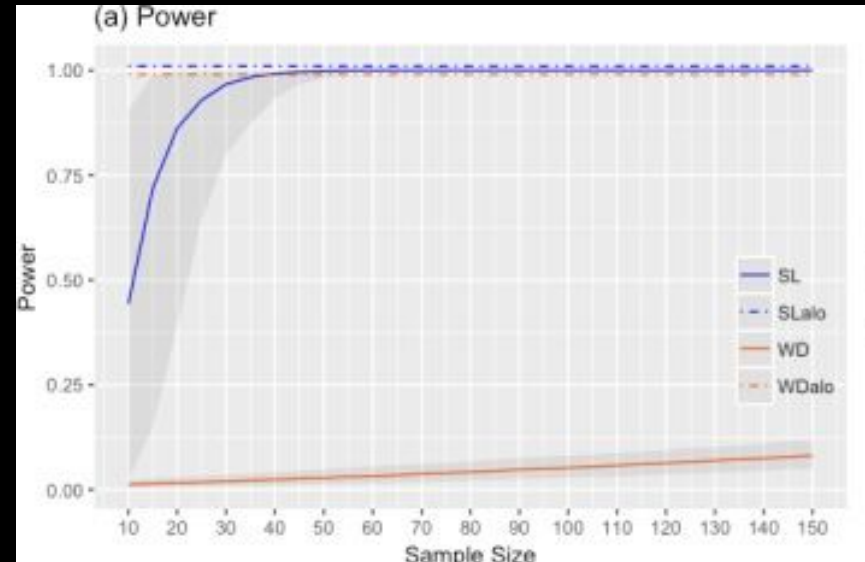
Where are we at - the bad

statistical power in brain imaging studies is low and has not improved

→ the combination of :

- ◆ a large number of dependent variables,
- ◆ a relatively small number of observations (subjects),
- ◆ and a need to correct for multiple comparisons

can decrease statistical power dramatically

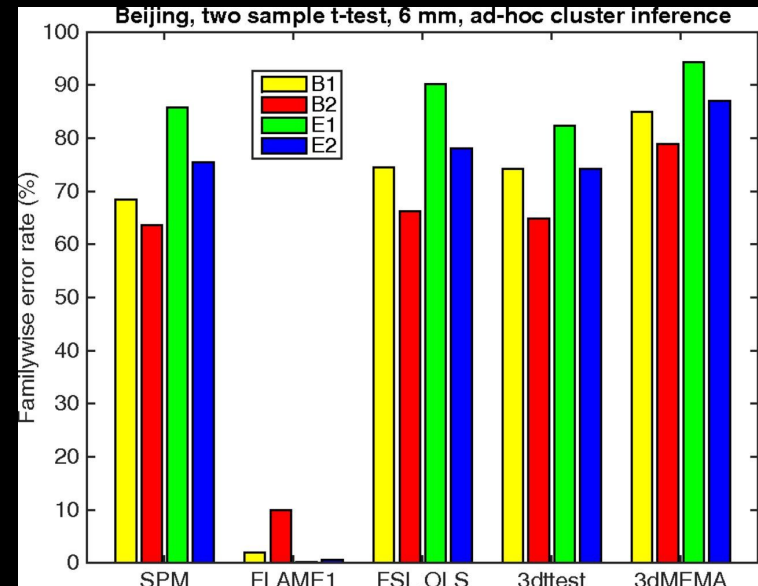


SL: Strong Localized effects, WD: Weak Diffuse effects

Where are we at - the bad

software errors in analysis tools are common and can go undetected for many years

most common software packages for fMRI analysis (SPM, FSL, AFNI) can result in false-positive rates of up to 70%



Where are we at - the bad

the open sharing of data, code, and materials is rare



Where are we at - the good

- renewed focus on reaffirming reproducibility, transparency, and openness as essential core values in psychological science and related fields
- emergence and rapid growth of data archives, meta-analytic tools, software pipelines, and research groups devoted to improved methodology
- greater clarity about what these values mean in practice (see last week ;))



What even is reproducibility?

What really is reproducibility - Methods reproducibility

different investigators are able to obtain the same results when applying the same tools and analytical procedures used in a study to the same (i.e., original) dataset



What really is reproducibility - Methods reproducibility

main problem:

- how to report the multitude of small individual choices involved in a typical analysis pipelines

Implemented solutions:

- standardised software packages (e.g. EEGLAB or SPM)
- promotion of data and metadata standards for research (e.g. The Big EEG Consortium)
- convergence on recommendations on what and how to report parameter-choices
- standard toolboxes for experimental tasks

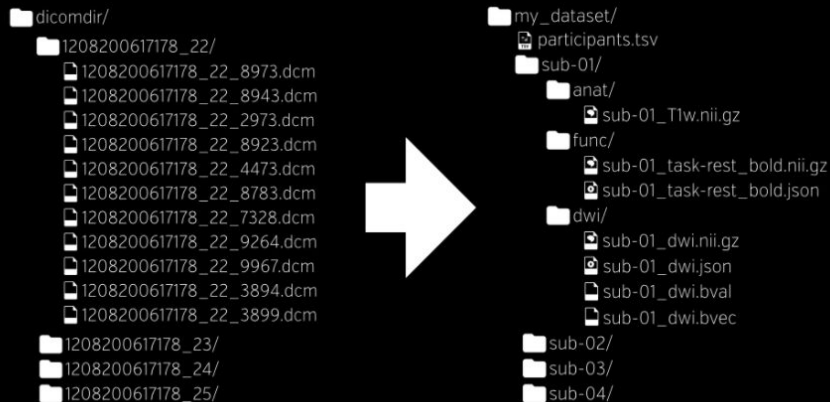
But:

- Not all these developments have become common practice (e.g. not reproducible graphical interfaces are still popular)
- there are untracked differences (e.g. the OS or dependencies)
- individualised task stimuli and designs are not often shared

Initiatives concerning **methods** reproducibility

Data format standardization

- BIDS format for BOLD data
 - Brain Imaging Data Structure
 - separate folders for different raw data types
 - derivatives folder
- easier for others to read, re-analyze or use it in a meta-analysis



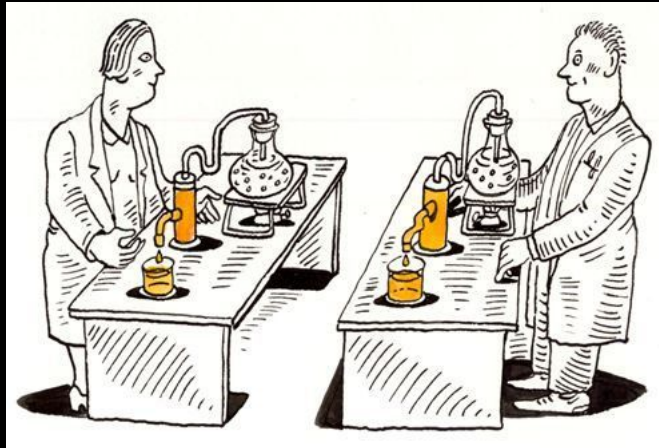
Initiatives concerning **methods** reproducibility

Behavioral method reproducibility

- Recording of videos
 - full communication of methods and a complete transparency
 - issue of privacy
 - could use an “ideal” or “fake” participant for this

What really is reproducibility - Results reproducibility

a new study with new data, collected following the original procedures as closely as possible, yields the same outcomes



What really is reproducibility - Results reproducibility

main problem:

- Results reproducibility highly depends on on the comparability of the different equipments and software and **results reproducibility is difficult to quantify**

Implemented solutions:

- highly convergent (imaging) techniques, that are capable of producing convergent results in different labs
- development of generalizable models for multi-site data collection
- systematic measurement and control of how fMRI signals vary across sites (fBIRN)
- Large-scale meta-analyses that show a marked degree of convergence on stable neural correlates for most major psychological processes

But:

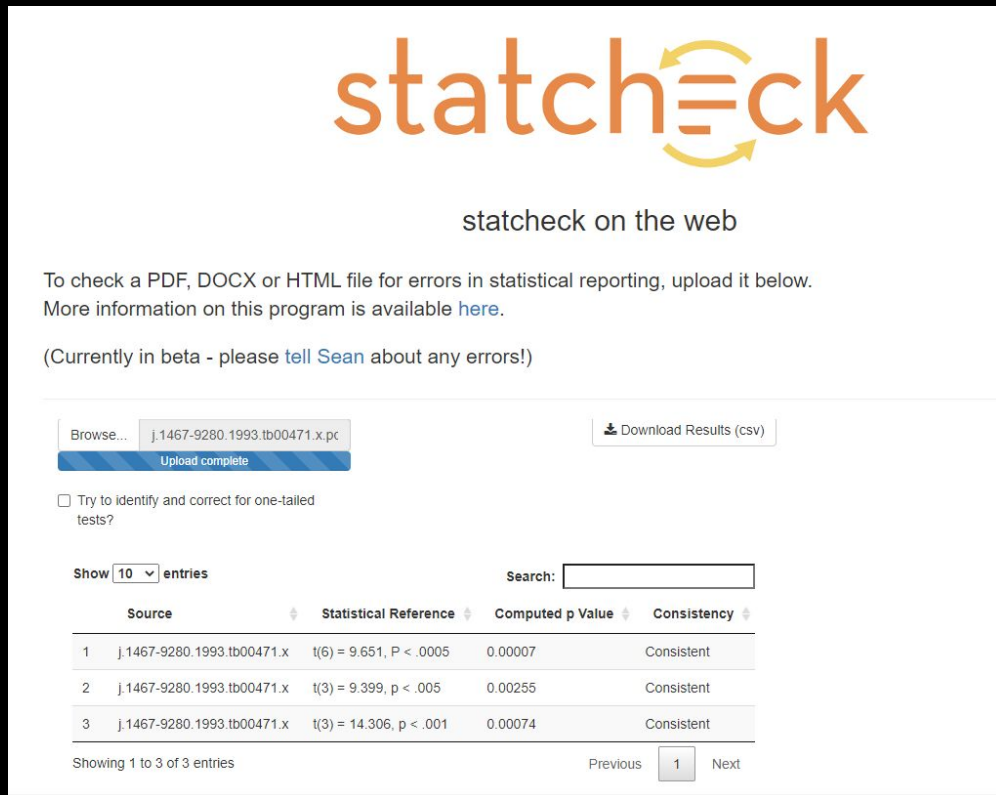
- initiatives promoting neuroimaging data sharing have historically met with limited success
- there is a lack of consensus about what quantitative measures should be used

Initiatives concerning **results** reproducibility

Stats

➤ Statcheck.io

- makes sure that the stats reported in the paper are correct
- does not actually redo the stats



The screenshot shows the Statcheck.io website. At the top is the logo "statcheck" in orange, with a circular arrow icon around the "e". Below the logo is the text "statcheck on the web". A paragraph states: "To check a PDF, DOCX or HTML file for errors in statistical reporting, upload it below. More information on this program is available [here](#)." Below this is a note: "(Currently in beta - please [tell Sean](#) about any errors!)". There is a file upload section with a "Browse..." button, a text input showing "j.1467-9280.1993.tb00471.x.pc", and a "Download Results (csv)" button. Below the upload section is a checkbox labeled "Try to identify and correct for one-tailed tests?". A table displays the results of the check, with columns for "Source", "Statistical Reference", "Computed p Value", and "Consistency". The table shows three entries, all marked as "Consistent". At the bottom, there is a pagination bar showing "Showing 1 to 3 of 3 entries" and "Previous 1 Next".

statcheck on the web

To check a PDF, DOCX or HTML file for errors in statistical reporting, upload it below.
More information on this program is available [here](#).

(Currently in beta - please [tell Sean](#) about any errors!)

Browse... j.1467-9280.1993.tb00471.x.pc Download Results (csv)

Upload complete

☐ Try to identify and correct for one-tailed tests?

Show 10 entries Search:

	Source	Statistical Reference	Computed p Value	Consistency
1	j.1467-9280.1993.tb00471.x	$t(6) = 9.651, P < .0005$	0.00007	Consistent
2	j.1467-9280.1993.tb00471.x	$t(3) = 9.399, p < .005$	0.00255	Consistent
3	j.1467-9280.1993.tb00471.x	$t(3) = 14.306, p < .001$	0.00074	Consistent

Showing 1 to 3 of 3 entries Previous 1 Next

Initiatives concerning **results** reproducibility

Neurovault

- Bunch of whole-brain statistical maps
- Let's take a look



A public repository of unthresholded statistical maps, parcellations, and atlases of the brain.

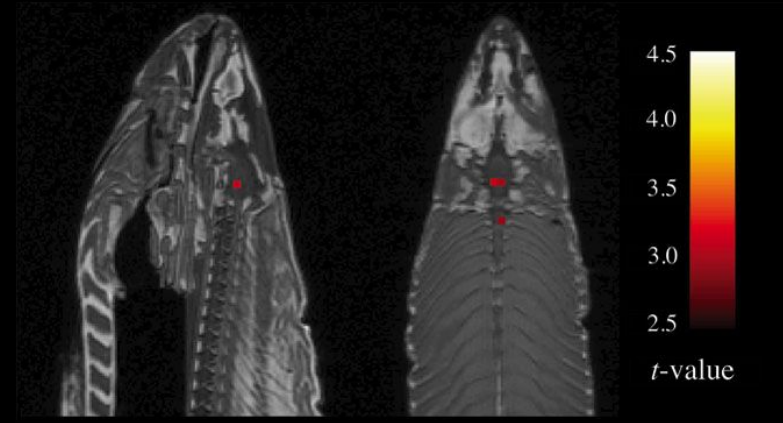
What really is reproducibility - Inferential reproducibility

independent researchers come to similar conclusions about what patterns of data mean, based on their own replication study or a reanalysis of a prior study

What really is reproducibility - Inferential reproducibility

main problems:

- low power in (fMRI) studies often leads to incorrect interpretations
- inflated false positives due to insufficient correction for multiple comparisons
- Researchers can disagree on the interpretation of even sound findings
 - what does the BOLD signal mean?
 - kind of information extracted from multivariate pattern analysis?
 - reverse inference
 - etc....



Initiatives concerning **inferential** reproducibility

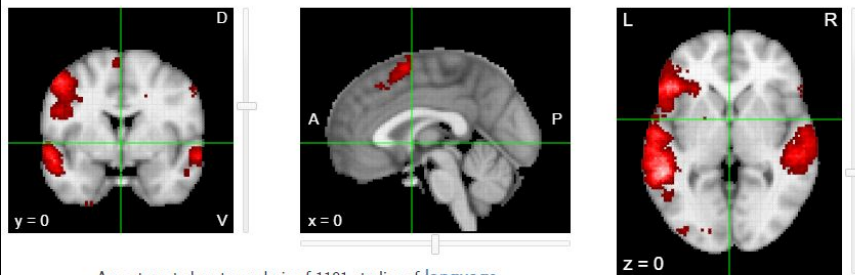
Neurosynth

- meta-analyses
- activation coordinates in MNI space and word frequencies text mined from papers on the web
- get ROIs for your dataset
- Let's take a look

neurosynth.org

Neurosynth is a platform for large-scale, automated synthesis of functional magnetic resonance imaging (fMRI) data.

It takes thousands of published articles reporting the results of fMRI studies, chews on them for a bit, and then spits out images that look like this:



An automated meta-analysis of 1101 studies of [language](#)

Initiatives concerning **inferential** reproducibility

Machine learning

- “new” philosophical way of thinking about explanations of neural data
- prediction of ‘out-of-sample’ dataset
- allows to get at information patterns than are too subtle or complex for standard summary statistics
- still an inferential problem
 - overfitting
 - above-chance decoding = ‘information’ in a region?

Take aways

- authors see (saw back in 2017) a bright tomorrow
- raise issue of who will store and pay for storage of large datasets
- distinction between the types of reproducibilities a bit artificial
- wonderful paper for pointing towards different resources/web-based apps etc.