Concepts and tools for protocol documentation and study pre-registration

Cyril Pernet
Robert Oostenveld
Melanie Ganz

What's the idea?

There are always discrepancies between

- Planned experiment and the actual data collection
- The experiment description and the implementation
- The planned analysis and the actual analysis

These create problems, in particular a lack of reproducibility.

**Why those discrepancies
lead to low reproducibility?
--- 5 min --**

Published peer-reviewed protocols

A research protocol is a detailed study design or set of instructions for carrying out a specific experimental process or procedure.

protocols.io

What is documented?

- background, rationale, objective(s), design, methodology, statistical considerations and organization of an experiment
- Step by step instructions

Keeping track of discrepancies

- Good old fashion lab paper notebook!
- Why not add digital log files as notes?
- Electronic lab notebooks → how to choose:
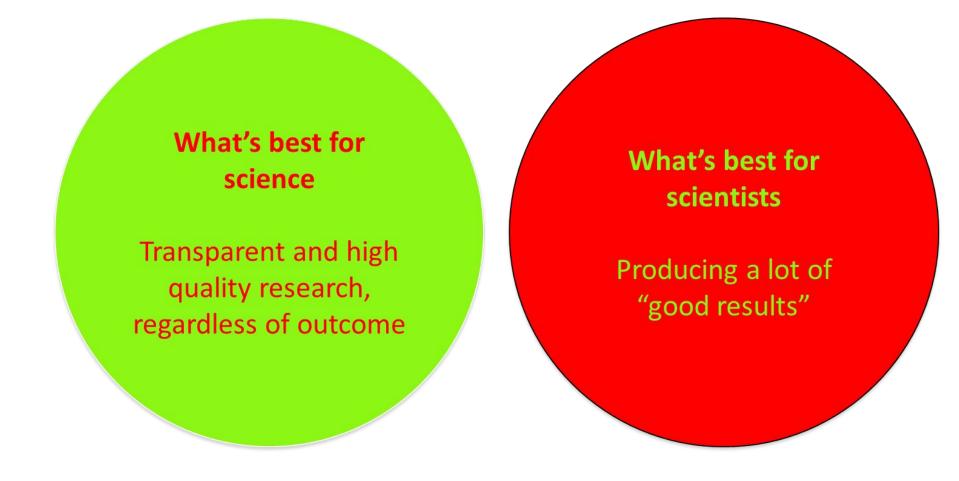 https://www.nature.com/articles/d41586-018-05895-3

They are shared between collaborators, they eliminate issues with poor handwriting and damaged paper notebooks and can prevent data from being lost when researchers move on.

Where is the information stored? re GDPR?

Study Pre-registration and Registered Reports

Tell me everything you know about registered report:

- What is it?
- What do the consist?
- Any misconceptions?
- etc

Pia Rotshteins; slide – OHBM 2016

What happens when researchers are pressured to get "good results"?

Pia Rotshteins; slide – OHBM 2016

Publication bias – *suppression of negative or complex findings*

Significance chasing – *"p-hacking", selective reporting*

HARKing – *hypothesizing after results are known*

Lack of data sharing – *no time, too hard, no incentive*

Low statistical power – *quantity of papers over quality*

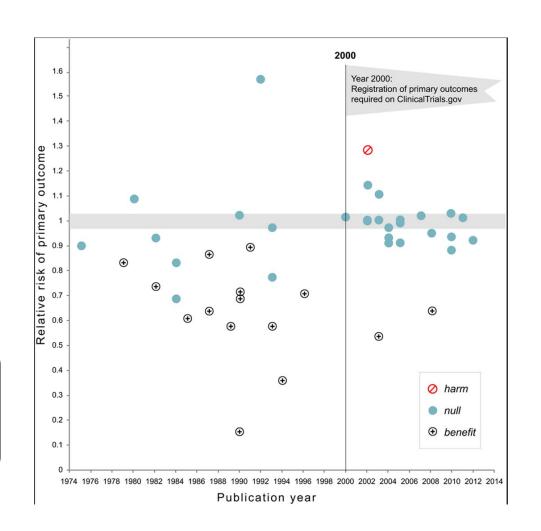**Lack of replication** – *seen as boring, lacking in intellectual prowess*

Registered reports

- Similar to protocol papers in clinical trials
- Editorial decision to take your paper before you have results based the QUESTION it asks and the QUALITY of the method it uses (and yes: being before collection, primary outcomes are defined a priori)

  → almost guarantee a publication, no matter the result!

Preregistration of clinical trials causes medicines to stop working!

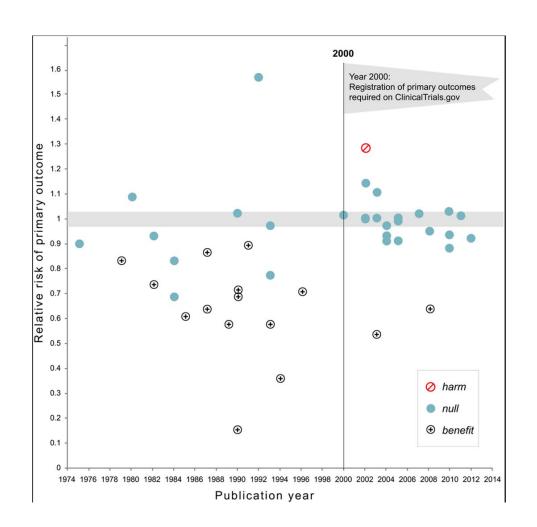http://chrisblattman.com/2016/03/01/13719/

Why do you think it 'stopped working'?

Preregistration of clinical trials causes medicines to stop working!

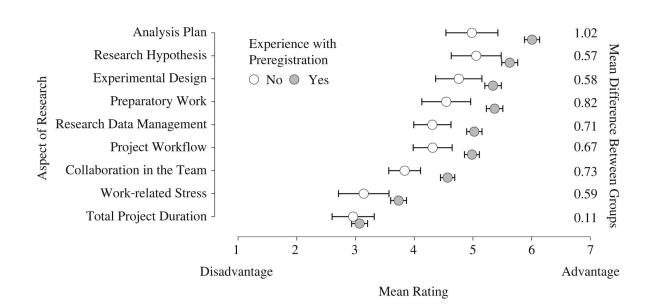http://chrisblattman.com/2016/03/01/13719/

- Publication bias (report negative results)
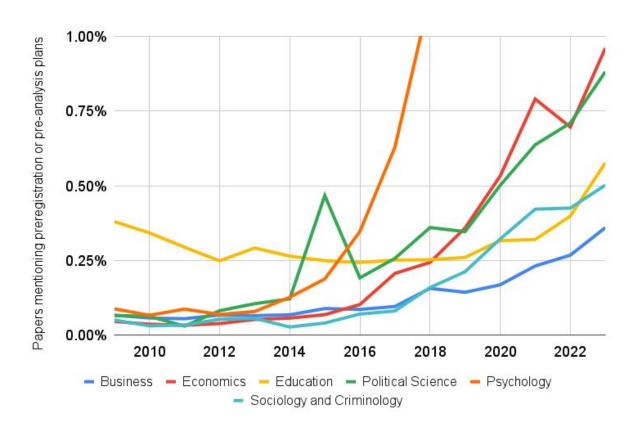- Primary outcomes 'cannot' be switched

A survey on how preregistration affects the research workflow: better science but more work

Alexandra Sarafoglou[†] ✉, Marton Kovacs, Bence Bakos,
Eric-Jan Wagenmakers and Balazs Aczel

It's picking up

Papers mentioning preregistration or pre-analysis plans

Legend: Business, Economics, Education, Political Science, Psychology, Sociology and Criminology

Good for reproducibility

- A study is accepted for publication based on design because
  - 1 – hypotheses are clear,
  - 2 – the method is sound, and
  - 3 – it is well powered
- Why studies are NOT reproducible? Partly because hypotheses aren't clear enough leading to HARKING (hypothesizing after results are known) and because it is not well powered leading to p-hacking (significance chasing by using multiple outcomes and/or methods) which is what pre-registered studies aim at fixing
- It is also good for science by reducing the publication bias, publishing positive and negative results alike, new and replication studies alike.

Good for reproducibility

Declaring statistical procedures before-hand protects from cognitive biases (systematic patterns of deviation from rational judgment)

➔ confirmation bias = interpret new evidence as confirmation of one's existing beliefs (if the results don't fit my expectation, there is a problem with data cleaning, or statistical test used)

➔ hindsight bias = tendency for people to perceive past events as having been more predictable than they were (if a result appear out of nowhere, we post-rationalize why it happen, maybe re-analyze, reinforcing chances to find get a false positive)