Preregistration

Daan Ornée
Charlotte Vrijen
Yoram Kunkels

info@openscience-groningen.nl
@OSCGroningen

Presentation from the Open Science Center: cos.io/prereg
Hypo-Deductive Model of the Scientific Method

Publish and/or conduct next experiment
Publication bias

Generate and specify hypothesis
Failure to control for bias

Interpret results
P-hacking

Design study
Low statistical power

Analyze data and test hypothesis
P-hacking

Conduct study and collect data
Poor quality control

Reference: http://www.nature.com/articles/s41562-016-0021#f1
The Problem

The combination of a strong bias toward statistically significant findings and flexibility in data analysis results can lead to irreproducible research.
The Problem

The combination of a strong bias toward statistically significant findings and flexibility in data analysis results can lead to irreproducible research.
http://127.0.0.1:8081/plosone/article?id=info:doi/10.1371/journal.pone.0010068

<table>
<thead>
<tr>
<th>Field</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Science (SP, N=104)</td>
<td></td>
</tr>
<tr>
<td>Geosciences (GE, N=127)</td>
<td></td>
</tr>
<tr>
<td>Environment/Ecology (EE, N=149)</td>
<td></td>
</tr>
<tr>
<td>Plant and Animal Sciences (PA, N=193)</td>
<td></td>
</tr>
<tr>
<td>Computer Science (CS, N=63)</td>
<td>a</td>
</tr>
<tr>
<td>Agricultural Sciences (AG, N=109)</td>
<td>l a</td>
</tr>
<tr>
<td>Physics (PH, N=71)</td>
<td></td>
</tr>
<tr>
<td>Neuroscience &amp; Behaviour (NB, N=143)</td>
<td></td>
</tr>
<tr>
<td>Microbiology (MI, N=140)</td>
<td></td>
</tr>
<tr>
<td>Chemistry (CH, N=95)</td>
<td></td>
</tr>
<tr>
<td>Social Sciences, General (SO, N=144)</td>
<td>s</td>
</tr>
<tr>
<td>Immunology (IM, N=145)</td>
<td></td>
</tr>
<tr>
<td>Engineering (EN, N=77)</td>
<td>a</td>
</tr>
<tr>
<td>Molecular Biology &amp; Genetics (MB, N=126)</td>
<td></td>
</tr>
<tr>
<td>Economics &amp; Business (EB, N=117)</td>
<td>a s</td>
</tr>
<tr>
<td>Biology &amp; Biochemistry (BB, N=113)</td>
<td></td>
</tr>
<tr>
<td>Clinical Medicine (CM, N=130)</td>
<td>l a</td>
</tr>
<tr>
<td>Pharmacology &amp; Toxicology (PT, N=142)</td>
<td></td>
</tr>
<tr>
<td>Materials Science (MS, N=105)</td>
<td>a</td>
</tr>
<tr>
<td>Psychiatry/Psychology (PP, N=141)</td>
<td>l s</td>
</tr>
</tbody>
</table>
The Problem

The combination of a strong bias toward statistically significant findings and flexibility in data analysis results can lead to irreproducible research.
The Garden of Forking Paths

Hypothesis: “Does X affect Y?”

Control for time?

Exclude outliers?

Median or mean?

Statistically significant result

Start here

Gelman and Loken, 2013
The Problem

The combination of a strong bias toward statistically significant findings and flexibility in data analysis results can lead to irreproducible research.
What is Preregistration?
What is Preregistration?

A time-stamped, read-only version of your research plan created before you begin data collection.
What is Preregistration?

A time-stamped, read-only version of your research plan created before you begin data collection.

It contains:

- Hypothesis
- Data collection procedures
- Manipulated and measured variables
- Statistical model
- Inference criteria
When the research plan undergoes peer review before results are known, the preregistration becomes part of a Registered Report.
What problems does preregistration fix?

1. The file drawer effect

2. **P-Hacking**: Unreported flexibility in data analysis

3. **HARKing**: Hypothesizing After Results are Known
What problems does preregistration fix?

Preregistration makes the distinction between confirmatory (hypothesis testing) and exploratory (hypothesis generating) research more clear.
Confirmatory vs. Exploratory Analysis

**Context of confirmation**
- Traditional hypothesis testing
- Results held to the highest standards of rigor
- Goal is to minimize false positives

**P-values interpretable**

**Context of discovery**
- Pushes knowledge into new areas/ data-led discovery
- Finds unexpected relationships
- Goal is to minimize false negatives

**P-values meaningless**
Confirmatory vs. Exploratory Analysis

Context of confirmation
- Traditional hypothesis testing
- Results held to the highest standards of rigor
- Goal is to minimize false positives
- P-values interpretable

Context of discovery
- Pushes knowledge into new areas/data-led discovery
- Finds unexpected relationships
- Goal is to minimize false negatives
- P-values meaningless

Presenting exploratory results as confirmatory increases the publishability of results at the expense of credibility of results.
Example workflow

Collect New Data

Confirmation Phase

Hypothesis testing

Discovery Phase

Exploratory research

Hypothesis generating

Create Preregistration

Theory driven, a-priori expectations
Incentives to Preregister

You can receive a **Preregistered Badge** for preregistering your research *before* you begin your study. Visit [cos.io/badges](http://cos.io/badges) for more information and to see which journals currently issue badges.
Let’s look at OSF form, real-time!

http://osf.io/prereg/

Examples of completed preregistrations:
https://osf.io/e6auq/wiki/Example%20Preregistrations/?view
Different formats for different studies

- Use of secondary data: https://osf.io/x4gzt/
- Qualitative Research: https://osf.io/dv8km/ or https://osf.io/j7ghv/
- Experience sampling (ESM) studies: https://osf.io/2chmu/
- fMRI studies: https://osf.io/6juft/
- Systematic reviews and meta-analyses:
  https://osf.io/preprints/metaarxiv/3nbea

Just some examples, there are many more
Want help? OSCG Q&A

Bi-monthly preregistration Q&A in 2021:

January 19, 13:00-14:00
March 1, 13:00-14:00
September 7, 13:00-14:00
November 8, 13:00-14:00

Please send an email to info@openscience-groningen.nl to receive the google meet link
Preregistration video tutorials

● Goal: offering information and step-by-step guides on pre-registration.

● Format: Short (5 to 10 minutes) video’s disseminated online.
Preregistration video tutorials

● Series: Looking at one major topic at a time and breaking it down into small, manageable chunks.

● Interactivity & revelancy: Easily skip to a part or video that is most relevant.

● Example:  
  https://www.youtube.com/watch?v=O6lL_VcSfE4
FAQs

Can’t someone “scoop” my ideas?
FAQs

Can’t someone “scoop” my ideas?

1. Date-stamped preregistrations make your claim verifiable.
2. By the time you’ve preregistered, you are ahead of any possible scooper.
3. Embargo your preregistration.
FAQs

Isn’t it easy to cheat?
Isn’t it easy to cheat?

1. Making a “preregistration” after conducting the study.
2. Making multiple preregistrations and only citing the one that “worked.”
FAQs

Isn’t it easy to cheat?

1. Making a “preregistration” after conducting the study.
2. Making multiple preregistrations and only citing the one that “worked.”

While fairly easy to do, this makes fraud more intentional.

Preregistration helps keep you honest to yourself.
Tips for writing up preregistered work

1. Include a link to your preregistration (e.g. https://osf.io/f45xp)

2. Report the results of **ALL** preregistered analyses

3. **ANY** unregistered analyses must be transparent
Thank you for attending!

1. POLL: Do you think you will attend one of the next Q&A sessions
THANK YOU!

Learn more: cos.io/prereg