Lessons in Open Science

Preregistration and Registered Reports

David Riedel
Dresden Service Center Research Data

7. Oktober 2021
Agenda

• Current Issues in Science
  • Credibility and trust
  • Prediction vs. postdiction
  • File Drawer, p-Hacking, HARKing, multiverse

• Definition and Merits
  • Benefits to science in general
  • Personal benefits
  • Adopting preregistration

• Process of Preregistration
  • Different methods
  • Registered Reports

• Common Obstacles and Concerns
Credibility of Science?

- Neither laypersons nor professional scientists trust in the majority of published results in the life sciences.

- Average statistical power is much too low (20-50%) considering the published positive rate (>90%) (Nosek, 2018).

Adapted from Baker, 2016

Is there a reproducibility crisis?
- 1576 Researchers surveyed -

- Don't know: 7%
- No, there is no crisis: 3%
- Yes, a slight crisis: 38%
- Yes, a significant crisis: 52%

Adapted from Baker, 2016
Replication Crisis in Psychology

- Replication of 100 studies
- Significant results decreased from 97% to 36%
- Reasons: coincidence/error, biases, questionable research practices

*Open Science Collaboration, 2015*
Hypothesizing After Results are Known

“HARKing and p-Hacking”

p-Hacking

“If you don’t reveal some insights soon, I’m going to be forced to slice, dice, and drill!”

https://ispgr.org/open-science-and-the-power-of-pre-registration/

“A publication bias exists if the probability that a study reaches the literature [...] depends on the results of the study.”

Adapted from Scargle, 1999
The Path to Incredibility

conduct a study

Did you find, what you predicted?

Did you find another interesting effect?

Can you find any effect, that is interesting?

File Drawer

HARKing

p-Hacking

YES

NO

YES

NO

YES

adapted from Brian Nosek, 2018: https://osf.io/9m6tx/
Multiverse

- outliers
- exclusion
- covariates
- transformation
- statistical model
- definition of threshold

Number of Different Analyses Based on Decisions

z-value

CC BY 4.0
Multiverse

- Outliers
- Exclusion
- Covariates
- Transformation
- Statistical model
- Definition of threshold

Number of Different Analyses Based on Decisions
## Prediction vs. “Postdiction”

<table>
<thead>
<tr>
<th>Property</th>
<th>Prediction</th>
<th>“Postdiction”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influenced by investigation and outcome?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Null hypothesis significance testing (p-values) is fully applicable?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Confirmatory or exploratory?</td>
<td>confirmatory</td>
<td>exploratory</td>
</tr>
<tr>
<td>Valuable to the scientific process?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

We only have to be able to differentiate between the two approaches.
Agenda

• Current Issues in Science
  • Credibility and trust
  • Prediction vs. postdiction
  • File Drawer, p-Hacking, HARKing, multiverse

• Definition and Merits
  • Benefits to science in general
  • Personal benefits
  • Adopting preregistration

• Process of Preregistration
  • Different methods
  • Registered Reports

• Common Obstacles and Concerns
What is preregistration?

- **A priori** documentation/publication of information on planned investigation, such as:
  - *study design, methods, hypotheses, analysis plan*

- **Independent** party registry
  - *E.g. Open Science Framework (OSF)*
  - *Embargo upon submission is possible*

- **Time Stamp**

- **Immutable**
  - *Can not be changed or deleted (without record)*
The Path to Credibility

- planning and preregistration of a study

  Conduct the study and analyze the data according to plan

  Were you able to stick to your plan?

    NO

  Are any changes justified and allow for confirmatory analyses?

    NO

    Label any exploratory analyses and provide justification transparently

    YES

    transparency

    YES

    Publish

File-Drawer
HARKing
p-Hacking
By making analytical decisions in advance, preregistration ensures arbitrary results within the multiverse.
Is preregistration effective?

Rate of positive results dropped from 57% to 8% with the introduction of obligatory preregistration.

Kaplan & Irvin, 2015

"The strongest factor associated with the false positive or true positive study outcome was if the study had a specific a priori hypothesis."

Swaen, Teggeler & van Amelsvoort, 2001
Benefits on an individual level

• Reduce implicit bias
  • Improve quality of research

• Distinction between confirmatory and exploratory investigations

• Increases focus on project management
  • Preregistration involves increased careful planning

• Initialization of collaborations
  • Involvement of all project partners

• Open Science Badges

https://osf.io/tyy7x/wiki/1.%20View%20%20the%20Badges/
Personal Benefits

Seven Selfish Reasons for Preregistration:

1. Take credit for your predictions.
2. Experience the excitement.
3. Prevent the data from taking you hostage.
4. Profit from online resources.
5. Increase your reputation and self-image.
6. Await your results without fear with in-principle acceptance.
7. Protect yourself against post-hoc critique.

Wagenmakers & Dutilh, 2016
Adopting preregistration

Christensen et al., 2020

Number of Preregistrations on OSF

Adapted from Nosek & Lindsay 2018
https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science
Adopting preregistration (Psychology)

Christensen et al., 2020

Hardwicke et al., 2021
Agenda

• Current Issues in Science
  • Credibility and trust
  • Prediction vs. postdiction
  • File Drawer, p-Hacking, HARKing, multiverse

• Definition and Merits
  • Benefits to science in general
  • Personal benefits
  • Adopting preregistration

• Process of Preregistration
  • Different methods
  • Registered Reports

• Common Obstacles and Concerns
Content and Templates

- Metadata
  - Title, Contributors, Subject, Tags...

- Study Design
  - Type, Blinding, Randomization...

- Sampling
  - Data collection, Subjects, Sample size...

- Analysis
  - Variables, Manipulation, Aggregation, Statistics, outlier/missing data policy...

https://osf.io
How to preregister

• Online registries
  • *e.g. OSF, AsPredicted, Prospero*

• Registered Reports

• Other publication methods
  • *Study protocol*
  • *Poster at conference*
Registered Report

Develop Idea → Design Study → Collect and Analyze Data → Write Report → Publish Report

Stage 1 Peer Review (constructive) - In Principle Acceptance

Stage 2 Peer Review (formal) - Adherence?

Adapted from https://osf.io/rr/
## Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Registry</td>
<td>• Easy and fast</td>
<td>• No quality assurance</td>
</tr>
<tr>
<td></td>
<td>• Most widely used</td>
<td>• Harder to discern between good and bad registration</td>
</tr>
<tr>
<td></td>
<td>• Stepwise approach easily applicable</td>
<td></td>
</tr>
<tr>
<td>Registered Reports</td>
<td>• Gold standard</td>
<td>• Takes much more time</td>
</tr>
<tr>
<td></td>
<td>• “In Principle Acceptance” status</td>
<td>• Less predictable</td>
</tr>
<tr>
<td></td>
<td>• Constructive review process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Counteracts File Drawer</td>
<td></td>
</tr>
<tr>
<td>Other publication of study protocol</td>
<td>• Additional publication</td>
<td>• Focusses on the dissemination of methods rather than decreasing degrees of freedom</td>
</tr>
<tr>
<td>(article, conference poster …)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Agenda

• Current Issues in Science
  • Credibility and trust
  • Prediction vs. postdiction
  • File Drawer, p-Hacking, HARKing, multiverse

• Definition and Merits
  • Benefits to science in general
  • Personal benefits
  • Adopting preregistration

• Process of Preregistration
  • Different methods
  • Registered Reports

• Common Obstacles and Concerns
Obstacles and Concerns

• Too much **extra** work!
  • *It is not going to be wasted, but worth it*

• Too restrictive, science needs **freedom**!
  • *You are free to explore, just label accordingly*

• It‘s not going to stop **fraud** anyway!
  • *No, but it is still helpful and might make it harder*

• Someone is going to **steal** my ideas!
  • *Embargo*
References


What funders can do

• Raise **awareness**
  • *Lectures, workshops, conference participation* ...

• **Guided Workflows**
  • *Manuals, templates, tutorials* ...

• **Policy**
  • *Guidelines, application evaluation, grants* ...

• Partnerships with **registries** as well as journals offering **registered reports**

Nosek et al., 2018
Key principles

• Avoid bias/influence by making important decisions and careful planning in advance

• Be as transparent as possible

• Changes and exploratory analyses are no problem, as long as you report and are able to justify them

Be aware of these principles and of what you are trying to achieve. If you are honest, preregistration will come naturally to you!
Obstacles and Concerns

“No, it’s too much extra work!”

• There is extra effort involved. Ask yourself if it could be worth it

• Most of it will have to be done at some point anyway

• Start slowly and incrementally, efficiency will increase

Nosek et al., 2018
Obstacles and Concerns

“No, I don’t want my freedom taken away!”

• You are free to explore
• Lack of freedom only with regards to lack of transparency
• Differentiation between confirmatory and exploratory investigations is vital
• “It’s a plan, not a prison” - https://www.cos.io/blog/preregistration-plan-not-prison

Nosek et al., 2018
Obstacles and Concerns

“No, I need to be able to change my plans throughout the project!”

• Justify those changes

• Be transparent

• Report as early as possible, potentially through another preregistration

• Sensible changes are no problem

Nosek et al., 2018
Obstacles and Concerns

“No, I want to use existing data!”

- Register before examining the data
- Try to blind as much as possible
- Clearly state how much you knew about the data before analyzing it
Obstacles and Concerns

“No, I have no predictions!”

• Exploratory investigation
• Label accordingly and adapt statistical analyses accordingly
• Consider hold out sample

Nosek et al., 2018
Obstacles and Concerns

“No, I need to look into the data, before I am able to decide on the analyses!”

• Be upfront about the decision you will have to make

• Consider the feasibility of a decision tree

• Consider stepwise preregistration

• Consider multiverse analysis
Obstacles and Concerns

“No, I don’t want my ideas stolen!”

• Time stamps can be valid without publication of your study plan
• Registrations can be embargoed (e.g. 4 years with OSF)
• It makes sense to publish your preregistration no later than your article

Nosek et al., 2018
Obstacles and Concerns

“No, it won’t stop fraud anyway!”

- No it will not, but it will make it harder and more explicit
- Helps to reduce implicit biases and to be honest to yourself
- Consider it a useful tool for scientists, who strive for better research
- Quality checks are necessary

Nosek et al., 2018