Pre-registration and OSF

Hannah Tarder-Stoll
Thursday, June 24, 2021

Roadmap

- What is preregistration?

- What is OSF?

- Why should I preregister?

- What information do I need to preregister?

- Work through an example

What is preregistration?

The idea behind preregistration is a simple one:

Publicly post your research question and hypotheses *in advance* (before collecting data)

Publicly post your specific study design, analysis plan, etc. *in advance*

What is preregistration?

DEVELOP IDEA → **DESIGN STUDY** → **SUBMIT** → **COLLECT AND ANALYZE DATA**
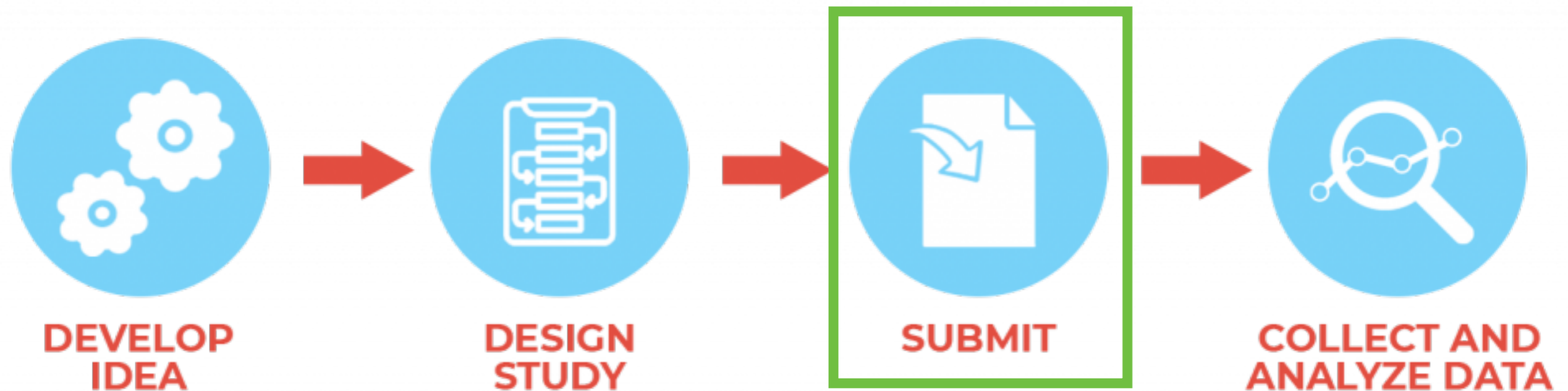
Submit preregistration before data collection!

If using pre existing data that you won't collect yourself, then post before peeking at it

What is OSF?

Open science framework (OSF) is a website where you can preregister your research

OSF provides easy to follow preregistration templates

Another option is aspredicted.org

What is OSF?

OSF also let's you do other cool things!

Name ∧ ∨	Modified ∧ ∨
📦 Does aging influence the use of episodic memo...	
− ⚙ OSF Storage (United States)	
− 🗁 Analysis	
+ 📁 Main Analysis	
+ 📁 Norming Analysis	
− 🗁 Testing Materials	
📄 Instructions.m	2021-06-08 05:23 PM
📄 Practice.m	2021-06-08 05:23 PM
📄 README.md	2021-06-08 05:23 PM
📄 Run.m	2021-06-08 05:23 PM
+ 📁 Schemes	
📄 Session.m	2021-06-08 05:23 PM
+ 📁 Stimuli	

Share files with collaborators (during the project) and the public (for transparency after publication)

What is OSF?

OSF also let's you do other cool things!

Share preprints of papers before publication

So what?

Preregistration, sharing materials, and preprints make the research process more transparent

But why do we need transparency in the first place?
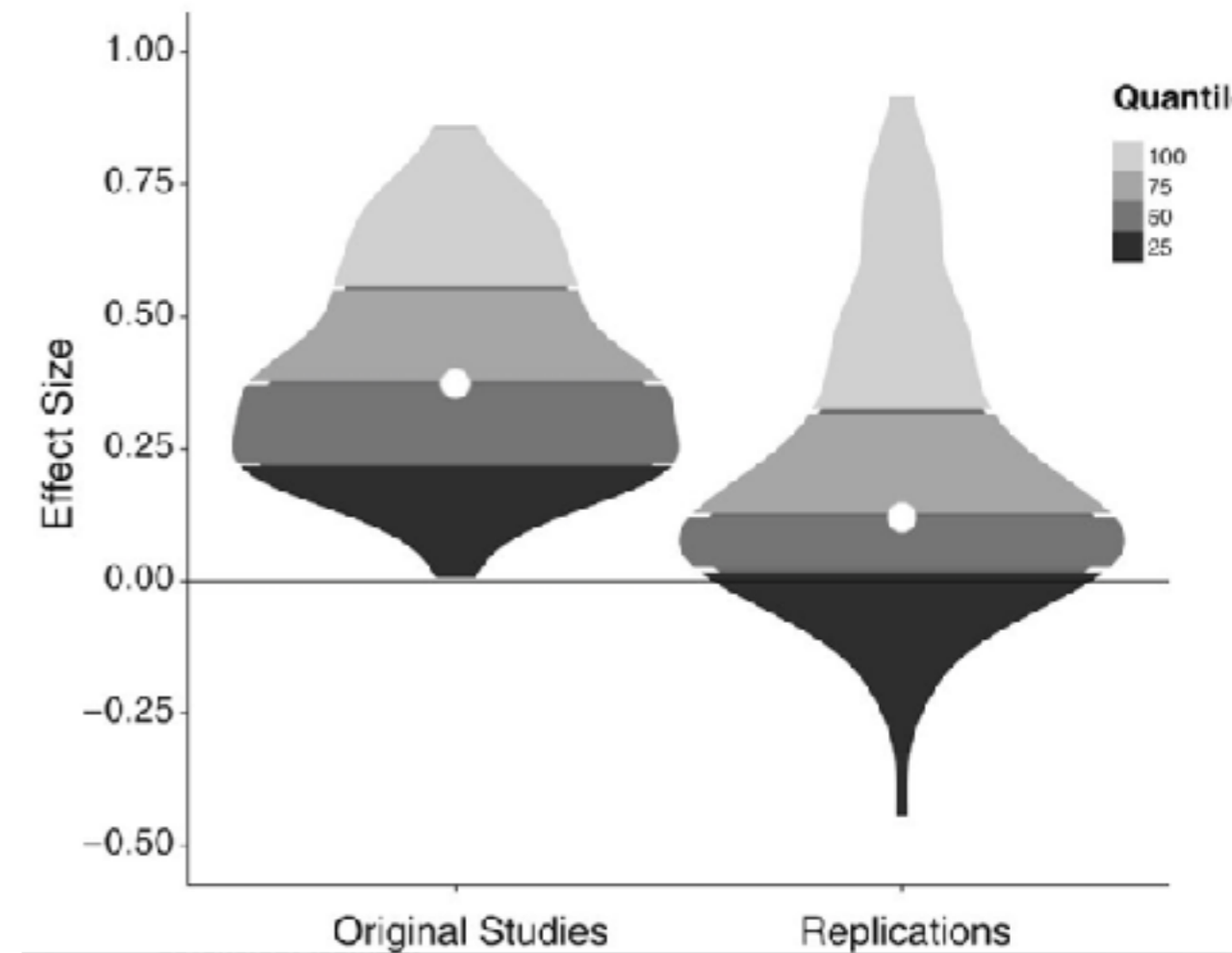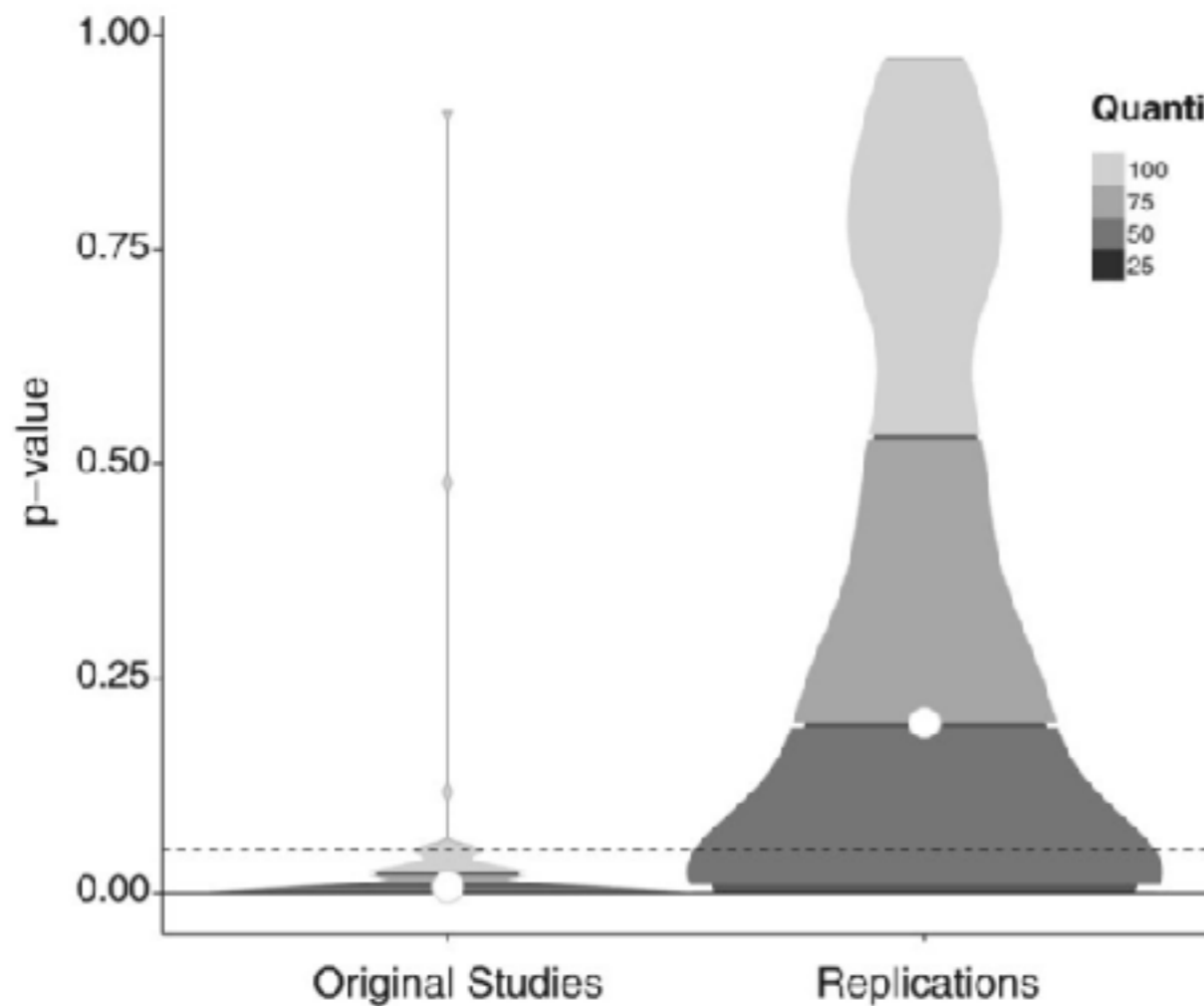
Why preregister?

Psychology has a **replication crisis**!

The results of many studies cannot be reproduced and/or their effect sizes are inflated

Replication Crisis

Out of 100 studies, only 36% of effects were reproducible
(Open Science Collaboration, 2015)

Why do we have a replication crisis?

Falsifying data/results is only part of the problem

Digging for significant findings and then publishing those findings selectively (p hacking)

Increases the likelihood of false positives (type 1 error) and are rarely replicable

Why do we have a replication crisis?

Peeking at the data as you're collecting it and stopping when the results are significant

Why do we have a replication crisis?

Excluding participants for various reasons (i.e. performance) until an effect is significant

Why do we have a replication crisis?

Hypothesizing after results are known (HARKing)

Why do we have a replication crisis?

Garden of forking paths:
"There is a one-to-many mapping from scientific to statistical hypotheses." (Gelman & Loken, 2013)

Issues arise when you report the analysis that best fits the hypothesis as strong evidence

How do we fix it?

Preregistration!

Open and transparent research practices

Document research plan before data collection

Publicly post data, analysis, and materials after data collection

What should I preregister?

1. Research Question and Hypotheses

2. Design/Method

3. Analysis Plan

What should I preregister?

Research Question and Hypotheses

- i.e. Does drug X and trait Y influence outcome Z?

- Numbered hypotheses

  1. Drug X will improve outcome Z

  2. Drug X will interact with trait Y to improve outcome Z

What should I preregister?

Method

- All independent variables and levels (within or between, orthogonal, etc.)
  - Drug x (treatment vs control, between subjects)
  - Trait y (continuous, between subjects)
- All dependent variables
  - Outcome z
- Covariates
  - Age

What should I preregister?

Method

- Pre selection criteria for sample
  - i.e. fluent in english

- Sample size justification
  - i.e. power analysis

- Stopping rule
  - i.e. n = 60 in each group

What should I preregister?

Method

- Exclusion criteria

  - Failing attention checks

  - Performance criteria

  - Outliers

  - etc.

What should I preregister?

Method

- Entire procedure/task design!

What should I preregister?

Analyses

- Preregister your analysis plan in as much detail as possible

- Preregistration makes a distinction between confirmatory and exploratory analyses

What should I preregister?

Analyses

Confirmatory

Test hypothesized effects

i.e. Drug X will improve outcome Z

i.e. Drug X will interact with Trait Y to improve outcome Z

Exploratory

Seek to discover effects

i.e. Drug X may interact with other biological variables

What should I preregister?

Confirmatory Analyses

- Variables and how they're calculated

- Statistical test (i.e. t test, ANOVA, multi level model, etc)

- Alpha (i.e. p > .05)

What should I preregister?

Confirmatory Analyses

- How will you handle missing data?

- Transformations? (i.e. log transform reaction time if not normally distributed)

What should I preregister?

Exploratory Analyses

"We may investigate the relationship between A and B but we do not have specific a priori hypotheses"

OSF Templates

Registration Forms and Templates

Registration Forms	Description	Templates
OSF Prereg* (citable from this preprint)	This is our standard, comprehensive, and general purpose preregistration form.	Google Doc, OSF Workflow, R Markdown by Frederik Aust, R Markdown by James Bartlett
Open-Ended Registration	Summary of registered work with a time-stamped snapshot of a research project. **Use this one** if **you are registering a completed project with data** or materials.	Word, GoogleDoc
Qualitative Preregistration* (Haven et al 2020)	Template for registering primarily qualitative work.	Word, GoogleDoc
AsPredicted Preregistration* form here	Eight questions derived from content recommended by AsPredicted.org.	Word, GoogleDoc
OSF-Standard Pre-Data Collection Registration	State whether data have been collected or viewed and other pertinent comments. **Use this one if your pre-analysis plan is uploaded on OSF as a doc**	Word, GoogleDoc

What if I deviate from my preregistration?

As a rule of thumb, just be transparent!

If you deviate a lot, you may have to preregister and run a follow up study

What do I do after data collection and analysis?

When ready to submit your study to a journal, upload data, analysis and materials to OSF

Provide a link to the preregistration and the OSF page containing materials

Once published you get a cool badge of approval

Ready for an example?

Use the `affect` dataset from the psychTools package in R

Imagine you're the researcher collecting this data set and begin writing a mock preregistration!

I don't expect you to finish the mock preregistration in this time. This exercise is meant to get you thinking about how to write a full one

SIPPS post-workshop survey access
(~1 minute survey)

Post-workshop survey (anonymous)

link: https://cumc.co1.qualtrics.com/jfe/form/SV_9HVYLAL5iXvjEmq

SIPPS post-workshop survey instructions
Preregistration and OSF (June 23)

1. Select "Research Skills" and press arrow

2. Select "Pre-registration and OSF (June 24)" and press arrow