# PREREGISTRATION
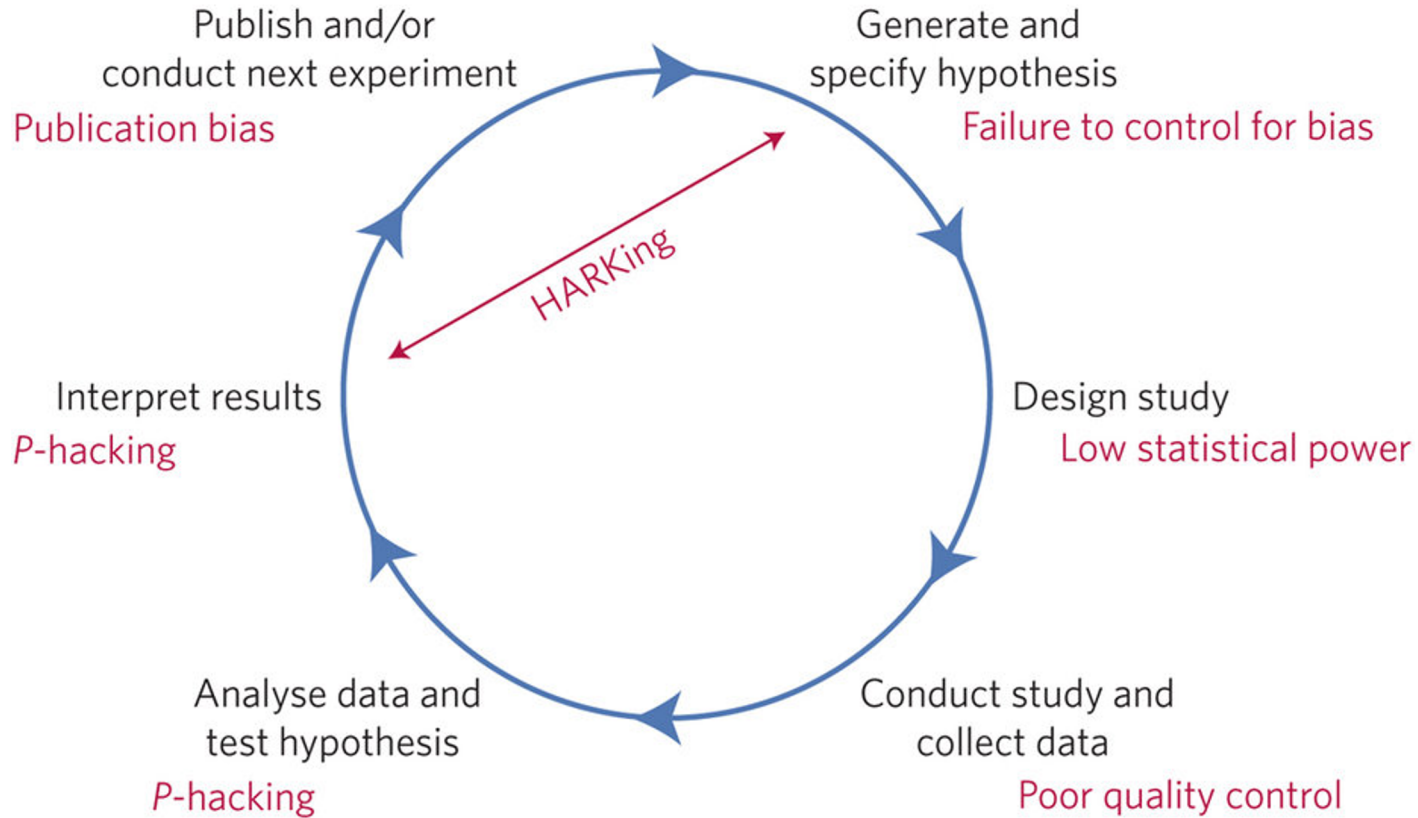
PREREGISTERED

Pamela Davis-Kean

University of Michigan

Methods Hour

10/8/2021

Slides available at cos.io/prereg

# Hypo-Deductive Model of the Scientific Method

# 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
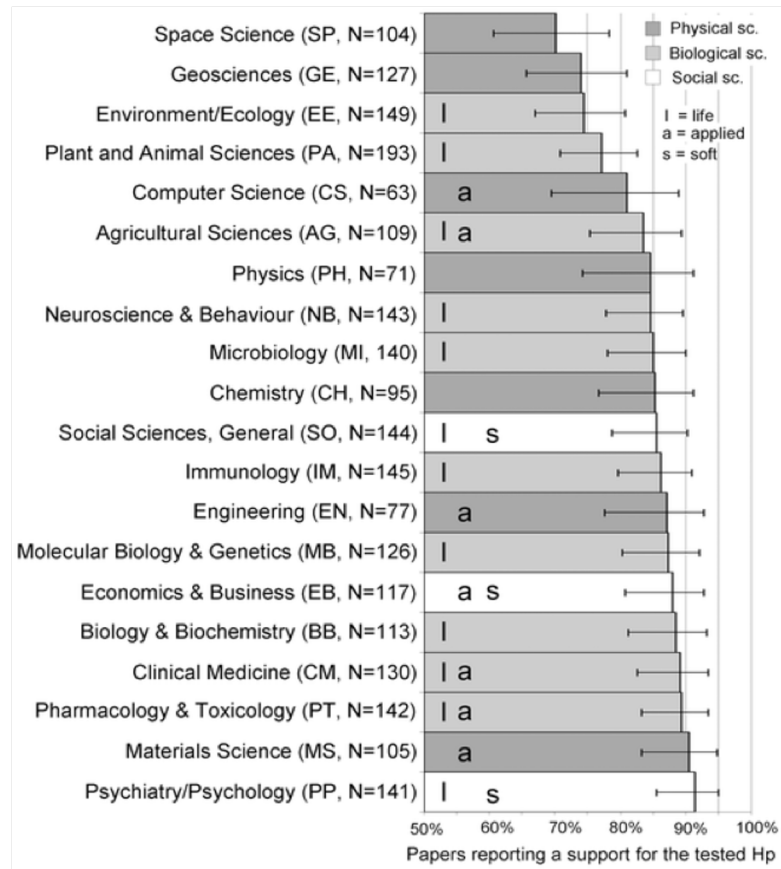
The combination of a <span style="color:red">strong bias toward statistically significant findings</span> and flexibility in data analysis results can lead to irreproducible research

Fanelli D (2010) "Positive" Results Increase Down the Hierarchy of the Sciences. PLoS ONE 5(4): e10068. doi:10.1371/journal.pone.0010068
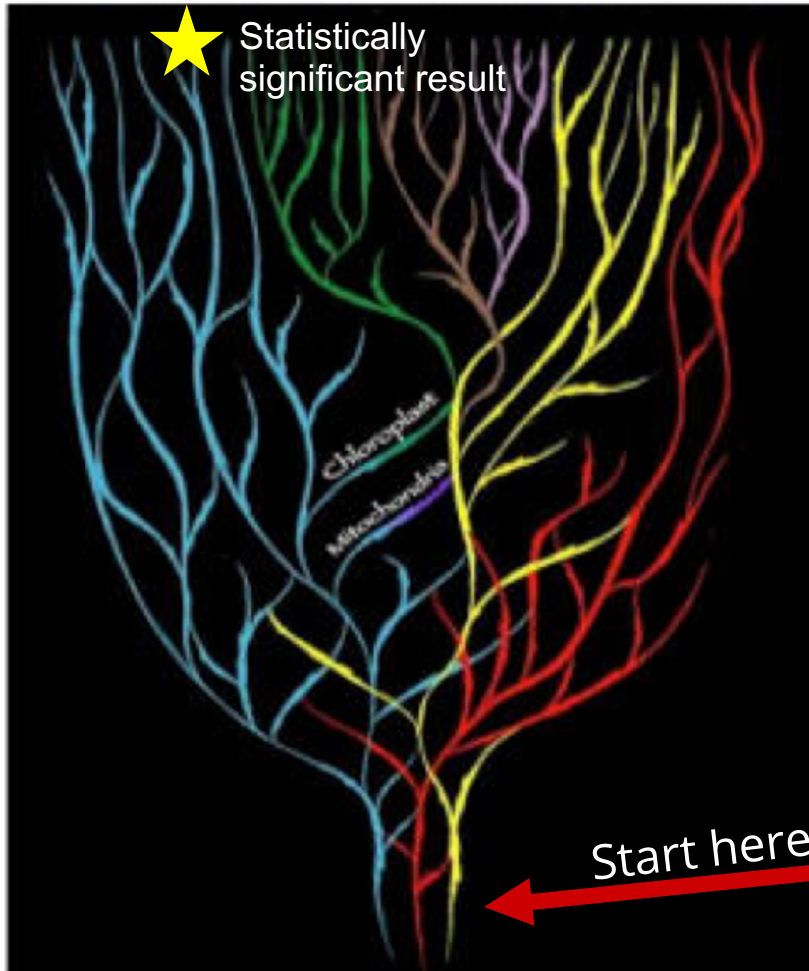
http://127.0.0.1:8081/plosone/article?id=info:doi/10.1371/journal.pone.0010068

# 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

Control for time?

Exclude outliers?
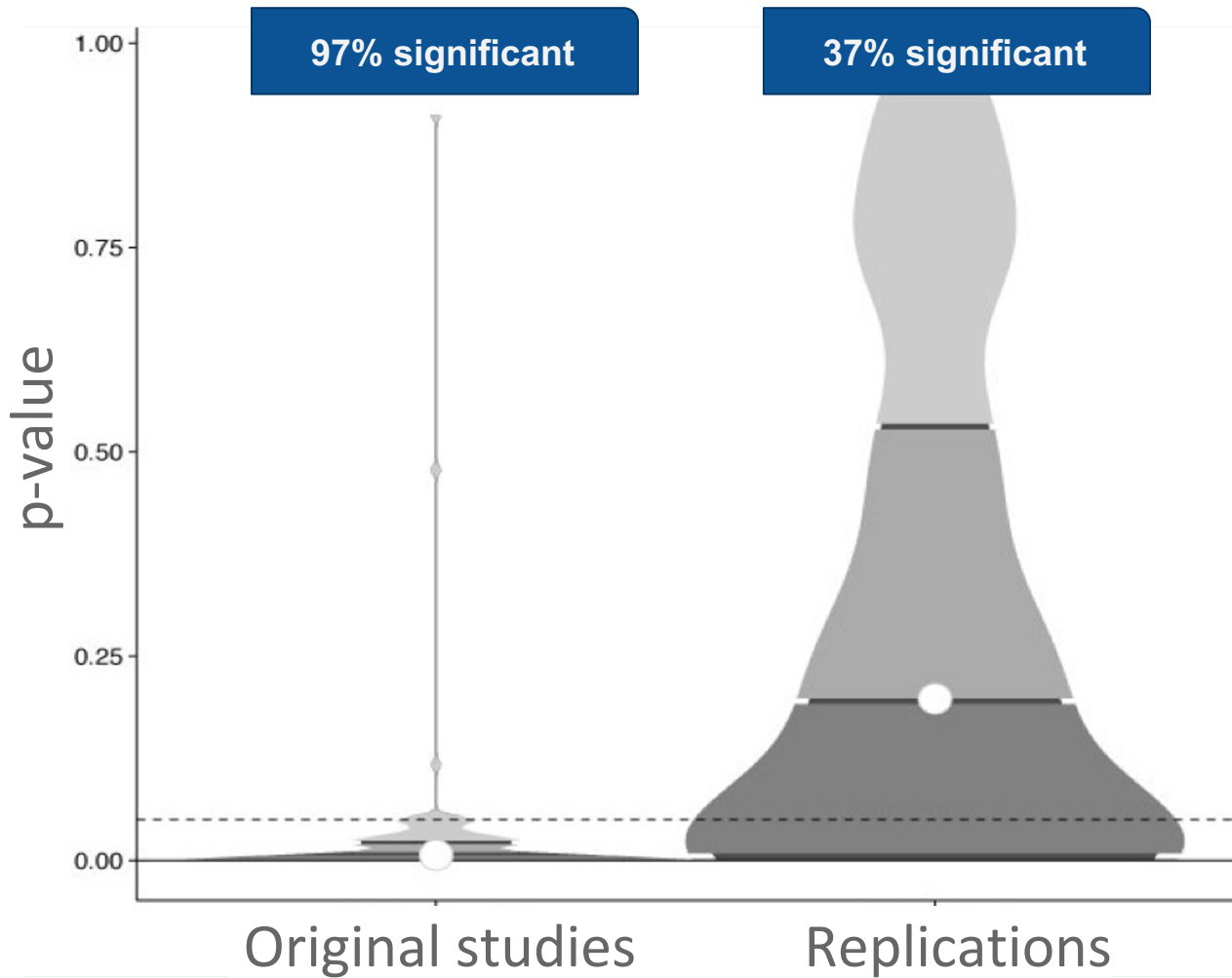
Median or mean?

**Hypothesis: "Does X affect Y?"**

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

Reproducibility Project: Psychology

97% significant

37% significant

p-value

Original studies

Replications

# 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
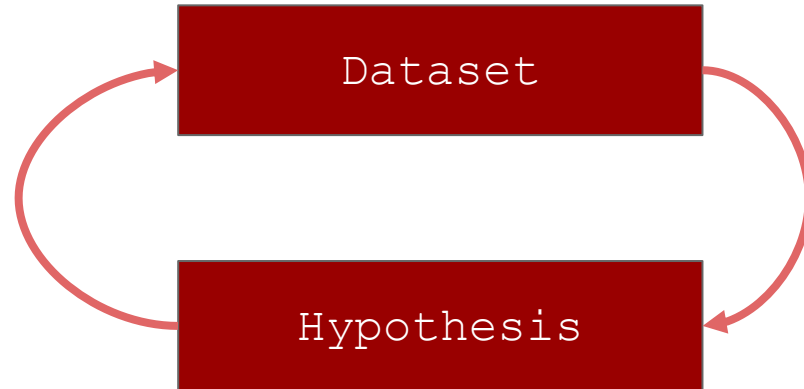
- Statistical model

- 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. **HARK**ing: 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

Presenting exploratory results as confirmatory increases the publishability of results **at the expense of credibility of results.**
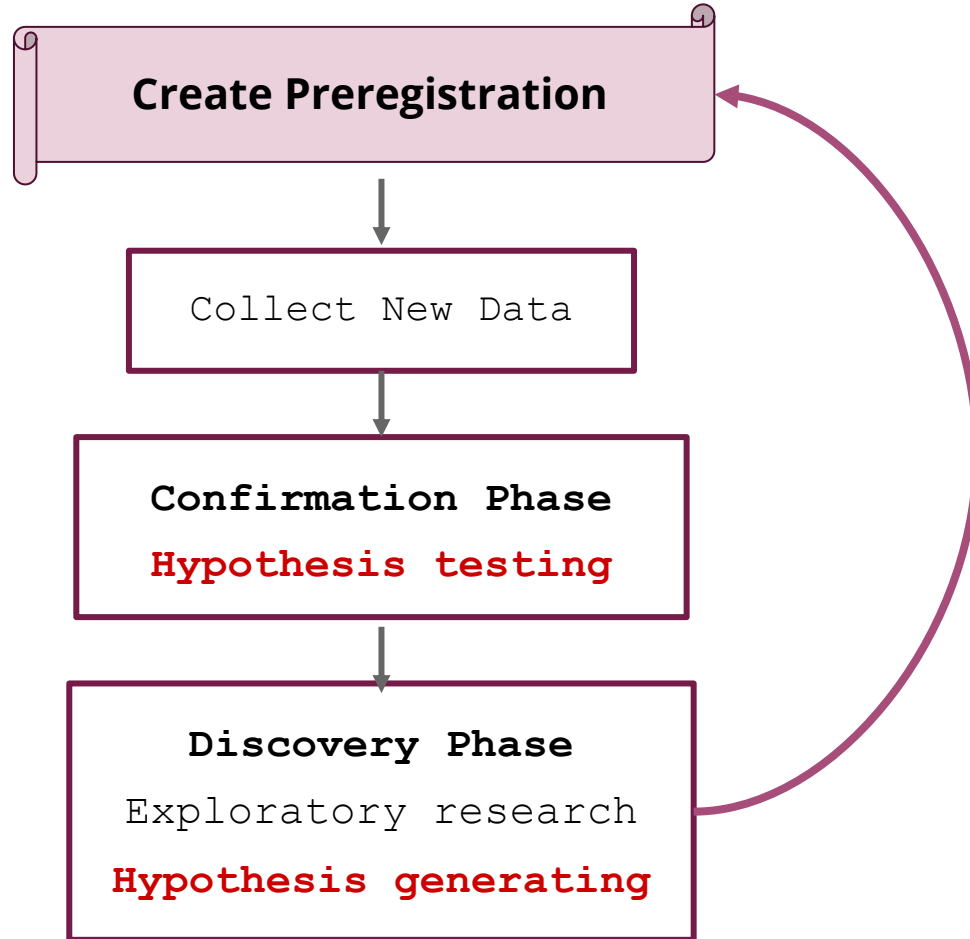
P-values interpretable

P-values meaningless

# Example workflow

Theory driven, a-priori expectations

# Prereg 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, FAQ |
| 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 |
| Replication Recipe (Brandt et al., 2013): Pre-Registration* | Register a replication study with a series of questions regarding the original work. | Word, GoogleDoc |
| Replication Recipe (Brandt et al., 2013): Post-Completion | Register a replication study after it has been conducted with questions regarding the outcomes of the replication. | Word, GoogleDoc |
| Pre-Registration in Social Psychology (van 't Veer & Giner-Sorolla, 2016)* | Preregister a research study outlining the hypotheses, methods, and analysis plan | Word, GoogleDoc, OSF |
| Registered Report Protocol Preregistration | Register your protocol AFTER having been given "in-principle acceptance" from a Registered Report journal | Word, GoogleDoc, OSF Workflow |
| Secondary Data Preregistration* | For preregistering a research project that uses an existing dataset. | OSF Page, Example, FAQ |

# **Prereg Templates**

These forms are not available as guided workflows on the OSF, but rather as template docs that you can fill out and register using the "Open-Ended Registration" form. We are always evaluating our guided workflows and will likely include one or more of these in the future.

| Registration Forms | Description | Templates |
|---|---|---|
| OLD "Qualitative Research Preregistration" | This is an earlier version of the form that is now included on OSF | OSF page |
| Cognitive Modeling (Model Application)* | Use when you wish to apply a cognitive model as a measurement tool to test hypotheses about parameters of the cognitive model | OSF Page and Preprint |
| fMRI Preregistration Template* | This project provides a detailed preregistration template for fMRI studies and provides some guidance for common difficulties that can occur for fMRI preregistration projects. | OSF page |
| Open Stats Lab and Project Tier* | This preregistration template is geared towards researchers who have little experience with preregistering studies | OSF Page |

# Evidence Prereg Increases Rigor?

- [Metascience 2021](#)

# Proportion of supported hypotheses
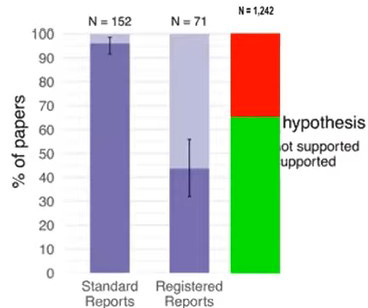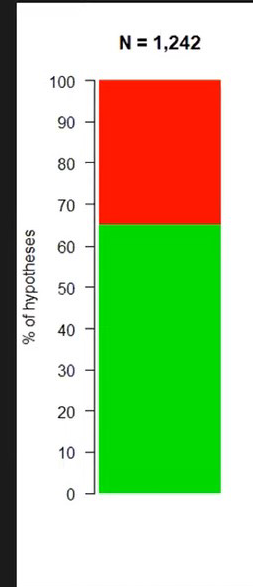
# Effect Size Correction
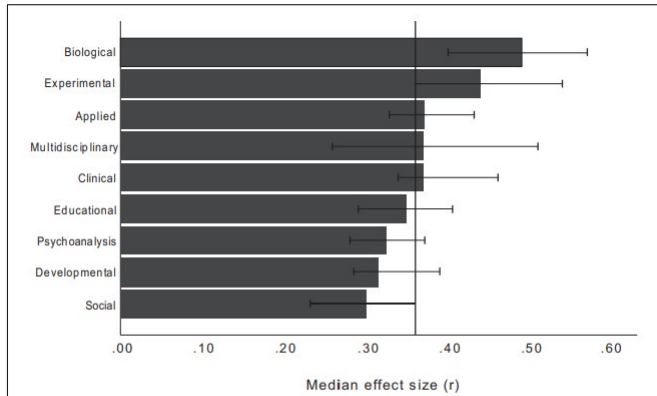


FIGURE 3 | Medians (with 95% bootstrap confidence intervals) of effects published without pre-registration (absolute values) for the nine Social Sciences Citation Index psychological sub-disciplines. The bars contain all effects that were extracted as or could be transformed into a correlation coefficient *r*. The vertical line is the grand median.
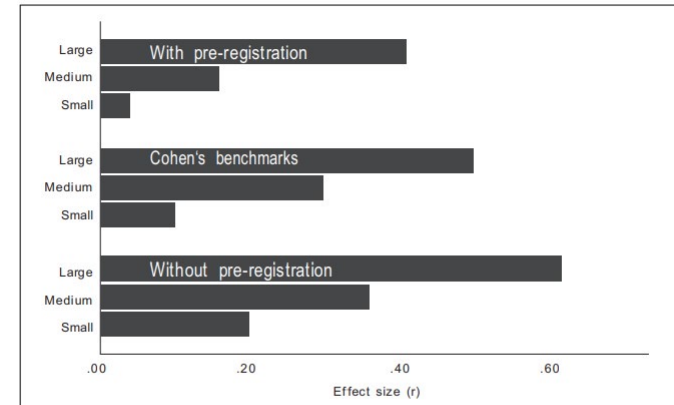
FIGURE 2 | Cohen's conventions for small, medium, and large effects compared with empirical effects (lower, grand, and upper median) from studies published with and without pre-registration.

Schäfer T and Schwarz MA (2019) The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. Front. Psychol. 10:813. doi: 10.3389/fpsyg.2019.00813

**TABLE 2 |** Median, mean, and SD of sample size, and percentage of significant effects for all studies where an effect size ($r$) was extracted or calculated.

| | All studies | | | | Between-subjects designs | | | | Within-subject designs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Mdn_N$ | $M_N$ | $SD_N$ | % sig. | $Mdn_N$ | $M_N$ | $SD_N$ | % sig. | $Mdn_N$ | $M_N$ | $SD_N$ | % sig. |
| **Studies with pre-registration** | | | | | | | | | | | | |
| All disciplines | | | | | | | | | | | | |
| | 268 | 1756 | 12424 | 64 | 358 | 2400 | 14730 | 63 | 71 | 181 | 295 | 65 |
| **Studies without pre-registration** | | | | | | | | | | | | |
| All disciplines | | | | | | | | | | | | |
| | 89 | 364 | 1729 | 79 | 82 | 299 | 849 | 84 | 89 | 415 | 2198 | 74 |
| Applied | | | | | | | | | | | | |
| | 190 | 524 | 1892 | 84 | 120 | 214 | 253 | 100 | 190 | 616 | 2147 | 79 |
| Biological | | | | | | | | | | | | |
| | 32 | 132 | 715 | 78 | 36 | 221 | 1038 | 86 | 30 | 52 | 74 | 71 |
| Clinical | | | | | | | | | | | | |
| | 90 | 217 | 461 | 86 | 74 | 278 | 620 | 87 | 96 | 165 | 254 | 84 |
| Developmental | | | | | | | | | | | | |
| | 82 | 232 | 518 | 80 | 80 | 228 | 481 | 84 | 88 | 236 | 554 | 76 |
| Educational | | | | | | | | | | | | |
| | 103 | 453 | 1703 | 68 | 70 | 323 | 770 | 71 | 107 | 560 | 2192 | 66 |
| Experimental | | | | | | | | | | | | |
| | 42 | 86 | 221 | 73 | 70 | 106 | 137 | 83 | 30 | 77 | 252 | 68 |
| Multidisciplinary | | | | | | | | | | | | |
| | 160 | 400 | 793 | 83 | 154 | 428 | 955 | 88 | 160 | 370 | 589 | 71 |
| Psychoanalysis | | | | | | | | | | | | |
| | 91 | 387 | 1141 | 79 | 91 | 536 | 1510 | 77 | 89 | 219 | 412 | 83 |
| Social | | | | | | | | | | | | |
| | 150 | 847 | 4153 | 84 | 128 | 207 | 254 | 89 | 191 | 1680 | 6341 | 65 |

$Mdn_N$ = Median of sample size; $M_N$ = Mean of sample size; $SD_N$ = SD of sample size; % sig. = Percentage of significant effects ($p < 0.05$). Note that studies with pre-registration were too few to divide them into sub-disciplines.

# FAQs

# 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?

# 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."

# 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!



Learn more: cos.io/prereg