Title

The preregistration revolution needs to distinguish between predictions and analyses

Journal

Proceedings of the National Academy of Sciences of the United States of America, 115(45)

ISSN

0027-8424

Author

Ledgerwood, Alison

Publication Date

2018-11-06

DOI

10.1073/pnas.1812592115

The preregistration revolution needs to distinguish between predictions and analyses

Alison Ledgerwood[1]

Nosek et al. (1) recently joined others in advocating for "widespread adoption of preregistration" as a tool for advancing science. But the language they use in making this important argument creates unnecessary confusion: Like many others discussing these issues, they seem to conflate the goal of *theory falsification* with the goal of *constraining Type I error.* This masks a crucial distinction between two types of preregistration: Preregistering a theoretical, a priori, directional *prediction* (which serves to clarify how a hypothesis is constructed) and preregistering an *analysis plan* (which serves to clarify how evidence is produced).

Indeed, philosophers of science have identified elements of both how a *hypothesis* is constructed and how *evidence* is produced that are important for scientifically valid inference (2-4). We can distill these to two key, separable questions.

Q1. Have these data influenced my theoretical prediction? This question is relevant when researchers want to test existing theory: Rationally speaking, we should only adjust our confidence in a theory in response to evidence that was not itself used to construct the theoretical prediction in question (3). Preregistering theoretical predictions can help researchers distinguish clearly between using evidence to inform versus test theory (3,5,6).

Q2. Have these data influenced my choice of statistical test (and/or other dataset-construction/ analysis decisions)? This question is relevant when researchers want to know the Type I error rate of statistical tests: Flexibility in researcher decisions can inflate the risk of false positives (7,8). Preregistration of analysis plans can help researchers distinguish clearly between data-dependent analyses (which can be interesting but may have unknown Type I error) and data-independent analyses (for which *p*-values can be interpreted as diagnostic about the likelihood of a result; 1,9).

Put differently, preregistration of theoretical predictions helps researchers know how to correctly calibrate their confidence that a study tests (vs. informs) a *theory*, whereas preregistration of analysis plans helps researchers know how to correctly calibrate their confidence that a specific *finding* is unlikely to be due to chance.

Conflating theoretical predictions and analyses is problematic for multiple reasons. First, it implies, erroneously, that pre-analysis plans can only help control Type I error when research is in a prediction-making /theory-testing phase (e.g., Theory Z predicts a gender difference in Trait X; 1)—in fact, pre-analysis plans can also be useful in the question-asking/discovery/theory-building phase (e.g., is there a gender difference in Trait X?). Second, it may lead people to preregister the wrong things (e.g., a researcher attempting to control Type I error records careful predictions but omits or only loosely specifies a pre-analysis plan). Third, it increases misunderstandings and backlash against preregistration as scientists discuss these issues in everyday life (e.g., students erroneously infer that their results are more robust if they correctly guess them ahead of time; skeptics understandably argue that recording one's prediction ahead of time has no effect on Type I error).

If we want clear communication, productive debates, and effective strategies for advancing science, we must first pull apart our tangled terminology. Preregistering theoretical predictions enables theory falsifiability. Preregistering analysis plans enables Type I error control.

[1] UC Davis

Author Note. The author thanks Brian Nosek, Simine Vazire, Paul Eastwick, and Roger Giner-Sorolla for their very helpful comments on earlier versions of this commentary**.**

Citation: Ledgerwood, A. (2018). The preregistration revolution needs to distinguish between predictions and analyses. *Proceedings of the National Academy of Sciences*, 201812592.

References

1. Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 201708274.
2. Mayo, D. G. (1991). Novel evidence and severe tests. *Philosophy of Science, 58,* 523-552.
3. Worrall., J. (2010). Error, tests, and theory confirmation. In D. G. Mayo & A. Spanos (Eds.), Error and inference: Recent exchanges on experimental reasoning, reliability,a nd the objectivity and rationality of science (pp. 125-154).
4. Rubin, M. (2017). When does HARKing hurt? Identifying when different types of undisclosed post hoc hypothesizing harm scientific progress. *Review of General Psychology*, *21*, 308-320.
5. Popper, K. (1962). *Conjectures and refutations.* New York: Basic Books.
6. van't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, *67*, 2-12.
7. Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*, 460-465.
8. Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359-1366.
9. de Groot, A. D. (2014). The meaning of “significance” for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han L. J. Van der Maas]. *Acta Psychologica, 148*, 188-194.