

The preregistration revolution needs to distinguish between predictions and analyses

Alison Ledgerwood^{a,1}

Nosek et al. (1) recently joined others in advocating for “widespread adoption of preregistration” as a tool for advancing science. The language they use in making this important argument, however, 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:

- i) 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).
- ii) 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; refs. 1 and 9).

Put differently, preregistration of theoretical predictions helps researchers know how to correctly calibrate their confidence that a study tests (versus 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 preanalysis 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; ref. 1). In fact, preanalysis 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 preanalysis 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.

Acknowledgments

I thank Brian Nosek, Simine Vazire, Paul Eastwick, and Roger Giner-Sorolla for their very helpful comments on earlier versions of this letter.

- 1 Nosek BA, Ebersole CR, DeHaven AC, Mellor DT (2018) The preregistration revolution. *Proc Natl Acad Sci USA* 115:2600–2606.
- 2 Mayo DG (1991) Novel evidence and severe tests. *Philos Sci* 58:523–552.
- 3 Worrall J (2010) Error, tests, and theory confirmation. *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science*, eds Mayo DG, Spanos A (Cambridge Univ Press, New York), pp 125–154.

^aDepartment of Psychology, University of California, Davis, CA 95616

Author contributions: A.L. wrote the paper.

The author declares no conflict of interest.

Published under the [PNAS license](#).

¹Email: aledgerwood@ucdavis.edu.

- 4 Rubin M (2017) When does HARKing hurt? Identifying when different types of undisclosed post hoc hypothesizing harm scientific progress. *Rev Gen Psychol* 21:308–320.
- 5 Popper K (1962) *Conjectures and Refutations* (Basic Books, New York).
- 6 van't Veer AE, Giner-Sorolla R (2016) Pre-registration in social psychology—A discussion and suggested template. *J Exp Soc Psychol* 67:2–12.
- 7 Gelman A, Loken E (2014) The statistical crisis in science. *Am Sci* 102:460–465.
- 8 Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22:1359–1366.
- 9 de Groot AD (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]. 1969. *Acta Psychol (Amst)* 148:188–194.