

## Letter

## Is Preregistration Worthwhile?

Aba Szollosi,<sup>1,\*</sup>David Kellen,<sup>2</sup>Danielle J. Navarro,<sup>1</sup>Richard Shiffrin,<sup>3</sup> Iris van Rooij,<sup>4</sup>Trisha Van Zandt,<sup>5</sup> andChris Donkin<sup>1</sup>

Proponents of preregistration argue that, among other benefits, it improves the diagnosticity of statistical tests [1]. In the strong version of this argument, preregistration does this by solving statistical problems, such as family-wise error rates. In the weak version, it nudges people to think more deeply about their theories, methods, and analyses. We argue against both: the diagnosticity of statistical tests depend entirely on how well statistical models map onto underlying theories, and so improving statistical techniques does little to improve theories when the mapping is weak. There is also little reason to expect that preregistration will spontaneously help researchers to develop better theories (and, hence, better methods and analyses).

To be clear about the scope of our argument, we do not discuss the role that preregistration plays in increasing the transparency of studies and methods in depth. Our critique is focused on the use of preregistration as an attempt to improve scientific reasoning and theory development.

## Fixing Science with Statistics

Scientific progress consists of iterative attempts to improve theories by trying to find and correct errors in them [2,3]. One useful technique for this process is to describe theories mathematically. In the social and behavioral sciences, such

mathematical formalisms are typically implemented as statistical models. Yet, the extent to which such models are useful depends entirely on how accurate the match between the theory and model is (Box 1). Many statistical models currently in use often approximate the underlying theory poorly (e.g., the general linear model). As such, development of theories, improving their implementation as models, and issues of experimental design should receive priority over the often irrelevant problems associated with poor statistical approximations [4,5].

Solving problems of poor statistical practice with preregistration does not compensate for weak theory [6]. Imagine making a random prediction regarding the outcome of an experiment. Should we observe the predicted outcome, we would not regard this theory as useful for making subsequent predictions. Why should we judge this theory as better if it were preregistered? In the absence of any theoretical reason for conducting a statistical test, solving statistical problems (such as increasing control over family-wise error rates) should be considered scientifically meaningless (Box 1). Researchers can preregister bad theories, the predictions of which are barely better than randomly picking an outcome, but which can nonetheless bear out in experiments.

Unlike *post hoc* comparisons in statistical inference, there is nothing inherently problematic about *post hoc* scientific inference when theories are strong. The crucial difference is that strong scientific inference requires that *post hoc* explanations be tested just as rigorously as ones generated before an experiment; for example, by a collection of *post hoc* tests that evaluate the many regularities implied by a novel theory [7]. There is no reason not to take such *post hoc* theories seriously just because they were thought of after or were not preregistered before an experiment was conducted.

## Preregistration as a Nudge?

Although preregistration does not require the improvement of theories, many argue that it at least nudges researchers to think more deeply about how to improve their theories (e.g., through better planning, with more precise operationalisation of constructs, clearer motivation for statistical analyses, etc.). Although this might sometimes be so, there is no clear explanation for why we should expect such improvements to happen.

One could argue that researchers are motivated to improve their theories when they encounter problems preregistering a study or when preregistered predictions are wrong. The problem with this line of

## Box 1. Scientific versus Statistical Inference

Because statistical inference is so often used to inform scientific inference, it is easy to conflate the two. However, they have fundamentally different aims.

Scientific inference is the process of developing better theories. One aspect of theory development is to test whether the implications of a theory are observed in experiments. We can use statistics to help decide whether such implications are realised. Used in this way, statistical models are simplified mathematical abstractions of scientific problems. Such simplifications can aid scientific inference, but only to the extent that the abstraction is appropriate [10,11].

Some problems arising from the statistical assumptions in such abstractions, such as family-wise error rates, only exist when hypotheses and statistical comparisons are effectively chosen at random. When statistical inference is used in scientific arguments, statistical models are just tools to test implications derived from theory. Therefore, such statistical problems become irrelevant because theories, not random selection, dictate what comparisons are necessary [12].

Problems in scientific inference, however, are important when the overall goal is theory development. Issues that prevent criticism of theory, such as poor operationalisation, imprecise measurement, and weak connection between theory and statistical methods, should take precedence over problems with statistical inference.

argument is that any improvement depends upon a good understanding of how to improve a theory, and preregistration provides no such understanding. In other words, although preregistration may prompt more thinking, it is not clear how it provides a path towards better thinking. We worry that, without this understanding, preregistration may be perceived as the solution.

Instead, better understanding of theory development requires thoughtful discussion and debate about what constitutes good scientific inference. For example, what are the best examples of good theories both within and outside of the social and behavioural sciences [8]? What are the characteristics of good theories [7]? How do we improve the link between psychological theory, measurement, methodology, and statistics? The answers to these questions are unlikely to come from nudging researchers with preregistration or some other method-oriented solution. They are likely to come from scientific problem-solving: generating, exchanging and criticising possible answers, and improving them when needed.

### Preregistration Is Nondiagnostic, and Potentially Harmful

Preregistration is not diagnostic of, nor does it necessitate, good science. When preregistration is hard, it is hard because it is based on good theory, and good theorising is hard. However, preregistration does not require that the underlying theory be strong, nor does it discriminate between experiments based on strong or weak theory. Because bad theories, methods, and analyses can also be preregistered, we should remain aware of the ways in which widespread adoption of preregistration could harm progress in our field.

Ultimately, we must decide whether any additional benefits outweigh the potential costs [9]. Transparency of methods and studies are important, but there are other solutions to solve such problems (e.g., asking researchers to disclose their

studies and methods when publishing). We conclude that, although preregistration should be an option for anyone who thinks it improves their research, requiring, rewarding, or promoting it (e.g., with badges, research funding, etc.) is not worthwhile.

### Acknowledgments

A.S. and C.D. are supported by Australian Research Council grants (DP130100124 and DP190101675). Authors thank Jared Hotelling and Ben R. Newell for useful discussion, and various others for their constructive comments on a preprint of the current paper (entitled 'Preregistration is redundant, at best').

<sup>1</sup>University of New South Wales, Kensington, Australia

<sup>2</sup>Syracuse University, Syracuse, NY, USA

<sup>3</sup>Indiana University, Bloomington, Indiana

<sup>4</sup>Radboud University, Nijmegen, The Netherlands

<sup>5</sup>Ohio State University, Columbus, OH, USA

\*Correspondence:

[aba.szollosi@gmail.com](mailto:aba.szollosi@gmail.com) (A. Szollosi).

<https://doi.org/10.1016/j.tics.2019.11.009>

© 2019 Elsevier Ltd. All rights reserved.

### References

1. Nosek, B.A. et al. (2019) Preregistration is hard, and worthwhile. *Trends Cogn. Sci.* 23, 815–818
2. Popper, K. (1959) *The Logic of Scientific Discovery*, Routledge
3. Deutsch, D. (2011) *The Beginning of Infinity: Explanations That Transform the World*, Allen Lane
4. Fiedler, K. (2017) What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspect. Psychol. Sci.* 12, 46–61
5. MacEachern, S.N. and Van Zandt, T. (2019) Preregistration of modeling exercises may not be useful. *Comput. Brain Behav.* 2, 179–182
6. van Rooij, I. (2019) Psychological science needs theory development before preregistration. *Psychon. Soc.* <https://featuredcontent.psychonomic.org/psychological-science-needs-theory-development-before-preregistration/>
7. Szollosi, A. and Donkin, C. (2019) Arrested theory development: the misguided distinction between exploratory and confirmatory research. *PsyArXiv Preprints* Published online September 21, 2019. <http://dx.doi.org/10.31234/osf.io/suzej>
8. Fiedler, K. (2018) The creative cycle and the growth of psychological science. *Perspect. Psychol. Sci.* 13, 433–438
9. Shiffrin, R.M. et al. (2018) Scientific progress despite irreproducibility: a seeming paradox. *Proc. Natl. Acad. Sci.* 115, 2632–2639
10. Navarro, D.J. (2018) Between the devil and the deep blue sea: tensions between scientific judgement and statistical model selection. *Comput. Brain Behav.* 2, 28–34
11. Kellen, D. (2019) A model hierarchy for psychological science. *Comput. Brain Behav.* 2, 160–165
12. Oberauer, K. and Lewandowsky, S. (2019) Addressing the theory crisis in psychology. *Psychon. Bull. Rev.* 26, 1596–1618

## Spotlight

# Semantic Search as Pattern Completion across a Concept

Sarah H. Solomon<sup>1,\*</sup> and Anna C. Schapiro<sup>1,\*</sup>



**What role does the hippocampus play in semantic memory? In a recent paper, Cutler et al. use a vector space model of semantics to characterize semantic search deficits in hippocampal amnesia. We relate their findings to properties of the hippocampal neural code and to controversies regarding hippocampal contributions to cognition.**

The hippocampus is known to play a crucial role in episodic memory, but its influence extends widely into other memory systems and cognitive domains [1,2]. This raises questions regarding the nature of hippocampal processing. Does the hippocampus contribute to online cognition only by retrieving stored hippocampal representations, or can it dynamically form new representations to support diverse cognitive functions? A version of this question is currently under debate in the episodic memory domain – is the hippocampus involved in remote episodic memory because it stores those memory traces, or because it helps to reconstruct cortically represented traces at the time of retrieval [3]?

In their recent paper, Cutler, Duff, and Polyn (hereafter, CDP) [4] ask a pertinent question in the semantic memory domain – how does the hippocampus contribute to the retrieval of remote semantic memories (i.e., our generalized knowledge of the world)? An earlier study by Klooster and Duff [5] reported that patients with hippocampal damage produce fewer responses than healthy controls on a semantic