Running Head: Pre-registration Is A Game Changer

Pre-registration Is A Game Changer.

But, Like Random Assignment, It Is Neither Necessary Nor Sufficient For Credible Science

Joseph Simmons
University of Pennsylvania

Leif Nelson
University of California, Berkeley

Uri Simonsohn
ESADE

Abstract

We identify 15 claims Pham and Oh (2020) make to argue against pre-registration. We agree with 7 of the claims, but think that none of them justify delaying the encouragement and adoption of pre-registration. Moreover, while the claim they make in their title is correct - pre-registration is neither necessary nor sufficient for a credible science – this is also true of many our science's most indispensable tools, such as random assignment. Indeed, both random assignment and pre-registration lead to more credible research. Pre-registration is a game changer.

Pham and Oh (2020) invited the editorial boards of the *Journal of Consumer Research* and the *Journal of Consumer Psychology* to complete a survey about pre-registration. As they report, 90.6% of the 181 researchers who opted to complete the survey opposed *mandating* pre-registration for all empirical research. This result might lead readers to believe that most researchers in our field are against pre-registering, but if the three of us had been invited to complete the survey, that percentage would have been *higher*. We don't think *all* empirical research should be pre-registered, because some research is exploratory, some research involves analyses of existing data sets, some research does not use inferential statistics, and so on. We, like Pham and Oh (2020), and the vast majority of researchers who completed their survey, oppose a universal pre-registration mandate.

Our agreement with Pham and Oh (2020) does not stop here. We carefully perused their arguments against pre-registration, and we identified 15 separate claims, which we briefly summarize (and respond to) in Figure 1. As shown in that Figure, we agree with seven of them, and we also see merit in several statements that we disagree with. In what follows we discuss in more detail the five claims we thought readers would be most curious about.

In the end, we conclude that, like so many effective practices, from using computers to analyze data to reading the literature prior to running new studies, pre-registration may not be necessary nor sufficient, but our research will be a lot better with it than without it.

Arguments by Pham & Oh (2020)	Do We Agree? Yes No		Response

Pre-registrations have frictions

		Yes	No	
1	It is costly to maintain a pre-registration depository		✗	The average AsPredicted pre-registration costs <5 cents.
2	It is onerous for researchers to create pre-registrations		✗	20,000+ people have used AsPredicted.org, despite the lack of. extrinsic rewards for pre-registering. It can't be that
3	It is onerous for reviewers to evaluate		✗	We believe they *reduce* the average reviewer's workload. (Read More)
4	It is easier to pre-register a study with a simple design	✓		Sometimes. But why is this a problem? (Read More) If you can run it, you can pre-register it.

Pre-registrations are insufficient

5	Not applicable to some research designs (e.g., CCT)	✓		Like MTurk, Qualtrics, and ANOVA, pre-registrations are mostly a tool for those who collect new and quantitative
6	Deviating from pre-registration defeats purpose		✗	Deviating is 100% totally fine. (Read More) Just say you did it and explain why.
7	Gaming pre-registration is possible	✓		Like literature reviews, peer review, and ANOVAs, (Read More) pre-registrations are gameable. But they are rarely gamed.
8	Pre-registration does not make studies more robust	✓		Robustness is not a goal of pre-registration.
9	Pre-registered studies may be fundamentally flawed	✓		In such cases, it will make sense for researchers to deviate from their pre-registration. And to explain why.

Pre-registrations are unnecessary

10	The nature of much of consumer research is exploratory		✗	If a finding relies on p-values, confidence intervals, or Bayes factors, it purports to be confirmatory.
11	Other methods can improve transparency (e.g., training, open data & materials)	✓		YES. Please also start posting your data and materials. ResearchBox.org makes it very easy.

Risks of pre-registration

12	Pre-registration discourages exploration		✗	No, it encourages researchers to properly label (Read More) confirmatory and exploratory analyses.
13	Pre-registration encourages studies that are likely to work		✗	Pre-registration should have no effect on what a researcher studies; it does make surprising findings more credible.
14	Pre-registration encourages easy to run studies		✗	Clinical trials are among the most difficult-to-run studies. Fortunately, they are all pre-registered.
15	Pre-registration encourages narrow theorizing		✗	Pre-registration enables (narrow or broad) theories to be built on true effects.

Fig 1. *Fifteen arguments raised by Pham and Oh (2020) and our responses to them. Arguments accompanied by a "Read More" icon are discussed in more detail in the text.*

Pham and Oh's (2020) Argument #3: Pre-registration Is Onerous For Reviewers.

We believe pre-registrations are likely to *reduce* the average reviewer's workload. Some reviewers are not interested in distinguishing confirmatory from exploratory analyses. Those reviewers will simply not open the pre-registration documents made available to them, and so their workloads will remain unchanged. Other reviewers are interested in that distinction, perhaps because they understand and care about the fact that inferential statistics (e.g., *p*-values, confidence intervals, and Bayes factors) are valid only when applied to confirmatory analyses.

What do *those* reviewers do now, when most consumer research manuscripts do *not* include a pre-registration? Some ignore this issue (i.e., fail at the job that they want to do). Others read and re-read the manuscript to try to guess which analytic decisions were planned vs. unplanned, and what the consequences of those analytic decisions may have been. When repeated reads are insufficient, reviewers pose questions to the authors – "Did you try a different outlier rule? Did you decide ahead of time to exclude participants who failed the attention check? What other dependent variables did you collect? Was the decision to control for gender made ex ante?" – and wait days, weeks, or months for a response.

When studies are pre-registered, reviewers do not have to try to read between the lines to figure out whether an analytic decision was made in advance, and they do not have to contact authors and the editor to get what might be insufficiently satisfactory answers to questions about the design or analysis. All they have to do is consult the time-stamped pre-registration document, and check whether the methods and analyses that are reported in the manuscript match the methods and analyses that were described in the pre-registration. Life for these reviewers is much easier, not harder.

Pham and Oh's (2020) Argument #4: Complex Designs Have Costlier Pre-registrations

Pre-registering complex designs does not *add* to the complexity involved in designing and analyzing them. Consider the single example Pham and Oh (2020) provide to make their argument: "Full specification of an analytical plan becomes rapidly more complex as the number of variables and constructs increases . . . As an illustration . . . [the] PROCESS macro for mediation analysis now includes no fewer than 80 different statistical models" (p. 7).

We agree that it would be a tremendous burden for a researcher to explain all of the 80+ PROCESS models available. Fortunately, this is not required in order to adequately pre-register a

study that will use *one* of those models. In general, the researcher's burden is not made greater by having to describe what they planned to do before running the study rather than afterwards. To pre-register the analysis in question, the authors could simply write, "We will run PROCESS Model 8". This is not prohibitively difficult.

While we believe that the vast majority of experimental studies are simple enough to pre-register analyses for, we also believe that exceptions might exist. In those situations, pre-registration can still specify the sample size, hypothesis, experimental conditions, measures, and exclusions. For the complex analysis itself, the pre-registration could describe the analytic *approach* without detailing each concrete analytical step, and readers will know that some analytic decisions were made after the study was run rather than before (and thus that the study's conclusions are more tentative).

Pham and Oh's (2020) Argument #6: Deviating From a Pre-registration Defeats Its Purpose

The purpose of pre-registration is not to commit to a course of action, but to distinguish between planned and unplanned courses of action. If an author realizes there is a problem with a pre-registered analysis, they should abandon that analysis, but this does not defeat the purpose of pre-registration. On the contrary, readers who look at the pre-registration will be adequately warned that the analysis was chosen after the data were collected, and they can do with that information whatever they want (just as they can do what they want with all other pieces of information included in a published paper). Similarly, if a reviewer requests an additional analysis, then there is no problem adding it to the article; the pre-registered analysis is unaffected by other analyses that supplement it.

In general, if an author decides to deviate from a pre-registration, it seems good practice to (1) make it clear to the reader that this is the case, (2) explain the reason(s) for the deviation, and (3)

report somewhere in the paper or supplement the results for the pre-registered analysis, even if deemed invalid or inferior to the one chosen. The authors had that information available when deciding what to report, so it seems sensible that a reader should have it available as well.

Pham and Oh's (2020) Argument #7: Pre-registrations Can Be Gamed

First, yes, everything can be gamed. Some researchers fake data, some voters vote twice, some spouses cheat, and yet, we still do research, vote, and get married. But let's consider the specific scenarios of abuse considered by Pham and Oh (2020). They write that authors "may preregister multiple studies or versions of the same study, then only report a selection" (p. 9). On AsPredicted.org this is not easy to do because algorithms flag and prevent similar submissions. When a new pre-registration resembles a recently created one, authors are asked to either bundle them or delete the earlier one. The OSF and SocialScienceRegistry combat this problem by requiring all pre-registrations to eventually be made public; thus, when readers are concerned another similar submission exists, they could wait for the embargo period to end (so that the pre-registrations become public) and then look for it via search queries. Technology aside, after receiving roughly 40,000 submissions on AsPredicted.org, and despite our quite actively looking for it during the first few years (before the algorithmic solution was implemented), we are unaware of any instance of this form of deception having been attempted.

A second concern raised is that researchers "may preregister a study only after running a long series of calibration tests that are not reported". We don't see a problem with this. Researchers can run pilot studies to try to calibrate their stimuli and instructions, and those pilot studies need not be pre-registered. And it is good, rather than bad, that those pilot studies will then be replicated prior to publication instead of incorrectly being passed on to readers as if they were confirmatory.

We agree that if the researcher then finds that the result works for those stimuli and only those stimuli, then that is a problem. A *generalizability* problem. But generalizability is not the problem that pre-registration is intended or designed to solve. Pre-registration is designed to ensure that if you run the exact same study again, you will get, within sampling error, the same result. It is designed to ensure that findings are *replicable*.

We absolutely share Pham and Oh's (2020) concerns about generalizability. More specifically, we too are concerned that some replicable research findings may not generalize to the real-world circumstances of interest or even simply to other stimuli. Furthermore, we believe that researchers almost always insufficiently sample their stimuli, and in our opinion, this is among the biggest and most underappreciated shortcomings of experimental work (see Wells & Windschitl, 1999). But lack of stimulus sampling is not a problem that is made worse (or better) by pre-registration.

Lastly, Pham and Oh (2020) express the concern that authors may "specify the hypotheses to be tested and analyses to be performed in loose terms, hence buying themselves flexibility." We believe this happens often. But, on the one hand, even vague pre-registrations *reduce* flexibility, and so are better than the alternative. And, on the other hand, a vague pre-registration makes the ambiguity of analytical decisions transparent and explicit, further serving the intended purpose of pre-registration. We believe researchers will over time become better pre-registration writers as reviewers start holding their pre-registrations to the higher standards outlined in our previous article (Simmons, Nelson, & Simonsohn, in press). Journals who reward pre-registrations that are properly written and followed, instead of rewarding the mere act of including a pre-registration, will ultimately make pre-registration better.

As Vazire (2019) eloquently put it ,"Transparency doesn't guarantee credibility; transparency and scrutiny together guarantee that research gets the credibility it deserves." Just as non-

preregistered studies deserve less credibility, so do studies that are accompanied by overly flexible pre-registrations.

Pham and Oh's Argument #12: Pre-registration Reduces Exploration

Because researchers are supposed to follow their pre-registration plans when analyzing their data, it is natural to presume that the act of pre-registration might reduce researchers' likelihood of carrying out exploratory analyses that were not planned. Nevertheless, this need not and should not be the case (see our section on this in our previous paper; Simmons et al., in press). Researchers who pre-register their analyses should conduct and report the results of exploratory analyses. Pre-registration does not prevent researchers from conducting exploratory analyses; it merely prevents researchers from labeling exploratory analyses as confirmatory.

Conclusions

We agree with Pham and Oh's (2020) contention that pre-registration is neither necessary nor sufficient for credible research. In fact, we would go further. *Nothing* is a sufficient condition for credible research, and exceedingly few conditions are necessary for credible research. For example, the use of random assignment does not ensure that a research finding is credible, nor is random assignment a necessary condition for credibility. But random assignment is extremely valuable, and it would be absurd to campaign against random assignment because it is neither necessary nor sufficient for researchers to use random assignment. Pre-registration provides researchers with a cost-effective way to satisfy a core mathematical assumption on which virtually all inferential statistical tools depend. Pre-registration is, like random assignment, an indispensable tool for all experimental scientists, and it is going to be a game changer for consumer psychology.

The sooner, the better.

References

Pham, M. T. & Oh, T. T. (in press). Preregistration is neither sufficient nor necessary for good science. *Journal of Consumer Psychology.*

Simmons, J. P., Nelson, L. D., & Simonsohn U. (in press). Pre-registration: Why and How. *Journal of Consumer Psychology.*

Vazire, S. (2019). Do we want to be credible or incredible? *APS Observer, 33*(1).

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*(9), 1115-1125.