Breaking free
How preregistration hurts scholars and science

Rose McDermott, *Brown University, USA*

Abstract. **Pre-registration has become an increasingly popular proposal to address concerns regarding questionable research practices. Yet preregistration does not necessarily solve these problems. It also causes additional problems, including raising costs for more junior and less resourced scholars. In addition, pre-registration restricts creativity and diminishes the broader scientific enterprise. In this way, pre-registration neither solves the problems it is intended to address, nor does it come without costs. Pre-registration is neither necessary nor sufficient for producing novel or ethical work. In short, pre-registration represents a form of virtue signaling that is more performative than actual.**

Key words: pre-registration, pre-analysis plan, research methods, experimental, creativity

In 1999, I attended a workshop at what was then called the Center for Behavioral Research in the Social Sciences, run by James Alt at Harvard University. He wanted to increase the utility of experimental methods in political science, and so he invited a small number of luminaries in the field to attend, including Shanto Iyengar, Rick Wilson, and a couple of notable economists. Later iterations included others such as Becky Morton and Alan Gerber. I was a first-year faculty member at Cornell University doing an Olin Fellowship that year. Steve Rosen told Jim Alt that I had some training in psychology and that I might be interested in attending, and so Jim graciously included me in those early meetings. Since, in some sense, I was present at the creation, in however minor and junior a role, I feel some sense of responsibility for the introduction of experimental methods to political science. In most ways, this has been a very useful and productive incorporation; in other ways, most notably ethical ones, I have some concerns about the monster it has become (McDermott & Hatemi, 2020).

Recently, additional methodological and procedural concerns have taken hold as well. One of these relates to the imperative of filing pre-analysis plans prior to publication. Daniel Rubenson raised this issue in the Fall 2021 issue of this journal. Obviously, there are many hills to die on, and while ethics clearly remains the most significant consideration, other points are worth fighting over. I choose to plant a flag against imposing a stranglehold of preregistration on experimental or other studies. To be clear, I am not arguing that preregistration is bad or wrong; neither am I claiming that it does not have a place. Rather, I claim that preregistration is not an unmitigated good; it can engender significant costs that are worth considering. Especially when it is adopted unquestioningly across the entire discipline, wider discussion should take place before preregistration is adopted as a widespread norm. The following discussion raises some of the possible negative consequences in more detail.

Perhaps some of the reservations discussed here simply reflect disciplinary differences in training and emphasis between psychology and economics, differences that find additional expression over issues such as deception. Once again, the risk is that political science remains driven by the norms of the economics profession that it considers higher in status without sufficient regard for the different scope of our own field. At the intersection with the life sciences, this risk is even greater because the imposition of preregistration risks alienating other disciplines that are less enamored of its value. However, the potentially negative implications of requiring preregistration hold

doi: 10.1017/pls.2022.4
Correspondence: Rose McDermott, Brown University, USA.
Email: *rose_mcdermott@brown.edu*

regardless of disciplinary background, and they deserve wider discussion before preregistration becomes required for publication in major journals in political science for some of the reasons explicated here, among others.

Again, I am not proposing that preregistration be banned; rather, I am suggesting that it not be required because it is neither necessary nor sufficient for the production of high-quality science. Anyone who doubts this might note that the United States recently launched a $10 billion experiment with the James Webb Space Telescope designed to look into the far reaches of the universe without a single preregistered hypothesis. Rather, the telescope was designed for pure discovery, without pre-established beliefs about what might be found. Furthermore, other tools remain available to advance the goals that advocates of preregistration claim it offers. At the very least, preregistration does not provide the only mechanism to advance goals of transparency, replicability, and scientific honesty.

The following discussion begins with a very brief background, and then proceeds to a consideration of some of the main limitations imposed by preregistration, including the disproportionate cost it imposes on junior and less resourced scholars, the ways it discourages exploration and creativity, and the direct and indirect mechanisms by which it diminishes the overall scientific enterprise.

Background

Preregistration is designed largely to address concerns with scientific transparency, reproducibility, and outright fraud (Singal, 2015). Oftentimes, these problems are grouped together under the heading of "questionable research practices" or QRPs (John et al., 2012). These are significant problems that deserve serious attention and institutional attempts to rectify, or at least to diminish them. The problem is that preregistration does little to address these underlying concerns. Rather, it simply serves as a form of virtue signaling, whereby those who participate are understood to be free of such problems, while those who do not are assumed to be somehow less honest or transparent in their work. In other words, preregistration provides a quick and easy heuristic for scholars, reviewers, and editors to rely on to determine credibility rather than having to undertake the harder work of actually determining whether a given paper suffers from the concerns we hope to prevent.

Rubenson (2021) does a fine job laying out the major concerns that pre-analysis plans are designed to correct. He correctly identifies three major issues: HARKing (Hypothesizing After the Results are Known), fishing, and the file drawer problem. He is right that these can be serious problems. However, as with burnout, the problem, and the solution, does not lie with the individual but with the organization—in this case, the journals that prefer to publish statistically significant results. Making scholars compensate for the failings of journals is not only unfair and unjust, it is unlikely to achieve effective change. In other words, the file drawer problem in particular is not a problem that individual scholars should be expected to solve; rather, the proclivity of journals to disproportionately publish particular kinds of studies should be resolved at the level of the field, among journal editors and through professional organizations. Some journals, such as the *Journal of Experimental Political Science*, are careful to point out that they are keen to publish null findings, but few other journals have had the courage to explicitly join that chorus.

The challenge posed by the introduction of pre-analysis plans is that it does not necessarily reduce the other problems raised by Rubenson, particularly those associated with HARKing, for reasons that will be explicated at greater length later. Rubenson (2021) ends his section on what the problem is by advocating the use of pre-analysis plans, stating that without them, "Put plainly, our ability as social scientists to advance knowledge is weakened" (p. 144). While I understand this argument, I contend this is exactly wrong: *preregistration* will hinder our ability to wonder about the world and discover the previously unknown in ways that are scientifically limiting as well as downright sad. This is true for several reasons delineated here.

Preregistration disproportionately harms more junior and less resourced scholars

The notion that preregistration exists on an equal playing field is wrong. This is because the easiest way to get around the handcuffs of preregistration is simply to do a bunch of studies first, as a form of pilot testing, that are not preregistered. As outcomes become known and predictions emerge, scholars can then register those hypotheses and undertake the "legitimated" preregistered form of the study. In extreme forms, this is known as PARKing (Preregistering After the Results are

https://doi.org/10.1017/pls.2022.4 Published online by Cambridge University Press

Known) (Yamada, 2018). While this practice is indeed suspect, although no more so than the HARKing practice outlined by Rubenson (2021), the line between extensive pilot testing and outright fraud can quickly become a slippery slope. Regardless, the ability to run a number of studies prior to engaging in formal registration is much easier for those who have more time, namely senior scholars, and those who have more resources, namely those at more elite and well-endowed universities.

This raises a larger issue with regard to proper experimental practice. The best experimental designs rely on replication to determine scope conditions. One experiment builds on another by expanding populations, changing stimuli or context, or further specifying variables in order to determine the extent of generalizability. Preregistration can restrict this ability by imposing disproportionate burdens on scholars with less time or fewer resources who would have the additional administrative burden of preparing additional documents with each iteration of a study. Anyone who has dealt with Institutional Review Boards knows how cumbersome such demands can be, and how quickly administrative requirements can become more performative than protective in nature. If a scholar finds surprising results, that person may be penalized by not being allowed to publish their data without additional hoops and delays. Indeed, they may be forced to conduct an entirely new study that is preregistered to render their findings legitimate. This not only adds to costs and delays, but literally reduces the diversity of work that a given scholar can undertake over time.

The issue of additional burdens does not simply fall on the scholars. It also falls on editors and reviewers. Who will be in charge of monitoring compliance? Is every editor going to have to go through and make sure every element of the final paper meets the standards specified in the preregistration plan? Alternatively, are editors more likely to outsource that work to reviewers, who already are terribly overworked and overburdened? Will this additional requirement make the already restricted reviewer pool even more reluctant to undertake such unrewarded and burdensome service requests? If not journal editors or reviewers, then what kind of monitoring mechanisms must be put in place to ensure compliance with the dictates associated with preregistration? The concern here includes cases in which tired reviewers or editors reject promising manuscripts out of hand simply because the final manuscript does not properly align with the preregistration plan in some trivial manner; after all, this kind of heuristic saves them time, effort, and work. There are literally no costs to the reviewer for taking this shortcut since the process remains anonymous; however, all the benefits of saved time accrue to the reviewer, who can feel sanctimonious on top of eliminating work. Regardless, such an outcome does greater disservice to more junior scholars, or those with fewer resources, who have less time and ability to simply rerun the study from start to finish to meet some administrative requirement imposed by preregistration.

Even Rubenson (2021) raises this issue when he writes that "preregistration of a PAP [pre-analysis plan] does not guarantee that the design is good, nor does the lack of one suggest the design is bad, by definition. As referees, readers, and editors, we have to always evaluate based on the design—not based on the presence of a PAP" (p. 146). If that is the case, and that seems very reasonable, then it is unclear what value the addition of the PAP imbues, aside from replicating a kind of elitism among scholars under the guise of ensuring a common value like transparency. In other words, high-status scholars with lots of resources can impose additional burdens on those less able to compete, all the while claiming that such costs are in service of a common value that simply ensures the retention of resources and status by those who already have it.

Preregistration crushes creativity

Many scholars begin their careers because they are interested in a particular topic or area and want to know more about it. They wonder what might cause some outcome they do not understand. They key word here is *wonder*. They are not sure; they want to know more. In other words, they are driven by a kind of curiosity. Indeed, a great deal of scientific discovery is driven by serendipity. Lucky scholars will recognize the kind of delight that occurs when their data surprise them and show them something they did not expect to see or find. Data can inform discovery. Not all legitimate forms of science are deductive in nature. That kind of inductive discovery is, and should be, at the heart of good science. For this reason, science should not discourage exploration. Yet that is exactly what preregistration entails and demands. Under the guidelines of preregistration, scholars are expected to know what they will find before they run the study; if they get findings they do not expect, they cannot publish them because the study will not be considered legitimate if it was not preregistered. However, if you already know what your study will show you,

https://doi.org/10.1017/pls.2022.4 Published online by Cambridge University Press

then why do you need to undertake it at all? Where is the wonder of discovery? Of course, most scholars have hypotheses or suspicions about what they expect to see in their data or what they hope will occur in their study, but they might be wrong. More significantly, they might be surprised.

Imposing preregistration standards on scholars effectively reduces the discovery of novel findings. As Pham and Oh (2021) argue,

> Notwithstanding a professed tolerance for exploration, a preregistration regime effectively suppresses exploration in two ways. First, it makes researchers apprehensive to undertake analyses of the data that have not been preregistered or to report any post hoc conceptualization of the observed data, even if such additional analyses and post hoc theorizing would be informative despite their exploratory nature. Second, by placing much higher evidentiary status on confirmatory as opposed to exploratory research, a preregistration regime undermines researchers' incentive to undertake and report exploratory investigations—a type of research that is critical for scientific progress. (p. 169)

In addition, if preregistration is required, an additional burden is imposed if the purpose of science—actual discovery—is achieved, because scholars would have to rerun the study after the answer is already known and the data have already been collected. When flexibility is reduced, creativity is suppressed. There may be some who are willing to make that trade-off, but that seems to undermine the entire value of the scientific enterprise. The wonder of discovery is removed from the entire enterprise, rendering the job much closer to that of a mechanic than an artist.

Additional collateral damage arises when the prospects of interdisciplinary collaboration are introduced. Many fields do not typically require preregistration. This may make the kind of interdisciplinary collaborations that are increasingly common and valued more difficult to sustain in practice, thus inadvertently limiting the scope of inquiry. Such costs incur even within the discipline of political science, by once again privileging the kind of quantitative work that allows for such forms of preregistration. Work that is inductive or interpretive proves much less amenable to such forms of pre-analysis. If such plans are required, this only serves to alienate further those who practice more qualitative forms of methodology or inquiry.

Preregistration hurts the pursuit of good science

Some of the negative externalities that accrue to preregistration are indirect. Specifically, preregistration changes the kind of science that is pursued. Pham and Oh (2021) write about this most eloquently when they note that "sound scientific thinking entails an openness to learn from observed data beyond a mechanical testing of whether the data support some a priori prediction" (p. 170). Furthermore, they argue, "Overall, such reductions in conceptual and empirical exploration, and inflexibility in methodology, while presumably reducing the probability of false-positive results (type-I error), tend to increase the total cost of science and likely increase the probability of genuine findings not being discovered (type II error)" (p. 170).

One of the ways this is likely to occur is that scholars, particularly junior ones under intense time pressure, will come to rely increasingly on canned data sets since it is much easier to know what to expect from that data. While there is nothing inherently wrong with this strategy, it severely restricts the possibility of discovering anything novel in the data, especially when it has been picked over repeatedly by others. In fact, the earlier work is precisely what makes such data sets valuable to those required to preregister; they are assured of the likelihood of findings that others have seen before, but very unlikely to discover something further that is hidden. Such a strategy also disincentivizes the collection of entirely new data sets whose utility may not be clear at the outset.

In many ways, this comes down to what kind of science is valuable to pursue. Should publications mostly devolve into the business of helping scholars obtain secure academic posts at various institutions? Perhaps that is already what we have largely achieved. After all, most manuscripts are rarely, if ever, cited. That is fine as long as our goal revolves primarily around teaching. However, if conducting research actually constitutes an important and valuable aspect of the academic enterprise, then shouldn't we strive for something more than cookbook recipes for publication? Shouldn't we hope to discover interesting, useful, and previously unknown aspects of the world? Or do we prefer to become resigned, as Anne Arvin, vice provost for research at Stanford University, once put it, to know more and more about less and less until we know everything about nothing?

https://doi.org/10.1017/pls.2022.4 Published online by Cambridge University Press

Conclusions

Preregistration may have a place in certain environments, and, of course, it should not be prohibited for those who wish to employ its benefits. However, scholars should not remain under the illusion that preregistration prevents the concerns it ostensibly prevents. Everyone can agree that transparency, reproducibility, and honesty are laudable scientific goals. It would be nice if preregistration could enhance, much less ensure, the accomplishment of those worthy objectives. However, it is not at all clear that it can do so, and pretending it achieves those objectives only further allows QRPs to proliferate. In other words, it helps accomplish exactly the goal it is designed to prevent.

Rather, as Pham and Oh (2021) state, "A major shortcoming of preregistration as a normative standard is that the increased transparency it provides may be more illusory than real" (p. 166). It gives scholars, editors, and reviewers a false sense of confidence that preregistration has solved the problems associated with QRPs as long as scholars follow the plan they ostensibly sent out in advance. However, that plan does not prevent pilot testing that helps develop hypotheses and ensures statistically significant results prior to formal registration. It certainly does not push journal editors to cease their preference for statistically significant results. Nor does it come without significant costs to the broader scientific enterprise, by disproportionately harming more junior and less resourced faculty, discouraging creativity, and biasing the kind of studies that are undertaken in subtle and perhaps even unconscious ways. This matters because it makes our products less interesting and useful to the real world by reducing the scope of our inquiry, restricting our ability to engage in productive interdisciplinary collaboration, and placing an undue and likely unrealized burden on journal editors and reviewers.

In other words, preregistration is more performative than real. As with the Transportation Security Administration, we can invest a huge amount of resources in making people feel secure without actually making them actually safer. Airports may be more secure from terrorists because of the massive investment in security infrastructure, although that is debatable, but it does not, for example, render school shootings less likely, because violence can exist in many forms and take place in many venues. Similarly, we can require an artifice like preregistration designed to ensure values that we all share, such as transparency, reproducibility and honestly, but it will not necessarily accomplish that goal, because other aspects of the discipline still structurally incentivize cheating. Maybe not a lot, but some people will cheat because the pressure is enormous. We can make it harder to cheat, but it is not clear that preregistration can even accomplish that goal; it will simply change the way people cheat, or when in the process they do so. Preregistration will not stop cheating. However, it may restrict the ability of the most creative and ambitious scholars from engaging in the process of exploration and discovery that can, and should, provide the most rewarding aspects of an academic life.

References

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532.

McDermott, R., & Hatemi, P. K. (2020). Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms. *Proceedings of the National Academy of Sciences*, 117(48), 30014–30021.

Pham, M. T., & Oh, T. T. (2021). Preregistration is neither sufficient nor necessary for good science. *Journal of Consumer Psychology*, 31(1), 163–176.

Rubenson, D. (2021). Tie my hands loosely: Pre-analysis plans in political science. *Politics and the Life Sciences*, 40(2), 142–151.

Singal, J. (2015, May 29). The case of the amazing gay-marriage data: How a graduate student reluctantly uncovered a huge scientific fraud. *The Cut*. http://nymag.com/scienceofus/2015/05/how-a-grad-student-uncovered-a-huge-fraud.html

Yamada, Y. (2018). How to crack pre-registration: Toward transparent and open science. *Frontiers in Psychology*, 9, 1831.

https://doi.org/10.1017/pls.2022.4 Published online by Cambridge University Press