Preregistration in the Context of Expertise Research: Benefits, Challenges, and Recommendations

David Moreau
School of Psychology and Center for Brain Research, University of Auckland, New Zealand

Correspondence: David Moreau, d.moreau@auckland.ac.nz

Abstract
A number of recent reforms in psychological science have centered around following best practices to improve the robustness and reliability of empirical findings. Among these, preregistration has become a fundamental component, on the rise in the last few years, yet it remains relatively uncommon in expertise research. In this paper, I point out the numerous benefits of preregistration, drawing on specific examples from the field of expertise. I then examine some of the challenges the field of psychology is currently facing to implement systematic preregistration, including many that are particularly exacerbated in expertise research. Specifically, I discuss widespread design characteristics such as small sample sizes, the lack of consistent definitions regarding what constitutes expert performance, and inherent difficulties in conducting replication studies with rare, elite populations. Finally, I make a number of recommendations to facilitate preregistration in expertise research, including tips to handle and report deviations from original plans, and discuss future directions toward more prevalent open science practices.

Keywords
open science, preregistration, elite performance, skill, abilities

Introduction
Recent reforms in psychological science have focused on best practices with respect to the design, analysis and reporting of empirical studies. Among these, open science practices have been particularly emphasized in order to promote transparency and facilitate reproducibility. A number of tools and incentives have been put forward to promote best practices, including dedicated platforms for preregistration, sharing data and materials (e.g., Open Science Framework), repositories to share manuscripts freely (e.g., PsyArXiv), and journal policies to signal transparency (e.g., open science badges, Kidwell et al., 2016). As a result, open science is on the rise—as of early 2019, about 35% of faculty researchers in psychology are embracing open science practices, compared with only 5% just five years earlier (Nosek, 2019). Importantly, this increase is evident in junior faculty as well as in more senior academics (Nosek, 2019), suggesting that the importance and value of open science practices are recognized across career stages.

Open science practices are typically broken down into three main components: preregistration, open materials, and open data. The importance of sharing materials and data has been highlighted in several contributions focusing on either philosophical or practical aspects (Gilmore, Lorenzo Kennedy, & Adolph,
2018; Houtkoop et al., 2018; Popkin, 2019; Soderberg, 2018). Here, I focus on preregistration, a way to transparently disclose intent ahead of a research project. Defined as “the specification of a research design, hypotheses, and analysis plan prior to observing the outcomes of a study” (Nosek & Lindsay, 2018), preregistration is a key component of open science, as it allows readers to better evaluate the credibility of a set of findings. However, and perhaps more so than open data and materials, preregistration comes with a number of challenges and forces us to rethink the way we do research (Ledgerwood, 2018; Nosek, Ebersole, DeHaven, & Mellor, 2018). In this article, I first outline the benefits of preregistration and advocate for more systematic preregistration in expertise research. I then discuss common challenges in preregistering studies, especially in the context of expertise research. Finally, I propose recommendations to facilitate preregistration in the field of expertise and discuss future directions.

**Benefits of Preregistration**

Preregistration comes with numerous benefits. First, it helps distinguish between exploratory research—research that is data-driven and aims to formulate explanations post hoc—and confirmatory research that uses data to test predefined hypotheses. This point is critical, given that presenting exploratory results as confirmatory can lead to substantial overconfidence in research findings. For example, preregistration prevents undisclosed flexibility in data collection and data analyses, which is detrimental to theory falsifiability (Ledgerwood, 2018) and can lead to a number of questionable research practices, including p-hacking (Head, Holman, Lanfear, Kahn, & Jennions, 2015; Simmons, Nelson, & Simonsohn, 2011) and HARKing (Kerr, 1998).

Preregistration is also key to reduce, and eventually suppress, publication bias. Publication bias arises because certain types of findings are easier to publish in scientific journals than others. Unfortunately, this is not only contingent upon the quality of the experimental design, or the importance of a research question, but also on the statistical significance of research findings. In this context, the available, published literature is often biased, in the sense that it does not provide an objective account of the existing pool of findings. This problem is especially exacerbated in the context of meta-analyses, where it often leads to overestimations of effect sizes, but remains valid when looking at individual studies: With over 90% of published findings being statistically significant, psychology was arguably the worst among all scientific fields pre-crisis (Fanelli, 2012). If all confirmatory analyses were systematically preregistered, one could check public records and get an unbiased view, irrespective of the significance of the findings, thereby providing a powerful remedy to the problem.

In addition, and perhaps counterintuitively, one important benefit of preregistration relates to raising awareness about what is not known or difficult to predict, that is, modeling uncertainty. Planning for what one will do if things do not go according to plan is one of the most important, yet difficult, aspects of a preregistration. Importantly, deviations from the initial plan are often unavoidable, and should not be seen as a weakness of a research project. In a survey of 27 preregistered articles published in Psychological Science from 2015 to 2017, Claesen and colleagues found that all studies deviated from the plan in at least one aspect, yet only one of them reported all deviations (Claesen, Gomes, Tuerlinckx, & Vanpaemel, 2019). Potential deviations should be anticipated if possible—for example, in the form of a decision tree accounting for uncertainty—but, if they cannot be foreseen, it is important to ensure that deviations are thoroughly documented and reported.

Finally, preregistration is extremely helpful in protecting researchers against themselves. Individuals are prone to a wide variety of biases in judgment and decision-making, including confirmation bias (Nickerson, 1998), anchoring effects (Furnham & Boo, 2011), and hindsight bias (Roese & Vohs, 2012). These might be exacerbated in expertise research, because one
can often easily relate to, or identify with, the population they study. For example, a researcher might believe that a specific finding makes sense given what they have themselves experienced in the same or a similar activity. These biases can lead us to feel a false sense of confidence in our findings, whereby we assume that we are not affected by circumstances outside our control or awareness. Preregistration thus also serves as a safeguard against our own biases, as it allows keeping a time-stamped record of our intent and predictions, against which to compare protocol and results after data analyses.

**Challenges of Preregistration**

Preregistration also comes with a number of challenges that can appear daunting at first. Whereas other aspects of open science such as sharing data and materials are often fairly straightforward, when it comes to preregistration, willingness to increase transparency is necessary but might not be enough: One also needs to be able to formulate precise hypotheses at the onset of a research project, and determine the analyses that will allow testing these specific hypotheses, typically before data have been collected. Although in line with common assumptions of the scientific method, rethinking the workflow of a research project in such a way may be at odds with traditional practices in the field of psychology. Regardless, a global shift toward more systematic preregistration is underway, with important consequences for research and academia (Nosek, 2019; Nosek et al., 2018).

Change has perhaps been a little slower in specific subfields of psychology, including expertise research, for a few reasons. First, expertise research sometimes reports on studies with relatively small sample sizes, with a number of important consequences, most notably low statistical power and large sampling error (McAbee, 2018; McAbee & Oswald, 2017). For example, in our meta-analysis on the relationship between deliberate practice and sports performance, the median sample size for studies reporting group comparisons was 29 across groups (Macnamara, Moreau, & Hambrick, 2016). If we assume that the observed effect size for this subset of studies was unbiased—in the sense that it was not inflated by publication bias (this is almost certainly untrue, which means that the situation is likely worse)—statistical power in the field is a mere .15 (for a two-sided group comparison, assuming α = .05). Although this situation is problematic irrespective of the specific area of interest (Fraley & Vazire, 2014; Marszalek, Barber, Kohlhart, & Holmes, 2011) and has been instrumental in the so-called “replication crisis” in psychology, the solution for most of psychology has been rather straightforward, with a push toward larger sample sizes facilitated by large-scale, multi-lab projects, such as Many Labs (Klein et al., 2014, 2018) and the Reproducibility Project: Psychology (Open Science Collaboration, 2015), or via platforms such as the Psychological Science Accelerator (Moshontz et al., 2018) or StudySwap (Chartier, Riegelman, & McCarthy, 2018). In expertise research, the answer might not be as simple: Elite individuals are by definition unique, and this aspect typically leads to difficulties in aggregating individuals with one another within a coherent sample. Because of this challenge, it is also difficult to put forward general recommendations that can hold across topics of expertise research. Nevertheless, expertise research can tremendously benefit from a move toward multi-labs projects that can provide larger, more representative samples and can strengthen the conclusions from this type of studies.

Second, expertise research has suffered from difficulties to replicate study protocols, rather than results, given that many findings rely on rare, elite samples. Not only are those individuals difficult to recruit, it is also unclear how comparable experts in one field or activity are to experts in another (Ackerman, 1988; Fleischman & Mumford, 1989). Oftentimes, studies in expertise research generalize from a given field to another, for example via general conclusive statements about what the studies mean for our understanding of expertise. This implies that skill acquisition is majoritarily uniform across fields, or at the very least that its
underlying mechanisms are comparable across domains. If, however, effects are not expected to replicate across a range of activities, say from violinists to soccer players, then it suggests that the findings of the initial study might hold only for violinists, or perhaps for musicians, but may not hold at a more general level or may not provide fundamental insight about general skill acquisition or elite performance. To complicate things further, there is evidence that expertise within broad domains, such as sports or music, is not a unitary construct (Hodges, Starkes, & MacMahon, 2006), with a number of moderating factors influencing the very nature of expertise (Macnamara, Hambrick, & Oswald, 2014). In the case of sports, these moderating factors might be characteristics of the activity itself, such as whether it involves teammates (Helsen, Starkes, & Hodges, 1998), whether it is internally or externally paced (Singer, 2002), or whether skills are majoritarily open or closed (Wang et al., 2013). In addition, the demands of certain roles or positions on the field might greatly differ from one another (Allard & Burnett, 1985), which in turn can induce substantial differences in the development of specific skills and expertise. Similarly, these differences are often correlated with performance on a range of cognitive ability measures (Moreau, 2013; Moreau, Mansy-Dannay, & Clerc, 2011). Together, these disparities are responsible for one major challenge in expertise research: studying individuals that are similar enough to enable precise and useful predictions, yet diverse enough to extract general mechanisms that can eventually facilitate a broader understanding of expertise across domains. Importantly, the study of expertise does not solely focus on experts or elite performance, but also includes skill acquisition more broadly defined; that is, all steps that are part of the learning process in humans, on the path to expert performance. Other research areas within the field of learning and skill acquisition may not suffer from the same challenges, assuming they focus on more replicable, less atypical, performance.

Finally, one central but perhaps sometimes overlooked challenge in preregistering expertise studies stems from the lack of consistent definitions such as on the topics of deliberate practice (Ericsson, 2016; Macnamara, Hambrick, & Moreau, 2016; Macnamara, Moreau, et al., 2016), sport performance (Piermatté, Lo Monaco, Reymond, Eyraud, & Dany, 2018; Swann, Moran, & Piggott, 2015), or skill acquisition (Baker, Wattie, & Schorer, 2015; Castles, Rastle, & Nation, 2018; Davids, Button, & Bennett, 2008; Treiman, 2018). Defining expertise across domains is no easy feat, but acknowledging that definitions are often lacking is a first step toward promoting theoretical work that can provide a fresh perspective on what defines or constitutes expert performance. Ultimately, work that suffers from a lack of definitions about the population of interest will tend not to replicate, as general characteristics and idiosyncrasies cannot be clearly distinguished in this case. In this context, a lack of successful replications might be grounded in failures to model and define the theoretical tenets of the question of interest. Indeed, a lack of clear predictions is often the sign of weak theoretical frameworks—if specific predictions are not possible, we need to either rethink theoretical frameworks, or explicitly acknowledge that a study is data driven and analyses are exploratory. Although they have proved challenging to implement, remedies are therefore straightforward; they include specifying definitions or lack thereof, recognizing specific limitations of theoretical frameworks, and explicitly labeling analyses that are exploratory.

**Future Directions**

Meaningful change takes time and effort. Preregistration, both generally speaking as well as in the context of expertise research, remains work in progress, with many challenges still lying ahead (Nosek et al., 2018). Perhaps more than ever, however, expertise research now has a comprehensive set of tools available to tackle these challenges with well-integrated online platforms to share, collaborate, and improve inferences.

Ultimately, preregistration is bound to become the norm, and journal badges may then
become redundant. Assuming preregistration is ubiquitous, a researcher simply needs to state which section of the analyses is confirmatory, and which section is exploratory. There can be no ambiguity in this scenario, because what is not preregistered is de facto exploratory. Many research articles already follow this template for clear, easy-to-navigate results. Badges, especially in their current, dichotomized implementation, will likely be superseded by more fine-grained ways to differentiate between different scenarios: Some analyses, or parts of the design, might have been preregistered, yet others perhaps changed from conception to implementation. This is relatively difficult to capture in the current system, in a dynamic that tends to ostracize exploratory analyses and to reward preregistration irrespective of its extent within a study or paper. Similar to the publishing landscape, for which the traditional distinction between published and non-published findings has become obsolete with the rise of a large number of outlets including low quality or even predatory journals (Bartholomew, 2014), badges might not capture the subtleties of open practices, especially in the context of preregistration. They were necessary to instigate the open science trend (Kidwell et al., 2016), but perhaps time has come to move on to non-binary displays of open science practices.

In conjunction with the democratization of Bayesian statistics in recent years (Andrews & Baguley, 2013), preregistration has the potential to become more empirically informed and theoretically committed, based on previous findings and one’s interpretation of them. This dynamic could also help alleviate many of the alleged weaknesses of Bayesian analyses, most notably with respect to the subjective aspect of priors (Berger, 2006), given that with preregistration priors can be clearly defined, in a transparent manner, before the analyses. Preregistration itself is a Bayesian process (Dienes, 2016), in which one declares intent and predictions beforehand, and how these will be combined with new information, the data, to generate more comprehensive, informed knowledge. This is an ideal that may not yet illustrate accurately how researchers typically use preregistration—for example with respect to common descriptions of research hypotheses via imprecise, verbal statements rather than mathematical formulations—but it represents a goal to strive for and one in which expertise research could potentially take the lead.

**Conclusion**

There has been a push for more transparency in research practices in psychology recently. Despite the specific challenges it faces, expertise research has the potential to benefit tremendously from open practices as well. In particular, preregistration has an important role to play to improve the reliability of empirical research investigating skill acquisition and expert performance. Even in instances in which open science is not an intrinsic priority, embracing open practices is beneficial to researchers at all stages of their career, helping increase visibility, impact, and citations (McKiernan et al., 2016). In this context, preregistration benefits science, scientists, and consumers all at once, in a dynamic enabling a greater overall impact on society.

**Acknowledgements**

David Moreau is supported by a Marsden grant from the Royal Society of New Zealand, funding from the Neurological Foundation of New Zealand, the Ministry of Science and Technology (MST) in Taiwan, a University of Auckland Early Career Research Excellence Award (ECREA) and the Faculty Research and Development Fund (FRDF).

**Author’s Declarations**

The author declares that there are no personal or financial conflicts of interest regarding the research in this article.

**References**


Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L.


preliminary-report/


Received: 23 July 2019
Revision received: 16 October 2019
Accepted: 18 October 2019