**Preregistration in ‘Open Mind’ Articles**

By:

Annika Oldach

[aoldach@uni-osnabrueck.de](mailto:aoldach@uni-osnabrueck.de)

University of Osnabrück

Toward Open Cognitive Sciences

Dr. phil. Timo Röttger

Submission Date:

16th October 2020

# Introduction

Recently, *Open Science* has begun to gain traction within the scientific community, and awareness of its importance and necessity has spread. As a whole, Open Science aims to make research more accessible, reproducible, transparent, and thus, more representative of reality.

If science is not open, then it is likely that the literature is biased. Usually biases in this area occur so that effects are found to be larger than they actually are, sometimes so much larger, that an insignificant effect becomes significant. It is especially problematic when meta-analyses aggregate results from many studies, as errors accumulate, leading to an incorrect representation of effects within the population (Scheel et al., 2020).

One aspect of Open Science is *Preregistration* and refers to the registration of studies on an online database, prior to the collection or viewing of data. This is primarily what this study is about. Another aspect is called *Open Data*, and it requires researchers to make their data and codebook available to other researchers and scientists. *Open Materials* and *Open Code* refer to the public posting of anything used for the study, and the public posting of the code used to conduct the relevant analyses, respectively. *Open Access* articles are free and available to the public, making them more accessible. Finally, *Registered Reports* is a relatively new format for journal articles, which provides a strict set of guidelines to which authors must adhere when preregistering their study (Standen, 2019). We will dive more into the topic of Registered Reports in following sections.

In this study, we investigate the percentage of preregistered studies in a relatively young journal: *Open Mind (MIT Press Journals, 2020)*. Open Mind is a journal that publishes articles related to discoveries in Cognitive Science and was founded in 2017, thus, it was born into the midst of the Open Science debate. As it only contains recent articles, we hypothesised that over 75% of included papers would have been preregistered. For our study, we looked at all articles published by the journal, so as to get a definitive and accurate result. This was well within the scope of such a course assignment, as the total corpus contains only 31 published articles.

# Method

## Journal

Open Mind is an Open Access journal, meaning that all published articles are free to the public. Furthermore, Open Mind engages in “blind” reviewing, which means that author’s identities are not revealed to the reviewers, in order to avoid potential biases towards the age or gender of the article’s author (Aslin, 2017). Particularly relevant to this study, is Open Mind’s statement on increasing the reliability and replicability of published findings. The editorial board agrees that there is a bias for positive results, and thus, also an increased probability of false positives being published, but they do not feel that every published article should have to be preregistered. They believe that “new and unexpected findings […] would never be discovered by the pre-registration model.” (Aslin, 2017). The editorial board also encourages Open Data but does not set those guidelines as a prerequisite for publishing.

## Corpus

After having decided on the journal and the corpus, all articles were downloaded onto the author’s hard drive (23.9.2020). From here, each article received a random number to be used for identification purposes. All articles published by Open Mind were included in the study, with the exception of the editorial in the first volume and a memorial article in volume three. This left us with a total of 31 articles, thus, meeting the assignment requirement of at least 30 articles.

Subsequently, we ran a word-search on each of the downloaded pdf files (using the Ctrl+F function) for the terms “prereg”, “pre-reg”, and “register”. With decided on using these terms, so that we would not only find preregistered studies, but also Registered Reports.

## Versions

All software versions used throughout this study were documented, so as to increase the level of reproducibility and replicability. Articles included in this study were published by Open Mind between 2017 and 23.9.2020. *Version 1.3.1073* of RStudio was used to conduct the analyses, and data was stored in Excel csv files using Version *Microsoft Excel for Microsoft 365 MSO (16.0.13127.20266) 32-bit*.

## Online Registrations

All changes throughout the study were committed to a [GitHub repository](https://github.com/aoldach/WrittenAss), and all files and details on the study were registered on the [Open Science Framework](https://osf.io/qhtdf/?view_only=6912fa3269324a1b9dcb333d47cb81a9).

## Relevance

The purpose of science is to estimate real effects within the population. However, this is only possible, if the scientific literature is unbiased. Cohen recommended two ways to assess the size of an effect, namely the *Comparison Approach* and the *Conventions Approach* (Cohen, 1988). The former relies on comparing one’s own result to results from other studies within the same field, while the latter recommends scientists to make use of a set of benchmarks, according to which the effect is either small, medium, or large (Cohen, 1988). Unfortunately, due to publication bias and the use of questionable research practices (QRPs), the requirements for these two approaches are likely not given (Schäfer & Schwarz, 2019).

There are many kinds of QRPs, though two of the most common ones are referred to as *HARKing* and *p-hacking*. HARKing is an acronym for “Hypothesising After Results are Known”, meaning that researchers look at the results and then slightly modify or even completely alter their hypothesis, so that they seemingly predicted the results all along. P-hacking can be done in many ways, but usually it refers to the selection of statistical analyses and tests, so that the study result is significant, i.e. p < 0.05, when in actual fact those significant results may just be a fluke for those hand-selected tests (Nosek et al., 2018).

Preregistration can quite successfully prevent these kinds of QRPs, as researchers submit a detailed analysis plan prior to data collection. If they do engage in QRPs, then reviewers can refer to the protocol and see where the actual and the planned analyses differ (Center for Open Science).

Publication bias consists of two issues: *reviewer bias*, and *file-drawering*. The former refers to the fact that reviewers want to publish papers that have interesting, positive, and large effects. This means that null-result articles or replication studies are rarely published. The latter aspect is a natural consequence of the first. Researchers know that null-result studies are difficult to get published, therefore, they keep them tucked away in a file-draw, and instead work on studies in which they hope for the interesting, positive, large effects that reviewers will want to publish (Scheel et al., 2020).

This is not fixed as easily. Here, we need the aforementioned Registered Reports. Like preregistered studies, researchers must submit a detailed protocol before data collection. However, this is where the procedure differs from a simple preregistration. Reviewers review the submitted protocol, and if they accept it, they promise the author to publish their paper, irrespective of the results. This fixes the issues of Reviewer Bias and File-Drawering, as the reviewers are blind to the data, and the authors will not keep the manuscript in their file-drawers, as they have already been promised an article in the respective journal (Center for Open Science).

## Analysis

All analyses were conducted using R in RStudio. The years that preregistered articles were published in Open Mind were collected in one vector, and the years that articles without preregistration were published in Open Mind were added to another vector. We then looked at the minimum and maximum values of both vectors and compared these. Subsequently, we looked at how many months younger the preregistered papers were on average, than the articles that were not preregistered. All scripts can be found in the GitHub and OSF repositories.

## Results

The results of the study do not support our hypothesis that over 75% of studies were preregistered. We found that merely 6.45% of investigated studies were preregistered, none of which were Registered Reports. To find out how the articles belonging to the two different categories differed from each other, we compared several statistics.

The oldest overall paper published in Open Mind was not preregistered and was published in 2016. In contrast, the oldest preregistered paper in Open Mind was published in 2018. As preregistration is still a relatively young idea, such a pattern would be expected. However, when looking at the most recently published Open Mind article that was not preregistered, we found that the year of publication was 2020, one year later than the newest preregistered paper, which was published in 2019. Our intuition that explains the former, does not explain this latter finding.

Having said that, the average years of publication for both categories does support the idea that, on average, preregistered articles are younger than articles that were not preregistered. The average year of publication of preregistered articles was 2018.5, while the average year of articles that were not preregistered was around 2017.8. This means that preregistered papers published in Open Mind are ca. 8.5 months younger than unregistered articles, on average.

When looking at these results, one must be aware that the preregistered articles numbered only 2, while 29 articles were unregistered (see **Figure 1**). Thus, the representativity and meaningfulness of such findings is questionable. One can say that there is a trend towards preregistered articles being more recently published than unregistered articles, but drawing definitive conclusions from such small sample groups is difficult.

Chart

Description automatically generated

**Figure 1** ID numbers of all analysed articles on the x-axis, Boolean of preregistration status on y-axis.

# Conclusion

In our study we found that surprisingly few articles published in the relatively young scientific journal Open Mind are preregistered. Contrary to our hypothesis, less than 75% of articles were preregistered, in fact not even 6.5% of articles were preregistered.

Being cautious of the small sample sizes, we argue that there may be a trend towards preregistered articles being younger than unregistered articles, but that more extensive research must be done to answer this question definitively.

References

Aslin, R. N. (2017). Editorial. *Open Mind*, *1*(1), 1–3. https://doi.org/10.1162/opmi\_e\_00007

Center for Open Science. *Preregistration*. https://www.cos.io/initiatives/prereg

Center for Open Science. *Registered Reports*. https://www.cos.io/initiatives/registered-reports

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

MIT Press Journals. (2020). *MIT Press Journals*. https://www.mitpressjournals.org/loi/opmi

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. *Frontiers in Psychology*, *10*, 813. https://doi.org/10.3389/fpsyg.2019.00813

Scheel, A. M., Schijen, M., & Lakens, D. (2020). *An excess of positive results: Comparing the standard Psychology literature with Registered Reports.* https://doi.org/10.31234/osf.io/p6e9c

Standen, E. (2019). *Open science, pre-registration and striving for better research practices*. https://www.apa.org/science/about/psa/2019/11/better-research-practices