Psychological Research on Urban Society

Volume 3
Number 1 *Vol. 3 No. 1 (2020): April 2020*

Article 8

4-30-2020

A Brief Reflection on Open Science

Henry Otgaar
Faculty of Psychology and Neuroscience, Maastricht University & Faculty of Law, KU Leuven,
henry.otgaar@maastrichtuniversity.nl

Recommended Citation

Otgaar, Henry (2020) "A Brief Reflection on Open Science," *Psychological Research on Urban Society*: Vol. 3: No. 1, Article 8.
DOI: 10.7454/proust.v3i1.85
Available at: https://scholarhub.ui.ac.id/proust/vol3/iss1/8

e-ISSN 2615-8582
p-ISSN 2620-3960

EDITORIAL NOTE

A Brief Reflection on Open Science

Psychological Research on Urban Society
2020, Vol. 3(1): 1-5

DOI: 10.7454/proust.v3i1.85
proust.ui.ac.id

Henry Otgaar[1,2]*, Anna Sagana[1], and Nina Tupper[1]

[1]Faculty of Psychology and Neuroscience, Maastricht University
[2]Faculty of Law, KU Leuven

A Brief Reflection on Open Science

In the past years, the psychological community witnessed an outbreak of failures to replicate well-known psychological findings. These failures have undermined the credibility of the psychological field and have shaken the trust of practicioners and the public to psychological reseach. Some have called this period the replication crisis (e.g., Maxwell et al., 2015). Others have termed this period the credibility revolution because these failures have caused psychological scientists to critically evaluate their used methodology and statistical apparatus and search for ways to improve them (e.g., Vazire, 2018). In this reflection, we give a short historical overview of important events that likely caused this credibility revolution. Furthermore, we describe a possible solution to address these failures to replicate psychological findings: Open Science. We discuss the basic idea behind Open Science and provide several recommendations for psychological scientists when designing and conducting psychological studies.

A Short History

Ioannidis (2005) was perhaps one of the first to explicitly mention that most published findings are false. He argued that research results are less likely to be true when, for example, there is much flexibility in designing and conducting studies, and analyzing data. His argument was that many published studies contain such flexibility, thereby potentially containing false positive findings (i.e., incorrect rejection of the null hypothesis). However, Bem's (2011) study is likely to be one of the most important events that catalyzed the credibility revolution.

Bem (2011) conducted nine experiments to examine the phenomenon of precognition, or the ability to anticipate what will happen in the future. In his paper, he claimed to have found convincing evidence that precognition exists. In some of his experiments, participants were presented with two pictures of curtains. They had to "feel" which curtain picture had another picture (e.g., erotic picture; Experiment 1) behind it. Behind the other curtain picture, no picture was present. What Bem found was that participants performed statistically significantly above chance level. Bem reasoned that these results could only be explained by phenomena such as precognition. Following Bem's (2011) paper, research groups attempted, but failed, to replicate the results of Bem (e.g., Ritchie et al., 2012; Robinson, 2011). When the failed replications were published, critical inquiry shifted to determining the likely cause of Bem's findings showing purported evidence of precognition. A probable candidate for Bem's findings was the flexibility he might have had in designing, conducting, analyzing, reporting, and interpreting his experiments.

Concrete evidence that this flexibility might lead to false positive findings was provided by Simmons et al. (2011). They argued that researchers, when conducting experiments, must oftentimes make many different decisions while data collection is ongoing or after it is complete. For example, researchers must frequently decide whether more data should be collected, whether certain observations should be disregarded, and/or whether irregular data should be transformed. Simmons and colleagues called this flexibility *researcher degrees of freedom*. In their experiments, they showed that when such

Corresponding Author:
Henry otgaar
Faculty of Psychology and Neuroscience, section Forensic Psychology, Maastricht University.
Email: Henry.Otgaar@maastrichtuniversity.nl

degrees of freedom are used, but not disclosed, false positive findings might emerge. Specifically, in two experiments, they examined whether listening to certain songs could change participants' age both subjectively (Experiment 1: *how old do you feel right now?*) and objectively (Experiment 2: *what is your birth date?*). Of course, for Experiment 2, no statistically significant effect should be demonstrated. However, Simmons et al. (2011) showed that in both experiments, certain songs could indeed change participants' age. These effects became statistically significant by performing a multitude of different analyses (e.g., combining conditions, controlling for gender, adding more observations). Because many published studies do not disclose all decisions they might have made before, during, or after conducting an experiment, it might well be the case that the psychological literature is riddled with false positive results.

To estimate the replicability of the psychological field, a large scale multi-site replication study was conducted (Open Science Collaboration, 2015). Specifically, 100 experimental and correlation studies published in three psychology journals were replicated using the original materials and high powered designs. Although 97% of the original studies showed statistically significant results, only 36% of the replications had statistically significant results. In other words, 64% of original findings could not be replicated. These failures to replicate studies published in the psychological field fuel the suspicion that the published psychological literature likely contains many false positive findings. Consequently, researchers began looking for solutions to the replicability crisis. The approach that gained the most momentum was Open Science practices.

Open Science

Open Science refers to the objective to increase transparency in the scientific community (e.g., McKiernan et al., 2016). As such, open science is not a set of specific rules, but a collection of practices that aim to increase rigor, reproducibility, and openness in science (Crüwell et al., 2019). Some of these policies include public access to the scientific publications, data sharing, and transparency in methods and techniques (McKiernan et al., 2016). For example, many researchers are -perhaps unintentionally- not fully transparent on the gamut of statistical analyses that they have conducted in a single experiment. If researchers only report the analyses showing statistically significant results and leave out the non-significant ones, it might well be the case that they are reporting false positive findings or overstating interpretations of their findings. Also, researchers do not always make their data and materials available for other researchers. This is problematic as it prevents the reproducibility of results and prevents successfully conducting replications. One important way to increase transparency in research is by preregistering studies.

When using preregistration, researchers specify which hypotheses they have and which analyses they are planning to perform before data collection (Nosek et al., 2018). One of the advantages of preregistration is that it can protect against several questionable research practices. For example, one questionable research practice is called HARKing (Hypothesizing After Results are Known) indicating that after analyses are conducted, researchers change their hypothesis in line with what they find in their data. Another questionable research practice is p -hacking, referring to, for example, unplanned statistical practices implemented to achieve statistical significance. It is important to emphasize that many questionable research practices are not intentionally performed, but might be the result of cognitive biases that researchers have during the execution of a study (e.g., confirmation bias). Preregistration can shield against such biases by clarifying which analyses were planned (confirmatory analyses) and which included later (exploratory). One might counterargue that preregistration prevents the use of exploratory analyses, but that is not the case. Preregistration, simply put, is a plan; during the execution of a study, researchers might choose to deviate from this plan. Because preregistration is a time stamped research plan, it permits other researchers to understand which decisions were made at which points during the execution of a study.

A variant of preregistration is Registered Reports. This is a form of preregistration that is submitted to journals before data collection. This preregistration will be sent to external reviewers who will evaluate the merits of the preregistra-

tion. If the preregistration is accepted, data collection can commence. The incentive of using Registered Reports is that, provided that the researcher follows their accepted preregistration, the data is accepted for publication. The net effect of such practice is that the data will be published even when no statistically significant effects are observed. Such practice might reduce publication bias, in which oftentimes only statistically significant results are published.

Recommendations and Concluding Remark

We propose several recommendations to improve psychological studies in the future. What these recommendations have in common is that they are mostly Open Science practices.

Preregistration

When designing new studies, we believe that preregistering these new studies might be a promising way to possible limit unintentional bias and decrease the use of questionable research practices. Although preregistration is not without limitations (e.g., Szollosi et al., 2019), the net effect is positive. Preventing questionable research practices might lead to better science and decrease the amount of false positives.

Open Materials and Data

To be fully transparent, we encourage psychological scientists to make their materials and data publicly available when possible. Doing so can increase confidence in the published work. More importantly, sharing the original materials and data will facilitate other researchers to conduct replications on published experiments. The Open Science Framework (https://osf.io) can, for example, be used as a platform for open materials and data practices. Also, there are several journals that award Open Science badges to promote open science practice. Such badges are provided when researchers demonstrate open science practices such as using pregistration and making their materials and data publicly available (Kidwell et al., 2016).

Replications

Conducting replications has often been regarded as a practice that was not novel and hence, should not be given priority. However, considering the fact that many psychological studies fail to replicate, it is vital that researchers conduct high powered replication studies (for a discussion, see Maxwell et al., 2015). Hence, executing replications should be given high value as it is the only way to examine whether certain effects are reliable. It is relevant to distinguish here between conceptual and direct replications. When conceptual replications are conducted, the original study is not exactly replicated in that, for example, other stimuli are used or different instructions are given. Direct replications attempt to exactly duplicate the original study and our recommendation is that direct replications should be favoured before resorting to conceptual replications.

Sample Size Justification

One of the reasons to conduct replication studies is that many original studies had low sample sizes and were therefore underpowered. To improve psychological experimentation, it is vital that researchers should justify their chosen sample size by, for example, resorting to a priori power analyses. We think that it is relevant that future studies in psychology should ideally be highly powered, thereby reducing the chance of false positive results.

Attention to Theory

Weak or absent theories can increase researcher degrees of freedom. Without a strict theoretical framework, researchers' decisions in the design, analysis, and interpretation of the data, may be driven- intentionnaly or unintentionally – by the quest for statistical significance. This is because significance often leads to publication. Well-formulated and internaly consistent theories are important because they provide the foundation for testable hypotheses. Therefore, stong theories set the boundary conditions for the analysis and interpretation of the data. This is not to say that we oppose exploratory analyses. However, we think that findings coming from expoloratory analyses should be treated with caution and put to the test anew, rather than accepted at face value.

To conclude, psychology is undergoing a

reform in the way research can be improved. We believe that this reform is needed to increase confidence in psychological science. Our take is that Open Science is an important reform that can radically improve psychological science. Of course, there might be good reasons for not completely adopting Open Science practices. For example, there might be reasons to not make data publicly available as the data concerns a sensitive topic (e.g., data on victims of sexual abuse). Equally important, apart from adopting such practices for individual researchers, scientific journals have a responsibility as well. Too often, journals almost exclusively publish significant results or novel research. It is therefore vital that journals start embracing Open Science practices as well. One promising candidate to counter the abundance of significant results is that journals should accept Registered Reports as part of their submissions. It is important to emphasize here that journals are increasingly adopting Open Science practices. For example, psychological associations such as the *Association for Psychological Science* and the American *Psychological Association* highly recommend the use of Open Science practices for their journals[1]. Also, an increasing number of journals are embracing Open Science practices[2]. Lastly, universities, as the occupational home of psychological researchers, could move to support Open Science practices as well. For example, direct replications could be built into educational curricula to train the next generation of researchers. In short, we argue that Open Science is the way forward for psychological science and that adopting it will benefit psychology and its societal impact.

Recommended Sources

The Open Science Framework: https://osf.io
Open Science Principles: https://opennessinitiative.org
Sample size calculation: http://www.psychologie.hhu.de/arbeitsgruppen/allgemeine psychologieundarbeitspsychologie/gpower.html or https://psyarxiv.com/baxsf/
Preregistration: https://osf.io or https://aspredicted.org

References

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. Journal of Personality and Social Psychology, 100(3), 407–425. https://doi.org/10.1037/a0021524

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., Orben, A., Parsons, S., & Schulte-Mecklenbeck, M. (2019). Seven Easy Steps to Open Science. Zeitschrift Für Psychologie, 227(4), 237–248. https://doi.org/10.1027/2151-2604/a000387

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. PLoS Medicine, 2(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T. M., Fiedler, S., & Nosek, B. A. (2016). Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. PLOS Biology, 14(5), e1002456. https://doi.org/10.1371/journal.pbio.1002456

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? American Psychologist, 70(6), 487–498. https://doi.org/10.1037/a0039400

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., & Yarkoni, T. (2016). How open science helps researchers succeed. ELife, 5, e16800. https://doi.org/10.7554/eLife.16800

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

Open Science Collaboration. (2015). Estimating the reproducibility of science. Science, 349 (6251), aac4716–aac4716. https://doi.org/1

[1]See for example: https://www.psychologicalscience.org/publications/open-science and https://www.apa.org/pubs/journals/resources/open-science

[2]https://www.topfactor.org

0.1126/science.aac4716
Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect. PLoS ONE, 7(3), e33423. https://doi.org/10.1371/journal.pone.0033423
Robinson, E. (2011). Not feeling the future: A failed replication of retroactive facilitation of memory recall. Journal of the Society for Psychical Research, 75, 142–147.
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science, 22(11), 1359–1366. https://doi.org/10.1177/0956797611417632
Szollosi, A., Kellen, D., Navarro, D. J., Shiffrin, R., van Rooij, I., Van Zandt, T., & Donkin, C. (2019). Is Preregistration Worthwhile? Trends in Cognitive Sciences, 24(2), 94–95. https://doi.org/10.1016/j.tics.2019.11.009
Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress. Perspectives on Psychological Science, 13(4), 411–417. https://doi.org/10.1177/1745691617751884