

Open Psychology: transparency and reproducibility

Psicología abierta: transparencia y reproducibilidad

Psicologia Aberta: transparência e reprodutibilidade

Maarten Derksen
ORCID ID: 0000-0003-1572-4709
University of Groningen, Países Bajos

Autor referente: m.derksen@rug.nl

Historia editorial
Recibido: 28/02/2019
Aceptado: 19/09/2019

ABSTRACT

I describe the attempt by a group of psychologists to reform the discipline into an Open Science (which I will call 'Open Psychology'). I will first argue that their particular version of Open Science reflects the problems that gave rise to it and that it tries to solve. Then I will describe the infrastructure that this group of people is putting in place to facilitate transparency. An important function of this infrastructure is to restrict what are called 'researcher degrees of freedom'. In Psychology, transparency is as much about closing down as it is about opening up. I will then focus on the flagship project of Open Psychology, the Reproducibility

Project. According to the Open Psychologists, the neglect of replication is at the core of Psychology's current problems, and their online infrastructure offers the perfect framework to facilitate replication and give it a place in the field's research process. But replication, I will argue, is not just an epistemological, methodological issue: it implies a particular ontology and tries to enact it. The Reproducibility Project, and Open Psychology generally, can be considered as social experiments, that attempt not only to reform Psychology, but also to perform a new psychological object.

Key words: Open science; replication; crisis; performativity

RESUMEN

En este artículo describo el intento de un grupo de psicólogos de reformar la disciplina en una Ciencia Abierta (la cual llamo 'Psicología Abierta'). Primero, argumentaré que su versión particular de Ciencia Abierta refleja los problemas que le dieron origen y los cuales ella intenta resolver. A continuación, describiré la infraestructura que este grupo de personas está implementando para facilitar esta transparencia. Una función importante de esta infraestructura es restringir los llamados 'grados de libertad del investigador'. En Psicología, la transparencia buscada se refiere tanto a delimitar cuanto sobre expandir las investigaciones. Posteriormente, me centraré en el proyecto emblemático de la Psicología

Abierta, el Proyecto de Reproducibilidad. De acuerdo con los Psicólogos Abiertos, el descuido acerca de la replicación está en el corazón de los problemas actuales de la Psicología, y su infraestructura en línea ofrece el soporte perfecto para facilitar la replicación y darle un lugar en el proceso de investigación de campo. Pero la replicación, argumentaré, no es sólo una cuestión epistemológica, metodológica: implica una ontología particular e intenta ponerla en acción. El Proyecto de Reproducibilidad, y la Psicología Abierta de forma general, pueden ser considerados experimentos sociales, cuyo esfuerzo reside no sólo en reformar la Psicología, sino también en realizar/performar un nuevo objeto.

Palabras clave: Ciencia abierta; replicación; crisis; performatividad

RESUMO

Descrevo a tentativa de um grupo de psicólogos de reformar a disciplina em uma Ciência Aberta (a qual eu nomeio 'Psicologia Aberta'). Primeiro, argumentarei que sua versão particular de Ciência Aberta reflete os problemas que lhe deram origem e os quais ela tenta resolver. Em seguida, irei descrever a infraestrutura que esse grupo de pessoas está implementando para facilitar a transparência. Uma função importante dessa infraestrutura é restringir os chamados 'graus de liberdade do pesquisador'. Em Psicologia, transparência é tanto sobre delimitar quanto sobre expandir. Posteriormente, focarei no projeto emblemático da Psicologia Aberta, o

Projeto de Reprodutibilidade. De acordo com os Psicólogos Abertos, a negligência acerca da replicação está no cerne dos problemas atuais da Psicologia, e sua infraestrutura online oferece o suporte perfeito para facilitar a replicação e dar a ela um lugar no processo de pesquisa de campo. Mas a replicação, argumentarei, não é apenas uma questão epistemológica, metodológica: implica uma ontologia particular e tenta operá-la. O Projeto de Reprodutibilidade, e a Psicologia Aberta de forma geral, podem ser considerados experimentos sociais, cujo esforço reside não apenas em reformar a Psicologia, mas também em performar um novo objeto psicológico

Palavras-chave: Ciência aberta; replicação; crise; performatividade

Crisis

It is a common argument in Science and Technology Studies that there is, in fact, no such thing as *science*. Science is multiple and diverse, rather than unified and homogeneous. Similarly, it may be better to speak of *Open Sciences*, in the plural, than of *Open Science*. According to their particular needs and circumstances, different scientific fields emphasise different tools and practices associated with Open Science. For example, in synthetic biology the focus is on databases for sharing information about the building blocks of bio-engineering. Because it is a field where academic interests meet those of the biotechnology industry, a *diverse ecology* of the open and the proprietary appears to be taking shape (Calvert, 2012; Kelty, 2012). The field of High Energy Physics on the other hand is a homogeneous community, making open access pre-publication (via arXiv) a viable solution to the demand for faster circulation of research results (Gunnarsdóttir, 2005). In Psychology, the Open Science initiative is to a large extent a reaction to a crisis. Fundamental problems in the field have come to light, psychologists are up in arms about them (or dispute the fact that there are any serious problems), and Open Science is put forward as the solution (or rejected).

The trouble emerged in a remarkable series of events between March 2011 and March 2012 – Psychology's *annushorribilis*. The most spectacular case was the fraud of the Dutch social psychologist Diederik Stapel, that came to light in September 2011. Two more cases of fraud, less spectacular but still embarrassing, were discovered a few months later. Earlier that year, there had been the publication, in a top social psychology journal, of a paper claiming evidence for *anomalous retroactive influences on cognition and affect*: precognition (Bem, 2011). The furor over that article was compounded soon after when the same journal refused to publish the report of a failed attempt to reproduce those results, on the grounds that it never publishes replication studies. Then, at the end of the year, two papers were published about so-called questionable research practices (QRP's) – for example: modifying your hypothesis after

you've seen the results of your experiment. The first paper showed that by fully exploiting the flexibility such QRP's offer, a researcher can produce the absurd result that people get on average one-and-a-half years younger by listening to The Beatles *When I'm Sixty-Four* (Simmons, Nelson & Simonsohn, 2011). The second paper was a report of a survey among academic psychologists that aimed to measure the prevalence of QRP's. The results were worrying: for example, 35% admitted to having reported an unexpected finding as having been predicted from the start (John, Loewenstein & Prelec, 2012). Finally, in March 2012 a row broke out when a group of researchers managed to get a non-replication published of one of the most famous experiments in recent social psychology, the *elderly walking study* (Doyen, Klein, Pichon & Cleeremans, 2012). Suddenly, a flourishing field of study in social psychology was in the spotlight, and people started to wonder whether its often spectacular results were solid, or due to QRP's.

By this time, the word *crisis* was starting to be used, and the contours of a community of critics began to emerge. The trouble in psychology according to these critics can be briefly summarised as follows: under pressure to produce a high *output* of articles in top journals, and faced with the requirement of those top journals that articles present an eye-catching story backed by clean results, researchers sometimes engage in questionable research practices (or worse) to manufacture just such results. Increasingly, journals are filled with spectacular but weak research (frivolous fluff rather than solid results). Because journals do not publish replications, researchers have very little incentive to check each other's work. Additionally, since negative results (experiments that 'do not work') are equally unpublishable, the field suffers from publication bias and an unwarranted confidence in its findings. As a result, the literature has become, in the words of two critics, 'a vast graveyard of undead theories' (the title of Ferguson & Heene, 2012).

The tools of transparency

To clean up the mess and prevent further disasters, the critics advocate transparency. Some have proposed that journals should demand that authors disclose exactly how data were collected and analyzed (Simmons, Nelson & Simonsohn, 2012). Others have sung the praises of data sharing (Stroebe, Postmes & Spears, 2012). The most comprehensive solution, and the one that has been most successful, is the Open Science Framework (OSF). The OSF started off in November 2011 as a Google Group devoted to discussing Psychology's problems and various open science solutions to them. Now, bolstered by large grants from a.o. the Laura and John Arnold Foundation, the activities are coordinated by a Center for Open Science (at the University of Virginia), and the Framework itself (a web application) is online.

In the OSF, after opening an account (user name, password), researchers can organise their work in *projects*. They can upload files to the project (such as stimuli used in the experiment, or data sets); they can add a wiki describing the project, its history and its current status (for example); they can check their statistics and see how many page views the project has had. In other words, researchers can move their whole workflow online, including the experiment if one uses for example Amazon's MTurk service of online workers. Most importantly, of course, each project can be made public, making the research process transparent for whoever cares to look.

The convergence between open source and open science that Willinsky (2005) has noted is clear in the OSF. The Framework is built in Python; it uses Git as a version control system, and offers users the option to add Github as a data and code repository. It has also borrowed the concept of *forking* from open source culture: any user of the OSF can fork any public project and thus create a clone of it in her own account, with a link to the original embedded in it. This can be used for example to conduct a replication, or extend someone else's study in a new direction. Furthermore, the Center for Open Science aims to support the community at the intersection of open

source and open science. However, although criticism of the incentive structure in academia is often ventilated, what is lacking so far is a political arm to the movement: there is as yet no counterpart to the Free Software Foundation, for example, no Richard Stallman of Open Psychology.

Closing down by opening up

The explicitly stated goal of the Center for Open Science is to align the values that scientists hold dear and the practices that they actually engage in. According to the diagnosis of the Open Psychologists, the current heavy emphasis on *output* in academia has made *getting it published* more important than *getting it right* (Nosek, Spies & Motyl, 2012). Despite their good intentions, scientists are drawn to biased reasoning and exploiting the loopholes in the system to produce publishable, but not necessarily accurate results. Thus, Open Psychologists work with a naturalized epistemology which views scientists as biased reasoners (Flis, 2018).

To counter these problems, Open Psychology offers both a carrot and a stick. Transparency is made attractive with alternative incentives. The Center for Open Science for example maintains a system of Badges for Open Practice: if the study that is reported in an article is based on open data, the journal publishing it can award it the Open Data Badge – a bit like the MSC-label for certified sustainable seafood. Thus, transparency itself is a reward for good conduct: now everyone can actually see your good behaviour.

But transparency is also a stick that keeps researchers on the straight and narrow by making their work visible. In the OSF, transparency is as much about constraint and coveillance as it is about freedom and creativity – another example of the fact that, as Chris Kelty (2012) noted, the distinction between *open* and *closed* science does not work. The OSF for example offers the option to *register* a project component: this produces a time-stamped copy of that file and saves it for later reference. The main

use of this feature is for registering data analysis plans, effectively preventing many QRP's that consist of adapting the analysis to the data. Thus, such *pre-registration* limits “researcher degrees of freedom” (Simmons et al., 2011, p.1). The timeline presented on each project page has a similar function, as it automatically records and presents the actual order of events in a research project, preventing self-serving reconstructions after the fact. Likewise, the data archiving and sharing that the OSF facilitates are intended on the one hand to help generate new research, but on the other hand are also presented as ways to increase the control that researchers have over each other's work.

Replication and the hardest science

The primary objective of Open Psychology appears to be methodological. It is about doing better science by making research transparent and thus tightening standards and limiting unwanted flexibility. There is, however, another aspect to this initiative that is ontological rather than epistemological. This concerns the emphasis on replication. To the advocates of Open Psychology, *Enhancing reproducibility* and *Open Science* are basically the same thing (see for instance Open Science Collaboration, 2017). Reproducibility, they say, is the essence of science. Unfortunately, the current problems in psychology (QRP's, publication bias, etcetera) have led to an unknown but great number of irreproducible results. Open Science solves the problems, and is thus the best way of enhancing reproducibility. In order to gauge the extent of the problem, the OSF hosted a special flagship project, announced on the Google OSF group the day after it had opened: the Reproducibility Project (RP). The RP was a crowd sourced, collaborative effort to conduct replications of 100 studies reported in the 2008 volumes of three psychological journals. It was a self-conscious attempt at big science – the project had more than 150 contributors. Its report appeared in 2015 and it concluded that, depending on how one defined a successful replication, only about 40% of the

original results could be reproduced.

Although this diagnosis of the state of the field – too many irreproducible results – is widely shared and many support the Open Science solution, at the same time it runs up against a conviction that is firmly ingrained in the minds of many psychologists: that human behaviour is so sensitive to context, that replication is especially difficult in this field. For this reason, Psychology is sometimes said to be *the hardest science* (Srivastava, 2009). Some have rejected the Reproducibility Project on this basis, because, they say, if a replication fails, it could be due to anything: the weather, the colour of the walls, the smile of the experimenter. A failed replication is therefore uninformative: it doesn't say anything (e.g. Stroebe & Strack, 2014). The Reproducibility Project is chasing a chimera. In Psychology, they insist, replication must take the form of *conceptual replication*: rather than precisely reproducing the original experimental conditions (a *direct replication*), a researcher derives a novel hypothesis from the same theory, and tests that. If she succeeds, the theory is replicated, though not the original result.

A model of surface and depth is operative here: the phenomena of social behaviour (the surface) are complicated and hard to predict, and therefore difficult to reproduce. The mechanisms that underlie this behaviour, however, are stable and universal and can be described by theory. Every time a new effect is produced that can be explained by the theory, it is confirmed. Thus, the 'elderly walking study' may not be reproducible, because the experimental effect is fragile and the social context on which it depends has changed. The theory behind the effect, however, has been confirmed in hundreds of studies ever since, all of them conceptual replications of that theory. Or so its defenders argue. To the advocates of Open Psychology, *conceptual replication* is an important part of the scientific process, but it is insufficient to assess the solidity of particular results or even the theory behind them. The main reason is that a conceptual replication is only used to confirm a theory: if it fails, it never counts against that theory.

The failure simply disappears in a file drawer, and is never heard from again. Conceptual replication is the incarnation of confirmation bias (Pashler & Harris, 2012).

Reproducible objects

Let us consider Open Psychology as an experiment. It is, first of all, an experimental scientific community, that tries to reform scientific practice from the ground up by doing science in a transparent way, that tries to build the tools and infrastructure to support itself and further its ideals, and tries to survive within the constraints of the incentive structure that the academic world imposes on it. Its chances may depend to an important extent on whether it is possible in Psychology to work transparently and still produce the copious output that an academic career seems to require.

But Open Psychology is also an experiment in the sense that it attempts to produce a new psychological object: reproducible instances of social behaviour. This was not its overt aim – it meant only to improve methodology – but that is what it in fact demands. By insisting on direct replication rather than (only) conceptual replication, Open Psychology is saying that Psychology should be about re-presenting overt phenomena as well as hidden mechanisms. If direct replication is to become a standard procedure in psychological research, then social behaviour must be made available for inspection in a reliably reproducible way. Rather than treating experimental effects as mere symptoms of an underlying reality of psychological mechanisms, psychologists must shift attention to those effects themselves and stabilize them.

I want to connect this point about ontology with two well-known arguments from STS. Michael Mulkay and Nigel Gilbert (1986), in their study of replication in science, found that the scientists they interviewed (biochemists) consider *mere replication* (what psychologists call *direct replication*) to be uninteresting. *Mere replication* is assumed to be easy, and therefore conceptual replications are preferred. Only when mistakes, artefacts or foul play are suspected, does it become interesting to do a direct

replication. Psychologists have, until recently, been equally uninterested in *mere replication*, but not because it is too easy, but because it is assumed to be too hard. Based on the ontological premise that the social world is very different from the natural world – complex, fragile, unpredictable, or even historically and culturally variable – a form of social science has been defended that is remarkably similar to other sciences in rejecting direct replication. This consensus is now being challenged: more and more psychologists are calling for direct replication to become a central concern. My point is that this is not just a matter of method, of epistemology, but also of ontology, because direct replication demands that the much-vaunted complexity of the social be made visible and calculable.

Secondly: with its emphasis on constraint, restriction of freedom, and faithful replication, Open Psychology may seem the perfect example of Latour's claim that the social sciences suffer under the mistaken belief that science is about mastery and control (Latour, 2000). According to Latour, their physics envy, combined with a lack of understanding of how physics really works, leads social scientists to seek control over their object, to manipulate people rather than stimulate their disobedience as they should. A scientist should allow the object to object, not seek to control it. From this perspective, Open Psychology, with its Calvinist rejection of frivolity and excess, will only send Psychology further down the wrong track. But there is another interpretation possible. The complexity, fragility, unpredictability of the social has long been a commonplace in Psychology, but it has largely remained a background assumption, not a topic in its own right. The consequences of that complexity – experiments that *do not work* – have remained hidden in the file drawer. With the increasing emphasis on direct replication, they will be dragged into the spotlight. The demand that experimental effects be reproducible in direct replications is not only challenging for the researchers, but it is also a challenge to the object of study itself: resist if you can! If social behavior is really so complex, fragile, and unpredictable, it will withstand direct replication. It is

precisely the attempt to make experimental effects in social psychology reproducible that allows those effects to show their variability. In that way, the object of Psychology may finally get a chance to object.

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), 407–25.
- Calvert, J. (2012). Ownership and sharing in synthetic biology: A “diverse ecology” of the open and the proprietary?. *BioSocieties*, *7*(2), 169–187. doi:10.1057/biosoc.2012.3
- Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral Priming: It’s All in the Mind, but Whose Mind?. *PLoS ONE*, *7*(1), e29081. doi:10.1371/journal.pone.0029081
- Ferguson, C.J, & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication bias and Psychological Science’s Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561.
- Flis, I. (2018). *Discipline Through Method: Recent history and philosophy of scientific psychology (1950-2018)*. (PhD thesis). Utrecht University. Retrieved from <http://dspace.library.uu.nl/handle/1874/373086>
- Gunnarsdóttir, K. (2005). Scientific Journal Publications On the Role of Electronic Preprint Exchange in the Distribution of Scientific Literature. *Social Studies of Science*, *35*(4), 549–579. doi: 10.1177/0306312705052358
- John, L., Loewenstein, G. F., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices with Incentives for Truth-Telling. *Psychological Science*, *23*(5), 524–532. doi:10.2139/ssrn.1996631
- Kelty, C. M. (2012). This is not an article: Model organism newsletters and the question

- of “open science”. *BioSocieties*, 7(2), 140–168. doi:10.1057/biosoc.2012.8
- Latour, B. (2000). When things strike back: a possible contribution of “science studies” to the social sciences. *British Journal Of Sociology*, 51(1), 107–123.
- Mulkay, M., & Gilbert, G. N. (1986). Replication and Mere Replication. *Philosophy of the Social Sciences*, 16(1), 21–37. doi:10.1177/004839318601600102
- Nosek, B.A, Spies J.R, & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, 7(6), 615–631.
- Open Science Collaboration. (2017). Maximizing the Reproducibility of Your Research. In S. O. Lilienfeld & I. D. Waldman (Eds.), *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*. New York: Wiley.
- Pashler, H., & Harris, C. R. (2012). Is the Replicability Crisis Overblown? Three Arguments Examined. *Perspectives on Psychological Science*, 7(6), 531–536.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology. *Psychological Science*, 22(11), 1359–1366. doi:10.1177/0956797611417632
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution. *Dialogue. The Official Newsletter of the Society of Personality and Social Psychology*, 26(2), 4–7.
- Srivastava, S. (2009, March 14). Making progress in the hardest science. *The Hardest Science*. Retrieved from <http://hardsci.wordpress.com/2009/03/14/making-progress-in-the-hardest-science/>
- Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, 7(6), 670–688. doi:10.1177/1745691612460687
- Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspectives on Psychological Science*, 9(1), 59–71. doi:10.1177/1745691613514450

Willinsky, J. (2005). The unacknowledged convergence of open source, open access, and open science. *First Monday*, 10(8). doi: 10.5210/fm.v10i8.1265

Declaración de contribución de autores

MD contribuyó en la totalidad del artículo.

Formato de citación

Derksen, M. (2019). Open Psychology: transparency and reproducibility. *Psicología, Conocimiento y Sociedad*, 9(2), 318-330. doi: <http://dx.doi.org/10.26864/PCS.v9.n2.13>
