Open Science in the MPS
Discussing strategies for systematic changes to foster transparent and accessible research

Open Science Working Group of the Max Planck PhDnet

Date of publication: June 16, 2021

Last updated on: June 24, 2021

Open Science Workgroup

The Open Science Workgroup of the PhDnet was founded in 2018 to strengthen the advancement of transparent and reproducible research in the MPS with a focus on doctoral researchers. Furthermore, its goal is to inform and teach scientists about OS practices.

Authors

Gisela H. Govaart (Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences)

Cornelia van Scherpenberg (Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences)

Acknowledgments

We would like to thank Nikolai Hoermann (Max Planck Institute of Neurobiology, OS WG) for co-authoring previous versions of this discussion paper. Moreover, we thank Sofie Valk (Max Planck Institute for Human Cognitive and Brain Sciences), Corina Logan (Max Planck Institute for Evolutionary Anthropology), Bernd Wirsing, Georg Botz (Max Planck General Administration), and Rashi Goel and Sinikka Paulus (OS WG) for their comments on earlier versions of this paper.

Abbreviations

Doctoral Researcher (DR), Early Career Researcher (ECR), General Administration (GA), Max Planck Institute (MPI), Max Planck Society (MPS), Open Access (OA), Open Science (OS)

1. Introduction

Open Science (OS) is an umbrella term covering a range of approaches that support transparent, rigorous, reproducible, and replicable science. Ever since the "replication crisis" in many scientific sub-disciplines surfaced [1], the credibility of scientific results has suffered, both within the scientific community itself as well as in the public eye [2].

The current COVID-19 pandemic highlights the importance of transparent science and exchange of research findings, which are core principles of OS [3]. Furthermore, OS practices help to maintain a high level of research quality necessary to retain trust in scientific advances. Here, we want to address three topics that we consider important for the science conducted in the Max Planck Society (MPS):

1. Transparent reporting of research workflows
2. Openly accessible research output
3. Creating incentives and support for OS practices

As the OS working group of the Max Planck PhDnet [4], we hereby address members of the MPS at all scientific and administrative levels to start a dialogue on the benefits of OS for the MPS and the implementation of OS practices.

2. Current status of Open Science in the Max Planck Society

The MPS is the driving force behind many research projects, innovations, and initiatives, among those some of the most significant Open Access (OA) initiatives. For example, the MPS has been involved in the movement of OA since the Berlin Declaration [5] in 2003. It hosted some of the Berlin Open Access Conferences and has spearheaded initiatives like Open Access 2020 and Project DEAL [6]. Furthermore, the MPS has supported Open Con [7], a conference focused on Open Access and Open Data. More recently, the MPS has restarted their Open Access Ambassadors program [8] to educate early career researchers (ECRs, i.e., postdocs and DRs) from every Max Planck Institute (MPI). However, the application of OA, and more broadly, OS practices within the society has been left to the interest of the individual researcher.

In 2019, the OS workgroup conducted a survey amongst MPS-DRs to assess their level of familiarity with and attitude towards OS practices. The survey revealed considerable differences between sections of the MPS, some knowledge gaps, but also positive attitudes towards Open Data, OA and other transparent research measures. A detailed report of the survey and its findings was recently published [9].

The survey results and several OS initiatives at MPIs demonstrate the motivation for using OS practices on all levels of the MPS. Therefore, we believe the time has come to establish collaborations between ECRs, directors, group leaders, and members of the general administration (GA) within the whole MPS. By sharing our knowledge and combining our efforts, the MPS can become an **international leader in transparent, rigorous, reproducible and replicable research**.

3. Incentives for the use of OS practices

There is increasing evidence that the use of OS practices is beneficial for individual researchers (DRs, postdocs, group leaders or directors), as well as for the scientific community, and society in general. We will discuss the benefits for these three groups below, focusing on both reproducibility & replicability and accessibility. An overview of the respective OS practices can be found in the appendix.

Reproducibility & replicability

For the individual researcher, a reproducible workflow helps to keep a well-organized and well-documented research pipeline, which makes the research less error-prone [10]. Articles with open data

have been found to be cited more often [11] and more and more journals are even instating a mandatory open data sharing policy (e.g., Cognition, Science, PLOS [12, 13]). Furthermore, open datasets can be cited when assigned a DOI. Regarding benefits for the scientific community, open data and code allow other researchers to check the work for errors, check its robustness by running different statistics on the data, and attempt to replicate the study [13]. This will lead to scientific results that are valued based on their verifiability instead of on blind trust in the researcher. By combining the openly available data on similar topics, the results can be combined to strengthen the claims of individual studies.

The individual researcher benefits from preregistration because it might increase editors' and reviewers' trust in one's methodological choices [14]. Moreover, it can protect against the reviewers' pressure to change hypotheses post-hoc (HARKing). Finally, an increasing number of journals encourage preregistrations (e.g., Psychological Science [15]) and it is expected that preregistrations will become more and more important in the evaluation of confirmatory research. Registered Reports have an added benefit: results will be published independently of the study's outcome. For the individual researcher, this means that even unexpected or null-results can be published. Moreover, preregistration and registered reports address the publication bias, a systematic problem in academia hampering the trustworthiness and sustainability of research.

Accessibility

Publishing preprints benefits the individual researcher because it can increase the citations of the article [16]. In addition, preprints often get helpful feedback by voluntary reviewers, especially if they are shared via Twitter/social media, which allows researchers to address critique before publication [17]. Moreover, the online availability of preprints allows committees to take them into account for job and grant applications. Generally, preprints help accelerate the speed at which information is distributed, which is especially important in cases when new discoveries are time-sensitive, such as during a pandemic or the climate crisis. At the same time, both researchers, science journalists and the public need to be aware that content published in preprints has not been peer-reviewed and should be subject to scrutiny before citation.

Publishing peer-reviewed articles OA has great individual, societal and ethical benefits. For the individual, the research gets on average more citations, mentions, and downloads than non-OA papers [18, 19]. Self-archiving of research results is often even a strict requirement by funding bodies. For the scientific community, OA can contribute to more equity, since it allows all researchers, regardless of their funding, to access all published research findings. What is more, the society as a whole can gain free and public access to scientific knowledge. This is important and ethical since research is, at least in Europe, mainly funded by taxpayers, but research institutions spend a lot of money on both subscriptions and Article Processing Charges (APCs). Publishing OA means that those who paid for the research to be conducted – the public, and those who did the work – the researchers, are ensured access to research output.

4. OS for ECRs

Even though engaging in OS practices has many benefits on the individual and societal level, especially ECRs can also experience some downsides. Pursuing a transparent and reproducible workflow along OS principles can be time-consuming, which is especially difficult for ECRs facing publication pressure and short-term employment contracts. Moreover, the current publishing and funding system values novel, positive results over rigorous cumulative science or replication studies. Even though efforts are being made to reward quality over quantity (e.g., sfDORA), we argue that changing incentive structures and research assessment practices is essential for ECRs, since it directly impacts their career chances. It is also needed for systemic change: to keep people with experience in OS in academia and make it attractive for researchers to engage in OS.

5. Strategies to accommodate OS practices in the MPS

In the following, we want to propose several strategies for the implementation of OS practices in the MPS. We acknowledge that while discussing these measures, the scientific freedom of the MPIs and their researchers has to be protected. We believe that only by combining a top-down and a bottom-up approach we will be able to systematically evolve the way science is conducted in the MPS, thereby sharing the responsibility and tasks.

In addition to concrete measures, we propose the initiation of a pilot project under central supervision, where individual institutes or departments that already implement OS practices can lead by example.

(1) Infrastructure

We recommend the appointment of an **Open Science Officer** (similar to the EO officer/ Sustainability officer) at every MPI to advise people on matters concerning OS or direct them to other sources. This position should ideally be held by a permanently employed person. Furthermore, we would encourage having a contact point for OS at the GA that together with the OS workgroup of the PhDnet can coordinate the strategies on an MPS-wide level. We believe that this network will make it possible to allow MPS researchers to transition to OS practices faster and more efficiently.

(2) Open Access Ambassadors (OAA)

With the OAA initiative, the MPS has created a network of interested ECRs with the potential to establish a profound knowledge base on OS and OA topics, spread this knowledge into the MPIs efficiently, and thus pave the way for a sustainable transition towards OA and OS in the long term. We recommend strengthening this **OAA program** in a collaborative effort of the PhDnet, MPDL and GA.

(3) Training

To facilitate the widespread use of OS practices, we suggest a **training** on OS practices for ECRs with a focus on the benefits for the individual researchers. Trainings could be coordinated with or hosted by the OAA to a certain extent (see (2) above). Moreover, the OS working group of the PhDnet together with the GA is planning to develop an introductory course that could be integrated into the Learning Management System of the Planck Academy. Furthermore, we are planning to create information packages on OS practices and tools that allow interested people to understand and apply them quickly and from the beginning of their time in the MPS.

(4) OA Publications, preprints and postprints

We endorse the efforts of the MPS to continuously support OA publishing within our society, last but not least by negotiating Project DEAL. In addition to this, we suggest that the MPS actively encourages scientists to upload their manuscripts to the preprint server in their respective research area. Furthermore, to utilize **Green OA**, postprints should be automatically added to MPG.PuRe [20], especially for publications in non-OA journals. To implement these suggestions, MPIs should collaborate with their librarians and the MPDL.

(5) Data management and storage plans

Researchers within the MPS collect large amounts of data during their affiliation with an MPI. Currently there is no systematic way of managing or storing these data to allow their future access or reanalysis by other interested researchers from inside or outside the MPS. A dedicated storage space for each MPI to upload **openly accessible data** (to the extent that it is ethically possible and in accordance with data protection regulations) is an important step to enable long term accessibility and use of a multitude of datasets. Currently, the MPS hosts the Open Research Data Repository Edmond [21] through the MPDL. We recommend improving the visibility and implementation of this service within the MPIs and to set up its infrastructure according to the FAIR Data Principles [22] (Findable, Accessible, Interoperable,

Reusable). Open Data can facilitate scientific exchange and collaborations between research groups, thus increasing the overall impact of the data collected in the MPS.

Pilot project: Leading towards transparent and accessible research by example

Currently, there are already some MPIs, departments, or research groups, which have implemented a wide range of OS practices. Here we would like to point out the MPI for Human Cognitive and Brain Sciences (CBS) as an example. The MPI CBS has a very active Open Science Initiative in which not only DRs but also postdocs and group leaders play an active role. In 2019, ECRs at the institute organized two events on Open Science, hosting internationally known keynote speakers, stakeholders of the academic system, and DRs from within and outside the MPS [23, 24]. Resulting from these events and initiatives, the institute as a whole has now endorsed local preregistrations of projects, making it mandatory to preregister neuroimaging projects within the institute database for future reference.

Building on examples like these, we propose that in a **pilot project**, OS pioneers amongst MPIs, departments or research groups should be publicly announced and endorsed by the MPS headquarters and rewarded for their efforts. At the same time, they should be subject to evaluation by a central supervising body – e.g., an Open Science Officer in the GA, a taskforce consisting of senior and junior researchers, and/or the scientific council – which monitors their implementation and consistent application of OS practices regularly. All in all, this pilot project should serve the purpose of highlighting actors within the MPS that are leading by example in transforming their research to more transparency, accessibility and reproducibility. This will increase the incentives for other institutes and researchers to follow this example, thereby, in the long term, helping the MPS as a whole to be a leading organization in open and accessible research.

Rewarding and assessing research quality over quantity

To increase the incentives of applying the measures we suggest above and taking part in the pilot project, efforts should be rewarded visibly. OS initiatives, individual researchers, group leaders, directors or even whole MPIs prominently advancing OS should be awarded with formal recognition. In addition, pioneer projects could receive additional funding (see the QUEST Center [25] or Fellow Programm Freies Wissen [26] for examples).

Moreover, recently, the ethics committee of the MPS has developed new Guidelines for Good Scientific Practice. Amongst other important topics that ensure responsible and trustworthy research practices, it is stated that concerning personnel evaluation, "performance must be assessed predominantly using **qualitative benchmarks**". We endorse this statement and suggest that applying OS practices should serve as an additional marker of evaluation of the science conducted in a research unit, be it during scientific advisory board meetings, the hiring of new directors or group leaders, or research group applications.

6. Conclusion

The MPS with its high international standing has the power to actively contribute to changing research incentives, such that practices that are beneficial for individual researchers benefit the community at large. We acknowledge the commendable efforts of the MPS in promoting and implementing OA, however, other areas of OS should be encouraged more. We, the OS workgroup, hope to start collaborating with the different organizational structures within the MPS to bring forth a systematic change towards a broader implementation of OS practices. Thus, we will not only maintain and even improve the high quality of research conducted in the MPS but can help lead the whole research community towards a more sustainable way of conducting science.

7. References

[1] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*; 349. Epub ahead of print 28 August 2015. DOI: 10.1126/science.aac4716.

[2] Wingen T, Berkessel JB, Englich B. No Replication, No Trust? How Low Replicability Influences Trust in Psychology. *Soc Psychol Personal Sci* 2019; 11: 454–463.

[3] Hoermann N, van Scherpenberg C, Goel R. Open doors with social distance - research opens up during Covid-19 pandemic. *The Offspring Blog*, https://www.phdnet.mpg.de/131182/2020-04-14_openscience-covid19 (2020, accessed 17 May 2021).

[4] Open Science Group, https://www.phdnet.mpg.de/phdnet/who/workgroups/open-science-group (accessed 21 January 2021).

[5] Berlin Declaration, https://openaccess.mpg.de/Berlin-Declaration (accessed 21 January 2021).

[6] Project DEAL, https://www.projekt-deal.de/ (accessed 21 January 2021).

[7] OpenCon. *OpenCon*, https://www.opencon.community/ (accessed 21 January 2021).

[8] Open Access Ambassadors, https://oambassadors.mpdl.mpg.de/ (accessed 21 January 2021).

[9] Toribio-Flórez D, Anneser L, deOliveira-Lopes FN, et al. Where Do Early Career Researchers Stand on Open Science Practices? A Survey Within the Max Planck Society. *Front Res Metr Anal*; 5. Epub ahead of print 2021. DOI: 10.3389/frma.2020.586992.

[10] Klein O, Hardwicke T, Aust F, et al. A practical guide for transparency in psychological science. *Collabra Psychol* 2018; 4: 1–15.

[11] Colavizza G, Hrynaszkiewicz I, Staden I, et al. The citation advantage of linking publications to research data. *PLOS ONE* 2020; 15: e0230416.

[12] Open Access Directory - Journal open data policies, http://oad.simmons.edu/oadwiki/Journal_open-data_policies (accessed 21 January 2021).

[13] Hardwicke TE, Mathur MB, MacDonald K, et al. Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal Cognition. *R Soc Open Sci*; 5: 180448.

[14] Lakens D. Will knowledge about more efficient study designs increase the willingness to pre-register? Epub ahead of print 14 March 2017. DOI: 10.31222/osf.io/svzyc.

[15] Preregistration of Research Plans. *Association for Psychological Science - APS*, https://www.psychologicalscience.org/publications/psychological_science/preregistration (accessed 21 January 2021).

[16] Abdill RJ, Blekhman R. Tracking the popularity and outcomes of all bioRxiv preprints. *eLife* 2019; 8: e45133.

[17] Maggio LA, Artino Jr AR, Driessen EW. Preprints: Facilitating early discovery, access, and feedback. *Perspect Med Educ* 2018; 7: 287–289.

[18] Hajjem C, Harnad S, Gingras Y. Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Eng Bull* 2005; 28: 39–47.

[19] Wang X, Liu C, Mao W, et al. The open access advantage considering citation, article usage and social media attention. *Scientometrics* 2015; 103: 555–564.

[20] MPG.PuRe, https://pure.mpg.de/ (accessed 21 January 2021).

[21] Edmond - the Open Research Data Repository of the Max Planck Society, https://edmond.mpdl.mpg.de/imeji/ (accessed 21 January 2021).

[22] FAIR Principles. *GO FAIR*, https://www.go-fair.org/fair-principles/ (accessed 21 January 2021).

[23] CBS Open Science Day, https://www.cbs.mpg.de/en/cbs-open-science/kick-off (accessed 21 January 2021).

[24] Doing Good - Scientific Practice under Review, https://www.cbs.mpg.de/doing-good (accessed 21 January 2021).

[25] QUEST Calls and Awards, https://www.bihealth.org/en/research/quest-center/calls-and-awards/quest-calls-and-awards (accessed 21 January 2021).

[26] Fellow-Programm Freies Wissen. *Stifterverband*, https://www.stifterverband.org/freies-wissen (2016, accessed 10 May 2021).

Appendix: Overview of Open Science practices

In the following, we address the two aspects of Open Science in more detail that have been discussed in the discussion paper: Reproducibility & Replicability and Accessibility. We are aware that there are many other aspects of Open Science that could be added. We particularly acknowledge the need for science to 'open up' to researchers from diverse backgrounds, because we believe that equity, diversity and inclusion should be integral parts of OpenScience. Within the PhDnet, this topic has been, and continues to be, highlighted and discussed by the Equal Opportunities Working Group [1, 2].

1. Reproducibility and replicability

A reproducible research project is documented and reported transparently enough such that it can theoretically be reproduced. A research project is replicable when, if the experiment is carried out another time, this replication attempt will yield the same or similar results. Note that a reproducible study is not necessarily replicable. Both reproducibility and replicability rely highly on transparency, both in the reporting of the workflow, as well as in the reporting of the researchers' choices during the experiment.

Reproducible workflow

Research is a process that involves many steps, from defining a research question, literature review, designing an experiment or planning field work, to collecting data or resources, analysing, synthesising and disseminating the results. Independent of the specific discipline, keeping a well-structured and transparent workflow is essential for assuring the reproducibility of one's work and to counteract human errors and biases.

There are multiple methods that can support researchers in adapting such a reproducible workflow. For empirical research, documenting the experimental set-up and data-collection process should be done by keeping a lab book and by meticulously documenting the research process in the Methods section of a manuscript, for example with the help of checklists; moreover, preregistration can both increase transparency and counteract human biases that impact reproducibility (see below). During data analysis, all code should be thoroughly commented and documented; for this, version control programs can be of great help. Moreover, after the project is completed, all data and code can be uploaded to open platforms and thereby be made available to reviewers, readers, or anyone interested in the analysis workflow. Such platforms are, for example, the Open Science Framework (OSF) [3], Dryad Digital Repository [4], Zenodo [5], or Edmond [6] (a service offered by the MPS). re3data [7] serves as a platform to search for repositories according to the user's requirements. For both data and code, researchers should follow the rules of good scientific practice in the annotation of their data and code in order to facilitate their reuse.

Preregistrations and Registered Reports

Preregistration is the practice of specifying the research plan (experimental design, hypotheses, and planned analysis) before starting data collection. In clinical research, it has been a longstanding requirement to preregister a study before conducting and publishing the study. In other fields, this practice is fairly new, but it is becoming more and more common [8], appreciated, and even expected by some journals [9]. Common platforms for uploading preregistrations are the Open Science Framework and AsPredicted [10]. Preregistrations can be published immediately or put under an embargo until the researcher decides to publish the preregistration (to avoid possible scooping). Preregistration is an important tool to increase transparency [11]. This reduces selective outcome reporting and can mitigate confirmation and hindsight bias [12].

An extension of preregistrations are Registered Reports (RRs). In this publishing format, a researcher submits a manuscript before data collection, i.e., without the results section and discussion, which is peer-reviewed and, if deemed suitable, in principle accepted by the journal. If the researchers follow their prespecified plan, the article is published regardless of the results of the study. RRs therefore have the added benefit of mitigating the publication bias.

Preregistration and RRs are important tools to make the research process and researchers' choices transparent, and therefore contribute to making research more reproducible and verifiable.

2. Accessibility

An important factor in doing transparent research is having access to knowledge and data. This allows researchers to design experiments more thoroughly, save resources, and avoid repetitions of errors; furthermore, it increases the use of published research findings. Therefore, it should be encouraged to share results, data and code, to enhance open practises within the community.

Preprints and Postprints

Published articles go through several stages before becoming available to the public. First, the authors submit the final version of the manuscript to a journal for peer review. At this stage, that is, before peer review, the manuscript can be uploaded as a 'preprint' on a preprint server, and immediately becomes publicly accessible. After the manuscript is adapted according to peer review and accepted to a journal, but before having been formatted, the manuscript is called a 'postprint'; this version can also be uploaded to the preprint server as an updated version of the preprint, or on an OA repository. The finalized article that is published in the journal is called the 'publisher's version'. Preprints are becoming more common and are a means for authors to disseminate their research before peer-review. Among the many 'preprint servers', the most well-known are: arXiv [13] or bioRxiv [14] (for a list of preprint servers, see [15]). Preprints available there can be freely downloaded. In addition, the number of publishers supporting the decision of authors to self-archive their postprint on repositories after publication is increasing, enhancing the accessibility of post-peer review research articles.

Publishing

The implementation of OA practices, in which journal articles or books are freely available to everyone, differs significantly between publishers and journals. Some journals (such as eLife, PLOS, or others) offer full OA publishing by charging 'article processing charges' (APC), which is called 'Gold OA'. In comparison, most traditional publishers have subscription-based models with the option to publish under OA conditions for an additional charge, called 'Hybrid OA'. Additionally, the publisher may allow 'self-archiving', meaning that the postprint can be archived after a certain time period on a repository of the author's choice, called 'Green OA' (specific journal guidelines can be checked in the Sherpa Romeo repository [16]). Finally, there are journal that are fully open access, without charching APCs, called 'Platinum/Diamond OA'. One important aspect to consider during publication is who owns the copyright. It is common that authors give away the full copyright to the publishers when signing the publishing contract, which limits the possibility of disseminating the research by the author. Instead, it would be recommendable that authors keep the right to distribute their research, at the very least in the form of self-archiving.

3. References

[1] Equal Opportunity Group of the Max Planck PhDnet, https://www.phdnet.mpg.de/phdnet/who/workgroups/equal-opportunity-group (accessed 1 March 2021).

[2] Mental Health Collective of the MPS, Medawar E, Vvedenskaya O, et al. Black Lives Matter - Black Minds Matter. *Max Planck PhDnet Offspring Blog*, https://www.phdnet.mpg.de/136128/20200611_BlackLivesMatter---BlackMindsMatter?c=22833 (accessed 1 March 2021).

[3] Center for Open Science. Open Science Framework, https://osf.io (accessed 1 March 2021).

[4] DRYAD. Dryad Digital Repository, https://datadryad.org/stash.

[5] Zenodo - Research. Shared., https://zenodo.org/ (accessed 1 March 2021).

[6] Edmond - the Open Research Data Repository of the Max Planck Society, https://edmond.mpdl.mpg.de/imeji/ (accessed 21 January 2021).

[7] re3data.org - Registry of Research Data Repositories, https://doi.org/10.17616/R3D.

[8] Nosek BA, Ebersole CR, DeHaven AC, et al. The preregistration revolution. *Proc Natl Acad Sci USA* 2018; 115: 2600.

[9] Preregistration of Research Plans. *Association for Psychological Science - APS*, https://www.psychologicalscience.org/publications/psychological_science/preregistration (accessed 21 January 2021).

[10] AsPredicted, https://aspredicted.org/ (accessed 1 March 2021).

[11] Navarro D. Paths in strange spaces: A comment on preregistration. Epub ahead of print 23 September 2020. DOI: 10.31234/osf.io/wxn58.

[12] Paul M, Govaart G, Schettino A. Making ERP Research More Transparent: Guidelines for Preregistration. Epub ahead of print 4 November 2020. DOI: 10.31234/osf.io/4tgve.

[13] Cornell University. arXiv.org e-Print archive, https://arxiv.org/ (accessed 1 March 2021).

[14] Cold Spring Harbor Laboratory. bioRxiv.org - the preprint server for Biology, https://www.biorxiv.org/ (accessed 1 March 2021).

[15] FOSTER. Sharing Preprints, https://www.fosteropenscience.eu/learning/sharing-preprints/#/id/5ac23bbcdd1827131b90e79d (accessed 1 March 2021).

[16] JISC. Sherpa Romeo, https://v2.sherpa.ac.uk/romeo/ (accessed 1 March 2021).