



## DATA NOTE

# REVISSED Applicability of open science practices to completed research projects from different disciplines and research paradigms [version 2; peer review: 3 approved, 1 approved with reservations]

Jürgen Schneider

University of Tübingen, Tübingen, Germany

**V2** First published: 11 Apr 2022, 11:408  
<https://doi.org/10.12688/f1000research.111383.1>

Latest published: 20 Jul 2022, 11:408  
<https://doi.org/10.12688/f1000research.111383.2>

## Abstract

The purpose of this data collection was to uncover the extent to which communities have emerged that cultivate a shared understanding of open science. In a cross-sectional survey, we assessed the applicability of 13 open science practices over different disciplines and research paradigms. Focusing on completed research projects, participants were able to draw informed evaluations concerning the applicability of open science practices. The total sample is N=295 researchers, with approximately equal numbers from six broad disciplines (between 42 and 52 participants per discipline). The survey included an attention check.

## Keywords

Open Science, Open Access, Open Data, Open Education, Open Evaluation, Open Methods, Open Participation, Open Policies, Open Software, Open Tools



This article is included in the [Research on Research, Policy & Culture](#) gateway.

## Open Peer Review

Approval Status

	1	2	3	4
<b>version 2</b>				
(revision)				
20 Jul 2022	<a href="#">view</a>		<a href="#">view</a>	<a href="#">view</a>
<b>version 1</b>				
11 Apr 2022	<a href="#">view</a>	<a href="#">view</a>		

1. **Suzanne Dumouchel** , CNRS, Paris, France
2. **Lauren Cadwallader** , Public Library of Science, San Francisco, USA
3. **Chuang Liu**, China Academy of Science, Beijing, China  
**Ke Gong**, North China University of Science and Technology, Tangshan, China  
**Xiangtong Jia**, Nankai University, Tianjin, China
4. **Henrik Bellhäuser** , Mainz University, Mainz, Germany

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding author:** Jürgen Schneider ([juergen.schneider@uni-tuebingen.de](mailto:juergen.schneider@uni-tuebingen.de))

**Author roles: Schneider J:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by Fellowship Open Science, from Wikimedia Germany, Volkswagen Foundation and Stifterverband.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2022 Schneider J. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Schneider J. **Applicability of open science practices to completed research projects from different disciplines and research paradigms [version 2; peer review: 3 approved, 1 approved with reservations]** F1000Research 2022, 11:408 <https://doi.org/10.12688/f1000research.111383.2>

**First published:** 11 Apr 2022, 11:408 <https://doi.org/10.12688/f1000research.111383.1>

**REVISED Amendments from Version 1**

In the data set we have now added a column with the discipline from each participant (according to the Frascati manual classification). Previously, the data set contained only cluster-level discipline classification information. We also provided more information on the generation of the 13 items for the Open Science Practices survey. In particular, we have now described the interview survey process in more detail.

**Any further responses from the reviewers can be found at the end of the article**

**Introduction**

“Open science” is not a new term in the repertoire of academic disciplines, it has developed over the last several decades. However, what researchers understand by open science and which open science practices they refer to in particular varies greatly between research projects. In some research projects, open science practices are closely related with aspects of the research paradigm’s philosophy of science (e.g., preregistration in critical rationalism), in other research projects, they are at odds (e.g., replicability in constructivism). Accordingly, the implementation of open science practices naturally varies in degree of implementation and stage of development. The purpose of this data collection was to reveal the extent to which subcommunities have emerged that share a common profile of implementing open science practices and how they are related to research paradigms and disciplines.

**Materials and methods****Data collection protocol**

The study was conducted as a cross-sectional survey using the survey tool *formr* (Arslan, Walther & Tata, 2020). In order to achieve a high ecological validity, the survey referred to the applicability of open science practices in concrete research projects instead of time periods (e.g., referring to the past year). We focused on evaluating *completed* research projects as participants were able to draw on their experiences from all phases of the project, which allowed them to make informed assessments of the factors that influenced research project decisions. Participants were recruited via the online access panel provider prolific.co. We used prolific’s built in filter to target researchers (“Industry Role = Researcher”). In the study description for prolific, we merely indicated that the study addressed practices in research projects; the focus on open science practices was avoided to reduce selection bias in the sample:

“In this study, you will indicate whether 13 practices are *potentially applicable* to a research project you were conducting. The survey contains *only 16 items* in total.

We are looking for participants who have conducted a *research project* associated with *one of the disciplines*

- Natural sciences
- Engineering and technology
- Medical and health sciences
- Agricultural and veterinary sciences
- Social sciences
- Humanities and the arts”

Participants took on average 3.98 minutes (median 3.43) to answer the survey and received USD 0.85 as compensation (approx. USD 12.81 per hour on average). The first participant started on August 20 2021, the last session on March 16 2022.

We aimed for a sample distributed across all research disciplines. For this reason, we drew on the classification of research fields from the OECD Frascati Manual (OECD, 2015). The Frascati Manual is an internationally acknowledged standard on the methodology of collecting and using research and development statistics, developed by the OECD. As a standard, it is the first choice for the definition and taxonomy of research disciplines. For each “broad classification” from the manual (natural sciences, engineering and technology, medical and health sciences, agricultural and veterinary sciences, social sciences, humanities and the arts) we aimed for n=50 participants, which would have led to a total sample of N=300 (limit of allocated financial resources). As soon as 50 participants from a discipline (broad classification) finished the

**Table 1. Count of participants from each discipline.**

Discipline (broad classification)	Count of participants
Natural sciences	52
Engineering and technology	50
Medical and health sciences	50
Agricultural and veterinary sciences	42
Social sciences	51
Humanities and the arts	50
$\Sigma$	<b>295</b>

survey, access to the survey closed for participants from that discipline (“cell closed”). For two broad classifications we exceeded the stopping rule (Natural Sciences, Social Sciences), as cells only closed after the last participant from that cell *finished* the survey, while further participants from that cell were still able to begin the survey (see Table 1). In addition, we were only able to recruit 42 participants for the agricultural and veterinary sciences despite several postings on prolific. This may not be surprising, since agricultural and veterinary sciences is a narrower field compared to the other broad classifications. This is also the reason why the data collection is spread over a longer period of time. After the start of data collection, participants of the other broad classifications could be collected within a few weeks. Due to the repeated invitation of researchers from the agricultural and veterinary sciences, the period of data collection stretched out, unfortunately only a few additional participants could be recruited (see codebook).

### Procedure and measures

On the first page, participants agreed to the declaration of consent.

On the second page they indicated the *discipline* in which the research project was based regarding which they would like to answer the following questions: “Discipline. On the next page you will answer questions regarding a previous research project. To which discipline is this research project most closely related?” In a dropdown menu, participants were able to choose from all 42 second-level classifications from the OECD Frascati Manual (OECD, 2015). In using the second-level classifications, we tried to avoid inconsistent assignments to the broad classifications by the participants. After that an attention check item was displayed (see below).

The third page gave a quick instruction on how to answer the items following on the next page: “When answering the items on the next page, please think of a research project of yours that you have already completed. Regardless of whether you actually applied the practices in this research project: Which of the practices would have been *potentially applicable*, given all the characteristics and circumstances of the project? This includes both scientific, and practical considerations in conducting the study.”

On top of the fourth page the following question was displayed: “To what extent are the following behaviors applicable in your research project?” Which was then followed up by 13 items on open science practices (item labels see Table 2). The practices were derived and synthesized using a top-down and bottom-up approach from the **FOSTER Taxonomy of Open Science** (top-down) and nine additional expert interviews from different disciplines (bottom-up). Through the top-down and bottom-up approach, blind spots were mutually exposed to ensure that the broadness of open science practices are reflected in the survey. The FOSTER taxonomy is the only taxonomy on open science that we know of. It was created as part of the FOSTER Plus project, an EU-funded project on Open Science. The goals in the project explicitly covered the generation of high quality training resources, which includes the taxonomy we use. For the bottom-up approach, we interviewed nine experts in open science. We recruited the experts from an open science fellows program in which they served as mentors. Following the theoretical sampling approach, we recruited mentors who came from a variety of disciplines (e.g., sociology, computer science, sinology) and applied different research paradigms (qualitative, quantitative, mixed methods, theoretical). In a focused interview, interviewees were given a narrative prompt to retrospectively consider open science practices in their field: “Please recall one of your most recently completed research projects. Thinking about the entire span of the project, from the initial idea to the completion of the project, what aspects of open science do you consider significant and how can they be exemplified in research projects?” The interviewer then asked follow-up questions about other practices: “Are there other aspects of Open Science that you consider significant in your research projects (i.e., potentially others as well)? If so, how could these be implemented?”. The interviewer also asked follow-up questions to clarify individual practices mentioned. Two trained coders transcribed and segmented the interview material around each open science practices mentioned. Disagreements in the coding process were resolved through discussion throughout the coding process. With the segmented material, the coders conducted a qualitative

**Table 2. The items of the 13 open science practices addressed in the survey.**

<p><b>Involving the non-academic public in the research process</b> The non-academic public is involved in the process of scientific research – whether in community-driven research or global investigations. Citizens do scientific work—often working together with experts or scientific institutions. They support the generation of relevant research questions, the collection, analysis or description of research data and make a valuable contribution to science. <i>“Citizen Science”</i></p>
<p><b>Publicly sharing project plans to encourage feedback and collaboration</b> Researchers make their project plans publicly available at an early stage (e.g., on social media, websites) to optimize the study design through feedback and to encourage collaboration. <i>“Open Collaboration”</i></p>
<p><b>Preregistering study plans</b> Researchers submit important information about their study (for example: research rationale, hypotheses, design and analytic strategy) to a public registry before beginning the study. <i>“Preregistration”</i></p>
<p><b>Publicly sharing the methodology of the research process</b> Researchers describe methods, procedures and instruments that are used in the process of knowledge generation and make them publicly available. <i>“Open Methodology”</i></p>
<p><b>Using Open file formats and research software</b> Researchers use software (for analysis, simulation, visualization, etc.) as well as file formats that grant permission to access, re-use, and redistribute material with few or no restrictions. <i>“Open File Formats and Research Software”</i></p>
<p><b>Publicly sharing research materials</b> Researchers share research materials, for example, biological and geological samples, instruments for measurement or stimuli used in the study. <i>“Open Materials”</i></p>
<p><b>Publicly sharing data analyses</b> Researchers make the procedure of the data analyses and their scripts (“code”) publicly available so that others are able to reach the same results as are claimed in scientific outputs. <i>“Open Code/Open Script”</i></p>
<p><b>Publicly sharing research data</b> Researchers publicly provide the data generated in the research process free of cost and accessible so that it can be used, reused and distributed provided that the data source is attributed. <i>“Open Data”</i></p>
<p><b>Generating open educational resources</b> Researchers produce and release teaching, learning and research materials in any medium that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions. Open educational resources include full courses, course materials, modules, textbooks, streaming videos, tests, images, software, and any other tools, materials, or techniques used to support access to knowledge. <i>“Open Educational Resources”</i></p>
<p><b>Deciding for openness in the peer review process</b> Researchers opt for some kind of openness in the peer review process, including making reviewer or author identities open, publishing review reports or enabling a broader community to participate in the process. <i>“Open Peer Review”</i></p>
<p><b>Publishing open access</b> Researchers publish their research paper online, free of cost with free reusability regarding copyright restrictions. This involves any form of open access (preprints, gold and hybrid open access, etc.). <i>“Open Access”</i></p>
<p><b>Providing open source code of software</b> Researchers make source code for a piece of software that was developed in the research process publicly available, along with an open source license permitting reuse, adaptation, and further distribution. <i>“Open Source”</i></p>
<p><b>Communicating research results to nonacademics</b> Researchers use appropriate skills, media, activities, and dialogue to produce one or more of the following personal responses to science: Awareness, Enjoyment, Interest, Opinions, Understanding. Science communication may involve science practitioners, mediators, and other members of the general public, either peer-to-peer or between groups. <i>“Science Communication”</i></p>

Note: The item format was a 4-point Likert scale with the answer format from “not applicable at all” to “highly applicable” on the question “To what extent are the following behaviors applicable in your research project?”.

content analysis. In two stages they abstracted the practices named by the interviewees to an equivalent level of abstraction. These practices thus obtained were finally compared and synthesized with the FOSTER taxonomy resulting in 13 items on open science practices.

On the bottom of the page we assessed the research paradigm the project was situated in: “Research paradigm. What was the project’s primary research interest and design?” with the single choice answer categories “mainly qualitative empirical”, “mainly quantitative empirical”, “explicitly mixed-methodological (equally qualitative and quantitative empirical)” and “nonempirical”.

For details on items and item statistics, see the codebook (created with the R package *codebook*; Arslan, 2019) in the *Extended data* (Schneider, 2022).

### Data validation

Participants had to pass an attention check at the beginning of the survey in order to be able to complete the other questions. The attention check looked as follows:

*“Please read the following scenario briefly and answer a question about it:*

A famine has broken out in your village. You and some others have been chosen to leave the village and search for food. It begins to rain heavily and soon there will be flooding. Participants in studies like this are sometimes not very attentive. We have included this question here to check if you have actually read the scenario. If you read this, leave the following question unanswered just click next.

*According to the scenario, would it be appropriate to take the raft and leave the others behind?”*

Followed by a seven-point Likert scale with the anchors “absolutely no” and “absolutely yes”. The attention check was considered “passed” if *nothing* was marked on the seven-point Likert scale (i.e. an NA value on this item). Overall, 20 participants eligible for participation failed the attention check and were thus excluded. These participants are not included in the data set that is available for download (Schneider, 2022). They jumped to the end of the survey after failing the attention check and therefore did not complete the 13 items on the open science practices.

As a limitation regarding data validation, it should be noted that we did not target a representative sample of researchers across disciplines. For the data set, it was important that we had variance in the backgrounds of the researchers. Any analyses comparing disciplines should therefore be interpreted with caution.

### Ethical approval and consent to participate

The present data collection received approval from the ethics committee of the Faculty of Economics and Social Sciences at the University of Tübingen (no approval number). Participants agreed to the consent details printed below before beginning the survey.

### Future analyses

In the future, the data will be analyzed to answer the questions whether there are different communities in the application of open science practices and to what extent the open science practice profiles of these communities are similar or different to each other. Are there open science practices that all communities share? Are there practices for which there are particularly strong differences between communities? In addition, the role of research disciplines and research paradigms will be explored.

### Data availability

#### Underlying data

Zenodo: Applicability of open science practices to completed research projects from different disciplines and research paradigms. <https://doi.org/10.5281/zenodo.6834569> (Schneider, 2022).

This project contains the following underlying data:

- osc\_data.RData (data set as RData-file)
- osc\_data.csv (data set as CSV-file)

### Extended data

Zenodo: Applicability of open science practices to completed research projects from different disciplines and research paradigms. <https://doi.org/10.5281/zenodo.6834569> (Schneider, 2022).

This project contains the following extended data:

- codebook.html (codebook report of survey and its items)
- STROBE-checklist-v4-cross-sectional.pdf (STROBE Statement: Checklist of items that should be included in reports of cross-sectional studies)
- Consent Statement.pdf (Consent Statement: Details of the Consent Statement the participants agreed to)

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

### Acknowledgements

We thank the nine interviewees from the pilot study. We acknowledge support by Open Access Publishing Fund of University of Tübingen.

### References

---

Arslan RC: **How to Automatically Document Data With the codebook Package to Facilitate Data Reuse**. *Adv. Methods Pract. Psychol. Sci.* 2019; **2**(2): 169–187.

[Publisher Full Text](#)

Arslan RC, Walther MP, Tata CS: **formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R**. *Behav. Res. Methods.* 2020; **52**(1):

376–387.

[PubMed Abstract](#) | [Publisher Full Text](#)

Schneider J: **Applicability of open science practices to completed research projects from different disciplines and research paradigms**. **[Dataset]**. 2022.

[Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:    

---

Version 2

Reviewer Report 26 September 2022

<https://doi.org/10.5256/f1000research.136169.r145712>

© 2022 Bellhäuser H. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Henrik Bellhäuser** 

Department of Psychology, Mainz University, Mainz, Germany

In this data note, the author gives access to and describes a data set collected from N=295 researchers from various fields that assessed whether a number of open science practices were applicable in their discipline.

I value the approach to make the data accessible and overall I think the author did a very good job in documenting all details relevant for this data collection. The data is both interesting and easily reusable based on this data note.

I have only few minor comments as I believe the previous reviewers and the author already solved the major issues:

- At first, I thought the declaration of consent would not be included. Later, I found it on Zenodo—you might want to mention on p.4 where the reader can find it.
- If possible, you might add the link to Zenodo already in the abstract.
- The codebook is great—makes me want to start form as well. If it would be possible to adjust the y-axes of all variables so that the size of the bars are comparable, it would be even better in my opinion.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**



Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** educational psychology

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 19 August 2022

<https://doi.org/10.5256/f1000research.136169.r146059>

© 2022 Gong K et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Chuang Liu**

China Academy of Science, Beijing, China

**Ke Gong**

College of Life Sciences/Center for Genomics and Bio-computing, North China University of Science and Technology, Tangshan, China

**Xiangtong Jia**

Nankai University, Tianjin, China

The research involves nearly all aspects of open science practices (including six broad disciplinary), the span of disciplines is large, the content is complex and the data is valuable. The work had been approved by an ethics committee and the data had been validated properly.

However, the paper does not mention similar research but no comparable analysis. The total number of the validated participants is 295, and 42-52 for each discipline. It is suggested to increase the number of participants and consider to have different number of participants for different disciplines according to the proportion of research projects of related discipline, so that it may be more reasonable.

The practice of open science involves complex socio-technical elements, including scientific communities, general public and companies and so on, participants with different backgrounds may have different opinions, this might be further considered.

I agree very much for the setting of attention check, to ensure the reliability of the survey and to avoid the random answers of the respondents. The questionnaire form could be concise and friendly, can even add some bogus items options.

Furthermore, I wonder if the author has considered privacy issues during the survey, how have these issues been treated? The survey asked participant to make a judgement between 1-4, why not chosen between 1-10?

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Electronic engineering and telecommunication, engineering education, scientific policies

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 01 August 2022

<https://doi.org/10.5256/f1000research.136169.r144742>

© 2022 Dumouchel S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Suzanne Dumouchel** 

TGIR Huma-Num, CNRS, Paris, France

I have read the answers to my review and seen the additions and it's fine for me. I approve the publication.

**Is the rationale for creating the dataset(s) clearly described?**

Partly

**Are the protocols appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

**Are the datasets clearly presented in a useable and accessible format?**

Partly

**Competing Interests:** No competing interests were disclosed.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

Version 1

Reviewer Report 27 May 2022

<https://doi.org/10.5256/f1000research.123102.r137032>

© 2022 Cadwallader L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Lauren Cadwallader** 

Public Library of Science, San Francisco, CA, USA

This research aims to gain insights into open science practices in six broad disciplinary areas by asking researchers about the potential to practice open science behaviours in relation to a past project. Thirteen different open science behaviours are included, which give a comprehensive view of open science rather than just publishing related activities.

The rationale for the study is good and asking researchers about past projects, rather than asking them to predict future work, is an interesting and valuable approach to assessing open science practices. The range of open science practices included in the survey is broad, which should give a good insight into all activities that contribute to the aims of being open. However, one aspect I think is lacking is the extent to which researchers were mandated by their funder or institution to carry out any of these activities, which would add context to the extent to which they view certain practices as applicable.

In terms of the demographics of the survey sample, it is good to see balanced sample sizes in the disciplines and the inclusion of whether the researcher's work is quantitative, qualitative or mixed methods. The methodology should include more information on which geographical areas were targeted for respondents and information on career stages of the respondents if possible, as these are both factors in adoption of open science practices.

I have concerns that the wording used in the Likert scale may have been confusing for respondents. The aim was to get researchers to report on their past research but the question about open science practices was in the present tense ("To what extent **are** the following behaviours applicable..." [emphasis added]). I'm also unsure of how to interpret the "not at all applicable" response - the codebook seems to suggest a "N/A" option was also offered but this doesn't seem to have been used by anyone. For example, were all respondents creating software but not all thought that sharing software openly applied to their project? Or were the respondents

who chose "not at all applicable" really saying they didn't create software? Including the survey instrument (i.e. a pdf copy of the survey questions as presented to the respondents) might help with understanding this better.

A couple of other specific points:

- There is no mention of data handling with respect to GDPR rules.
- The codebook gives an identical explanation for two variables - osp\_sco and osp\_cit. Presumable the latter is Citizen Science and not Science Communication.

The dataset in Zenodo is well presented and easy to use (aside from the labelling mistake mentioned above). As mentioned earlier, I would recommend adding the survey instrument to the dataset to aid understanding and also enable others to repeat this approach.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

**Are the datasets clearly presented in a useable and accessible format?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Scholarly communications, publishing, open science.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 14 Jul 2022

**Jürgen Schneider**, University of Tübingen, Tübingen, Germany

Thank you as well for the constructive comments. I will respond to your comments chronologically:

1. We agree that it would have been interesting to have data on the mandates of funders or institutions. Unfortunately, we can't collect this data a posteriori. However, there is first evidence that mandates at the level of the research institution do not necessarily have the greatest impact on researchers' open science practices. For example, the existence of a research data policy at a research institution is not related to higher data sharing: <https://doi.org/10.31234/osf.io/9yhcz>
2. We agree that information on the geographical areas and career stage of participants

would complement the data set. Unfortunately, we had to keep the survey short due to financial resources. Therefore, we don't have these data.

3. I agree that the wording of the items in the present tense is unfortunate. Participants were encouraged to ask us questions via the survey-internal direct messaging system if any questions arose during the course of the questionnaire. The participants did indeed use this opportunity, and we received 27 direct messages. However, these inquiries were all aimed at the attention check that was placed at the beginning of the questionnaire, no questions arose about the subsequent items.
4. We have been careful not to collect personalized data in the survey for which the GDPR would be relevant.
5. Thank you for the pointing out the mistake in the codebook. We corrected this and uploaded a new version of the codebook.

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 24 May 2022

<https://doi.org/10.5256/f1000research.123102.r137033>

© 2022 Dumouchel S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Suzanne Dumouchel** 

TGIR Huma-Num, CNRS, Paris, France

The process of the survey is clearly described such as the tools that have been used and the reasoning behind them. The selection of the researchers is also well explained and the average per discipline is relevant for this kind of study. However, it could be interesting to be precise with any information regarding the geographic coverage of the researchers and their career stage. It can contribute to better understand of the OS practices. This might have been taken into account however it is not mentioned in the paper.

It seems particularly relevant to offer the possibility to specify the discipline and not to stay at the level of the broad classification. On this topic, further explanation of the reasoning behind the broad classification could be useful, especially to understand why "agricultural and veterinary sciences" are not part of natural sciences. What is the reasoning behind this?

The selection of the Frascati Manual seems to be relevant but more details can be provided to mention maybe other options and to explain why this one has been selected and upon which criteria?

This precision can be expected as well for the FOSTER Taxonomy of Open Science.

The role of the nine additional expert interviews is not clear enough. It is difficult for the reader to understand how practices have been synthesised from both the FOSTER Taxonomy and the expert interviews. The paper could be stronger by explaining 1) the goal and process of the interviews; 2) the selection of these experts (for instance regarding the disciplines as it is the main focus in the paper); 3) how the practices have been built from these interviews and the taxonomy. For instance, it can be interesting to highlight the potential bias or reasoning to synthesise in one way or another.

It is interesting to see that the survey has run between August 2020 and March 2022. Regarding the increase of knowledge and practices related to Open Science, the answers might have evolved between both. Even if it can be quite tricky to determine this aspect, it can be relevant to mention or highlight any difference if some have been noticed.

Lastly the reader could expect a short conclusion about the main outputs of the survey. And if the goal is to provide it in a dedicated paper, some highlights can be useful or at least it can be announced at the end of the paper.

In any case, this kind of survey is particularly relevant and should contribute to implement the OS practices in the different disciplines as well as to better identify the needs for the researchers. With these small precisions, we believe the paper (and the data collected) contributes to the OS reflexion.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

**Are the datasets clearly presented in a useable and accessible format?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** research infrastructures, social sciences and humanities, FAIR data, EOSC, Open Science

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 14 Jul 2022

**Jürgen Schneider**, University of Tübingen, Tübingen, Germany

Thank you very much for your constructive feedback. I will respond to your comments chronologically:

1. We agree that information regarding the geographic coverage and their career stage would complement the data set. Unfortunately, we had to keep the survey short due to financial resources. Therefore, we don't have these data.
2. We agree that specifying the disciplines might be a valuable information for data users. Therefore, we have now included the disciplines in the data set as well as the Codebook.
3. Concerning the Frascati manual, we added the following paragraph to the methods section: "The Frascati Manual is an internationally acknowledged standard on the methodology of collecting and using research and development statistics, developed by the OECD. As a standard, it is the first choice for the definition and taxonomy of research disciplines."
4. Concerning the FOSTER taxonomy, we added the following paragraph to the methods section: "The FOSTER taxonomy is the only taxonomy on open science that we know of. It was created as part of the FOSTER Plus project, an EU-funded project on Open Science. The goals in the project explicitly covered the generation of high quality training resources, which includes the taxonomy we use."
5. We agree that more information on the synthesis of the FOSTER taxonomy and the interviews would help readers understand the construction of the survey. We therefore added the following paragraph to the methods section: "Through the top-down and bottom-up approach, blind spots were mutually exposed to ensure that the broadness of open science practices are reflected in the survey. [...] For the bottom-up approach, we interviewed nine experts in open science. We recruited the experts from an open science fellows program in which they served as mentors. Following the theoretical sampling approach, we recruited mentors who came from a variety of disciplines (e.g., sociology, computer science, sinology) and applied different research paradigms (qualitative, quantitative, mixed methods, theoretical). In a focused interview, interviewees were given a narrative prompt to retrospectively consider open science practices in their field: "Please recall one of your most recently completed research projects. Thinking about the entire span of the project, from the initial idea to the completion of the project, what aspects of open science do you consider significant and how can they be exemplified in research projects?" The interviewer then asked exmanent follow-up questions about other practices: "Are there other aspects of Open Science that you consider significant in your research projects (i.e., potentially others as well)? If so, how could these be implemented?". The interviewer also asked intrinsic follow-up questions to clarify individual practices mentioned. Two trained coders transcribed and segmented the interview material around each open science practices mentioned. Disagreements in the coding process were resolved through discussion throughout the coding process. With the segmented material, the coders conducted a qualitative content analysis. They

abstracted the practices named by the interviewees to an equivalent level of abstraction in two stages. These practices thus obtained were finally compared and synthesized with the FOSTER taxonomy resulting in 13 items on open science practices."

6. We agree that the choice of the timeframe may raise some questions. The reason for the long time frame was our desire to recruit more researchers from the broad classification of "agricultural and veterinary sciences". Accordingly, we have added a paragraph: "This is also the reason why the data collection is spread over a longer period of time. After the start of data collection, participants of the other broad classifications could be collected within a few weeks. Due to the repeated invitation of researchers from the agricultural and veterinary sciences, the period of data collection stretched out, unfortunately, only a few additional participants could be recruited (see codebook)."
7. We agree that information on the output of the survey would be interesting information for the readers. We included the following paragraph in a newly created "Future analyses" section: "In the future, the data will be analyzed to answer the questions whether there are different communities in the application of open science practices and to what extent the open science practice profiles of these communities are similar or different to each other. Are there open science practices that all communities share? Are there practices for which there are particularly strong differences between communities? In addition, the role of research disciplines and research paradigms will be explored."

**Competing Interests:** No competing interests were disclosed.

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**