Swinburne Open Science Task Force Survey (2019)

Dr Jennifer Beaudry, A/Prof Jordy Kaufman, Prof Tom Johnstone,
Prof Lisa Given & Michelle Vincent

09 September 2020

1 Abstract

In late 2018, then Deputy Vice Chancellor - Research & Enterprise, Professor Aleks Subic, officially formed the Swinburne Open Science Task Force in order to generate recommendations for how Swinburne as an institution can most productively engage with open sciences practices. To best inform these recommendations the core working group of the Task Force surveyed our researchers' current use of open science practices, their perceptions of these practices, and their perceived barriers to adopting these practices. Specifically, we were interested in their attitudes and experiences with (i) study preregistration, (ii) open materials and/or code, (iii) open data, (iv) pre-publication archiving, and (v) open access publishing. Over 200 researchers (n = 239) completed our online survey in September 2019. This document presents an overview of their responses to the various quantitative questions. The key finding was that researchers overall perceived most open science practices as important for their fields; however, most researchers engaged in these practices infrequently or not at all. Our data also revealed some key barriers that researchers felt impeded their use of open science practices.

2 Methodology

Across the university, 239 researchers started the online survey (run through Qualtrics). Not all respondents completed all questions; thus, we report the sample size for each question. This document summarises their responses to the various quantitative questions. We also asked a number of open-ended questions, which we are currently analysing. The full survey is available at https://osf.io/a4bwp/.

2.1 Definitions

In the survey, we explained that the umbrella term "open science" ensapsulated various practices, including (i) study preregistration, (ii) open materials and/or code, (iii) open data, (iv) pre-publication archiving, and (v) open access publishing. As they progressed through the survey, we also offered more detailed definitions of each of these open science practices.

2.1.1 Preregistration

We defined study preregistration as documenting and submitting to a journal or public repository one's research questions, methodological design, and anlysis plan prior to analysing the data. This time-stamped document is made openly available by the time the research is published so that any deviation from the original research plan is visible to the scientific community.

2.1.2 Open materials and/or code

We defined open materials and/or code as researcher-created resources used while collecting or analyzing data (e.g., survey questions, video stimuli, vignettes, algorithms, coding schemes, analytic code, etc.) that are made openly available to the research community.

2.1.3 Open data

We used Foster's definition of open data: "online, free of cost, accessible data that can be used, reused and distributed provided that the data source is attributed and shared alike" (Foster Open Science).

2.1.4 Pre-publication archiving

We definied pre-publication archiving as making a manuscript openly available *before* it undergoes peer review in an academic journal or other outlet. Generally, this is achieved by uploading the manuscript to an archive such as arXiv (physics, maths), bioRxiv (biology), SocArxiv (sociology), etc. We explained that pre-publication archiving is also known as preprint archiving.

2.1.5 Open Access Publishing

Finally, we explained that there are two basic models for open access publishing. First, researchers could pay their publisher a fee (an Article Processing Charge) to make the *final version of record* available for open access in a peer-reviewed journal. Second, researchers could publish under a normal contract in a peer-reviewed journal, and deposit the *final accepted version of the manuscript* (peer-reviewed but not journal-formatted) in an open access repository (e.g., institutional repository or an external subject repository).

3 Demographics

In terms of respondents' demographic information, we did not collect information about gender or age to provide more anonymity to our respondents. We did, however, ask about their academic level and their Discipline (based on 2-digit and 4-digit Field of Research codes).

3.1 Academic Levels

Of the 142 respondents who provided details about their academic level, 32% were PhD students, 16% were Professors, and 12% were Senior Lecturers. Table 1 reports the academic levels for respondents who answered the questions; 97 respondents chose not to report their academic level.

Table 1: Academic Levels of Respondents (n = 142)

Academic Levels	Frequency	Percentage
Professor	23	16
Associate Professor	10	7
Senior Lecturer	17	12
Senior Research Fellow	1	1
Lecturer	14	10
Research Fellow	9	6
Postdoc	11	8
PhD Student	45	32
Masters Student	7	5
Research Assistant	2	1
Other	3	2

3.2 Disciplines

In terms of Field of Research [FOR] codes, 138 respondents classified themselves as belonging to one of the two-digit FOR codes from 01 to 20, one person indicated "Other", and 100 respondents did not specify their FOR code. We did not have any responses from researchers in Earth Sciences (04), Agricultural and Veterinary Sciences (07), History and Archeology (21), and Philosophy and Religious Studies (22).

For ease of interpretation and to reduce the identifiability of the respondents, we combined similar divisions with smaller numbers of respondents into disciplinary groups (see Table 2). See Appendix 1 for details of how we combined FOR codes into disciplinary groups.

Table 2: Discipline Groupings of Respondents (n = 239)

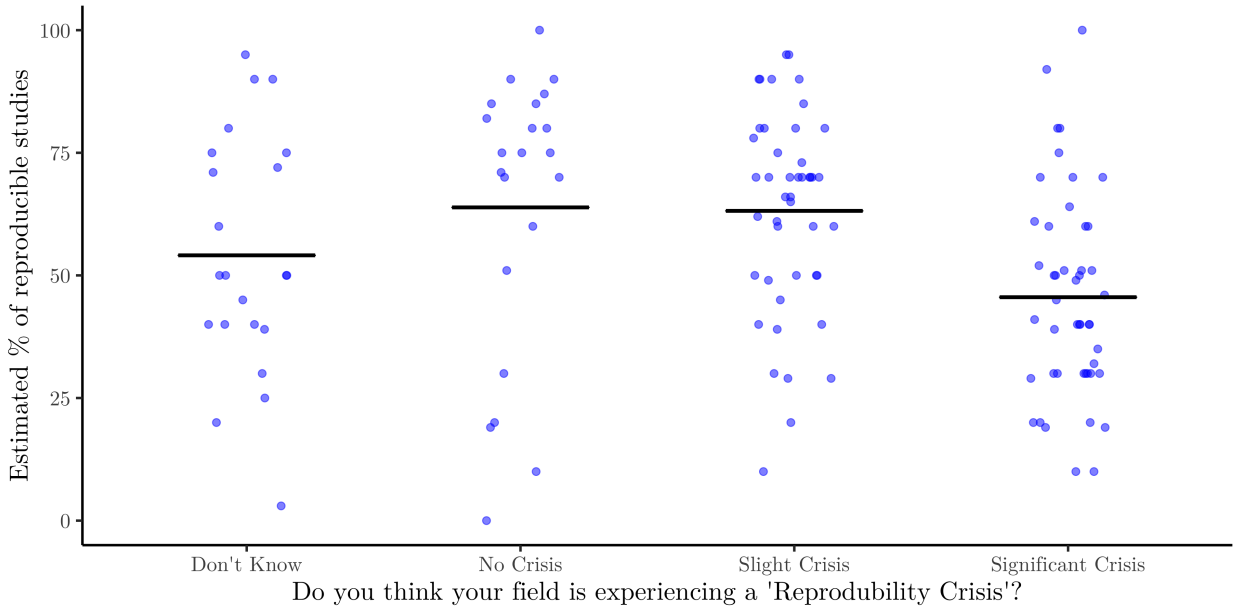
Disciplines	Frequency	Percentage
Math, Chem, Enviro, & Bio Sciences	13	5
Physical Sciences	14	6
Tech & Comp Sciences	7	3
Engineering	17	7
Medical & Health Sciences	13	5
Arts, Soc Sciences, & Humanities	12	5
Business & Law	11	5
Psyc & Cog Sciences	45	19
Other	7	3
Not Specified	100	42

4 Estimates of Reproducibility by Perceived Crisis

In line with Baker (2016), we asked participants if they believed their field is experiencing a "reproducibility crisis". Of the 138 respondents who answered this question, 17% indicated that they didn't know if there was a crisis, 16% indicated there was no crisis, 33% indicated that there was a slight crisis, and 35% that there was a significant reproducibility crisis in their field.

We also asked participants to estimate the percentage of research publications in their field that are reproducible. In Figure 1, we plotted participants' reproducibility estimates according to their perceived crisis in their field (n = 136). The black bar is the mean estimate of reproducibility by the perceived crisis levels.

Figure 1: Estimates of reproducibility by perceived crisis in field



5 Views and Use of Open Science Practices

5.1 Overall Experience with Open Science Practices

Before asking participants about individual open science practices, we asked about their experience with open science practices, in general (see Table 3). Of the 207 participants who answered this question, 25% reported that they were unaware of open science practices and only 6% reported that they had extensive experience with open science practices.

Table 3: Overall Experience with Open Science Practices (n = 207)

Reported Experience	Frequency	Percentage
Unaware	51	25
Aware, But Not Used	75	36
Some	69	33
Extensive	12	6

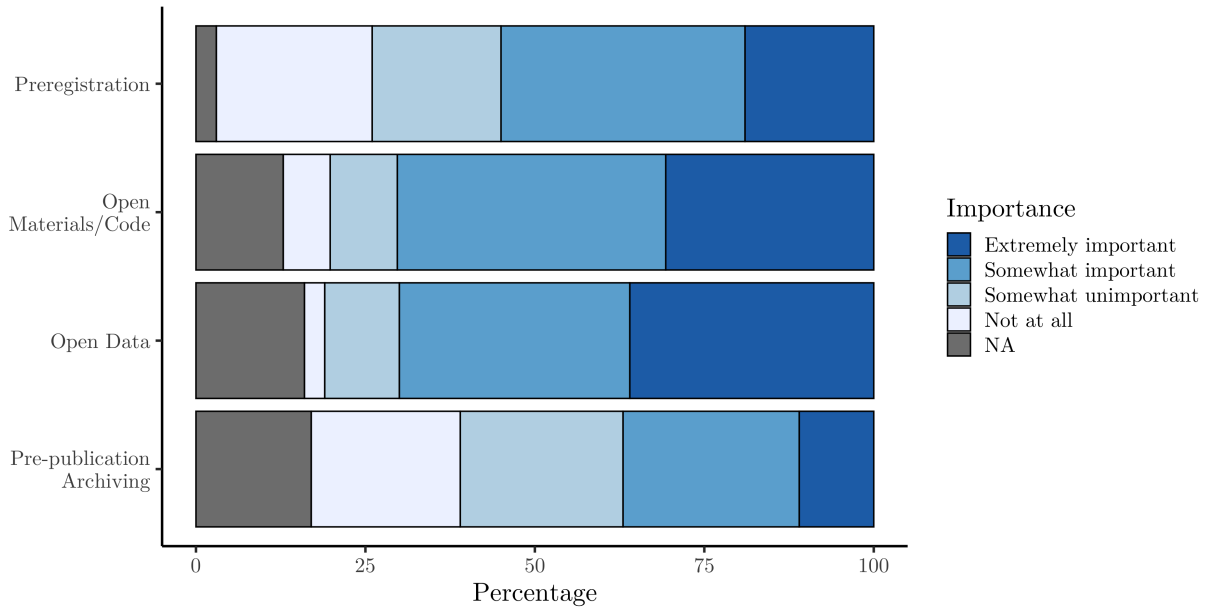
5.2 Perceived Importance of Open Science Practices

Throughout the survey, respondents separately rated the four main open science practices in terms of how important they are for their field. Specifically, we asked them about (i) preregistering their studies in

advance of data collection, (ii) making their study material and/or code openly available, (iii) making their data openly available, and (iv) making their work available on pre-publication archives.

Figure 2 illustrates the responses from the 194 respondents who answered the perceived importance question for at least one of the four practices. Open materials and/or code and open data were the most important practices for respondents' fields; nearly 75% indicated that these practices were important (either extremely important or somewhat important) for their field. Pre-publication archiving was the least important practice; only 36% indicated that it was important for their field.

Figure 2: Perceived importance of open science practices

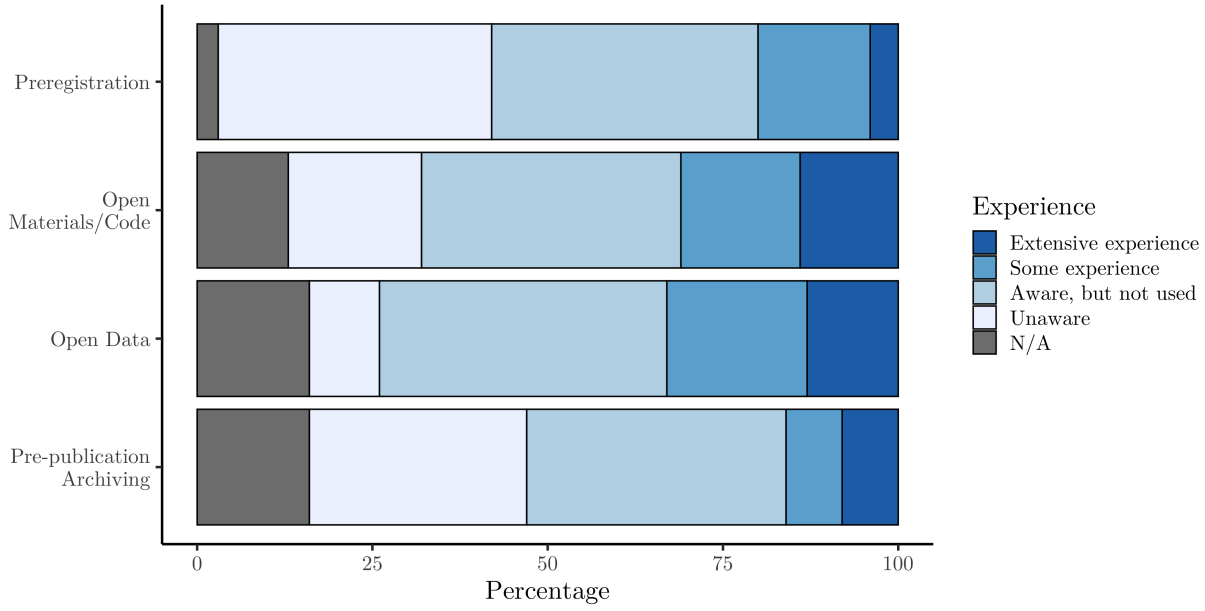


5.3 Experience with Open Science Practices

Participants also reported their own experience with each of the four main open science practices: (i) preregistering their studies in advance of data collection, (ii) making their study material and/or code openly available, (iii) making their data openly available, and (iv) making their work available on pre-publication archives.

Figure 3 includes the data from the 191 respondents who answered the experience question for at least one of the four practices. Our respondents reported very little experience with any of these open science practices. Respondents had the most experience with open materials/code and open data, but even with those practices, the most common response (37% for materials/code and 41% for data) was that they were aware of the practice, but had not personally used it. With regard to preregistration, respondents were largely unaware of the practice (39%) or were aware, but had not used it (38%). Only 16% of respondents reported any experience with pre-publication archiving.

Figure 3: Reported experience with open science practices



For participants who reported some experience or extensive experience with each of the four main open science practices, we asked them follow up questions to better understand how they have used the practice in their research activities.

5.3.1 Use of Study Preregistration

For those participants who reported that they preregister their studies, we asked them for more information about how they have used preregistration in their previous work. Respondents were allowed to select as many uses as applied. Of these 36 researchers, the most popular response (67%) was that they had preregistered at least one study that was *not* a clinical trial. A small percentage of respondents (8%) reported that they had submitted a Registered Report—a new type of manuscript format where the proposed method and analytic plan are peer-reviewed ***before*** data collection (for more information see the Center for Open Science).

Table 4: Reported Use of Study Preregistration (n = 36)

Reported Use	Frequency	Percentage
I have preregistered at least one study that was not a clinical trial	24	67
I have preregistered at least one clinical trial on a clinical trials database	10	28
I have published a research protocol prior to carrying out the study	10	28
Other	4	11
I have submitted a Stage 1 Registered Report to a journal	3	8

5.3.2 Use of Open Materials and/or Code

For those participants who reported that they use open materials and/or code (that is, those who reported some use or regular use), we asked them for more information about how they have used open materials and/or code in previous work. Respondents were allowed to select as many uses as applied. Of these 59

researchers, the most popular response (71%) was that they used other researchers' open materials and/or code in their own research. Few researchers reported using open materials and/or code when peer reviewing other researchers' work (14%).

Table 5: Reported Use of Open Materials and/or Code (n = 59)

Reported Use	Frequency	Percentage
I have used other researchers' open materials and/or code in my research	42	71
I have made my materials and/or code openly available	37	63
I have used open materials and/or code when reading other researchers' work	22	37
I have used other researchers' open materials and/or code when peer reviewing their work	8	14
Other use	2	3

5.3.3 Use of Open Data

For those participants who reported that they use open data (that is, those who reported some use or regular use), we asked them for more information about how they have used open data in previous work. Respondents were allowed to select as many uses as applied. Of these 62 researchers, the most popular response (81%) was that they used open data from other sources in their own research. Few researchers reported using open data when peer reviewing other researchers' work (13%).

Table 6: Reported Use of Open Data (n = 62)

Reported Use	Frequency	Percentage
I have used open data from other sources (e.g., researchers, government, etc.) in my research	50	81
I have made my research data open	30	48
I have used open data when reading other researchers' work	21	34
I have used open data when peer reviewing other researchers' work	8	13
Other use	1	2

5.3.4 Use of Pre-publication Archiving

For those participants who reported that they use pre-publication archiving (that is, those who reported some experience or extensive experience), we asked them for more information about how they have used these archives. Respondents were allowed to select as many uses as applied. Of these 30 researchers, the most popular response (87%) was that they read an article from, searched, or browsed a pre-publication archive. That is, they used other people's work posted to pre-publication archives for their own purposes. The next most common use by these researchers is to upload their own work to a pre-publication archive before submitting it to a journal (77%).

Table 7: Reported Use of Pre-publication Archiving (n = 30)

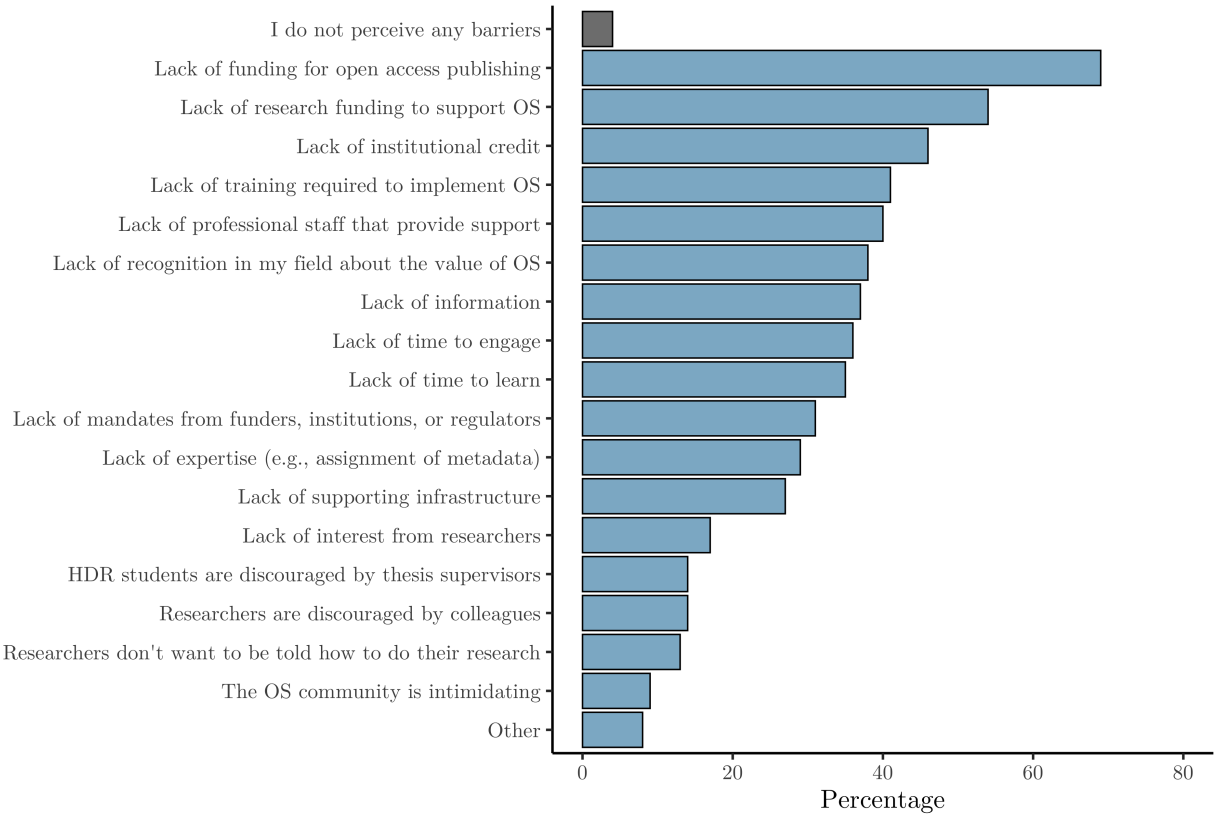
Reported Use	Frequency	Percentage
I have read an article from, searched, or browsed a pre-publication archive	26	87
I have uploaded a manuscript to a pre-publication archive before submission to a journal	23	77
I have cited a manuscript from a pre-publication archive	20	67
Other use	2	7

6 Perceived Barriers and Reported Concerns

6.1 Barriers to Engaging in Open Science Practices

We provided the respondents with a list of possible barriers to the uptake of open science practices in general. We asked them to endorse any statement that they agreed is a barrier in their field. Figure 4 shows the percentage of the 140 respondents who selected each response.

Figure 4: Barriers to the uptake of open science in your field

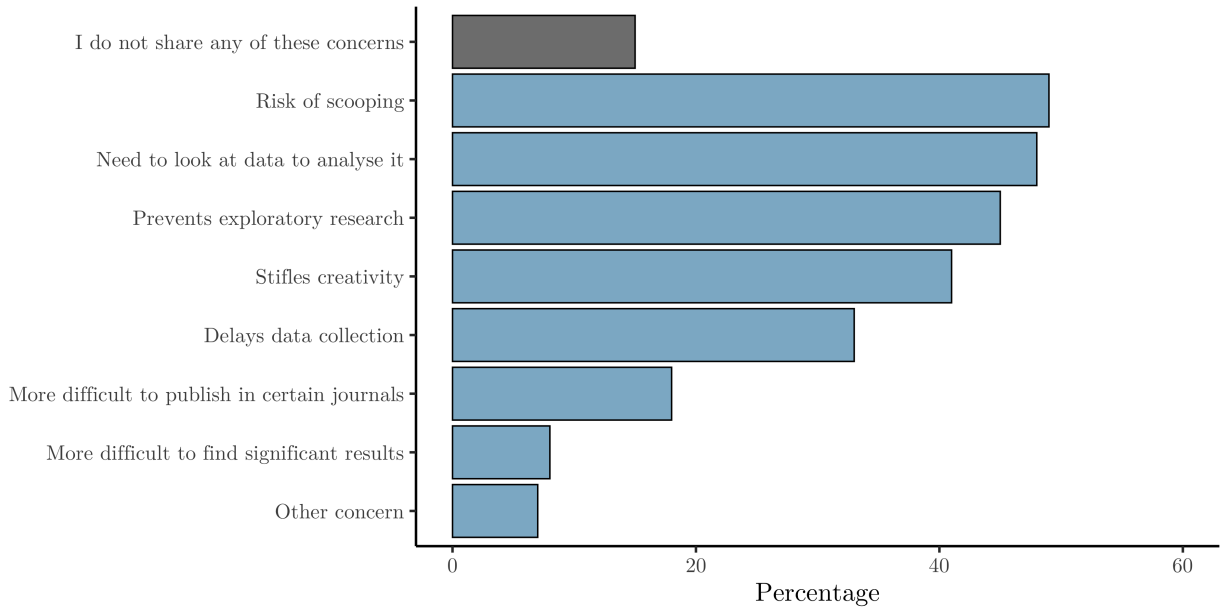


For each of the four main open science practices, we also asked respondents about their concerns about each of the practices. For each of the practices, we asked respondents to share their concerns in a free text format. Then, we provided a list of possible concerns that researchers could have about the relevant practice. From this list, respondents were able to select as many concerns as were relevant.

6.2 Concerns about Preregistration

Of the 168 respondents who answered the question about their concerns for preregistering research studies, 15% reported that they do not share any of the listed concerns. Figure 5 presents, of those who answered this question, the percentage who selected each of the concerns.

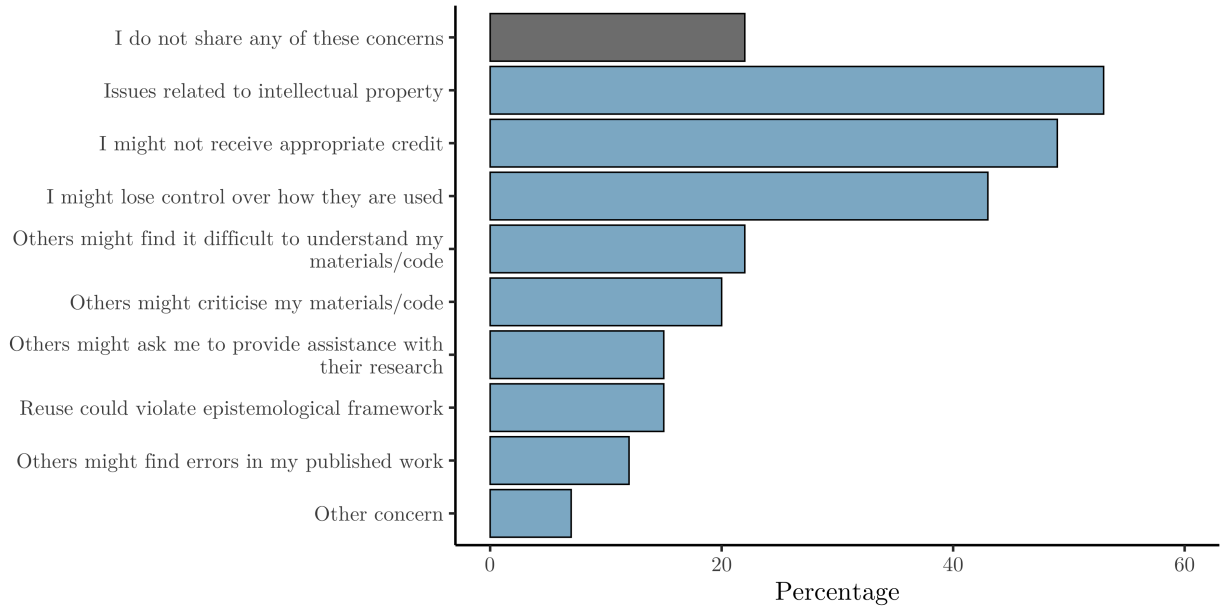
Figure 5: Concerns about preregistering studies



6.3 Concerns about Open Materials and/or Code

A total of 157 respondents answered the question about their concerns about making research materials and/or code open. Of those who responded, 22% reported that they do not share any of the listed concerns. Figure 6 presents, of those who answered this question, the percentage who selected each of the concerns.

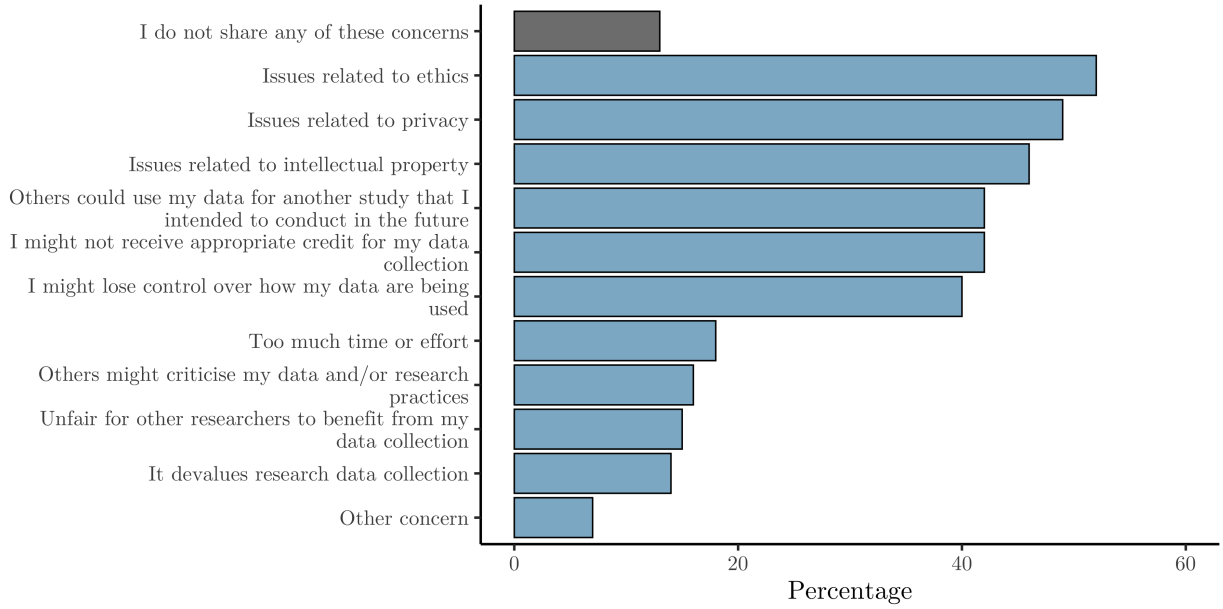
Figure 6: Concerns about open materials and/or code



6.4 Concerns about Open Data

A total of 153 respondents answered this question about concerns they had about making their data open. Of those who responded, 13% reported that they did not share any of the listed concerns. Figure 7 presents, of those who answered this question, the percentage who selected each of the concerns.

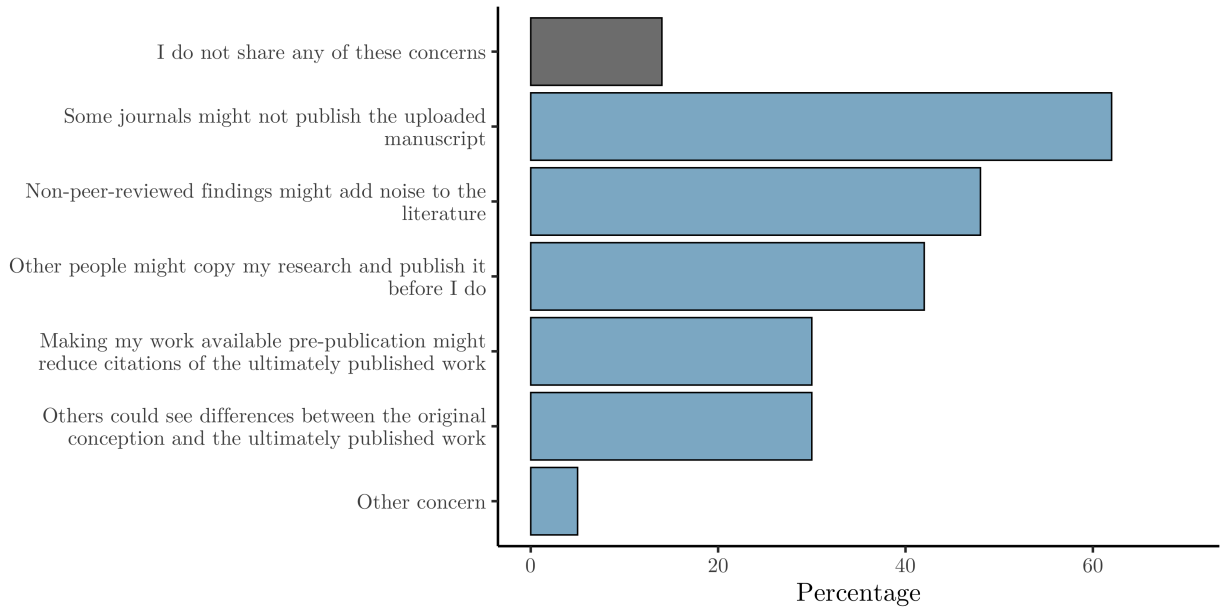
Figure 7: Concerns about open data



6.5 Concerns about Pre-publication Archiving

A total of 152 respondents answered this question about concerns they had about uploading a manuscript to a pre-publication archive before submitting it for peer review. Of those who responded, 14% reported that they do not share any of the listed concerns. Figure 8 presents, of those who answered this question, the percentage who selected each of the concerns.

Figure 8: Concerns about pre-publication archiving



7 Additional Open Science Issues

7.1 Open Access Publishing

The next section of the survey addressed Open Access publishing. We first explained the two different kinds of open access publishing: (1) paying an open access fee to make the final version of your published article open access, and (2) deposit the final accepted version of your manuscript (peer-reviewed but not journal-formatted) in an open access repository (e.g., Swinburne Research Bank or bioRxiv).

7.1.1 Proportion of Publications that are Open Access

We asked participants to indicate approximately what proportion of their publications from the last 5 years are open access (see Table 8). Of the 157 participants who answered this question, the most common responses were that some (33%) or none (29%) of their publications were open access. A small proportion (13%) indicated that they did not know what proportion were open access.

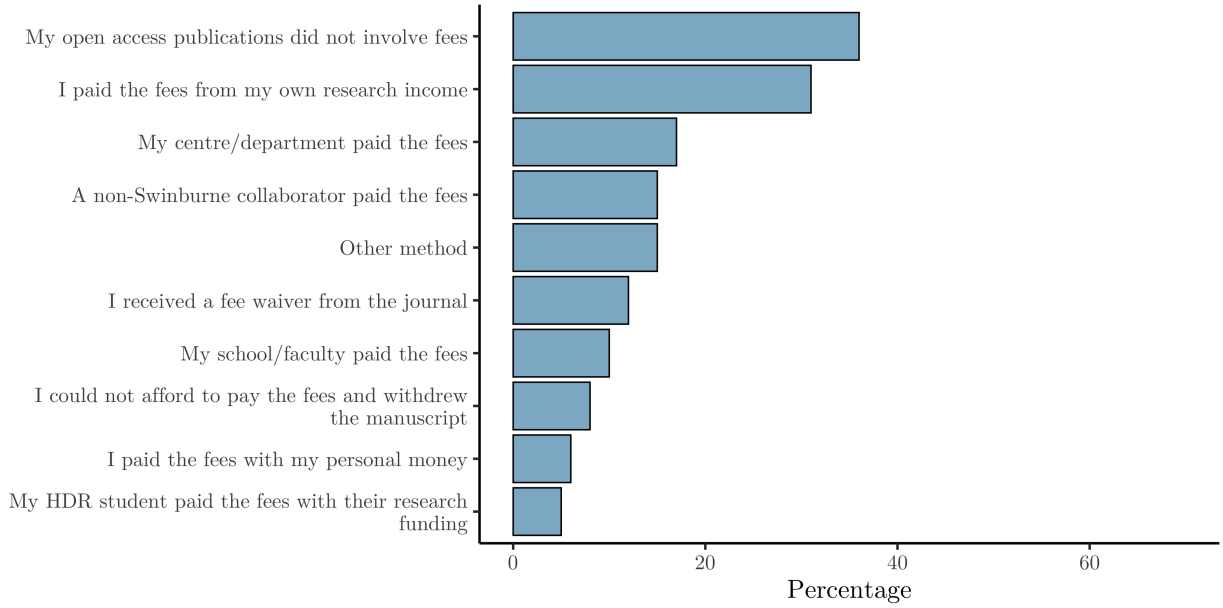
Table 8: Proportion of Publications that are Open Access (n = 157)

Response	Frequency	Percentage
I don't know	21	13
None	45	29
Some	52	33
Half	12	8
Most	17	11
All	10	6

7.1.2 Open Access Fees

We asked respondents, except those who said that none of their publications in the last 5 years were open access, how they paid for article processing charges that many journals charge for open access publishing. We provided a list of possible payment approaches and respondents were able to select as many concerns as they liked. Figure 9 presents, of those who answered this question (n = 100), the percentage who used each of the different payment approaches.

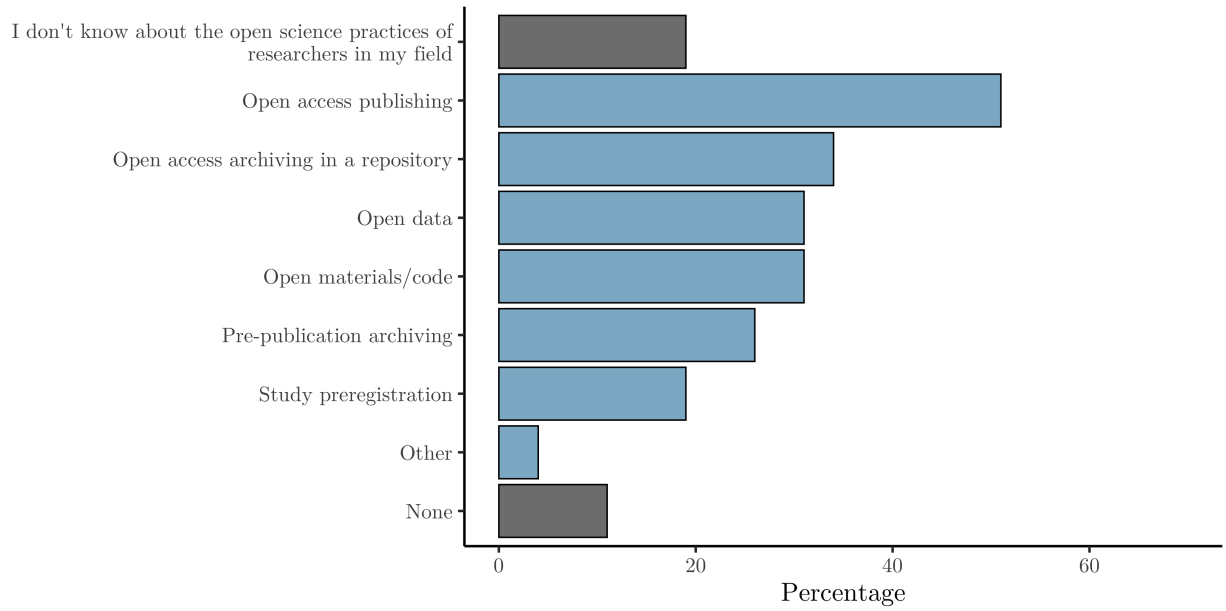
Figure 9: How respondents paid for open access publishing



7.1.3 Most Common Open Science Practices in Their Field

We asked respondents which of the following open science practices (if any) were commonly used in their field. They could tick all that applied. Figure 10represents the percentage of the 151 respondents who selected each response. Given that information about some of these open science practices are publicly available and often associated with the published product (e.g., the final paper regularly includes links to repository storing data, materials, and/or code, as well as a link to the preregistration, if relevant), it is disconcerting that nearly a fifth of respondents reported that they did not know about the open science practices of researchers in their field.

Figure 10: Most common open science practices in respondents' fields



8 Appendices

8.1 Appendix 1: Discipline Breakdown

As mentioned in Demographics, we merged FOR disciplines into disciplinary groups. Specifically, the new discipline grouping of Arts, Social Sciences, and Humanities (ASSH) included Education (FOR 13, n = 2); Studies in Human Society (FOR 16, n = 4); Studies in Creative Arts and Writing (FOR 19, n = 2); and Language, Communication and Culture (FOR 20, n = 4).

The new discipline grouping of Business & Law included Economics (FOR 14, n = 2); Commerce, Management, Tourism and Services (FOR 15, n = 7), and Law and Legal Studies (FOR 18, n = 2).

The new discipline grouping of Technology & Computer Sciences included Information and Computing Sciences (FOR 08, n = 4), and Technology (FOR 10, n = 3). We allocated Built Environment and Design (FOR 12, n = 6) to the "Other" category.

We retained the FOR divisions of Engineering (FOR 09, n = 17); Medical and Health Sciences (FOR 11, n = 13); Physical Sciences (FOR 02, n = 14); and Psychology and Cognitive Sciences (FOR 17, n = 45) as their own discipline groupings. Table 9 presents the breakdown for the number of respondents according to discpline groupings.

Table 9: Reported Two-Digit Field of Research Codes for Respondents (n = 239)

FOR Code	Field of Research Division	Frequency	Percentage
01	Mathematical Sciences	2	1
02	Physical Sciences	14	6
03	Chemical Sciences	4	2
05	Environmental Sciences	2	1
06	Biological Sciences	5	2
08	Information and Computing Sciences	4	2
09	Engineering	17	7
10	Technology	3	1
11	Medical and Health Sciences	13	5
12	Built Environment and Design	6	3
13	Education	2	1
14	Economics	2	1
15	Commerce, Management, Tourism and Services	7	3
16	Studies in Human Society	4	2
17	Psychology and Cognitive Sciences	45	19
18	Law and Legal Studies	2	1
19	Studies in Creative Arts and Writing	2	1
20	Language, Communication and Culture	4	2
99	Other	1	0
NA	Not Specified	100	42

8.2 Appendix 2: Perceived Importance Tables

Table 10: Perceived Importance of Preregistration (n = 188)

Perceived Importance	Frequency	Percentage
Not at all	45	24
Somewhat unimportant	37	20
Somewhat important	69	37
Extremely important	37	20

Table 11: Perceived Importance of Open Materials and/or Code (n = 169)

Perceived Importance	Frequency	Percentage
Not at all	13	8
Somewhat unimportant	19	11
Somewhat important	77	46
Extremely important	60	36

Table 12: Perceived Importance of Open Data (n = 162)

Perceived Importance	Frequency	Percentage
Not at all	6	4
Somewhat unimportant	21	13
Somewhat important	66	41
Extremely important	69	43

Table 13: Perceived Importance of Pre-publication Archiving (n = 161)

Perceived Importance	Frequency	Percentage
Not at all	42	26
Somewhat unimportant	46	29
Somewhat important	51	32
Extremely important	22	14

8.3 Appendix 3: Reported Experience Tables

Table 14: Experience with Study Preregistration (n = 185)

Reported Experience	Frequency	Percentage
Unaware	74	40
Aware, But Not Used	72	39
Some Experience	31	17
Regular Use	8	4

Table 15: Experience with Open Materials and/or Code (n = 167)

Reported Experience	Frequency	Percentage
Unaware	37	22
Aware, But Not Used	71	43
Some Use	32	19
Regular Use	27	16

Table 16: Experience with Open Data (n = 160)

Reported Experience	Frequency	Percentage
Unaware	20	12
Aware, But Not Used	78	49
Some Use	38	24
Regular Use	24	15

Table 17: Experience with Pre-publication Archiving (n = 161)

Reported Experience	Frequency	Percentage
Unaware	60	37
Aware, But Not Used	71	44
Some Experience	15	9
Extensive Experience	15	9

9 References

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature News, 533* (7604). Retrieved from https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970.

Center for Open Science (nd). Registered Reports. Retrieved from https://www.cos.io/our-services/registered-reports

Foster Open Science (nd). Open data definition. Retrieved from https://www.fosteropenscience.eu/taxonomy/term/110.