

Empirical Study of PhilSci Archive Postings from Three Journals

Ananya Chatteraj, Brian Hanley, and C. Kenneth Waters
University of Calgary

We collected preliminary data to investigate three questions concerning the posting of preprints of articles appearing in academic journals. Our analysis of the data indicates possible strategies to increase preprint postings of articles accepted for publication in prominent journals.

Our questions:

1. What proportion of articles published in mainstream philosophy of science journals have been posted (as preprints) on *PhilSci Archive*? The Board keeps track of how many preprints are posted but does not have data relevant to this question.
2. Have editors' recommendations to post preprints of articles accepted for publication in their journals increased postings? Several years ago, the board decided to ask journal editors to recommend to authors whose manuscripts have just been accepted for publication to post preprints of their articles on *PhilSci Archive*. A number of editors agreed. The Board has not kept track of whether editors are continuing to make this recommendation to authors and we have not collected data to determine whether editors' recommendations have increased postings.
3. Does posting preprints on *PhilSci Archive* increase citations of published articles? Again, the board does not have data to address this question.

We decided to collect data from a small number of journals to make the preliminary collection of data manageable. We chose three prestigious journals: one representing journals that publish across philosophy of science generally (*Philosophy of Science*); one representing journals specializing in the philosophy of a particular area of science (*Biology and Philosophy*), and a third representing general epistemology journals that regularly publish articles in philosophy of science as well as articles outside philosophy of science (*Synthese*). One could ask whether these journals are indeed representative or whether the results would differ for articles published in journals representing particular philosophical perspectives (e.g. the history and philosophy of science or the history of philosophy of science). But we decided to limit our data collection to three journals and examine the results before deciding whether the project should be extended.

We contacted the editors of the three journals to check whether they have been continuously recommending that accepted offers post preprints. They responded positively. The editors of all three journals started to make their recommendations in the beginning of 2017. But anecdotal evidence about *Synthese* seems to suggest that some authors writing articles, which we consider to

fall within the Archive's scope, have not received recommendations to post their preprints. Furthermore, at least one of these authors believes the *Synthese* policy requires authors to wait a year after the appearance of their article in the journal before posting a preprint. Perhaps the joint editorship of *Synthese* and/or the fact that only a portion of the accepted articles fall within the scope of the Archive's scope has led to inconsistencies in communication. Anecdotal evidence also suggests that editors' recommendations to post preprints is going unnoticed by authors.

We collected information about every article published in the years 2013 – 2019 for each journal.

- For articles published in *Philosophy of Science*, we collected information about the apparent subarea of the article (e.g. philosophy of physics), we checked to see whether a preprint of the article was posted on *PhilSci Archive*, we drew on *PhilSci Archive's* Plum Analytics to determine the number of times the preprint was download (if the preprint was posted), and we used Google Scholar to estimate the number of times the journal article was cited (regardless of whether a preprint was posted).
- For articles published in *Biology and Philosophy*, we assumed all articles were in the subarea of philosophy of biology. Otherwise, we collected data as we did for the *Philosophy of Science* articles concerning whether a preprint of the article was posted, the number of downloads of preprints if posted, and the number of times the article was cited.
- *Synthese* publishes many articles in areas of philosophy outside the scope of the *PhilSci Archive*. So for each article published in *Synthese*, we quickly assessed whether it fell into an area included in the *Archive*. We used this assessment as a filter and collected data for only the articles we deemed fell within the scope of the *Archive*. For these articles, we collected data as we did for the *Philosophy of Science* articles concerning subarea of the article, whether a preprint of the article was posted, the number of downloads of preprints if posted, and the number of times the article was cited.

Our data is contained in an accompanying excel file labelled "PhilSci Archive Data Collection.xlsx".

Our results indicate that the proportions of published articles that were uploaded to *PhilSci Archive* are relatively low. The journal *Philosophy of Science* shows the highest upload rate and was the only of the three journals to have years with upload rates greater than 50%. The rate of upload did not dramatically differ across different subareas of philosophy of science. This was a surprise because the proportion of preprints in the subarea of philosophy of physics is significantly greater than of those in other subareas (including philosophy of biology). The rate of upload to *PhilSci Archive* increased slightly in the latter part of 2017 for *Philosophy of Science* and *Synthese*. This is correlated with the timeframe when editors began recommending to authors that they upload their papers to the *Archive*.

The results also indicate that papers uploaded to the archive regularly receive many, even hundreds, of downloads within the first year of the paper being uploaded to the *Archive*. Of course, the number of downloads increases over time.

Our data also reveals that there is a correlation between the upload of preprints and the number of citations articles receive. In most years we surveyed, both the mean and median number of citations for papers uploaded were greater than the mean and median number of the citations for papers not uploaded. These citation rates did not differ greatly among the three journals. Of course, correlation is not causation. In addition, even if it's true that posting preprints increases citations of articles published in these three journals, it might not be true of articles published in other journals. On the other hand, posting preprints of articles published in less prominent journals might have greater impact on citation rates than posting has for articles published in these three prominent journals. Perhaps further investigation is indicated.

The data also revealed an interesting outlier. The April 2019 issue of *Biology and Philosophy* had an unusually high posting rate of nearly 70% (eleven of sixteen articles). It turns out that this issue included a special collection of nine articles on the topic of Paleobiology and Philosophy and preprints of all nine of these articles were posted. We contacted Adrian Currie, who organized the special collection. He reported that he asked each author to post their preprint and then followed up with a second request afterwards.

We believe these results suggest strategies for increasing the number of preprint postings.

- It appears that journal editors' recommendation to post preprints is having a positive effect. Perhaps the Archive Board should keep track of whether this practice is being continued and contact new editors to congratulate them on becoming editors and to call attention to the journal's ongoing practice of recommending posting. Currently the Board does not keep track of editors' recommendation to post preprints.
- Perhaps we should make the results of our data collection available in promotional materials for authors in general, and in special communications with journal editors. Authors might be motivated by the statistics on download rates for preprints and correlation between posting and citation rate. Journal editors might be motivated by the correlation between posting and citation rate. We might submit a poster on the archive highlighting these results to conferences. (We have presented posters on the archive in the past.) We might consider contacting editors of a number of journals to find out whether they are (still) recommending posting preprints and to share the results of this study.
- The 100% posting rate of nine articles appearing as a special collection in *Biology and Philosophy* suggests that following up on an editors' recommendations to post preprints might be a very effective way to increase postings. Representatives of the *Archive* might try to work with individual editors of a few key journals to explore ways to systematically follow up on an editor's initial recommendation that authors post preprints of their forthcoming articles. If we did this, we should see whether the rate of postings increases dramatically to determine whether the follow-up procedures should be continued with these journals and whether we should encourage additional journal editors to follow this lead.