

The hidden map of science: Pre-publication history of articles tells us that rejection leads to higher citations

 blogs.lse.ac.uk/impactofsocialsciences/2012/10/25/calcagno-prepublication-history-citations/

10/25/2012

*No-one wants to have their paper rejected by a top journal, but is there a silver lining to an initial disappointment? **Vincent Calcagno** finds that papers that are resubmitted to a second or third choice journal enjoy a 'benefit of rejection' and are more likely to receive a higher number of citations when published.*



Every scientific author or editor has his/her own (limited) experience of what the publication process is like. It is always frustrating to be rejected by a journal, having to reformat a manuscript for another journal, and the reasons for rejection are not always clear/understood/accepted. I am no exception and as a PhD student, I had a piece of research I liked rejected 6 or 7 times in a row by different journals. How frequent is this? Is it worth it? I wondered.

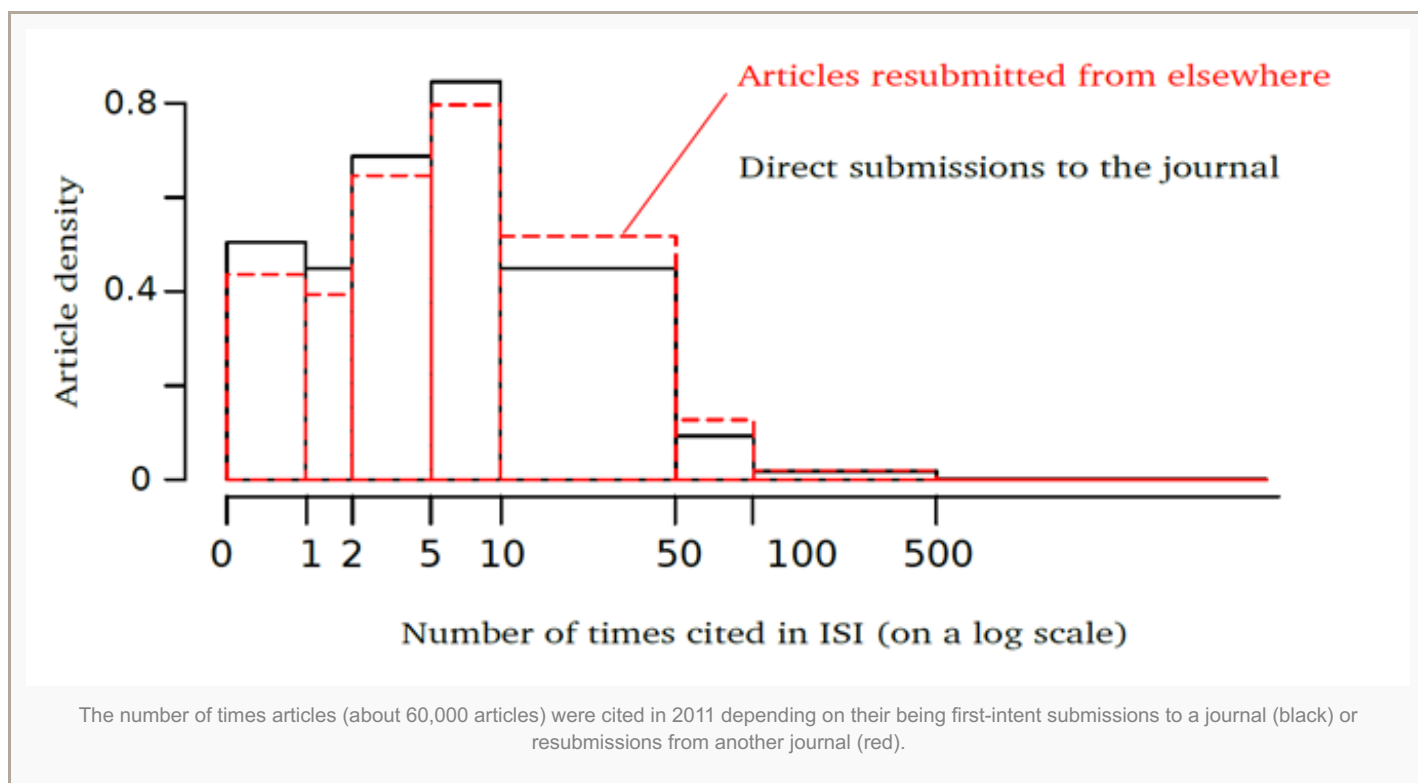
I quickly realized that we lacked systematic data on the pre-publication life of research contributions: journals themselves do not know where received manuscripts come from, or the fate of the manuscripts they reject. Only authors know, and this is why I decided to ask them. With a little bit of programming, I downloaded from ISI WoS all references published by journals in a selection of 16 categories of the Biological Sciences (my own field). This made for almost 1,000 journals and over 200,000 articles published in years 2006-2008. I then sent a customized email to each corresponding author, asking simply, "Was it the first journal you had tried? If not, what journal before?" This plain emailing approach certainly complicates subsequent analysis, but I wanted to maximize the response rate by requiring minimal effort to respondents. Maybe it worked, since we got about 100,000 emails back by August 2009 and the flood of responses temporarily crashed the email servers of the university I was in.

After more programming to analyze the answers we retrieved the pre-publication history of a little over 80,000 articles. As an ecologist, I like to see the jungle of scientific publishing as an intricate food-web in which journals are dreary predators competing for the valuable food items that research findings and manuscripts are. They sometimes catch the food directly from the source, and sometimes recycle food-items previously rejected by another predator. With our data, we thus reconstructed the "food web" of scientific journals, where trophic links really are publication of an article previously rejected by another journal. This is the first "submission network" of this sort: bibliometric analyses that have become common in the past 10 years typically use citation links between journals (who-cite-whom), or co-authorship patterns, but submission links had remained hidden. The results of this study have just been published in [Science](#).

Clearly, feedback from participants confirmed that individual experiences differed greatly among scholars: typical comments ranged from, "Of course it was the first journal. In more than 15 years I have never had a paper rejected", to "We had tried XX before but as usual were rejected". I personally expected, from experience and a vague notion of rejection rates, a majority of published articles to be resubmitted from another journal before successful publication. But, surprisingly, we found that about 75 per cent of all articles we declared to have been submitted to the publishing journal on first intention. Even assuming that, for some reason, authors were less likely to respond in the case of a resubmission, we still find that a majority of published articles are first-intent submissions. This suggests that authors are, overall, quite apt at targeting a proper journal and, conversely, that journals make sure they have a sufficient public: no journal was found to be entirely dependent on resubmissions from others.

What else can we learn from the submission network? Well, network analysis techniques ('science mapping') revealed that journals were connected in a way consistent with the subject categories of ISI. This was expected, obviously. Another expected finding was that the citational impact of journals (ISI IF) influenced submission patterns: journals more central in the submission network were also those with higher impact factor. Several factors can

explain this result. One component is of course that authors struggle for high-impact journals.



The other, less intuitive component, is that high impact journals tended to publish *fewer* first-intent submissions (and thus more articles that had been previously sent elsewhere). This fraction goes from about 90 per cent down to about 50 per cent from IF 1 to IFs 10-15. A naive “struggle-for-impact” argument would lead the opposite prediction: high-impact journals are tried first and thus publish many first-intent submissions, whereas low-impact journals are more often second-choices and publish more resubmissions. Our result means that, although preferred, high-impact journals are also in stronger competition for research manuscripts: they are connected to more competing journals, and the rejection rates in their neighborhood are higher. Overall, they are thus more likely to receive (and publish), resubmissions from another journal. This topological network effect is enough to revert the expected trend. Exceptions are Nature and Science, who do attract and publish mostly first-intent submissions as we would expect, even though they exchange a significant number of manuscripts and thus remain well below 100 per cent (90 and 80 per cent respectively).

We should stress that impact factor does not explain most of the variation that we observed among journals, suggesting that, although it is a valid metric of journal importance and it does influence authors’ submission behaviors, there is much more involved.

An intriguing possibility would be that submission-history affects the post-publication impact of articles. In 2011, i.e. 3-5 years post-publication, we compared the citation count (from ISI) of first-intent submissions and resubmissions from other journals.

Considering the extreme variability of citation count data, and the thousands of (hard-to-predict) factors affecting them, I honestly did not expect to detect anything. Interestingly, we did find significant differences: in a given journal and a given year, an article that had been resubmitted from another journal was on average more cited than a first-intent submission. Resubmissions were less likely to receive zero or one citation (about 15 per cent less, controlling for publication year and journal) and more likely to receive several (e.g. 10 and 50) citations, shifting the mean to higher values. This intriguing result suggests a “benefit of rejection”. The simplest explanation would be that the review process and the greater amount of time spent working on resubmitted manuscripts does improve them and makes them more cited, although other mechanisms could be invoked.

Another influence of pre-publication history was that resubmissions between journal groups (i.e. journals that are tightly connected) were less cited than resubmissions within the same journal group. This suggests that crossing boundaries is a risky move when resubmitting, with interesting implications for multi-disciplinarity. These first results indicate that manuscript submission history can have non-trivial consequences and could be a valuable datum when studying the process of science making.

Note: This article gives the views of the author(s), and not the position of the Impact of Social Sciences blog, nor of the London School of Economics.

About the author:

Vincent Calcagno is an evolutionary biologist and ecologist working at the French Institute for Agricultural Research (INRA).

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.