

Introduction : From challenges to solutions in book evaluation

Galleron, Ioana^{1 2}

¹ Université Sorbonne Nouvelle, France

² COST Action 15137

DOI [10.5072/zenodo.1205027](https://doi.org/10.5072/zenodo.1205027)

TO CITE

Galleron, I. (2023). Introduction : From challenges to solutions in book evaluation. *Proceedings of the Paris Institute for Advanced Study*, 4. <https://doi.org/10.5072/zenodo.1205027>

PUBLICATION DATE

16/05/2022

ABSTRACT

Evaluation of Social Sciences and Humanities in Europe. Hcéres Colloquium Proceedings - Paris IAS, 16-17 May 2022 - Session 1 "Books and Monographs" - Books in Evaluation - Introduction

Once upon a time, when hiring junior or senior professors in a French university, in history, linguistics or literary fields, the members of the jury were required to have strong arms, good backs and high speed in reading. The reason was that, previous to the interviews, they were supposed to receive samples of the scientific production of the candidates, and these samples were, quite often, books. In those not so remote times, one could, of course, send offprints of separate chapters, and candidates often did so, but it wasn't rare to receive quite voluminous mail boxes with several books. When performed on hard copies of the research output (as required, for instance, in older sessions of REF scheme, see <https://www.ref.ac.uk/>), evaluation of research units or of higher education institutions rendered the problem even more acute, with several cubic meters of prints to be gathered and read by the panelists.

The digitization of the publications took care of the weight problem, but left intact the need – and the difficulty – to read increasing volumes of text in short periods of time. This difficulty is not solely related to books, since journal articles have also multiplied over time, but books are maybe the most prominent part of the problem. On the one hand, the overall book production increased, and this not only in the UK, as shown in

the Crossik report quoted by Pierre Mounier. On the other hand, some book formats, such as the doctoral dissertations, traditionally seen as a first book in several disciplines, tended to become more and more voluminous. Sections 9, 10, 21st and 22nd of the French CNU have received, in recent years, dissertations of more than 1000 pages. Even if they remain rather exceptional, such figures are symbolic about the inflationist trend of the scientific publications in general, and of books in social sciences and humanities in particular.

Faced with these realities and short timeframes for accomplishing their missions, most evaluation agencies and exercises cannot but resort to proxies for book evaluation. While panels can be directed to read selected papers and abstracts, it is unrealistic to expect them to carefully browse hundreds of pages written by one or several members of an evaluated research unit, all the more so as these units tend, in turn, to become more populated, for various reasons such as the lengthening of the careers and the demographics of the research sector. In these conditions, book evaluation becomes more a matter of filters and hierarchies, as Gunnar Sivertsen puts it in his contribution to this volume, than of an informed judgement made on the basis of a close and personal involvement with the text in question.

Both filters and hierarchies come with complex questions to answer and challenges to be solved. A first problem any panel faces is defining what is a book, with four main criteria (authors, formats, contents and public) displaying a large variety of situations. Should monographs be exclusively taken into consideration, while conference proceedings, collective works, scholarly editions and other types of texts are to be considered separately? To a certain extent, monographs are a kind of “crown output”, considering the amount of work they require and the breadth of reflection they display (see Sivertsen in the following pages); there is also some evidence that, in the ever-evolving ecosystem of scholarly publication, they tend to remain quite stable over the time, representing some 7% of the total output ([Williams et al., 2018](#)), and offering therefore an interesting basis of comparison. But what is a monograph, in this case, knowing that co-authored books are more and more numerous, even in disciplines where scholarly work was traditionally a more individual (and even solitary) activity? And how long should it be, if two (hard) covers and an ISNI number are not to be taken as an indicator of scientific quality? A corollary question is that of the measuring unit, with

some disciplines and countries counting the number of characters, others the number of words, and some the number of typographic sheets (ex. 8 typographic sheets = a book). In addition, are all publications satisfying the requirements in terms of authors and length to be taken into consideration as books, even if they target students or the large public?

Once these preliminary questions answered, the evaluation of books does not become straightforward. In most cases, a book is considered to be good because it has been accepted and marketed by a specific publishing house, or because it received certain distinctions, or, less frequently, on the basis of the reviews it received. The rationale behind this is that, within a short timeframe, it makes sense to get back to ex ante evaluation, or to existing ex post evaluation, rather than performing it on the volume itself.

Unfortunately, as shown by Elea Gimenez Toledo in various studies (see, amongst others, [Giménez-Toledo et al., 2013](#)), publisher lists are often reputation based, rather than rooted in a thorough examination of the national scholarly publishing systems, for which she advocates in her present contribution. Panelists knowledge remains limited about existent or inexistent policies in publishing houses with regards to the transparency of manuscript selection, the evaluation process and criteria, the procedures in place for coping with contestations and suspicions of scientific misconduct, and so on ([Giménez Toledo et al., 2014](#)). Services offered by publishers do not seem either to be taken into consideration when establishing what a “good publishing house” is, even if scholars are numerous to complain about having to accomplish typesetting tasks previously taken care of by the specialists in the publishing industry. National surveys and peer review labels, such as the one put into place in Belgium or Finland (see [Kulczycki et al., 2019](#)), seem, in this respect, the way forward, if publishers are to remain an acceptable proxy for quality.

At the other end of the spectrum, reviews put other types of problems when used as proxies for quality in an evaluation exercise. In both cases, the main challenge is their scarcity, and the timeframe for their production. Not all books are reviewed, not for the same journals, and not with the same set of criteria in mind. Several months, if not years, separate the publication of a book and its review in a journal. Also, ex post

reviews, necessarily not blind, are characterized by specific linguistic traits, adopted because of the need to maintain a certain collegiality (see, for instance, [Itakura, 2013](#)). Criticism is therefore less direct, and the general tendency is towards descriptivism and praise. These specificities render even more difficult the establishment of hierarchies, all the more so as, in most cases, evaluation is not about distinguishing between excellent and bad books, but between various types of middle ground, scholarly solid, and all in all good books.

In recent years, other proxies for quality have been found in the number of on-line consultations, or downloads, of digital books ([Hammarfelt, 2014](#)). Without getting into any detail, let's remind that such measures are at least as questionable as impact factors for papers in journals. A more promising approach consists in organizing a real post-publication peer review in the digital sphere, or at least in offering the possibility to add comments to digital books. It seems, however, premature to count on these for a quantitative aid to a qualitative, peer conducted research evaluation of books.

Challenges posed by the books when taken into account in evaluation exercises are not limited to the few prominent problems listed above. In her case study, Natasa Jerman details other questions, and the solutions brought within the Croatian system. Overall, as shown by a survey organized within ENRESSH COST action (<https://enressh.eu/>), responses to these various challenges, given by national frameworks and evaluation agencies, span from very detailed and strict guidance to a complete lack of specifications. The ideal mix between common specifications and criteria, and the exercise of an individual judgement by peers is probably still to be found, but the worse is refusing to face the problems, or solving them with a top-down and bureaucratic approach. In depth studies about motivations, such as the one described by Geoffrey Williams in his contribution, are needed to better understand how to stimulate the production of scholarly, and not career motivated, books. They have to be completed by surveys about national publishing systems, and by a more refined understanding of peer review practices.

All in all, as shown in the following contributions, books are not an oddity of SSH disciplines, nor an invasive species. With their specificities and challenges, they are a chance for evaluation frameworks and systems to better think and adapt their

procedures, so as to contribute to the maintenance of vibrant and successful research communities.

Bibliography

Giménez Toledo, E., Fernández Gómez, S., Tejada Artigas, C., & Mañana Rodríguez, J. (2014). *From Book Publishers to Authors: Information Transparency in Web Sites*.
<https://doi.org/10.3138/jsp.46.1.004>

Giménez-Toledo, E., Tejada-Artigas, C., & Mañana-Rodríguez, J. (2013). Evaluation of scientific books' publishers in social sciences and humanities: Results of a survey. *Research Evaluation*, 22(1), 64–77.
<https://doi.org/10.1093/reseval/rvs036>

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics*, 101(2), 1419–1430. <https://doi.org/10.1007/s11192-014-1261-3>

Itakura, H. (2013). Hedging praise in English and Japanese book reviews. *Journal of Pragmatics*, 45(1), 131–148. <https://doi.org/10.1016/j.pragma.2012.11.003>

Kulczycki, E., Rozkosz, E. A., Engels, T. C. E., Guns, R., Hołowicki, M., & Pölönen, J. (2019). How to identify peer-reviewed publications: Open-identity labels in scholarly book publishing. *PLOS ONE*, 14(3), e0214423. <https://doi.org/10.1371/journal.pone.0214423>

Williams, G., Basso, A., Galleron, I., & Lippiello, T. (2018). More, Less or Better: The Problem of Evaluating Books in SSH Research. In A. Bonaccorsi (Ed.), *The Evaluation of Research in Social Sciences and Humanities: Lessons from the Italian Experience* (pp. 133–158). Springer International Publishing. https://doi.org/10.1007/978-3-319-68554-0_6