This item is the archived peer-reviewed authors' version of:

Taking scholarly books into account, part II: a comparison of 19 European countries in evaluation and funding.

Elea Giménez-Toledo; Jorge Mañana-Rodríguez, Tim C. E. Engels; Raf Guns; Emanuel Kulczycki; Michael Ochsner; Janne Pölönen; Gunnar Sivertsen; Alesia A. Zuccala

Taking scholarly books into account, part II: a comparison of 19 European countries in evaluation and funding.

Elea Giménez-Toledo; Jorge Mañana-Rodríguez, Tim C. E. Engels; Raf Guns; Emanuel Kulczycki; Michael Ochsner; Janne Pölönen; Gunnar Sivertsen; Alesia A. Zuccala

Affiliations:

Elea Giménez-Toledo and Jorge Mañana-Rodríguez. Research Group on Scholarly Books (ILIA). Institute of Philosophy (IFS). Spanish National Research Council (CSIC)

Tim C. E. Engels and Raf Guns : Centre for R&D Monitoring (ECOOM), Faculty of Social Sciences, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium

Emanuel Kulczycki: Scholarly Communication Research Group, Faculty of Social Sciences, Adam Mickiewicz University in Poznań, Poland

Michael Ochsner. Swiss Center of Expertise in the Social Sciences, FORS, University of Lausanne, Switzerland and GESS, ETH Zurich, Switzerland

Janne Pölönen. Federation of Finnish Learned Societies, Snellmaninkatu 13, 00170 Helsinki, Finland.

Gunnar Sivertsen. Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway.

Alesia A. Zuccala. Department of Information Studies, University of Copenhagen, Copenhagen, Denmark

MSC Classification: 00-02

JEL Classification: C00

Abstract

In May 2016, an article published in *Scientometrics*, titled '*Taking scholarly books into account: current developments in five European countries*', introduced a comparison of book evaluation schemes implemented within five European countries. The present article expands upon this work by including a broader and more heterogeneous set of countries (19 European countries in total) and adding new variables for comparison. Two complementary classification models were used to point out the commonalities and differences between each country's evaluation scheme. First, we employed a double-axis classification to highlight the degree of 'formalization' for each scheme, second, we classified each country according to the presence or absence of a bibliographic database. Each country's evaluation scheme possesses its own unique merits and details; however the result of this study was the identification of four main types of book evaluation systems, leading to the following main conclusions. First, countries may be differentiated on the basis of those that use a formalized evaluation system and those that do not. Also, countries that do use a formalized evaluation system either have a supra-institutional database, quality labels for publishers and/or publisher rankings in place to harmonize the evaluations. Countries that do not use a formalized system tend to rely less on quantitative

evaluation procedures. Each evaluation type has its advantages and disadvantages; therefore an exchange between countries might help to generate future improvements.

Keywords: Scholarly books, book publishers, evaluation processes, classification. Research evaluation, Social Sciences; Humanities, book series

Introduction

Within the Social Sciences and the Humanities, books and book chapters have an outstanding role in the publication patterns. Different analyses of outputs show how relevant they are in terms of number (e.g. Michavila, 2012; Engels, Ossenblok & Spruyt, 2012; Kousha, Thelwall & Rezaie, 2011). This fact makes relevant to study criteria and procedures for evaluating books in research evaluation processes.

A comparative study carried out two years ago allowed having information on how books were being evaluated in Denmark, Finland, Flanders (Belgium), Norway and Spain (Giménez-Toledo et al., 2016). In the present study this comparison has been extended to include Croatia, Czech Republic, France, Israel, Italy, Latvia, Lithuania, Montenegro, Poland, Portugal, Serbia, Slovakia, Slovenia, and Switzerland.

The European Network for Research Evaluation in the Social Sciences and Humanities (abbreviated ENRESSH, COST action CA15137) has provided a solid structure for developing the present study. It is through ENRESSH that we have been able to: a) gather a relevant network of experts, b) obtain data from the different countries, c) develop comparisons between the selected countries, and d) generate a discussion of the issues, methodologies and practices concerning evaluation of books in the SSH across Europe.

Countries that already have a defined evaluation system in place are in a good position to facilitate new approaches to implementing evaluation systems in other countries. Our hope is that such exchanges will become fructiferous, not only through the provision of this overview, but also through the identification of common patterns and potential convergences in methodologies and results.

Objectives

The main objective of this work is to analyze how books are evaluated at the level of publishers, book series and/or individual titles in Europe. Participating countries are listed in the first column of Table 1. We refer to Flanders (Belgium) as a country, even though it is legally a region. It simplifies the nature of this study because in Belgium the regions are autonomous in terms of education and research policy.

Other objectives of this work are:

- To group the different approaches to the evaluation of scholarly books
- To examine the possibilities and limitations attached to the different evaluation processes regarding the role of books.

Methodology

This study is based on a structured questionnaire (see Appendix I), which was sent to the representatives of each ENRESSH country who attended two general meetings of the network, and subsequently accepted an invitation to become a participant. It is relevant to mention that the representatives are specialized in SSH evaluation and, in many cases, experts on the subject of scholarly book evaluations. The first question of the questionnaire was used to identify the kind of evaluation process in place for evaluating scholarly books in each country, both in terms of data infrastructures and tools (e.g., Current Research Information Systems or CRISs, and approved list of publishers, categorizations of publishers, quality labels, etc.). Once the evaluation processes were identified, the respondents were asked for further information about their features - i.e., the name of the model or system, description of the evaluation process, URL where additional information can be found, level of maturity of the system (consolidated, incipient, theoretical, applied, etc.), aggregation level (publisher, book series or title), fields in which the evaluation process is used, the type(s) of evaluations it is used for, who generates or feeds the evaluation system with information, and the advantages and disadvantages of the evaluation process.

Following our analysis of the results we have data to report concerning the relative weights that each country gives to books, including an overview of the different aims attached to each country's evaluation system. Some are similar in that they are directly linked to a Performance-Based Research Funding System (PRFS), while in other cases, a specific country's evaluation process is applied primarily to individual researchers for tenure and promotion. At the individual level, evaluative information is also used as a source of information to help improve upon their academic performance.

The representatives of the different countries completed the questionnaire, while it was available at http://ilia.cchs.csic.es/encuestas/EuropeEval/ during the first semester of 2017 (with the exception of Croatia, Latvia and Slovenia for which responses were collected in the last months of 2017). The answers were then recorded and extracted from a MySQL database and analyzed, by the participants, in order to determine how diverse evaluation processes might be grouped and presented.

The results reflect the variety of models for evaluating books. After processing the questionnaire responses, our first observation was that a high number of variables were found and could serve as a basis for the country evaluation system groupings and characterizations. Considering this, and the fact that many of the variables may be said to be central to the conditioning of the evaluation system itself, (e.g., the existence of a CRIS or not), two optimal groupings were finally developed and discussed: what we call 'descriptive approach' in the summary of results, consisting of a description of the main features of each evaluation system and a grouping based on the main types identified and a 'Theoretical approach', based on the existence of formalized evaluation systems on the one hand and publishers lists or labels on the other.

Results

Our analysis reveals the complexity and diversity of solutions used in the 19 different European countries. Each may be described and presented in various complementary ways, on the basis of unique combinations of features, and the conditioning of their model. We begin, therefore(with a general overview of the scholarly book evaluation instruments in the analyzed countries, and then suggest two approaches to their classification: The first one involves presenting the scholarly book evaluation process according to two main dimensions: (1) having a formalized or non-formalized evaluation process for scholarly books (the term 'formalized' refer to evaluation processes that have established some type of classification of publishers, quality labels, or formal criteria which have to be met by books under evaluation) and (2) using book publisher lists or quality labels. Secondly, a classification based on geographical location of the evaluation systems has been carried out.

1. Common features of book evaluation systems

In general terms, the book evaluation systems that we analyze in this article take into consideration three main elements[1,2].

1.1 The type of books. Usually, all evaluation processes specify the types of books that are included as well as those that are excluded. Research monographs and edited volumes are the types of books that, clearly, are within the scope of an evaluation process. Other types, such as critical editions or documented exhibitions catalogues, are sometimes considered as well (for example, in the case of Spain) but there are types of outputs that are not considered to be scholarly output, such as textbooks.

1.2 The existence of peer review, which is a factor that is related to the first element. All evaluation systems take this into account as one of the deciding factors for whether or not a book is included in the evaluation process. But defining peer review is not an easy task, due to the variety of ways it may be considered and described (e.g., Giménez-Toledo, Sivertsen & Mañana-Rodríguez (2017). Often, peer review is used as a demarcation line between scholarly and non-scholarly output. In this context, it is worth

---

[1] Note that these elements apply only to those countries that have a formalized book evaluation system.
[2] A fourth element, common to the majority but not all of the evaluation procedures studied here is the formal requirement of the existence of ISBN codes for published items prior to their inclusion in databases, labels or as a requirement for evaluation.

noting that in the publication process of scholarly books, a variety of manuscript selection processes can be used, including but not limited to traditional peer review (like in the case of scientific journals). The manuscript selection process that is chosen depends on factors like discipline, language, geography, and traditions associated with individual publishing houses.

1.3 The publication channel. Most evaluation systems attach some importance to the publisher, since knowing a book's publisher provides indirect information on what type of book it is. In all cases the publisher is a factor in the evaluation of books. Usually it is combined with other types of indicators (manuscript selection procedures, specialization, etc.), with quality labels for books or book series or with the assessment of expert panels.

Table 1 show the different information instruments used for performing scholarly book evaluations in each country.

Country	Country abbreviation	Comprehensive database covering scholarly outputs at supra-institutional level (i.e., national or regional)	Comprehensive database covering scholarly outputs at institutional level	Categorization or ranked list of book publishers	Quality label for book series or individual titles	No specific method or system for evaluating books.	Others (Different types of expert panels)
Countries with a formalized evaluation process							
Croatia	HR	✓					✓
Czech Republic	CZ	✓	✓				
Denmark	DK	✓	✓	✓			
Finland	FI	✓	✓	✓	✓		
Flanders, Belgium	BE/FL	✓	✓	✓	✓		
Latvia	LV						✓
Lithuania	LT	✓	✓				
Montenegro	ME	✓	✓				
Norway	NO	✓	✓	✓			
Poland	PO	✓	✓				
Slovakia	SK	✓	✓	✓			✓
Slovenia	SL	✓		✓			
Spain	ES		✓	✓	✓		
Countries without a formalized evaluation process							
Serbia	RS					✓	

France	FR					✓	
Israel	IL						✓
Italy	IT					✓	
Portugal	PO					✓	
Switzerland	CH		✓			✓	✓

Table 1. Resume of the main descriptors of the different evaluation scheme

2. General overview

2.1 Current Research Information Systems (CRIS) or their absence

One of the main differentiating features for each of the studied evaluation systems is the existence of a national or regional database for recording research outputs (Sīle, Guns, Sivertsen & Engels, 2017). Such information systems can vary in terms of integration with other research-related data, from tightly integrated CRIS to standalone database. The databases underlying the evaluation systems can provide information about: a) the total scientific output from a given sector (e.g. universities, all higher education, all research performing organizations) in that region or country, b) the classification or categorization of publishers or journals in which the researchers of that country / region have published, and therefore allows for c) evaluations in context, taking into account the whole output of the country, not only the parts covered by international databases. This is especially important with regards to books, because the coverage of books in the international databases is weak and not sufficient for evaluative purposes (Gorraiz, Purnell & Glänzel, 2013). In the Social Sciences and Humanities, books are among the most relevant forms of output. Several of the countries participating in the questionnaire count with CRIS (Table 1).

2.2 Formalized / non-formalized evaluation systems

Amongst the 19 countries examined, a distinction can be made between 'formalized' and 'non-formalized' evaluation processes. Formalized systems are in place in Croatia, Czech Republic, Denmark, Finland, Flanders (Belgium), Lithuania, Montenegro, Norway, Poland, Slovakia, Slovenia and Spain. The latter non-formalized evaluation processes are those in which the decisions pertaining to the assessment of books are taken by a local expert panel or a committee in the absence of supporting information sources (and established at the national level). This is the case for Serbia, France, Italy, Latvia, Israel, Portugal and Switzerland.

Countries with a formalized approach to the evaluation of scholarly books use data to support the qualitative judgments by experts, whereas countries that do not utilize formalized data carry out the evaluation process by expert judgment (note: this does not imply that they cannot take indicators into account, it simply means that they are not systematized at the national level).

2.3 Quantitative / qualitative approaches to the evaluation of scholarly books

The evaluation of scholarly books as a form of research output in Europe is based on a variety of methodological approaches. These approaches are directly affected by each country's general system although some of the methodological axes are commonly shared.

Among the national evaluation systems included in this study, some seem to be more qualitative and others more quantitative in their approach to scholarly books. This major distinction tends to be related to whether or not the country has a performance-based research

funding system. Broadly speaking, all of the system-types may be placed in four general categories (Sivertsen, 2017).

Countries situated in the first category (i.e., Italy, Lithuania, and Portugal), evaluate with the purpose of funds allocation. The evaluation is organized at intervals of several years. This approach is characterized by the use of qualitative evaluation methods in which the content of each publication is taken into consideration. It is typical for countries with a qualitative approach not to have national CRIS.

With category 2 countries (i.e., Croatia, the Czech Republic, Denmark, Flanders (Belgium), Finland, Norway or Poland), the funding allocation is based on a set of indicators that represent research activities. Bibliometrics may be part of this set of indicators. The indicators are used annually and directly in the funding formula. In these countries, a more quantitative approach to the evaluation of scholarly books is often found. The publication channel of the book (publisher and/or book series) may be evaluated instead of the individual title. Moreover, it is typical for countries that take a quantitative approach to have national CRIS, although there are exceptions such as Spain (quantitative indicators are used but there is no national CRIS). The quantitative approach mentioned here is not based on citation counts, but on other types of indicators. "Quantitative approach" is understood to mean that a systematic way of evaluating books is taken (categorizations, quality labels, etc.), whereby qualitative judgments of books publishers, book series, etc., are utilized in combination with quantitative indicators.

Some countries, for example France (category 3), have evaluation procedures that are similar to those of the first 'qualitative' group, but they are not connected to performance-based institutional funding but to the evaluation of individual researchers.

And finally, countries that fit within category 4, such as Switzerland, neither have a national research evaluation system nor a performance-based funding system with bibliometric indicators (formative evaluation is among the objectives of the assessment carried out at the department level).

Norway is an example where both approaches, quantitative and qualitative, are used but it still differs slightly from the countries in category 2. It has adopted a qualitative approach to scholarly book evaluation for its national research evaluation system (as pointed out in Sivertsen, 2017, p.3: '*It* [Norway] *also has a UK-inspired research assessment exercise (Geuna and Martin, 2003). It is not used for funding allocation. The purpose is to provide recommendations on how to increase the quality and efficiency of research*.' However, this system is separate from the performance-based institutional funding system, where there is a quantitative approach to the so-called "Norwegian model". This procedure is not exclusive to Norway as it applies also to other systems, like that which has been implemented in Flanders. Also in other countries that support the PRFS with a database, ranking or labels, more qualitative and less formalistic approaches to books may take place in other evaluation contexts: universities' internal research assessment, funding allocation, recruitment, promotion, personal performance and bonus systems, as well as in project evaluation by public and private research funds. A highly formalistic and quantitative approach has generally been chosen for the funding-scheme at the national level.

3. Classification of countries according to two dimensions

For each of the analyzed countries, Table 1 shows the different evaluation approaches. Note that the various approaches may be categorized according to two dimensions. The first dimension is the level of formalization of the evaluation. The second dimension pertains to the use of publisher lists and/or quality labels for publishers.

Based on these two dimensions the countries included in this study may be further divided into three groups. Figure 1 shows, on the one hand, that Croatia, the Czech Republic, Lithuania, Montenegro, and Poland have a formalized model of evaluation but do not use publisher lists whereas Denmark, Finland, Flanders, Norway, Slovakia, Slovenia, and Spain combine a formalized model with book publisher lists. On the other hand, a non-formalized model without publisher lists is used in France, Israel, Italy, Portugal, Serbia, and Switzerland. We do not find countries where a non-formalized model is used together with publisher lists and/or quality labels, possibly because the use of lists or labels already implies a certain degree of formalization.

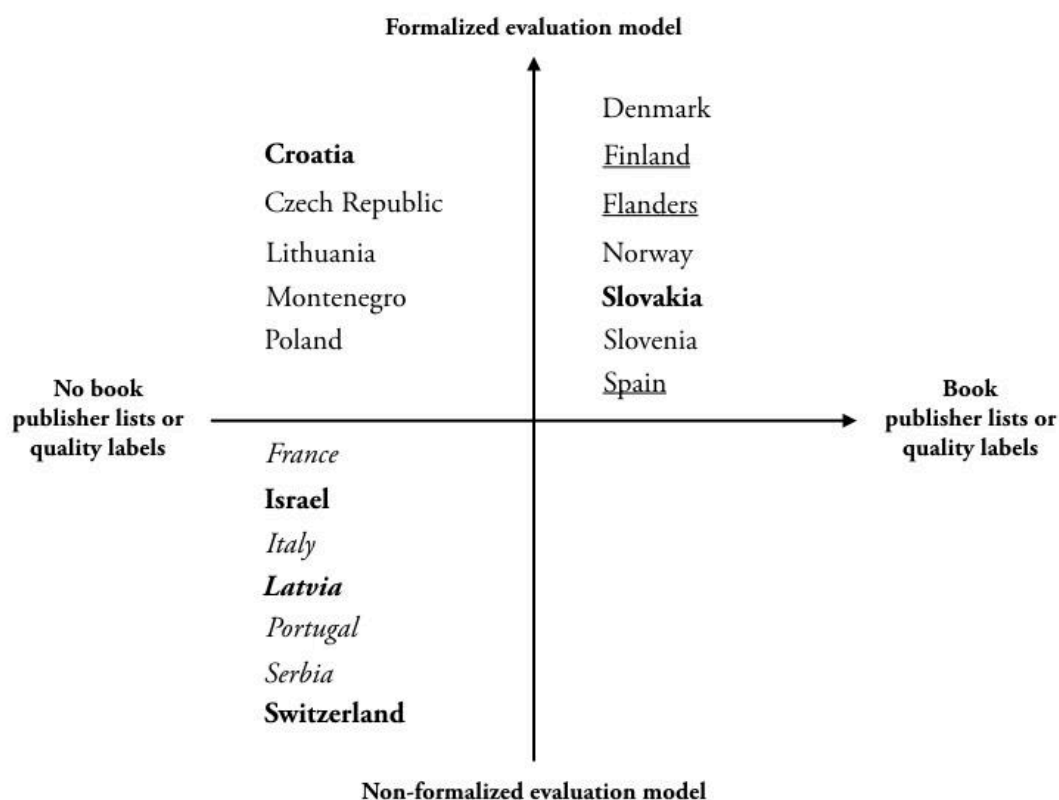**Figure 1.** Two dimensions of evaluation processes in ENRESSH countries.
Note: Countries are sorted in alphabetical order in each cluster. Italicized names represent countries with no specific method or system for evaluating books. Underlined names represent countries in which quality labels for book series of individual titles are used. Bold names represent countries in which expert panels are implemented.

Summary of the results

Evaluation systems and the role of the book in such systems are diverse across European countries. However, it is possible to describe the differences and commonalities between them.

There are several ways to systematize the way books are considered in research evaluation in the studied countries. In this article, we suggest that there may be two ways of clustering the countries; first, a **grouping along two dimensions**: 1) the use of quality labels for books or publisher rankings; second, a grouping of countries describing the different approaches to evaluation methods[3] 2) **degree of formalization of the evaluation process**.

Each of these clustering approaches has its own value in terms of providing answers to different questions. A comparison of the clustering using the two different approaches reveals the following (see Table 2).

In all approaches, there is a clear distinction between the countries having a formalized evaluation procedure in place and those that do not have such a formalized system. However, there are also differences between the approaches: First, let us focus on the countries with a formalized evaluation system. The formalized vs. lists and labels approach consequently groups the countries with a formalized system according to whether they have a qualifying list in place. The descriptive approach suggests a different solution: Spain switches places from Cluster 1 (formalized) to Cluster 3 (non-formalized) because there is no national database and the formalized evaluation happens at the individual rather than the national level.

	Type 1: Database and label/ranking	Type 2: Database	Type 3: No specific book evaluation	Type 4: Peer review based book evaluation
Theoretical	BE/FL, DK, ES, FI, NO, SI, SK	CZ, HR, LT, ME, PL	CH, FR, IL, IT, PT, RS	
Descriptive	BE/FL, DK, FI, NO	CZ, HR, LT, ME, PL, SI, SK	ES, FR, IT, PT, RS	CH, IL,

Table 2. Comparison of the two classifications.

While the approaches are different, the results are rather similar, showing that there are, in fact, different types of book evaluation in place. We therefore formulate four ideal types[4] of book evaluation systems (in parentheses the countries that best represent the ideal type, i.e. are consistently classified in this cluster across approaches): The first type uses complete data using a supra-institutional database and relying on labels or rankings to evaluate books in their

---

[3] Additionally, an empirical clustering using Joint Correspondence Analysis using the six variables created from the questionnaire is proposed as a complementary approach (See Appendix II). The JCA approach, suggests a differentiation between countries having a book quality label in place and those who have not while suggesting that a differentiation of quality levels of publication channels such as publishers be a less differentiating feature.

[4] We refer to "ideal types" in the Weberian sense as theoretical concepts that can take (slightly) different forms in reality, rather than in a normative sense.

formalized evaluation procedures (BE/FL, FI). The second type uses data from supra-institutional databases in formalized evaluation procedures while not using a qualifying approach to books using lists and rankings (CZ, LT, ME, PL). The third cluster applies no systematic evaluation of books (FR, IT, PT). The fourth cluster does not apply formalized evaluation procedures but has some peer review evaluation processes in place in which books are considered (CH, IL).

The descriptive approach (and the country selection representing best the ideal types) makes visible that the role of books in evaluations of scientific work is at least partially linked to regions. The first type coincides to a great extent to Nordic countries, the second type to Eastern European countries, and the third type to Southern European countries. The fourth type, however, does not find a regional counterpart. However, by proceeding further, we also see that the book evaluation process corresponds to the general evaluation system in a country. These systems are influenced by different situations in which the country's research system finds itself concerning the scientific policy agendas in each of them. After examining the different models, it can be said that in the quantitatively oriented book evaluation process of clusters 1 and 2, there are three elements which are taken into account in the evaluation of books: a) the nature of the publication, i.e., what counts as a scholarly book or a research book; b) the manuscript selection process (peer review); and c) the relevance of the publisher, often categorized using quality labels or publisher lists.

In countries with non-formalized systems, the evaluation process for books relies on peer review of the item itself. The different classifications show that having a formalized evaluation in place calls for either a comprehensive supra-institutional database or rankings or labels to support the evaluation process[5]. On the other hand, if a non-metric peer review process is applied, evaluations are not formalized at a supra-institutional level. The advantage of a formalized, data-driven approach to (book) evaluation is the smaller extent of subjectivism in the evaluation and the comparability between institutions; while the advantage of not having a formalized approach lies in its formative potential and the opportunity to customize the evaluation procedure to the institution's mission down to the evaluated unit.

Conclusions

Amongst all of the book evaluation systems that we have evaluated here the common denominator is that regardless of whether the focus is on categories of publishers or an individual approach to evaluation, the evaluation process itself is still carried out by experts.

All evaluation processes have advantages and disadvantages. There is no perfect process and that might be one of the issues that must be accepted as a rule in research evaluation. However, knowledge about different types of book evaluation systems makes it possible for countries utilizing the same type of evaluation approach to exchange knowledge and benefit from each other's experiences in a way that might help to improve certain procedures.

---

[5] A well-known exception is the UK's Research Excellence Framework (REF). It is formalized and based on peer review panels that attribute a score to a scholarly work/unit to formalize the evaluation. However, the reoccurring investigation on how to increase the use of indicators in the REF (see e.g. Adams, 2009; Wilsdon et al., 2015; for a historic overview, see Williams & Grant, 2018) points to the fact that a formalized system calls for a more technocratic approach than formative evaluations.

Knowing the advantages and disadvantages as well as the limitations of an evaluation process can support specific decisions pertaining to a future processes, for example, making a choice about its new function (distributive or formative) or selecting the unit (institution or individual) that will be under evaluation. A quantitative approach is often associated with national funding-schemes for institutions, while a qualitative approach tends to be present in institutional assessments as well as research evaluations at the individual level. One of our aims with this article is to reduce certain limitations to a minimum or, at least, facilitate opportunities for selecting a model that possesses the least pronounced limitation.

Our review of the pros and cons of the different evaluation processes clearly reveals the complementariness of the quantitative and qualitative approaches. Assigning a publisher to a category in a ranked list can provide useful but not absolute information for a quantitative evaluation process, especially if it can be complemented with other types of information concerning the monograph or book series. This should be the case, at least, for an evaluation processes that would best be carried out at an individual level, rather than at the institutional level. The difficulty lies in deciding how to apply one of the two approaches to a usually large amount of scientific output generated within a limited time span, in a way that is sensibly attributed to research activity.

When both approaches are used in combination the evaluation process might improve both its precision and acceptance of results by the scholarly community. For instance, if a single approach is used there may be specific limitations associated with assessing certain types of books that are more complex than others (e.g., Rens Bod's 'A New History of the Humanities. The Search for Principles and Patterns from Antiquity to the Present': This book is a scholarly output, but its impact has gone beyond academia as can be observed by the non-academic nature of some of the media on which it has been reviewed. The evaluation of a book with several types of impact is more complex than that of a book exclusively oriented towards an academic readership). The approach might be compromising and erratic, for example, if a highly prestigious book publisher is known for publishing non-scientific books or even scientific books that have not been subjected to a rigorous validation. Also, it can be difficult for an expert to recognize the relevance of a book that, personally, is not of interest to him/her if he or she does not appreciate or share the author's methodological principles or views simply because the book's author is an academic rival. With this example, we see why a combination of approaches might lead to a more balanced and fair judgment pertaining to a particular book.

Quantitative indicators employed at the level of publisher or series can provide a certain degree of objectivity and comparability, but they do not help with the identification of a book's unique characteristics. By contrast, a qualitative approach to the evaluation of a book can support a deeper appraisal of its content and value, but by with an intrinsically subjective process, biases may be introduced, making it difficult to draw effective comparisons between different evaluation approaches, particularly with respect to models that combine PRFS with research evaluation (e.g., the Research Excellence Framework in the UK) (Sivertsen, 2018). Consequently, the use of both approaches is more promising for the development of an adequate and fair evaluation.

References

Aagaard, K. (2015). How incentives trickle down: Local use of a national bibliometric indicator system. *Science and Public Policy, 42*(5), 1–13.

Aagaard, K., Bloch, C, Schneider, J, Henriksen, D., Kjeldager Ryan, T. & Lauridsen, (P 2014). *Evaluering af den norske publiceringsindikator*. Aarhus: Aarhus Universitet. http://www.uhr.no/documents/Evaluering_af_den_norske_publiceringsindikator.pdf.

Adams, J. (2009). The use of bibliometrics to measure research quality in UK higher education institutions. *Archivum Immunologiae Et Therapiae Experimentalis, 57*(1), 19–32. doi:10.1007/s00005-009-0003-3

Ahlgren, P., Colliander, C., & Persson, O. (2012). Field normalized rates, field normalized journal impact and Norwegian weights for allocation of university research funds. *Scientometrics, 92*(3), 767–780. doi:10.1007/s11192-012-0632-x.

ANVUR, 2017. http://www.anvur.org/index.php?lang=it

Calasso, R. *The Art of the Publisher*. Penguin, 2015.

Camiz, S., & Gomes, G. C. (2013). Joint Correspondence Analysis versus Multiple Correspondence Analysis: A solution to an undetected problem. In *Classification and Data Mining* (pp. 11-18). Springer, Berlin, Heidelberg.

Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics, 93*, 373–390.

Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T. C. E., Ingwersen, P., Sivertsen, G., Verleysen, F. T., & Zuccala, A. A. (2016). Taking scholarly books into account: Current developments in five European countries. *Scientometrics, 107*(2), 685-699. doi:10.1007/s11192-016-1886-5

Giménez-Toledo, E., Mañana-Rodríguez, J., & Sivertsen, G. (2017). Scholarly book publishing: Its information sources for evaluation in the social sciences and humanities. *Research Evaluation, 26*(2), 91–101. doi:10.1093/reseval

Giménez-Toledo, E., Sivertsen, G., & Mañana-Rodríguez, J. (2017). Peer review as a delineation criterion in data sources for the assessment and measurement of scholarly book publishing in social sciences and humanities. *16th International Conference on Scientometrics and Informetrics*. Wuhan (China).

Giménez-Toledo, E., Tejada-Artigas, C., & Mañana-Rodríguez, J. (2012). Evaluation of scientific books' publishers in social sciences and humanities: Results of a survey. *Research Evaluation*, *22*(1), 64-77.

Gorraiz, J., Purnell, P. J., & Glänzel, W. (2013). Opportunities for and limitations of the Book Citation Index. *Journal of the American Society for Information Science and Technology, 64*(7), 1388–1398. doi:10.1002/asi.22875

Greenacre, M. (2007). *Correspondence Analysis in Practice. Second Edition*. Boca Raton, FL: Chapman & Hall.

Hasgall, A., Lanarès, J., Marion, A., & Bregy, J. (2018). *The programme 'Research performance in the humanities and social sciences'. Report*. Bern: swissuniversities.

Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. *Journal of the American Society for Information Science and Technology, 62*(11), 2147–2164.

Kulczycki, E., Korzeń, M., & Korytkowski, P. (2017). Toward an excellence-based research funding system: Evidence from Poland. *Journal of Informetrics, 11*(1), 282-298.

Ministerio de Educación, Cultura y Deporte. Resolución de 23 de noviembre de 2017, de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se publican los criterios específicos aprobados para cada uno de los campos de evaluación. https://www.boe.es/boe/dias/2017/12/01/pdfs/BOE-A-2017-14085.pdf

Michavila, F. (ed.). (2012). *La Universidad española en cifras*. Madrid: CRUE. http://www.crue.org/Publicaciones/Documents/UEC/LA_UNIVERSIDAD_ESPANOLA_EN_CIFRAS.pdf

Molas-Gallart, J. (2012). Research Governance and the Role of Evaluation: A Comparative Study. *American Journal of Evaluation, 33*(4), 583-598.

Pölönen, J & Wahlfors, L. (2016). Local use of a national rating of publication channels in Finnish universities. Poster presented at the *21st Nordic Workshop on Bibliometrics and Research Policy*. URL: https://figshare.com/articles/Local_Use_of_a_National_Rating_of_Publication_Channels_in_Finnish_Universities_NWB_2016_poster_/4246541

Royal Netherlands Academy of Arts and Sciences. (2011). *Quality Indicators for Research in the Humanities*. Royal Netherlands Academy of Arts and Sciences: Amsterdam, The Netherlands.

Sello de calidad en edición académica (CEA/APQ). http://www.selloceaapq.es/

Sīle, L., Guns, R., Sivertsen, G., & Engels, T.C.E. (2017) European Databases and Repositories for Social Sciences and Humanities Research Output. Antwerp: ECOOM & ENRESSH. https://doi.org/10.6084/m9.figshare.5172322Sivertsen, G. & Schneider, J. (2012). Evaluering av den bibliometriske forskningsindikator. Oslo: Nordisk institutt for studier av innovasjon, forskning og utdanning. URL: https://brage.bibsys.no/xmlui/bitstream/handle/11250/280879/NIFUrapport2012-17.pdf?sequence=1&isAllowed=y

Sivertsen, G. (2016). Publication-Based Funding: The Norwegian Model. In M. Ochsner, S. E. Hug & H.-D. Daniel (eds.), Research Assessment in the Humanities: Towards Criteria and Procedures (pp. 79-90). Cham: Springer International Publishing.

Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective. Palgrave Communications 3, 17078.

Sivertsen, G. (2018). Why has no other European country adopted the Research Excellence Framework? Available at the London School of Economics and Political Science blog: http://blogs.lse.ac.uk/politicsandpolicy/why-has-no-other-european-country-adopted-the-research-excellence-framework/

Verleysen, F. T., & Engels, T. C. E. (2013). A label for peer-reviewed books. *Journal of the Association for Information Science and Technology*, 64(2), 428-430.

Verleysen, F. T., Ghesquière, P., & Engels, T. C. E. (2014). The objectives, design and selection process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW). In W. Blockmans, L. Engwall & D. Weaire (eds.), *Bibliometrics. Use and Abuse in the Review of Research Performance* (pp. 117–127). Portland Press.

Weller, M. (2012) Calling open access academic book publishers: how author and publishers could make a modest profit. LSE Impact Blog. Available at http://blogs.lse.ac.uk/impactofsocialsciences/2012/12/10/weller-calling-open-access-publishers/ (Accessed 2 April 2016)

Wilsdon, J. et al. (2015). *Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. doi:10.13140/RG.2.1.4929.1363

Williams, K., & Grant, J. (2018). A comparative review of how the policy and procedures to assess research impact evolved in Australia and the UK. *Research Evaluation*, Advance Access, published January 11, 2018. doi:10.1093/reseval/rvx042

Appendix I. Access interface and text version of the online questionnaire

SCHOLARLY BOOKS EVALUATION IN SSH ACROSS EUROPE

Survey for developing a comparative study in the framework of ENRESSH COST Action (WG3)

Identification data

- ° Name:
- ° Country:
- ° Institution:
- ° Role:

1. **Kind of system for evaluating books**

Please mark the option that fits better. More than one can be chosen.

☐ Comprehensive database covering scholarly outputs at national level, including books, book chapters and/or edited volumes and well defined criteria for evaluating them

☐ Comprehensive database covering scholarly outputs at regional level, including books, book chapters and/or edited volumes and well defined criteria for evaluating them

☐ Comprehensive database covering scholarly outputs at institutional level, including books, book chapters and/or edited volumes and well defined criteria for evaluating them

☐ Categorization or ranked list of book publishers as an independent tool for its use in evaluation process at different levels

☐ Quality label for book series or individual titles

☐ There is no specific method or system for evaluating books. Commercial databases such as Book Citation Index or Scopus as well as/or expert panels who evaluate each individual contribution are the basis for taking decisions on quality of books.

☐ Google scholar

- Others, please specify

  ……
  ……
  ……
  ……

2. **When some specific system exists for evaluating books (If you marked several in the previous section, please copy/paste the following fields for each case):**

   a. **Please, indicate the title**

………
………

   b. **Please, briefly explain how books, book chapters or edited volumes are evaluated or how the system works**

………
………
………
………
………
………
………
………

……

………ո»

c. Is there any public interface for accessing the system? Please indicate the URL

……

………..

d. Can it be considered as a well established system (i.e. it existed for many years and/or has been supported by laws or regulations for a long time) for taking into account scholarly books in evaluation process:
- **Yes**
- **No**
- **Comments:**

   ……

   ……

   ……

   ……

e. It is used for
- all fields, but with differences among them (Please, specify the most relevant differences)

- all fields, without differences among them (Please, specify the most relevant differences)

- It is used just for Social Sciences and Humanities
- It is used just for Social Sciences
- It is used just for Humanities

f. Which is the level for evaluation?
- Book publisher
- Book series
- Individual titles
- Please, if necessary, describe your answer in more detail:

   ……

   ……

   ……

g. It is used
- For institutional research evaluation
- For assessment at the individual level (e.g. promotions, etc.)
- For monitoring research activities
- For resources allocation

h. Who is in charge of feeding the system? You can mark several options

- Researchers
- Information managers, librarians, etc.
- Administrators
- Others. Please, specify

………

………

………

i. Could you point out pros and cons of the system?

Pros

………

………

………

………

………

………

Cons

………

………

………

………

j. Is there any further development / evolution foreseen for the system? Please, specify

………

………

………

………

………

Appendix II. Clustering by Joint Correspondence Analysis (JCA) (see e.g. Greenacre, 2007 or Camiz and Gomes, 2013 for a comparison with multiple correspondence analysis)

The empirical clustering along the variables presented in Table 1 shows that the variables can be reduced to two dimensions quite comprehensively: The first dimension reflects the degree

of formalization and use of metrics; the second dimension represents the use of rankings and labels.

Figure 2 shows the map of the JCA. The position of the variables shows that a strong relationship between the existence of a national or institutional database and a formalized evaluation system (i.e., very close to each other in the map), and that the non-existence of a specific book evaluation system lies just opposite to the former. These variables define the x-axis. The y-axis is defined by the existence of a quality label or a publisher ranking on the one hand and by the use of peer review panels on the other hand. The rankings and peer review panels, however, are also spread on the x-axis.

The countries form four clusters. In the top right quadrant, we find France, Italy, Portugal and Serbia. These countries do not have a specific book evaluation procedure in place. In the bottom right quadrant, Israel, Latvia and Switzerland form a cluster of countries having no formal evaluation procedure for books in place and do not have supra-institutional databases. Books are evaluated using peers in ex-post evaluations. There are small differences between the countries: Switzerland has databases on the institutional level, Israel includes books in evaluations systematically while in Latvia neither is the case. In the bottom left quadrant, the Czech Republic, Denmark, Lithuania, Montenegro, Norway, Poland, Slovenia and Slovakia form a cluster of countries having a formal evaluation system as well as comprehensive databases on a supra-institutional level in place. There is, however, a divide of this cluster into three groups, the first is having a ranking in place (DK, NO, SI), the second does additionally evaluate using peer review panels (SK) while the third has neither a ranking in place nor uses peer review (CZ, LI, ME, PL). Finally, on the top left quadrant, Finland, Flanders (BE) and Spain build a cluster having a formalized evaluation system, a quality label and a publisher ranking in place. Croatia reveals itself as a special case, because it stands for a formal evaluation system and a supra-institutional database. However, there are no other metric instruments used and books are evaluated using expert panels, pulling Croatia into the middle of the x-axis.

The empirical clustering thus reveals different insights than the theoretical clustering represented in Figure 2. While the axes represent similar dimensions, it groups the countries differently. The emphasis lies more in the differentiation of quantification and peer review, for example. Nevertheless, the results are, not surprisingly, similar. Rotating Figure 1 by 45 degrees results in a similar solution.
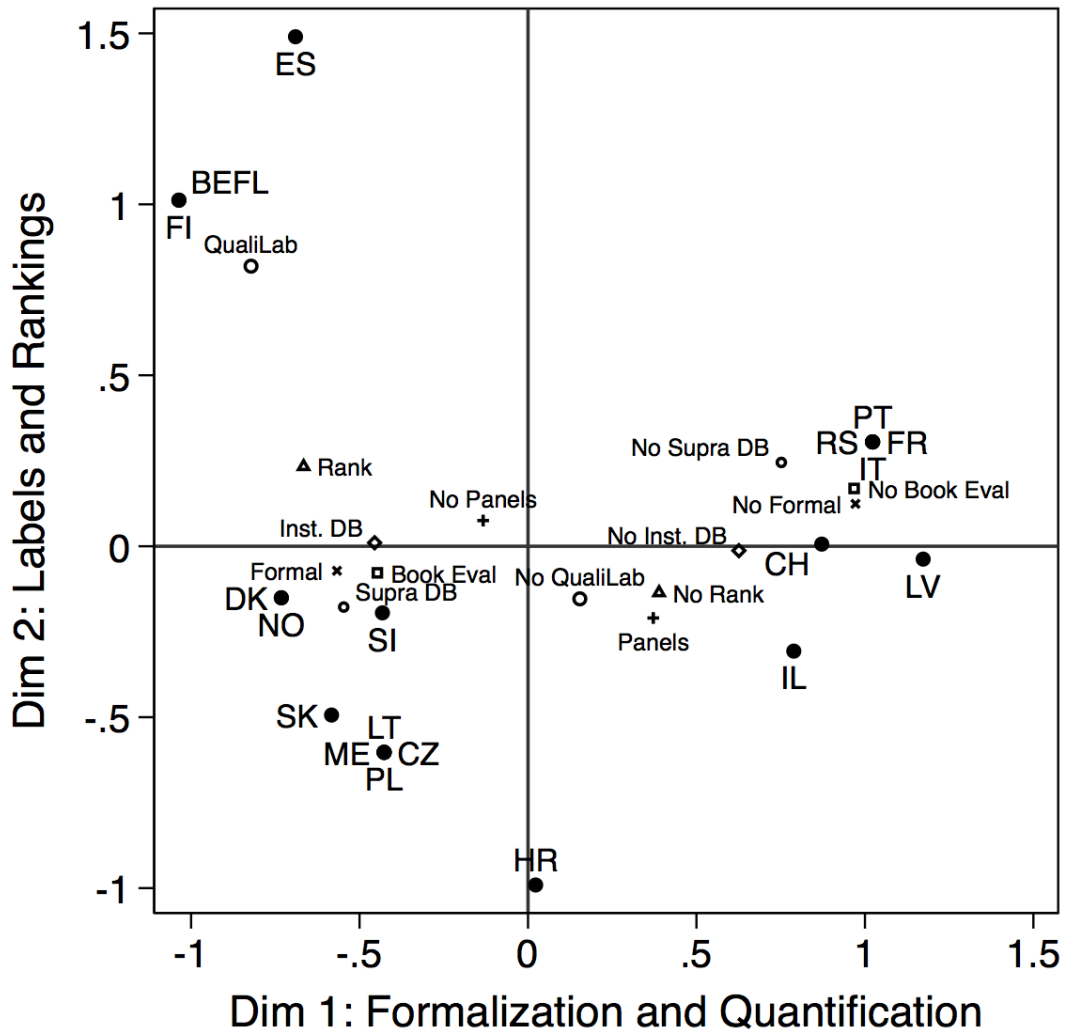
Figure 2

Figure 2. Two dimensions of evaluation processes in 19 countries using Joint Correspondence Analysis.
Note: Formal stands for formalized evaluation procedure, Supra DB stands for supra-institutional database, Inst. DB stands for institutional database, Rank stands for publisher ranking, QualiLab stands for quality label, Panels stands for expert panels, and Book Eval stands for specific procedure for book evaluation (see Table 1). [ES: Spain; FI: Finland; BEFL: Flanders (Belgium); LV: Latvia; DK: Denmark; NO: Norway; SK: Slovakia; LT: Lithuania; ME: Montenegro; CZ: Czech Republic; PL: Poland; IL: Israel; CH: Switzerland; RS: Serbia; PT: Portugal; FR: France; IT: Italy; HR: Croatia; SI: Slovenia]