

How international is co-authorship outside of Web of Science? The case of social sciences and humanities in Flanders, Belgium

Peter Aspeslagh¹ and Raf Guns²

¹*peter.aspeslagh@uantwerpen.be*, ²*raf.guns@uantwerpen.be*

Centre for R&D Monitoring (ECOOM), Faculty of Social Sciences, University of Antwerp, Middelheimlaan 1,
2020 Antwerpen (Belgium)

Abstract

This paper studies the question to what extent international research collaboration in the social sciences and humanities takes place in publications that are not indexed in Web of Science (WoS). We compare the internationality of co-authors of Flemish SSH publications in WoS (BOF-WoS) to those not indexed in WoS (BOF-VABB). While publications with 2+ authors in BOF-WoS are overall more international than those in BOF-VABB (60.0% versus 48.1%), the difference is smaller than expected and varies across disciplines and publication types. Main partner countries are the Netherlands, UK and US. Many nonprofit, company, and government organizations have co-authored Flemish SSH research but are as of yet not included in ROR.

Introduction

In this paper, we examine the degree of international research collaboration in the social sciences and humanities (SSH) beyond the coverage of the Web of Science (WoS). It is well-known that research in the SSH is generally not well-covered in bibliometric databases like WoS or Scopus (Mongeon & Paul-Hus, 2016; Ossenblok et al., 2012). One of the purported reasons for this disparity is that these databases orient towards internationally oriented research, whereas much of the SSH is more focused on local issues. However, equalizing indexation in a major database with internationality is problematic and far too simplistic (Guns & Hołowiecki, 2022; Tennant, 2020). In this paper, we study one dimension of internationality, through the internationality of their authors: if a database like WoS indeed captures all international research from the SSH, one would expect to find only national collaboration in non-indexed publications.

The dataset we use derives from the Flemish performance-based research funding system (PRFS). In order to implement an internationalization parameter in the PRFS, an extensive author affiliation data collection operation for peer-reviewed SSH publications not included in WoS was required. This has recently been completed and makes it possible (for the region of Flanders, Belgium) to comprehensively assess the degree of cross-border co-authorship for all peer-reviewed publications, within and outside of WoS.

This issue has, to the best of our knowledge, not yet been addressed in the literature. Indeed, earlier studies that looked at international co-authorship in the SSH were based on WoS (Henriksen, 2016; Larivière et al., 2006), while studies of co-authorship that started from a comprehensive set of SSH publications (e.g., Ossenblok et al., 2014) did not consider the question of international co-authorship.

In the remainder of this paper, we will use the terms ‘collaboration’ and ‘co-authorship’ interchangeably. This is a simplification, since co-authorship is only a partial indicator of collaboration (Katz & Martin, 1997); still, we believe that the results are indicative of where and how international collaboration in the SSH takes place.

Background

As a federal state, regions in Belgium have a high degree of independence; for instance, research policy in Flanders, the Dutch-speaking region of Belgium, is mainly decided on the regional level. The Flemish PRFS takes into account publications indexed in the WoS as well as peer-reviewed publications from the SSH outside of WoS; the latter publications are included in the *Flemish Academic Bibliographic Database for the Social Sciences and Humanities* (VABB-SHW; Verleysen et al., 2014). In order to be eligible for inclusion in the VABB-SHW, publications must be peer-reviewed, in addition to other, more formal criteria. A panel of senior scholars, the so-called GP, is responsible for the scientific management of the VABB-SHW.

In 2019, the distribution key of the PRFS was extended with parameters relating to internationalization (Luwel, 2021). Fifteen percent of the so-called BOF key (Bijzonder Onderzoeksfonds/University Research Fund) incorporates an international dimension in research activities, based on two components: EU funding (7.5%) and international publications (7.5%).

Publications are considered to be international if they have at least two authors and have at least one address outside of Belgium. For instance, if an author of a publication is affiliated to the University of Amsterdam, which has an address in the Dutch capital, then the publication counts toward the internationalization parameter. On the other hand, this excludes co-publications with researchers from, e.g., the Walloon region. While WoS incorporates affiliation and address data, the VABB-SHW did not contain any affiliations. Hence, the collection of affiliations was required for the implementation of the parameter. GP-selected publications do not appear in WoS, nor are they all structurally included in alternative bibliographic data sources like CrossRef or Scopus. Multiple strategies were deployed for the retrieval of their author affiliations, but a large share had to be collected manually and individually. As the number of publications that needed to be coded was large, affiliated organizations were assigned on publication level, rather than the more fine-grained publication-author level. From the next iteration on – only dealing with publications published in 2021 - the publication-author level will be applied.

Where possible, affiliated organizations were registered using an identifier from the Research Organization Registry (ROR). As ROR did not capture all affiliated organizations, new records with unique (non-ROR-) organization identifiers were added to an internal database (see Results section for more details).

Data and methods

We will compare international collaboration in two data sets:

- *BOF-WoS*: the set of publications from 2011–2020 that count toward the WoS parameter of the PRFS (only those authored by Flemish SSH scholars);
- *BOF-VABB*: the set of (GP-approved) publications from 2011–2020 that count toward the VABB-SHW parameter of the PRFS, authored by Flemish SSH scholars.

BOF-WoS data were retrieved from the in-house copy of WoS maintained by ECOOM-KU Leuven. BOF-VABB data are derived from the VABB-SHW database maintained by our own group, ECOOM-UAntwerp.

Table 1 summarizes the size and characteristics of both data sets. For BOF-VABB data, after selecting the publications with at least one co-author, author affiliation data for 23,079 publications had to be retrieved. This extensive retroactive data collection operation was completed in 2022, and we now dispose of affiliation data for 95.3% of these publications.

Since some of them do not contain affiliation data, a set of 21,279 publications remains eligible for analysis of international collaboration (92.2% of all co-authored BOF-VABB publications).

Table 1. Descriptive statistics, comparing the numbers of BOF-WoS and BOF-VABB papers

	<i>BOF-WoS</i>	<i>BOF-VABB</i>
Publications	42,421	44,790
Co-authored publications	37,964	23,079
- with affiliations	37,859	21,279
- national collaboration	15,124	11,054
- international collaboration	22,735	10,225

We compare shares of international collaboration between the BOF-WoS and BOF-VABB dataset for the whole, as well as according to the following variables:

- Publication type: Since VABB-SHW works with a limited set of high-level publication types, that is also what we use here. The consequence is that some finer distinctions present in WoS, such as the difference between article and review, cannot be represented.
- Discipline: each publication in our data has been classified into one or more disciplines based on the organizational affiliation of its Flemish authors. That is, a publication authored by Flemish researchers from the departments of sociology and history is classified under sociology and history (Guns et al., 2018).

The fine-grained data available for BOF-VABB allow us to study the main partner countries and organizations of that set in more detail. The anonymized dataset is publicly available at <https://doi.org/10.5281/zenodo.7707949>.

Results

Comparison of international collaboration in WoS and non-WoS publications

The overall degree of co-authorship in non-WoS publications is substantially lower than that in WoS publications (51.5% versus 89.8%), which confirms earlier findings by Ossenblok et al. (2014). If we limit ourselves to co-authored publications that mention affiliations, we find 60.0% international collaboration in WoS, compared to 48.1% outside of WoS (Table 1). In other words, the difference between BOF-WoS and BOF-VABB is far more pronounced when comparing co-authorship to single authorship than when comparing international to national co-authorship. This might indicate that differences between both data sets coincide with other variables, such as publication type and discipline.

Publication type

VABB-SHW distinguishes between five publication types: journal articles, books/monographs, edited volumes, book chapters and conference proceedings, whereas the part of WoS used in the PRFS only indexes journal publications and conference proceedings. While in BOF-WoS the share of international collaboration is markedly lower in conference proceedings than in journal articles, this is roughly the same in BOF-VABB (Table 2).

Table 2. Share of international collaboration by publication type, BOF-WoS and BOF-VABB

<i>Publication type</i>	<i>BOF-WoS publications</i>			<i>BOF-VABB publications</i>		
	<i>Co-authored</i>	<i>International</i>		<i>Co-authored</i>	<i>International</i>	
		<i>n</i>	<i>%</i>		<i>n</i>	<i>%</i>
Journal articles	36,555	22,202	60.7	10,562	4,529	42.9
Monographs	-	-	-	446	277	62.1

Edited volumes	-	-	-	1,747	1,164	66.7
Book chapters	-	-	-	6,816	3,543	52.0
Conference proceedings	1,304	533	40.9	1,708	712	41.7
<i>Total</i>	<i>37,859</i>	<i>22,735</i>	<i>60.0%</i>	<i>21,279</i>	<i>10,225</i>	<i>48.1%</i>

The most surprising finding here is the high share of international collaboration in book publications. Two thirds of the edited volumes are international, closely followed by monographs. Even book chapters have a share of international collaboration 10 percentage points above journal articles, despite the many formal characteristics they share. It is, however, important to keep in mind that single-authored publications are not taken into account in Table 2.

Discipline

Figure 1 shows the share of internationally co-authored publications (compared to co-authored publications with affiliation data) for the BOF-WoS and BOF-VABB data sets. As one might expect, the majority of disciplines has a larger share of international collaboration for BOF-WoS, than for BOF-VABB. However, this is not a universal pattern: we find greater shares for the BOF-VABB set in both social sciences (Political Sciences, Communication Studies) and humanities (Linguistics, Theology, Literature). More generally, the shares of international collaboration in the two sets appear to be only weakly correlated. Particularly large differences can be found for Social Health Sciences and Law. We believe this is due to the size difference between BOF-WoS and BOF-VABB in these disciplines. On the one hand, Social Health Sciences is strongly oriented toward publishing in WoS-indexed journals, leading to a situation where mainly locally oriented publications are published outside of WoS. On the other hand, Law is often mentioned as a discipline that is almost by definition focused on a local (legal) context and mainly publishing locally; hence, Law scholars are unlikely to publish in WoS journals unless they are collaborating with colleagues abroad (typically about issues of wider geographic interest, such as international law).

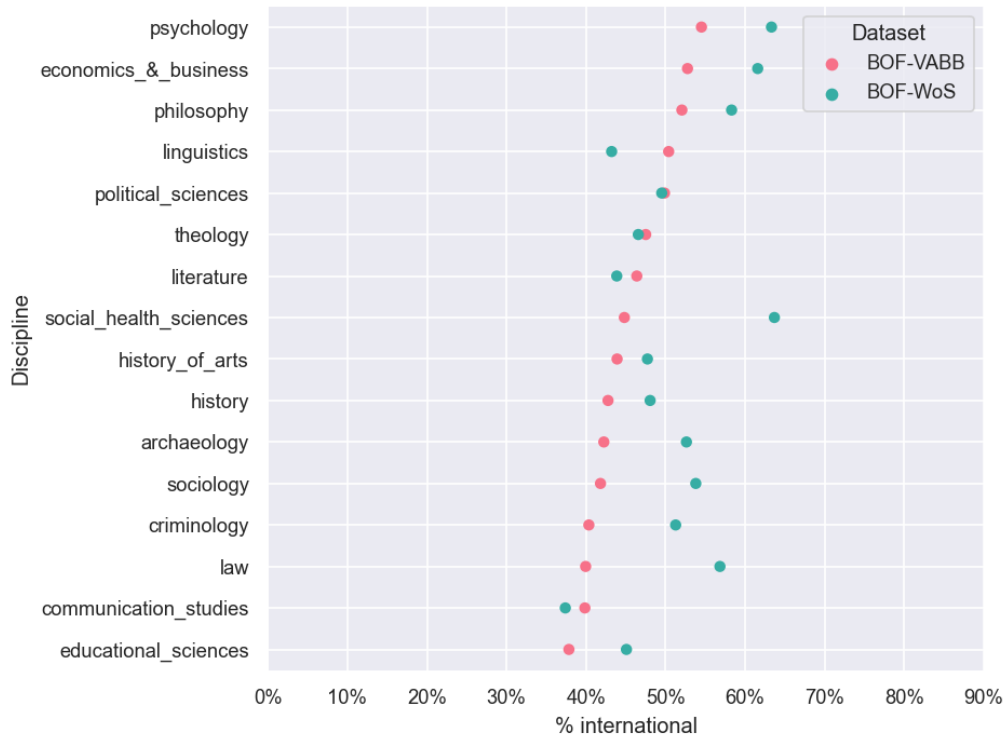


Figure 1. Share of international co-publications by discipline, for BOF-WoS and BOF-VABB

Analysis on organization level

The assignment of unique identifiers to the affiliated organizations allows additional insight in the nature of international collaboration of Flemish authors. These identifiers are in this dataset only available for BOF-VABB publications, which makes a comparison with affiliations in the BOF-WoS set impossible at the moment.

In total, the authors of the 21,279 BOF-VABB publications are affiliated to 5,671 different organizations. Of those organizations, 1,967 were not available in ROR (34.7%). This means that common organization databases only cover two thirds of the affiliated organizations for the non-WoS-subset. If only the publications with international collaboration are taken into account (n=10,225), then the share of non-ROR organizations decreases to 27.8%. In other words, the majority of non-ROR organizations are local/national ones.

Geographical distribution. International co-authors of the BOF-VABB publications are affiliated to organizations in 144 different countries (Figure 2). The largest number originates from the Netherlands (NL; 3,264 publications, distributed over 449 different organizations); the Netherlands is a neighbouring country, with the same language (Dutch) as Flanders. It is followed by the United Kingdom (GB; 1,834 publications, 328 organizations) and the United States (US; 1,444 publications, 666 organizations). Germany (DE), France (FR) and Italy (IT) each account for more than 500 publications. 58 countries account for only five or less co-publications. Note that the number of different organizations authors are affiliated to in the United States is substantially higher than the Netherlands and the United Kingdom, although both have a (much) higher number of co-authored publications.

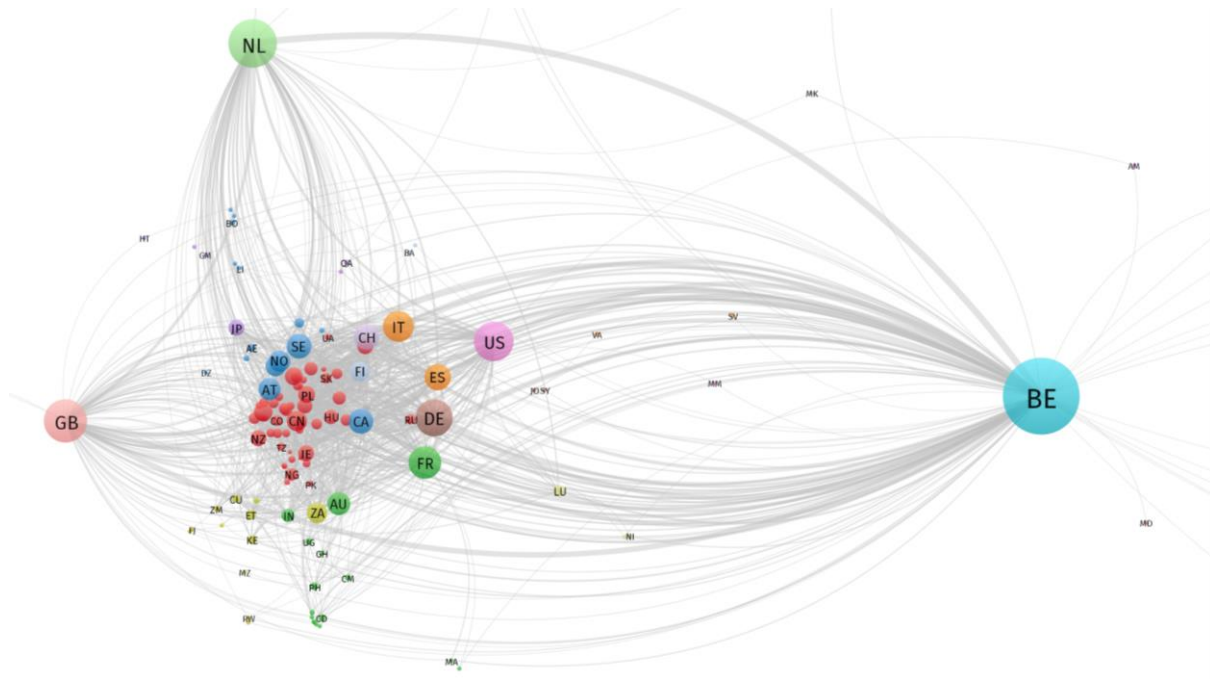


Figure 2. Collaboration network of Flanders (BE) with other countries in BOF-VABB

Type. ROR assigns a type to each of the organizations in the database. Allowed types are *education*, *healthcare*, *company*, *archive*, *nonprofit*, *government*, *facility* and *other*. Organizations added by ECOOM were coded according to the same scheme, enabling cumulative statistics as well as a comparison between ROR and non-ROR organizations. Publications with a co-author affiliated to a non-Belgian organization are distributed over 5,037 unique organizations (Figure 3). 43.2% of these are assigned to the *education* type, which, in practice, means universities or colleges. The remaining types each account for about 10% of the organizations, except *other* and *archive*, both of which are responsible for less than 5%. Although 27.8% of the individual organizations were added by ECOOM, it is remarkable that in absolute numbers the ROR organizations are surpassed by the new ones for the *nonprofit* and *company* type, and almost equal for *government*. Further in-depth analysis is required, but numbers suggest that these additional organizations are of local (Belgian and Dutch) origin. For both Low Countries – sharing the same native language – more unique organizations were added by ECOOM than picked from the ROR database.

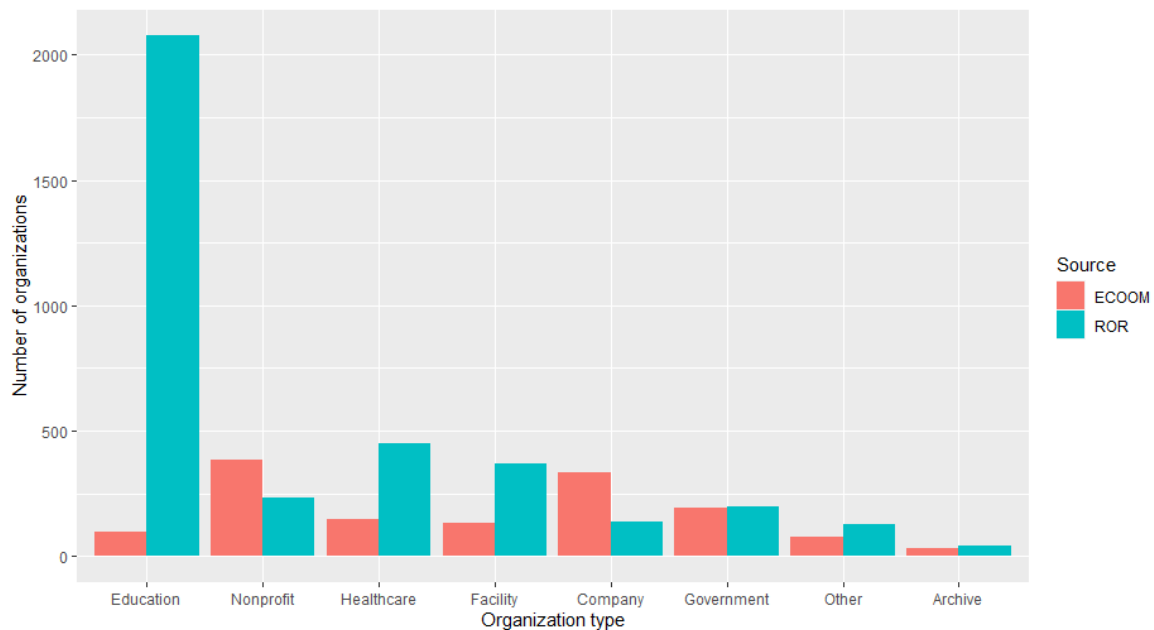


Figure 3. Number and type of organizations with which Flemish SSH researchers collaborate

Discussion

Our results demonstrate that a substantial share of SSH research in Flanders is internationally co-authored, also beyond the scope of WoS. This share varies by discipline and publication type, with the high share of international collaboration for book publications as one of the most remarkable findings. This is empirical evidence for the important role played by books in SSH communication. All in all, the results underscore that SSH research is often relevant beyond one single country, regardless of its indexation in a major bibliometric database like WoS.

We end the paper with our plans for future research. The preliminary results on organization level leave room for more detailed research and refined coding. As from the next data collection cycle on, author affiliations will be stored on publication-author-level, which will uncover more precise data about inter-institutional collaboration for non-WoS SSH publications. Furthermore, the organization database may also be coded in a more detailed way, as the ROR types are very broad. It is also our aim to deploy the extended database for other academic and administrative purposes, which can link author affiliations to other relevant datasets (Aspeslagh et al., 2022). Last but not least, the affiliated organizations in the BOF-WoS publications will be converted to a ROR-/ECOOM-identifier, which will enable comparison with non-WoS publications on a more fine-grained level.

References

- Aspeslagh, P., Engels, T., & Guns, R. (2022). Building on ROR. Enriching and customizing multi-purpose organization databases. *26th International Conference on Science and Technology Indicators, STI 2022*. <https://doi.org/10.5281/zenodo.6974621>
- Guns, R., & Hołowiecki, M. (2022). Journal lists in social sciences and the spectrum of quality standards. In T. C. E. Engels & E. Kulczycki (Eds.), *Handbook on Research Assessment in the Social Sciences* (pp. 262–277). Edward Elgar Publishing.
- Guns, R., Sīle, L., Eykens, J., Verleysen, F. T., & Engels, T. C. E. (2018). A comparison of cognitive and organizational classification of publications in the social sciences and humanities. *Scientometrics*, *116*(2), 1093–1111. <https://doi.org/10.1007/s11192-018-2775-x>
- Henriksen, D. (2016). The rise in co-authorship in the social sciences (1980–2013). *Scientometrics*, *107*(2), 455–476. <https://doi.org/10.1007/s11192-016-1849-x>

- Katz, & Martin, B. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18.
[https://doi.org/10.1016/s0048-7333\(96\)00917-1](https://doi.org/10.1016/s0048-7333(96)00917-1)
- Larivière, V., Gingras, Y., & Archambault, E. (2006). Canadian collaboration networks: A comparative analysis of the Natural Sciences, Social Sciences and the Humanities. *Scientometrics*, 68(3), 519–533. <https://doi.org/10.1007/s11192-006-0127-8>
- Luwel, M. (2021). Performance-based Institutional Research Funding in Flanders, Belgium. *Scholarly Assessment Reports*, 3(1), Article 1. <https://doi.org/10.29024/sar.29>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Ossenblok, T. L. B., Engels, T. C. E., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science—A comparison of publication patterns and incentive structures in Flanders and Norway (2005–9). *Research Evaluation*, 21(4), 280–290.
- Ossenblok, T. L. B., Verleysen, F. T., & Engels, T. C. E. (2014). Coauthorship of journal articles and book chapters in the social sciences and humanities (2000-2010). *Journal of the Association for Information Science and Technology*, 65(5), 882–897. <https://doi.org/10.1002/asi.23015>
- Tennant, J. (2020). Web of Science and Scopus are not global databases of knowledge. *European Science Editing*, 46, e51987. <https://doi.org/10.3897/ese.2020.e51987>
- 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.
<http://www.portlandpress.com/pp/books/online/wg87/087/0117/0870117.pdf>