

Comparison of classification-related differences in the distribution of journal articles across academic disciplines: the case of social sciences and humanities in Flanders and Norway (2006-2015)

Linda Sîle¹, Raf Guns¹, Frédéric Vandermoere², and Tim C. E. Engels¹

¹*Linda.Sile@uantwerpen.be, Raf.Guns@uantwerpen.be, Tim.Engels@uantwerpen.be*
University of Antwerp, Faculty of Social Sciences, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1,
2020 Antwerp (Belgium)

²*Frederic.Vandermoere@uantwerpen.be*
University of Antwerp, Faculty of Social Sciences, Department of Sociology, Sint-Jacobstraat 2-4, 2000
Antwerp (Belgium)

Abstract

Even though bibliometric analyses often rely on different disciplinary classifications, it is not known to what extent the choice of classification influences bibliometric findings. Here we explore differences in the distribution of articles across disciplines using four different classifications. We use data on social sciences and humanities (SSH) journal articles from comprehensive bibliographic databases in Belgium (Flanders) and Norway (2006-2015). In our analysis we use the original classifications used in VABB-SHW and Cristin, the Flemish and the Norwegian databases, Web of Science subject categories, and Science-Metrix journal classification.

Preliminary findings show that different classifications lead to considerable differences in the total number of SSH journal articles. For example, the percentage difference in the number of SSH publications for Norway is 17,5% when comparing the Science-Metrix and the original classification. This implies that there is a substantial number of publications in disciplinary terms residing on the boundaries between SSH and other knowledge domains. In contrast, on discipline level the differences due to the classification are small (the average difference in share is less than 2 p.p.). This might mean that if one employs a scheme with relatively broad categories then the choice of classification is of minor importance.

Introduction

Over time, multiple methods have been developed to identify academic disciplines to which publication sets belong (for an overview see Gläser, Glänzel, & Scharnhorst, 2017). One can use content-based classification approach (Bensman, 2007), classify publications (or journals) on the basis of citations (Carley, Porter, Rafols, & Leydesdorff, 2017; Leydesdorff, Bornmann, & Zhou, 2016), use text-based algorithmic approaches (Eykens, Guns, & Engels, 2019) or a hybrid text and citation-based approach (Janssens, Zhang, Moor, & Glänzel, 2009). Finally, it is possible to use a more pragmatic approach and rely on already existing journal classifications (e.g. Leydesdorff & Rafols, 2009; Wang & Waltman, 2016).

Not all of these approaches, however, are suitable for bibliometric studies of research within the social sciences and humanities (SSH). While the citation-based approach is severely limited due to the low coverage of SSH in WoS and Scopus (Leydesdorff, Hammarfelt, & Salah, 2011), text-based approaches are challenged by the fact that SSH literature is scattered across multiple disciplinary databases (such as PhilPapers, ERIC, PsycNET). Knowledge of their comprehensiveness and comparability is limited. Furthermore, such specialised databases are not available for all disciplines within SSH. An alternative is to use data from the more comprehensive national databases (e.g. VABB-SHW in Belgium, Cristin in Norway).

National databases, although more suitable for bibliometric studies of SSH, often employ local classifications (Guns, Sîle, Eykens, Verleysen, & Engels, 2018; Kulczycki et al., 2018; Ossenblok, Engels, & Sivertsen, 2012). Considering the literature on classifications in general (e.g. Bowker & Star, 2000), we know that classifications carry traces from the contexts in which they originate. It is possible that in one context a journal is understood as belonging to a discipline X while in another it is perceived as belonging to a discipline Y. This aspect of

classifications has not been explored in relation to (local) disciplinary classifications for SSH journals. This, consequently, is a challenge since it is not known how valid is the use of such classifications for the calculation of bibliometric indicators in comparative settings. Accordingly, the purpose of this article is to explore how the choice of classification influences the distribution of articles across disciplines in SSH. To do so, we use data on SSH journal articles from Flanders (Belgium) and Norway (2006-2015) acquired from two national bibliographic databases. In our analysis we employ four different classifications: (a) the VABB-SHW cognitive classification (VABB-OECD), (b) the classification used in the Norwegian list of Scholarly Journals (NPU) (c) Web of Science subject categories (WOS-SC) and (d) Science-Metrix classification for journals (SM, Archambault, Beauchesne, & Caruso, 2011). In what follows, we begin with a brief description of our data and methods. Second, we continue with preliminary findings. We present preliminary findings from our comparisons of the total number of SSH articles and the disciplinary structure for SSH. Finally, we draw links between our findings and implications for the use of bibliometrics in policy settings.

Data and methods

In this study we use data from two national bibliographic databases (VABB-SHW in Flanders and Cristin in Norway; for details on both databases see Sīle et al. 2018) and Web of Science (in-house database at ECOOM-Leuven, dataset retrieved on 23/07/2018). The analysis is conducted using eight datasets (A, B, C, D: four datasets for each country) delineated as follows. Data set A consists of peer-reviewed journal articles in SSH (2006-2015) by authors affiliated to universities (the 5 universities in Flanders and the 8 universities in Norway). The data sets B, C, and D are subsets of A (see Table 1). The data set B is limited to articles from the data set A that are indexed in the three main indices in WoS. The identification of WoS-indexed articles was carried out on article level using datasets retrieved the ECOOM-Leuven in-house WoS database and a string matching approach that allows for small differences in the matched references (for details see Sīle & Guns, 2019). The third data set C is limited to articles in journals which are included in a classification of journals developed by Science-Metrix (identified on the basis of ISSNs). The dataset D contains articles in journals that are both indexed in WoS and listed in the Science-Metrix classification.

Table 1 Samples and criteria used to delineate data subsets

Sample and criteria used to delineate data subsets	Dataset			
	A	B	C	D
Flanders	29648	12575	11542	9988
Norway	26007	9444	10399	7783
Articles indexed in Web of Science (SCIE, SSCI, AHCI)	-	Yes	-	Yes
Journals included in Science-Metrix list	-	-	Yes	Yes

To identify the extent which the choice of academic discipline classification influences bibliometric indicators for SSH, we, first assign each article to an academic discipline using four classifications. For comparability, all classifications are mapped to OECD FORD classification. Then we explore how the distribution of publications across disciplines changes depending on the classification. This exploration is carried out on the basis of percentage difference and arithmetic difference in share comparisons. All analyses are carried using fractionalised counts at the author level. The use of fractionalised counts is more appropriate given our use of bibliometric indicators as proxies of research activities.

Percentage difference is acquired by, first, dividing the difference between the number of publications acquired using one classification (V1) and the number of publications acquired using another classification (V2) by the average of the value, and then multiplying by 100. This

equation is deemed more suitable for the analysis presented here since we do not prioritise and assume one of the classifications as correct (as it would be when calculating, for example, percentage error).

Also we point out a limitation of this study that is related to the comparative nature of the analysis. We use data from two different national databases. Even though both databases are assumed to be comprehensive databases for peer-reviewed scholarly publications (for Flanders, only for SSH), we are aware that there are some differences in databases setups that might alter our results. Given the focus of the study, we assume that the acquired level of accuracy is acceptable.

Preliminary findings

Preliminary findings of this study indicate that the choice of disciplinary classification for (SSH journals) has a modest influence on bibliometric representations of SSH research: the total number of articles in SSH is over- or underestimated. However, there is practically no influence on the disciplinary structure—the distribution of articles across disciplines. The greatest difference is in the total number of SSH publications rather than difference in share of articles in specific disciplines (Table 1).

We find, for example, that the percentage difference in the total number of SSH publications for Norway is 17,5% when comparing SM with the original classification. 1201,1 journal articles that are assumed to belong to SSH according to the NPU classification, are not assumed as such using SM classification. For Flanders, comparison of SM and the original classification reveals a percentage difference of 11,5%. Other comparisons reveal smaller differences. Percentage difference between WoS and the original classification is 3,5% and 8,5% respectively. Differences in the total number and the share of SS and H publications are minor and range from practically no difference (H, SM versus original classifications, both countries) to a difference of 6% (SS, WoS versus original classification, Norway). In all cases, differences are slightly greater for Norway.

Percentage differences on article level vary considerably by discipline (range: 1,5%-154%; $M=31\%$; $SD=35\%$; $Md=18\%$). The percentage difference is especially high for the category ‘Other social sciences’ (comparing with WOS-SC: 154% for Flanders; 44% for Norway; comparing with SM: 143% for Flanders; 52% for Norway). Using the local classifications, the number in this category is lower. This might be an indication of a more conservative tendency in journal assignment to disciplinary categories: interdisciplinary research or research from new disciplines is perceived as belonging to one of the more established disciplines. This interpretation is supported also by the very low differences for categories ‘Languages and literature’ (comparing with WOS-SC: 2% for Flanders and Norway) and ‘Educational sciences’ (comparing with SM: 4% for Flanders and 5% for Norway). These low differences, however, are not consistent for the compared classifications (WOS-SC and SM) thus pointing out that also WOS-SC and SM carry assumptions on journal assignment to disciplinary categories that can be more or less in alignment to what is used in national contexts. These differences can certainly be partly explained by the uneven distribution of articles across disciplines: percentage difference for categories with a low total number of articles ($N<100$) will appear more substantial than for categories with high number of articles ($N>1000$). Nevertheless, these findings indicate that there is more agreement on journal-discipline pairs for some disciplines (e.g. Languages and Literature, Arts) than others (e.g. Other social sciences, Social and economic geography, Media and communications) and the distribution of differences is not equivalent for both countries. These findings are in line with our theory-guided expectation that classifications carry traces from the contexts they are embedded in.

Overall, these findings suggest that even though differences due to classification can be observed in the total number of articles in SSH and in the total number by discipline, these differences do not substantially influence the disciplinary structure.

Discussion

Preliminary findings of our study show that the influence of the choice of disciplinary classification for journals in SSH is considerable when, firstly, delineating SSH publications and, secondly, when calculating research performance indicators based on absolute counts of publications per discipline. In contrast, when the focus is on research representations that are based on the relative number of articles by discipline (e.g. the disciplinary structure), the differences due to the choice of classification are minor.

These findings relate to those found in science mapping literature. For example, Rafols and Leydesdorff (2009) find substantial differences in journal classifications yet the structure of science maps that is acquired using these classifications is similar. It might be that the differences in percentage change and the absence of differences in the disciplinary structure we observe can be explained statistically (as in Leydesdorff & Rafols, 2009). The disciplinary structure is not only affected by the number of publications in each discipline, but also by the number of publications *in relation* to other disciplines. For major differences in the disciplinary structure, changes in the absolute number due to classification would need to be much greater than the ones observed here.

We found a difference when comparing the total number of SSH publications identified as such using different classifications. On the one hand, this means that a considerable number of SSH publications (and journals) are residing on the boundaries between SSH and other knowledge domains (e.g. Medical fields, Environmental sciences). This might have implications when larger knowledge domains are used in indicator construction or in the choice of evaluation approach (e.g. the case of Italy described by Ancaiani et al., 2015). On the other hand, the small differences in the disciplinary structure seems to suggest that the choice of academic discipline classification is of no importance since the results are altered only to a minor extent. However, two points can be highlighted. First, even though the differences we identify appear small (on average 2.2 p.p.), they can turn out crucial if linked to some reward mechanism (e.g. funding allocation or promotion). In such contexts, even a difference of 2 p.p. may have consequences especially for small yet highly specialised knowledge domains. Second, these small alterations might be a consequence of the mapping activity employed in this analysis. As noted, all the classifications that we used were mapped to OECD FORD to improve comparability. However, for SSH OECD FORD is limited only to 12 either established (e.g. Psychology) or broad disciplines (e.g. Media and Communications) and 2 residual categories (Other social sciences and Other humanities). Such structure carries a risk that research in more specific disciplines that have designated categories in some disciplines (e.g. History of Ideas or Science, Technology, and Society studies) is rendered invisible.

References

- Ancaiani, A., Anfossi, A. F., Barbara, A., Benedetto, S., Blasi, B., Carletti, V., ... Sileoni, S. (2015). Evaluating scientific research in Italy: The 2004–10 research evaluation exercise. *Research Evaluation*, 24(3), 242–255. <https://doi.org/10.1093/reseval/rvv008>
- Bensman, S. J. (2007). Garfield and the impact factor. *Annual Review of Information Science and Technology*, 41(1), 93–155. <https://doi.org/10.1002/aris.2007.1440410110>
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: classification and its consequences* (First paperback edition). Cambridge, Massachusetts London, England: The MIT Press.
- Carley, S., Porter, A. L., Rafols, I., & Leydesdorff, L. (2017). Visualization of Disciplinary Profiles: Enhanced Science Overlay Maps. *Journal of Data and Information Science*, 2(3), 68–111. <https://doi.org/10.1515/jdis-2017-0015>

- Eykens, J., Guns, R., & Engels, T. C. E. (2019). Article Level Classification of Publications in Sociology: An Experimental Assessment of Supervised Machine Learning Approaches. *Proceedings of ISSI 2019*.
- Gläser, J., Glänzel, W., & Scharnhorst, A. (2017). Same data—different results? Towards a comparative approach to the identification of thematic structures in science. *Scientometrics*, *111*(2), 981–998. <https://doi.org/10.1007/s11192-017-2296-z>
- 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>
- Janssens, F., Zhang, L., Moor, B. D., & Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. *Information Processing & Management*, *45*(6), 683–702. <https://doi.org/10.1016/j.ipm.2009.06.003>
- Kulczycki, E., Engels, T. C. E., Pölonen, J., Bruun, K., Dušková, M., Guns, R., ... Zuccala, A. (2018). Publication patterns in the social sciences and humanities: evidence from eight European countries. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2711-0>
- Leydesdorff, L., Bornmann, L., & Zhou, P. (2016). Construction of a pragmatic base line for journal classifications and maps based on aggregated journal-journal citation relations. *Journal of Informetrics*, *10*(4), 902–918. <https://doi.org/10.1016/j.joi.2016.07.008>
- Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the American Society for Information Science and Technology*, *62*(12), 2414–2426. <https://doi.org/10.1002/asi.21636>
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, *60*(2), 348–362. <https://doi.org/10.1002/asi.20967>
- 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. <https://doi.org/10.1093/reseval/rvs019>
- Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, *60*(9), 1823–1835. <https://doi.org/10.1002/asi.21086>
- Sīle, L., & Guns, R. (2019). Article-level matching of Web of Science to a local database in a comparative context. *Proceedings of ISSI 2019*.
- Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. *Journal of Informetrics*, *10*(2), 347–364. <https://doi.org/10.1016/j.joi.2016.02.003>

Table 2 Summary of differences in the total number of SSH, SS, and H journal articles for Flanders and Norway (2006-2015)

Value	Classification		Percentage difference, %				Difference in share, p.p.				
	I	II	Dataset B	Dataset C	Dataset D	Dataset B	Dataset C	Dataset D	Dataset B	Dataset C	Dataset D
The total number of SSH articles	VABB-OECD	WoS	3	-	4	-	-	-	-	-	-
	NPU	WoS	8	-	9	-	-	-	-	-	-
	VABB-OECD	SM	-	11	12	-	-	-	2,7	-	-
	NPU	SM	-	17	18	-	-	-	2,7	-	-
	WoS	SM	-	-	F: 8 N: 10	-	-	-	-	-	-
The total number of SS articles	VABB-OECD	WoS	1	-	2	1,6	-	-	-	-	1,8
	NPU	WoS	6	-	7	1,6	-	-	-	-	1,9
	VABB-OECD	SM	-	14	14	-	-	-	-	-	1,9
	NPU	SM	-	20	21	-	-	-	-	-	1,8
	WoS	SM	-	-	F: 12 N: 14	-	-	-	-	-	F: 3,7 N: 3,6
The total number of H articles	VABB-OECD	WoS	3	-	7	1,9	-	-	-	-	2,2
	NPU	WoS	4	-	7	3,0	-	-	-	-	2,6
	VABB-OECD	SM	-	10	14	-	-	-	0,1	-	0,5
	NPU	SM	-	19	19	-	-	-	0,3	-	0,1
	WoS	SM	-	-	F: 20 N: 26	-	-	-	-	-	F: 2,6 N: 6,6

Table 3 Summary of differences in the disciplinary structure for Flanders and Norway (2006-2015) (F – Flanders, N – Norway)

Classification	Cosine similarity				Difference in share, p.p.								
	Dataset B	Dataset C	Dataset D	Dataset B	Dataset C	Dataset D	Dataset B	Dataset C	Dataset D	Dataset B	Dataset C	Dataset D	
VABB	0,99	-	0,99	-1	-	-1,2	-0,9	-	-	-1,2	1,3	-	1,5
NPU	0,98	-	0,98	-1,7	-	-1,8	-1	-	-	-1,1	2,2	-	2,5
VABB	-	0,97	0,98	-	0,7	0,6	-	0,5	-0,1	-	-	2,6	2,6
NPU	-	0,98	0,99	-	0,2	0,1	-	0,4	-0,1	-	-	1,5	1,3
SM	WoS	-	F: 0,98 N: 0,97	-	-	F: 1,7 N: 2,0	-	-	F: 1,5 N: 1,2	-	-	-	F: 2,1 N: 2,8