Integrating metrics to measure research performance in social sciences and humanities: the case of the Spanish CSIC

Adrián A. Díaz-Faes¹*, María Bordons¹, Thed N. van Leeuwen²

¹adrian.arias@cchs.csic.es, maria.bordons@cchs.csic.es
IFS, Spanish National Research Council (CSIC), Albasanz 26-28, Madrid 28037 (Spain)

*Corresponding author: adrian.arias@cchs.csic.es. Tel. +34 916022895

²leeuwen@cwts.leidenuniv.nl
CWTS-Centre for Science and Technology Studies, Leiden University, PO Box 905
2300 AX Leiden (the Netherlands)

Abstract: Knowledge dissemination in the social sciences and humanities (SSH) is characterized by an assorted set of publication channels and a more prevalent use of local languages, so international bibliographic databases do not provide a practical study source by themselves. This paper aims at laying the foundations for a comprehensive study of the research performance of SSH CSIC researchers from a micro-level perspective. Both the WoS and an internal CSIC database (‘ConCiencia’) are used in combination with a set of indicators describing the activity profile of researchers and the prestige of publication channels. Differences between the publication patterns of researchers in both domains are examined and so is the relationship between their research performance and their personal features. For researchers in humanities, the higher their academic rank and age the greater their activity in books and non-WoS articles was found to be; whereas for those working in the social sciences fields, a higher academic rank does not prefigure whether a researcher’s scientific activity is more likely to be national or international in scope. Finally, a series of recommendations to consider output beyond the boundaries of WoS articles, and develop more accurate indicators for assessing the prestige of publications channels is expounded.

Keywords: publication patterns, social sciences & humanities, prestige indicators, micro level, Spain.

Acknowledgments

Adrián A. Díaz-Faes is granted with a JAE predoctoral fellowship by the Spanish National Research Council (CSIC). A draft version of this paper was presented at the ISSI 2015 conference (Istanbul).
Introduction

Outlining the scholarly performance of scientists in social sciences & humanities (SSH) is often regarded as a challenge for bibliometric purposes since the predominant publication channels in these fields are not well covered by large bibliographic databases such as the Web of Science (WoS) or Scopus (Hicks, 2004; Nederhof, 2006; den Hertog et al., 2014). Therefore, it is quite clear that dealing with journal publications does not suffice when SSH is the subject matter under scrutiny (Archambault et al., 2006; Sivertsen & Larsen, 2012, van Leeuwen et al, 2016) since books and book chapters remain a major communication channel in these domains, and particularly so, in the humanities domain. Moreover, due to the more local orientation of SSH research, knowledge dissemination in these domains typically involves a more prevalent use of national journals and local languages (Bordons et al., 2013). This holds true notwithstanding the fact that a growing trend of publishing in international journals and in English has been recently identified in these domains, partly due to the influence of the parameters used in performance-based funding systems (Ossenblok, Engels, & Sivertsen, 2012). The bearing of these factors on our case of study may be due to the different approach of SSH towards society when compared to that of the natural, life and biomedical sciences, since SSH scholars address social and cultural issues that may have a direct influence on society (societal impact) (van Leeuwen, 2013). Bearing in mind these specific features pertaining to the SSH fields, a number of initiatives have been undertaken to expand SSH coverage in bibliographic databases. In the past few years, the WoS and Scopus have respectively increased the number of SSH journals by 22% and 39% (Hicks & Wang, 2011). In addition, the Book Citation Index was launched in 2011, but it suffers from the same coverage shortcomings identified in the journal databases, namely, although 61% of the total share of entries belongs to SSH, there is a strong bias in favor of English-written output (96% of the books) and editors located in specific countries, such as the UK and the US (Torres-Salinas et al., 2014). Therefore, despite said efforts, the national dimension is still inadequately covered in large commercial bibliographic databases.

On the other hand, even though there has been a certain trend to consider SSH as a whole, differences between the behavior of scientists working in the social sciences and humanities domains have been described in the literature (Mañana-Rodríguez & Giménez-Toledo, 2013; den Hertog et al., 2014). Although collaborative research is a growing practice in SSH (Ossenblok, Verleysen, & Engels, 2014) and informal collaboration may play a significant role (Olmos-Peñuela, Molas-Gallart, & Castro-Martínez, 2014), individual authorship still prevails among researchers in the humanities, whilst those working within the scope of social sciences show a more collaborative pattern. This becomes far more evident by comparing a discipline such as Psychology, which has become more quantitative and experimental over time, with Philosophy, where scientists work relatively isolated (Cronin, Shaw, & La Barre, 2003), reflecting differences in prevailing socio-cognitive structures and work practices among disciplines themselves. Even within the scope of humanities, several types of research have been identified and differences between “traditional” and “modern” research have been pointed out (Oschner, Hug, & Daniel, 2013). Thus, the SSH fields need to be considered as a heterogeneous domain which comprises disciplines and researchers with diverse scholarly communication habits (Verleysen & Weeren, 2016).
Thus, the above factors hinder the potential capacity of traditional bibliometric analyses to provide a reliable picture of the scientific activity of SSH researchers. As a consequence, the development of a Current Research Information System (CRIS) to obtain full coverage of publications in SSH was suggested (Martin et al., 2010). This type of database includes various kinds of peer-reviewed publications (books, book chapters and journals, including both WoS and non-WoS publications) and has been developed in countries such as Norway, Denmark, Finland and Belgium (Flanders). It has been motivated by the need to monitor the performance of university scholars in line with the implementation of performance-based funding models of university research, which has created powerful incentives within university systems (Hicks, 2012). In Norway, for instance, the Cristin database collects all bibliographic information from Norwegian universities, research institutes and hospitals in all fields of research since 2005. Plus, a performance-based indicator designed for scholarly output (books, book chapters and journals) at institutional level, involves the conduct of a qualitative assessment of publication channels and the establishment of the general consensus among peers in order to provide a comparable measurement tool and foster publishing through the most prestigious publication channels (Sivertsen, 2010). In the Belgian case, the Flemish government supported the creation of the VABB-SHW database to achieve total coverage of SSH peer-reviewed publications including journal articles (both WoS and non-WoS publications), books, book chapters and conference proceedings published since 2000 by researchers affiliated to Flemish universities (see Engels, Ossenblok, & Spruyt, 2012).

As for other countries, such as Spain, where no policy measures with a view to create a national bibliographic database ensuring full coverage have been implemented, providing a comprehensive picture of the SSH fields poses a difficult challenge. However, since many institutions have created their own databases for evaluative purposes, it would be practicable to study the research activity of SSH scientists by combining large bibliographic databases and institutional ones (de Filippo et al., 2011). This paper is focused on the scientific activity of SSH researchers at the Spanish National Research Council (CSIC), the largest public institution dedicated to research in Spain with more than 4,000 researchers and 125 affiliate institutes spread countrywide. CSIC institutes are organized in eight thematic areas, being social sciences and humanities one of them, which comprised 17 research institutes in 2007. This study’s aim is to build a comprehensive picture of the publication behavior of SSH CSIC researchers from a micro-level perspective combining different sources of information. For that purpose, an assorted set of publication channels including books, book chapters, international and national journal articles have been considered and specific indicators to assess the prestige of the different publication channels have been applied. Differences between the publication patterns of SSH researchers have been examined, and the relationship between their research performance and personal features such as age and academic rank has been analyzed. Additionally, the influence of scientific performance on career advancement is addressed.

Methodology

This study analyzes the scientific output of 257 active researchers affiliated to the CSIC in 2007 including both permanent researchers and postdoctoral research fellows. The
period under analysis is 2007-2011. Scientists work on different disciplines which include: art, archeology, history, linguistics, literature, philosophy and religion (in the humanities) and economics, geography, library and information science and sociology (in the social sciences). Scientists were assigned to social sciences (74 scientists, 29%) or humanities (183, 71%) according to their institutional affiliation (6 institutes in SS and 11 in H). For the whole SSH fields, male researchers represented 58% of the total sample, the average age of researchers was 50, and half the researchers were in the lowest permanent scientific category (tenured scientist). Postdoctoral research fellows accounted for only 6% of researchers in the SSH research domains. Small differences between humanities and social sciences can be observed in Table 1.

Table 1. Personal features and academic rank of SSH researchers.

<table>
<thead>
<tr>
<th></th>
<th>Humanities (N=183)</th>
<th>Social Sciences (N=74)</th>
<th>Total (N=257)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>109</td>
<td>40</td>
<td>149</td>
</tr>
<tr>
<td>Female</td>
<td>74</td>
<td>34</td>
<td>108</td>
</tr>
<tr>
<td><strong>Rank 2007</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-doc</td>
<td>10</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Tenured scientists</td>
<td>95</td>
<td>41</td>
<td>136</td>
</tr>
<tr>
<td>Research scientists</td>
<td>44</td>
<td>13</td>
<td>57</td>
</tr>
<tr>
<td>Research professors</td>
<td>34</td>
<td>15</td>
<td>49</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>50 ± 9</td>
<td>49 ± 10</td>
<td>50 ± 9</td>
</tr>
<tr>
<td></td>
<td>(28-70)</td>
<td>(32-70)</td>
<td>(28-70)</td>
</tr>
</tbody>
</table>

Note: age expressed as average ± standard deviation (min-max).

Publications were collected from two different sources: the Web of Science (WoS) (SSCI+AHCI+SCIE), which was used to download the articles with international scope, and an internal CSIC database called ‘ConCiencia’, to provide for the collection works published under modalities not covered by the WoS (books, book chapters and non-WoS journal articles). The total counting method was used. Ad hoc software was used to cope with name inconsistencies and to achieve the proper allocation of publications to authors. A manual revision of the assembled output, and in particular, that obtained from the ‘ConCiencia’ system, was conducted. Furthermore, some basic criteria were established, most notably in the case of book chapters. A minimum length of four pages was established as a threshold for book chapters (Engels et al., 2012), while certain typologies such as forewords or epilogues were only retained when they were at least ten pages long. Conference proceedings were only considered when they included an ISBN number, were part of published books and were at least four pages long. Based on the information retrieved, the following indicators were calculated:

a) Activity profile of researchers:

- **Share of papers by publication channel**: proportion of books, book chapters, WoS articles and non-WoS articles published by a researcher with regard to his/her total number of publications.
- **Sum of publications**: the total number of publications published by each researcher, including authored/edited books, book chapters and journal articles.
- **Average number of authors/paper:** this indicator measures the average number of authors per publication for the total output of a given researcher.

- **% International collaboration:** share of the total output of each researcher co-authored with researchers affiliated to one or more foreign institutions.

- **% English:** proportion of a researcher’s output published in English.

- **Diversity of publication channels.** Simpson index of diversity (1-D) was computed to compare the diversity of publication channels of researchers between social sciences and humanities. Simpson index (Simpson, 1949) is a measure of diversity used in ecology to quantify the diversity of species. It has been also applied in other domains such as demography or information science. It ranges between 0 and 1. The greater the index the greater the diversity.

\[
\text{Simpson index of diversity} = 1 - \frac{\sum n(n-1)}{N(N-1)}
\]

Where \( n \) is the total number of items of a particular publication type and \( N \) is the total number of publications (all publication types).

b) Prestige of publication channels: a set of various indicators of the impact/prestige of publication channels were applied as summarized in Table 2. It should be noted that all the scientific output of scientists is included in the ‘Conciencia’ system, with no quality selection; therefore considering the prestige of publication channels in the analysis of data becomes particularly important. Moreover, measuring the prestige of publication channels was considered a better choice than measuring citations received by individual publications since a long citation window is needed in the SSH to obtain reliable data due to the tendency of researchers to cite old literature in this area (Huang and Chang, 2008, Martin et al, 2012).

- **Top books and chapters (pptop10% Books & Chapters):** proportion of books and chapters of a given researcher published by the top 10% publishers according to the Scholarly Publisher Indicators Project (SPI) (Giménez-Toledo, Tejada-Artigas & Mañana-Rodríguez, 2013). This project describes the Indicator of Quality of Publishers according to Experts (ICCE) which is based on a quality assessment of publishers rated by Spanish researchers in a national survey. The SPI Project provides a general ranking of publishers as well as different rankings with regard to each of the 16 SSH disciplines considered under the project. Since many institutes show a multidisciplinary profile and their publications are spread across a wide range of disciplines, the general ranking was used, after checking it covered the most prestigious publishers within each discipline (on average 8 out of the 10 top publishers within each discipline were included among top publishers of the general ranking). Finally, the distribution of books and book chapters by type of publisher (commercial, institutional and others) was addressed. University Presses were considered in the institutional type, due to their academic nature and not-for-profit status.

- **Share of WoS papers in the most prestigious journals within each field.** The share of WoS papers published in the highest impact factor journals within each subject category might have been used, but impact factor is not provided by Thomson Reuters for Arts & Humanities journals. Accordingly, the SCimago Journal Rank (SJR) (Guerrero-Bote & Moya-Anegón, 2012), which is available also for Arts &
This is a postprint version of:

Humanities journals, was considered. The SJR is an indicator of journal prestige which takes into account both the number of citations received by a journal and the importance or prestige of the journals where such citations come from. The share of papers published in the top 25% journals of the ranking of journals in descending order of SJR by subject category\textsuperscript{v} was calculated (source: www.scimagojr.com).

- **Share of papers in top non-WoS journals (pptop10% non-WoS articles):** percentage of non-WoS papers published in top journals according to the Integrated Scientific Journal Classification (CIRC) (Torres-Salinas et al., 2010). CIRC is a proposal for a categorization of SSH journals developed by a group of experts in bibliometrics in Spain. It defines four categories of journals (A, B, C and D) according to their visibility measured by integrating the results of different journal classifications and assessments tools. For the purposes of this paper, “top journals” are those included in categories “A” and “B”.

<table>
<thead>
<tr>
<th>Type of publication channel</th>
<th>Indicators of impact/prestige</th>
<th>Source (indicator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WoS Articles</td>
<td>SJR (25% top journals by subject category)</td>
<td></td>
</tr>
<tr>
<td>Non-WoS Articles</td>
<td>CIRC (journals in A and B categories)</td>
<td></td>
</tr>
<tr>
<td>Books/Book Chapters</td>
<td>SPI (10% top publishers by expert opinion)</td>
<td></td>
</tr>
</tbody>
</table>

- c) Personal data: age, gender and academic rank (P=postdoctoral research fellow, TS=tenured scientist, RS=research scientist and RP=research professor) both at the start and the end of the period. This information was provided by the CSIC.

- d) Weighted productivity: since publishing patterns differ by discipline, a weighted productivity score similar to that proposed by Sivertsen (2010)\textsuperscript{iii} was obtained for each scientist. This score is calculated as a compound of two aspects of publications: on the one hand, the type of publication channel (WoS article, non-WoS article, book and book chapter), and on the other hand, two levels of publications are defined within each publication channel (top and non-top). Publications are thus weighted according to their prestige and type of publication channel (Table 3). An analysis of the relation between the research performance of scientists and promotion reward rates by the end of the period (i.e. 2011) was conducted. Note that this part of the study was limited to 180 scientists, since the following were not considered: Post-docs, due to the temporary nature of their position; Research Professors, already in the top rank category; and scientists who had left the institution by 2011, either by retirement or job change.

<table>
<thead>
<tr>
<th>Table 3. Weighted productivity scores.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>WoS Articles</td>
</tr>
<tr>
<td>Non-WoS Articles</td>
</tr>
<tr>
<td>Books</td>
</tr>
<tr>
<td>Book Chapters</td>
</tr>
</tbody>
</table>

Non-linear Principal Component Analysis (NLPCA) (see, for example, Guttman, 1941) was used to explore the relationship between the personal features of researchers
and their research performance. This method is a non-linear equivalent of the traditional (linear) PCA which aims to shrink a large number of continuous variables into fewer uncorrelated underlying variables (principal components). By using a few components, main patterns of the dataset can be brought to the foreground. NLPCA allows for the incorporation of nominal and ordinal variables by means of optimal scaling. The aim of optimal scaling is to maximize the proportion of variance accounted for the quantified variables (see Linting et al., 2007). Statistical analyses were performed by means of SPSS (v.20).

**Findings**

**Publication profile of researchers**

A total of 2,993 documents were published by CSIC’s SSH active researchers during 2007-2011. Differences between humanities and social sciences in the main publication types used are observed (Figure 1): WoS articles predominate in social sciences (41%) while book chapters (48%) are the most frequent publication channel in humanities (Figure 1). Percentages of articles in national journals are fairly similar for both domains (25%). In addition, although books do not present high share values (15% vs. 10%), their importance in humanities is unquestionable.

![Figure 1. Share of publication channels by domain.](image-url)

The diversity of publication channels in the output of researchers is the norm in SSH research. Roughly speaking, one third of the researchers presented output in all four different publication types: articles covered by WoS, non-WoS articles, books and book chapters. Three and two types of publication channels were observed, respectively, in 41% and 18% of the researchers, while results of just a single type were found for only
8% of researchers. There are some differences between social sciences and humanities worth noting: researchers who disseminate research through all four different publication channels are more frequent in humanities (37% vs. 27%), while using a single channel is more likely to occur in the social sciences domain (see Table 4). Moreover, using only WoS-covered journals is more common for social sciences researchers (15% vs. 2%). A higher diversity is observed in humanities when compared to social sciences (Simpson index of diversity=0.61 in H vs. 0.53 in SS). Since researchers with less than 4 publications could never reach four publication channels, the index was also calculated for researchers with at least 4 publications and higher diversity was also noted for those in humanities (0.61 vs. 0.56).

### Table 4. Distribution of researchers by number of publication channels and area.

<table>
<thead>
<tr>
<th>Number of publication channels</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humanities</td>
<td>37%</td>
<td>43%</td>
<td>16%</td>
<td>4%</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>27%</td>
<td>35%</td>
<td>22%</td>
<td>16%</td>
</tr>
<tr>
<td>Total</td>
<td>34%</td>
<td>41%</td>
<td>18%</td>
<td>8%</td>
</tr>
</tbody>
</table>

As for the publication in top channels, Figure 2 shows some remarkable differences between scientists by domain. A higher rate of social sciences scientists publishes their research in top SJR journals when compared with their humanities colleagues (46% vs. 15%). On the other hand, according to the SPI, humanities researchers publish books and book chapters via top publishers more often than social sciences scholars (36% vs. 24%).

![Figure 2](image.png)

**Figure 2.** Usage of top publication channels by domain.

Furthermore, a look at the type of book publisher reveals some interesting insights (Table 5). National books and book chapters predominate over international
ones in both social sciences and humanities (75% vs. 25%). Commercial publishers are most often used at international level (almost 60% of international ones) while institutional publishers play the dominant role at national level (more than 50%) in both domains. The ‘Other’ row includes a majority of scientific reports published by non-profit organizations or corporations. In general, research released through commercial or institutional publishers is more likely to be published by high prestige book publishers, while research grouped under the ‘other’ category shows a negligible rate of top publishers. In social sciences, the highest use of top channels corresponds to commercial publishers both at national (e.g. Gredos, Alianza) and international levels (e.g. Elsevier, Sage). As far as humanities are concerned, commercial publishers also claim the highest share of top channels at international level, but institutional publishers prevail in the case of national publications.

### Table 5. Share of books and book chapters by type of publisher.

<table>
<thead>
<tr>
<th>Type</th>
<th>Humanities</th>
<th>Social Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (National)</td>
<td>N (International)</td>
</tr>
<tr>
<td>Commercial</td>
<td>639 (32.9)</td>
<td>234 (56.8)</td>
</tr>
<tr>
<td>Institutional</td>
<td>877 (58.5)</td>
<td>157 (38.1)</td>
</tr>
<tr>
<td>Other</td>
<td>126 (8.5)</td>
<td>21 (5.1)</td>
</tr>
<tr>
<td>Total</td>
<td>1,642</td>
<td>412</td>
</tr>
</tbody>
</table>

### Research performance of scientists

Main statistics concerning the research performance of SSH scientists are shown in Table 6. The average number of total publications observed is higher for researchers in the humanities domain than in social sciences (14.9 vs. 11.3) due, in particular, to their high production rate of book chapters. Humanities researchers show a higher presence with top publishers for books and book chapters while social sciences scientists present a higher share of articles in top international journals (%top SJR articles). Finally, a higher co-authorship index is observed for social sciences scientists, who also display a more internationally-oriented activity according to their more prevalent use of English (39% vs. 13%) and a higher rate of collaboration with foreign partners (25% vs. 16%).

### Table 6. Description of the research performance of SSH researchers.

<table>
<thead>
<tr>
<th>Type</th>
<th>Humanities</th>
<th>Social Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>No. Books</td>
<td>2.1</td>
<td>2.5</td>
</tr>
<tr>
<td>No. Chapters</td>
<td>7.1</td>
<td>5.7</td>
</tr>
<tr>
<td>No. WoS Articles</td>
<td>1.8</td>
<td>4.3</td>
</tr>
<tr>
<td>No. Non-WoS Articles</td>
<td>3.9</td>
<td>4.8</td>
</tr>
<tr>
<td>Sum of Publications</td>
<td>14.9</td>
<td>12.2</td>
</tr>
</tbody>
</table>
NLPCA was used to explore the possible relations between the personal features of researchers and their performance. Two different studies were conducted to address the potential behavioral differences of scientists in each domain. In both cases, variance for the first factorial plane is about 41%. The loading plots show the correlation among variables and the principal components for an estimate of the extent to which each variable contributes to explain the new uncorrelated variables (principal components). Besides, since the cosines of the angles approximate Pearson correlations between indicators some interesting patterns have been found. In the case of humanities, the factorial plot (Figure 3a) shows that scientists of higher rank and age present a stronger activity in books and non-WoS articles and a higher share of top books and book chapters. However, a high number of total publications seems to be in closer relation to a high number of book chapters than to the researcher holding a higher rank. On the other hand, the publishing pattern of the WoS, which is a particular feature of a group of scientists placed in the fourth quadrant, is characterized by higher collaboration (as measured through the average number of authors per paper), and more internationally-oriented research (as measured through the share of English-written papers and internationally-coauthored publications). The scientists included in this group are younger and hold a lower academic rank. PC scores (Figure 3b) reveal that many researchers show a high number of publications due to the high proportion of book chapters in the total output (third quadrant) whilst there are two small groups of researchers favoring the publication of books (first and second) and articles in international journals (fourth).
In turn, for social sciences scientists, the factorial plot reveals two remarkable patterns (Figure 4a). A higher academic rank is linked to internationally-oriented scientific publications (high share of WoS articles) and a high share of SJR top papers (first quadrant); but a nationally-oriented scientific pattern also happens to be in connection with a higher rank (second quadrant). This type of scientist is usually older, publishes in top non-WoS journals, and yields the highest number of publications. No differences in terms of co-authorship rates have been found between these two patterns. PC scores (Figure 4b) show that nationally-oriented scientists are more numerous (second quadrant) than those mainly involved with research with an international scope (first quadrant).

**Figure 4.** NLPCA results for social sciences. Loading plot (4a) and scores for researchers performance (4b).

**Research performance and promotion**
The distributions of the weighted productivity score of researchers who got a promotion (71 researchers) and those who did not (109 researchers) are shown in the box-plots in Figure 5. In the humanities domain, there is not a clear relationship between research performance and ensuing promotion. Interestingly, some researchers who did not promote obtained outstanding performance results as attested by the presence of several outliers in the distribution of non-promoted scientists. With regard to the social sciences domain, median values of weighted productivity are clearly higher for scientists who were promoted in the two ranks eligible for this analysis.

**Discussion and conclusions**

Our study provides a comprehensive picture of the activity of SSH CSIC researchers from a micro-level perspective. The activity of CSIC researchers in different life and natural sciences disciplines has been analyzed in previous studies on the basis of data collected from the Web of Science (see for example, Costas, van Leeuwen, & Bordons, 2010). Nonetheless, this study introduces a couple of major methodological novelties to address the specific features of the communication habits of SSH scientists: a) the use of an institutional database (‘Conciencia’) in an attempt to attain a thorough coverage of the scholarly output of researchers, including not only WoS publications but also books, book chapters and non-WoS journal articles; and b) the application of a series of indicators to measure the prestige of the different publication channels.

**The ‘Conciencia’ system**

The ‘Conciencia’ system was created at CSIC to provide an extensive view of research activity in its affiliate institutions, enhance research assessment processes, and facilitate the production of annual reports. With this aim, the database includes different types of scientific output (articles, books, book chapters and proceedings) as well as other type of outputs (not considered in this study) related to the dissemination and transfer of
scientific knowledge, and teaching activities. The ‘Conciencia’ system presents a valuable collection of data, but in the course of this research several constraints hindering the use of this source in bibliometric studies were identified. In this context, the presence of uncompleted bibliographic references, duplicate records or missallocation of document types, and the confusion between authored-books and edited-books were observed in some cases and are worth noting. To solve these problems, more rigor in both the input of data (carried out by researches themselves) and in the screening and validation processes (by the institution) would be advisable and essential to increase the accuracy and reliability of the resulting studies. Moreover, recording in the database whether a publication has been peer-reviewed or not would be welcome as an significant improvement.

Range of publication channels

In our work, the relevance of books and book chapters in the SSH fields, and in particular, in the humanities domain, is confirmed, since they roughly account for 60% of the publications in humanities, as described by Piro, Aksnes, & Rorstad (2013), and about 35% of the output in social sciences as against 50% according to said reference study. With regard to WoS articles, they outnumber non-WoS journal articles in social sciences, but the opposite occurs in humanities (also described by Engels at el., 2012). In fact, WoS articles represent only 41% of the output of CSIC researchers in social sciences and 12% in humanities. Hence, focusing only on WoS articles would shrink meaningfully the visibility of these researchers. Besides, around 19% of social sciences researchers and 41% of those working in the humanities domain would be rendered invisible in a WoS-based study since they do not appear in any publication covered by this database in the period under study. Accordingly, including other types of output beyond the scope of WoS-articles is essential to obtain a true and accurate picture of the activity of SSH researchers.

As previously noted in the literature, SSH is a heterogenous domain (Nederhof, 2006; Verleysen & Weeren, 2016). Some relevant differences between social sciences and humanities are identified in our study in connection with their main publication channels, the national/international orientation of research and team research practices. Social sciences are found to stand closer to the natural sciences not only in their methodological approach (more empirical and ‘objective’ methods are applied), but also in terms of publication patterns (greater use of journal articles), scope of the research (higher internationalization) and collaborative practices (stronger than in the humanities domain).

Publication behavior and academic rank

The study of the relationship between the publication behavior of scientists and their academic rank enables us to explore which publication patterns are more clearly linked to academic promotion in SSH at the Spanish CSIC. In the case of humanities, academic rank correlates with the share of authored books, in particular by having books published by top publishers, and also with the share of non-WoS journal articles, mainly in non-English languages, thereby hinting at the importance granted to nationally-oriented research. In addition, academic rank is found to rise with the age of researchers, which means that long academic careers are rewarded. The total number of
publications does not correlate with rank, probably because, for scientists with a high share of published books, it is difficult to obtain a high total number of publications since writing books takes usually longer than other types of scientific works. In turn, in social sciences, a higher academic rank does not prefigure whether a researcher’s scientific activity is more likely to be national or international in scope. For some authors, an internationally-oriented profile is described, since they publish mainly WoS-journal articles, tend to select the most prestigious journals (top SJR), use English as publication language and are inclined to collaborate with foreign colleagues. However, there is also a significant number of researchers favoring a more national approach to their research, focusing on non-English papers in top journals not covered by the WoS.

The findings of our study reveal that internationally-oriented research is more often related with higher academic rank in social sciences than in humanities, which is consistent with the higher importance given to internationality by scientists themselves in social sciences (Hemlin, 1993). Although the growing internationalization of research has been pointed out for the SSH domains as a whole in recent years (Engels et al., 2012), humanities are trailing behind social sciences in this process. Internationality is highly rewarded by research assessment processes at the institutional (CSIC, 2009) and national level in Spain (BOE, 2014) in order to enhance the visibility of research and foster the mainstreaming of its centers into the international scientific community. As a consequence, an increasing share of the output produced by Spanish authors are directed to WoS journals (Bordons et al., 2013; Gómez, Sancho, Bordons, & Fernández, 2006; Jiménez-Contreras, Moya-Anegón, & López-Cozar, 2003) a trend which has also been found to be the case in other countries such as Germany and Belgium (Flanders) (van Leeuwen, 2013; Engels, et al., 2012). Interestingly, in our study, a higher degree of internationalization is observed for younger scientists and in social sciences compared to humanities. Young researchers seem to be more vulnerable to externally driven publishing strategies (also described by Hammarfelt & de Rijcke, 2014) probably because they are especially aware of the need to comply with evaluation criteria in order to secure access to resources and/or promotions.

The variety of publication channels used by SSH scientists stems from their need to reach different audiences: the international community, national scholars and non-scholarly public (Nederhof, 2006; Huang & Chang, 2008). SSH research has a more local focus than natural sciences which may fluctuate by discipline and by research subject, being theoretical research more likely to be published via international channels than works on applied topics of more local interest (Hicks, 2004). Although research with an international orientation is currently being strongly promoted under the assumption that it may reduce fragmentation in the SSH fields (Hicks & Wang, 2011) and thus shore up the trend to publish through the most significant channels (Sivertsen, 2010), the value of research in topics of local concern should also be recognised and efforts should be made to strike a balance between these different types of research. In the long run, if local research is not given a fair assessment in the fields where it prevails over the internationally-oriented works, the research agendas of SSH researchers in these fields may just get geared to topics better suited to the taste of international audiences (López-Piñero & Hicks, 2015). Our results suggest that, at the CSIC, the highest academic rank can be attained both by nationally and internationally-oriented researchers, although the observed prevalence of the international orientation among younger scientists can be a reflection of the international bias enshrined in current policy trends. However, the conduct of a comprehensive analysis by discipline,
for which a far larger sample would be required, may be convenient to shed light on the specific orientation of each community and provide a basis for the design of research assessment policies.

The higher weighted productivity of promoted scientists in social sciences suggests that the weighted score system underlying this calculation is consistent with the criteria applied by evaluation panels who decide promotion at the CSIC in this domain. The same cannot be said in the case of humanities, where no differences in weighted productivity by rank were observed. The underlying reasons for this might be the lack of consensus on the definition of research quality in humanities as well as on what indicators may be most appropriate for its assessment, although such indicators may not always be shared by the disciplines included under the humanities umbrella (Hug, Ochsner, & Daniel, 2013). In addition, the limitations of the indicators applied here to identify “top” channels could be stronger in the humanities domain since book-related indicators have been introduced only recently and are not well established yet (i.e. SPI). The selection of top book publishers used in this study includes those that were best ranked in a survey to faculty members and researchers in Spain, according to whom, having evaluations through external reviewers was a decisive factor in the quality of publishers (Giménez-Toledo et al., 2013). However, the perception of researchers could be influenced by several factors, such as the size of the publisher’s portfolio, its characteristics, and its market positioning. In addition, a leading publisher in any given domain may be irrelevant in other discipline, and a single ranking was used in this study. To overcome these problems, it would be advisable to include more objective indicators on the quality of publishers, even descending at the series level, since there may be differences in quality across series of a given publisher. The creation of a “quality label” for individual books in order to guarantee a peer-review standard as set up in Flanders in 2010 (Verleysen & Engels, 2013) stands out as an interesting idea. The introduction of a similar label in Spain for university publications is currently under discussion. In these models, publishers are required to demonstrate the evaluation of books following standardized procedures and thus the establishment of such label provides a useful indicator for research assessment in the SSH fields while fostering publishers to improve (and/or make transparent) their editorial processes.

Limitations

This study is subject to several limitations that must be set forth. Firstly, although the study focuses on four different types of academic output, it does not consider non-scholarly literature which may play an important role in SSH (societal impact) due to the national/regional scope of research in this domain. Secondly, only differences between social sciences and humanities have been analyzed, although differences by discipline could also exist. In fact, in a recent study by Verleysen & Weeren (2016) a clear divide between social sciences and humanities concerning specific publication patterns was observed, but also considerable diversity between disciplines and among authors within disciplines was revealed. Both aspects deserve further scrutiny in the future.

Main conclusions
Taking into consideration an assorted set of publication channels, including but not limited to WoS articles, is essential to obtain a comprehensive and reliable picture of SSH scientists’ activity. Although the use of institutional databases, such as CSIC’s ‘Conciencia’ system, may be an interesting option in an attempt to cover the full production of scientists, a high level of rigor as regards the input of data and the validation process is required to obtain accurate and reliable results. Quality indicators concerning SSH publication channels are less developed than in the natural and biomedical fields. In this study, a set of indicators contemplating books and non-WoS journals is presented, but it is not exempt from certain limitations. The redoubling of efforts to further develop and keep updated indicators and classifications of journals and book publishers should be a major concern. In spite of all the remarkable efforts made by Spanish bibliometric groups to develop such indicators, a science policy at the national level emerges as the only way forward in order to establish a national framework ensuring the comprehensive coverage of SSH over time, regardless of whether it is based on the current tools or on new indicators. In the wake of the creation and success of CRIS in the Nordic countries, national efforts in the same direction should be encouraged since it would make the development of national studies and comparisons with other countries on a large scale possible.

References


This is a postprint version of:


Martin, B. et al. (2010) Towards a Bibliometric Database for the Social Sciences and Humanities - A European Scoping Project. Research report produced for DFG, ESRC, AHRC, NWO, ANR and ESF.


This is a postprint version of:


i Scientists who have issued at least two publications in the period were considered as active.

ii A total of 313 subject categories according to the Scopus Classification are considered.

iii The weighted productivity system described in the Norwegian model (NM) is used for reference, but it is not strictly applied. Three document types are considered under the NM (articles, books and book chapters) while, under our system, articles are disaggregated into WoS articles and non-WoS articles (enabling us to increase the range of possible ratings). Besides, the weight of books is reduced in our study (3 vs 5 in NM for non-top, 5 vs 8 in NM for top books) because the material included in the book category has not undergone such an exhaustive revision as that conducted under the NM. The NM involves the implementation of a quality assurance system to control that only scholarly publications with peer review are considered, while this filter has not been applied in the Spanish case.

iv Editing a book has been sometimes undervalued for academic careers, but edited books may have a significant value to their fields, as they offer “multiple perspectives, a broad scope and a wide variety of methodologies” (Ossenblok, Guns, & Thelwall, 2015). In fact, a higher citation rate has been described for edited books as compared to monographs (Torres-Salinas et al. 2013). Both edited and authored books need to be considered -although separately- in the evaluation of research performance.