Scientometrics manuscript No.
(will be inserted by the editor)

Alphabetic order of authors in scholarly publications: A bibliometric study for 27 scientific fields

João M. Fernandes · Paulo Cortez

Received: date / Accepted: date

Abstract Paper authorship and author placement have significant consequences for accountability and assignment of credit. Moreover, authors in different scientific fields tend to follow distinct approaches towards their ordering in scholarly publications. This manuscript presents a bibliometric study aiming to characterize the trends in the adoption of alphabetically ordered lists of authors in scholarly publications for 27 scientific fields. The study is supported by two different datasets (with 83 and 32 thousand papers that have two or more authors) and uses two indicators that measure the degree of order of the authors list of a set of articles. The main results show that three fields (Economics; Mathematics; and Business, Management & Accounting) have a strong alphabetic ordering usage, while other five scientific areas present some tendency to use lists of authors in alphabetic order.

Keywords bibliometrics · scientific authorship · authors order · scholarly publication

1 Introduction

Whenever there are two or more authors, the authorship order becomes a relevant aspect of scholarly publications. This is becoming an increasingly pertinent issue, since diverse studies have shown a continuously increasing trend in the average number of authors per publication (Broad 1981; Grant 1989; Onwude et al. 1993; Persson et al. 2004; Greene 2007; Wuchty et al. 2007;

J.M. Fernandes (ORCiD: 0000-0003-1174-1966)
ALGORITMI Center, Department of Informatics, University of Minho, 4710-057 Braga, Portugal. E-mail: jmf@di.uminho.pt.

P. Cortez (ORCiD: 0000-0002-7991-2090)
ALGORITMI Centre, Department of Information Systems, University of Minho, 4804-533 Guimarães, Portugal. E-mail: pcortez@dsi.uminho.pt.

Fernandes 2014; Henriksen 2016; Fernandes and Monteiro 2017). In some scientific fields, such as Medicine, authors seem to follow a relatively clear and known set of authorship rules that stipulate how to position their names in publications (Baerlocher et al. 2007). Moreover, there have been some suggestions on how to solve authorship issues, such as recommended by Strange (2008). Despite of this, many authors still follow their own rules (thus ad hoc), although there are implicit rules that are often followed in practice and that are discussed in the next paragraphs.

Typically, the first author is considered the main author, the one that contributed the most to the intellectual effort of the paper. As argued by Peidu (2019), the first author is clearly the one with the highest contribution or responsibility. When there are two or more co-authors that have contributed equally, it is becoming more common to indicate several “equal first authors” (Hu 2009). This can be applied, for instance, when several research teams collaborate. In such cases, the leaders of each team can assume the role of corresponding authors.

Another implicit rule is to set the order of the authors based on the descending contributions to the contents of the paper. This approach sounds fair, but it implies that it is possible to measure the individual contributions, which often is not easy (e.g., long research project with a large team). Whenever this measurement is not possible or easy, the simplest solution is to use an alphabetical order by taking into consideration the surnames of the co-authors. This alternative may sound unfair, as co-authors may feel that their position does not reflect their relative contribution. Additionally, only the first author of papers with more than three co-authors appears in the bibliographic references when these are abbreviated as (first) “author et al.” There can also be hybrid solutions that use a mixture of degree of contribution and alphabetical based orders. For example, choose the first and last authors and order the rest alphabetically.

It should be noted that the choice of the position of the authors is not irrelevant and can have a great impact in the researcher career and her institution. In fact, there are some proposals to assign relative values to the co-authors of a scientific publication according to the relative positions of each one (Trueba and Guerrero 2004; Hagen 2014; Vavryčuk 2018; Bornmann and Osório 2019). For instance, the harmonic authorship credit method uses the following formula to distribute the one-unit point among the N co-authors of a scientific paper:

$$Credit\,(i) = \frac{1/i}{\sum_{j=1}^{N} 1/j}$$

where i and j denote the positions of the authors (i, j $\in \{1, ..., N\}$) in the publication. This formula progressively assigns a higher value to the first co-authors and lower ones to the last ones. For example, when $N = 4$ the first author ($i = 1$) gets a credit of 0.48 points, the second author ($i = 2$) gets half of this score (0.24), the third author is credited with a 0.16 score, and the last author receives just 0.12 points. Clearly, these relative scoring formulas only make sense when researchers from a scientific community or field tend to adopt the descending contribution order.

In this manuscript, we present a bibliometric study that targets a total of 27 scientific fields, aiming to characterize what is the prevalence of an alphabetic ordering of co-authors in scholarly articles. This manuscript uses two different datasets and adopts two ordering indicators that measure the degree of order of the authors list of a set of articles: percentage of fully ordered articles (PFOA) and degree of alphabetic ordering (DAO).

The paper is organized as follows. Section 2 describes the state of the art. Then, the adopted research methodology is presented (Section 3). Next, the obtained results are presented and analyzed. Finally, limitations are discussed in Section 5 and conclusions and future work are presented in Section 6.

2 Related Work

The decision on the order of the authors of a scholarly publication can follow several approaches, as indicated by Peidu (2019):

1. by amount of contribution;
2. alphabetical order;
3. multiple first author or multiple last author;
4. by seniority or reverse seniority;
5. by raffling or lottery system; and
6. by negotiation or mutual understanding.

The choice of the approach to follow depends on some facts. One decisive factor to fix the positions of the authors of a scholarly publication is the research field, because the authors order has different importance in different research fields. In some scientific areas, the first author in a multi-authored paper is considered to be the most important contributor. Thus, in those areas, authors are not typically listed according to an alphabetical order. Other scientific disciplines consider that the order of the authors is not important, since it is assumed that all have contributed equally or similarly. In such cases, authors are more commonly listed in an alphabetical order. Despite this reality, studies on the authors order across diverse disciplines are not abundant. We next describe the main results found in studies that address issues related to the authors order in different fields.

Peffers and Hui (2003) compared, in the field of information management systems, the percentages of papers with alphabetically ordered author lists in journals with high impact factors with the corresponding ones in journals with median or low impact factors. Their conclusion was that in median or low impact factor journals the alphabetical order of authorship tends to disappear.

Maciejovsky et al. (2009) analysed 38,000 journal articles from the fields of economics, psychology, and marketing, and concluded that the three fields have different author ordering practices.

Frandsen and Nicolaisen (2010) presented a study related to the credit assignment practices in the fields of economics, high-energy physics, and information science. They have shown that the practices of alphabetization of authorship are different among the three fields. A slight increase was found in the economics field during a 30-year period (1978–2007). In information science, a significant decrease was found to have occurred during the same 30-year period. High-energy physics, during the period 1990–2007, has witnessed a high and stable percentage of alphabetically ordered authors lines.

Waltman (2012) observed that in 2011, the authors of less than 4% of all publications intentionally chose to list their names alphabetically. Mathematics, economics (including finance), and high-energy physics were the fields where the use of alphabetical order in the authors list was more prevalent (Marušić et al. 2011; Waltman 2012). Also, it was found that publications with a large number of authors, often known as kilo-papers, tend to adopt an alphabetical order.

Sauermann and Haeussler (2017) pointed out the probability of error when deducing contributions based on the position of the author. Their paper discussed the data related to articles published in the period 2007-2011 in PLOS ONE, a journal primarily focused in the biological and life sciences. This periodical requires all its articles to disclose the types of contributions made by each co-author, using predefined categories. Sauermann and Haeussler have conducted two studies, being the first one related to the author order and the respective contribution statements. They concluded that in some cases the author order was not always aligned with the respective contribution statements. In particular, the author order was considered a less reliable indicator of the authors' contributions when there was a high number of co-authors.

Weber (2018) argued that alphabetical order gives an unfair advantage to researchers whose last name initials are at the beginning of the alphabet. Weber provided evidence that there was an alphabetical discrimination and that researchers often react to it, for example, by avoiding collaborations with other authors.

The two major differences between our research and the ones previously mentioned are next described. Firstly, this manuscript is the only one that uses an indicator that measures the degree of ordering of a list of authors. This is important since, contrarily to the binary indicator (ordered or not ordered), it provides a better measurement of the level or ordering that was adopted by the authors. Secondly, we cover all the subject areas considered in Scimago, thus this study embraces all major scientific fields, providing a more global overview. Only Waltman (2012) has performed a similar global analysis, by considering 25 fields for an older time period (2007 to 2011).

3 Methodology

This study aims to perform a comprehensive coverage of scientific areas, as reflected in terms of journal articles. Thus, we selected all the 27 subject areas that are listed in SCImago website (`www.scimagojr.com`), as consulted in May 2020 and shown in Table 1. SCImago is a publicly portal, backed by the Scopus scientific database and that is often used to rank the quality of journals (Falagas et al. 2008). It should be noted that in certain fields, such as Computer Science and Engineering, publications in conference proceedings are as prestigious as in journals (Glänzel et al. 2006; Lisée et al. 2008; Vardi 2009; Vrettas and Sanderson 2015). However, in order to adopt an uniform criterion for all scientific areas, this study only considers journal articles.

3.1 Research goal

The research approach we have used in our study is the Goal, Question, Metric (GQM) methodology (Basili 1992). Following the GQM goal template, the goal of this study is defined as to systematically identify issues and trends related to multi-authored papers, namely which scientific fields adopt the alphabetical order to list the authors. To tackle this goal, the following research question (RQ) is taken into account:

RQ: How is the use of the alphabetical order of authors characterized for all scientific areas?

3.2 Data related with scientific publications

To answer the RQ, we consider two different datasets, each related with different queries used to fetch the authors ordering data of journal articles. A semi-automated retrieval method was adopted to fetch the paper metadata, which involved a manual selection of the target journals per scientific area, executed via the known International Standard Serial Number (ISSN). Then, the metadata of the associated articles was collected using the Scopus engine, as downloaded in May 2020. Dataset 1 (DS1) is related with all papers that were published in a prestigious journal of a subject area, assuming the SCImago Journal Rank (SJR) index for 2018. Dataset 2 (DS2) contains the metadata of a minimum of 1000 articles published in one or more top journals of a given subject area. To further differentiate the datasets, the DS2 sample includes recent articles, published in the years of 2018, 2019 and 2020.

Table 2 shows the journals that were considered for DS1. In the majority of the cases, DS1 includes the first ranked journal, according to the SCImago Journal Ranking indicator. There are a few exceptions to this rule. The first notable exception is subject area #20 (Multidisciplinary), for which two prestigious journals are considered: Science and Nature. The rule was also not considered when the top journal for a given subject area is listed in several (let us say three or four) subject areas. In these cases, we consider that the journal has a multidisciplinary coverage and thus it is excluded from DS1, since the aim is to select journals that are representative of a subject area. Thus, in these cases, the journal was replaced by the highest ranked journal that is highly related with the scientific area under consideration. This situation occurred, for example, for #2 (Arts & Humanities), due to its wide scope. The selected journal is Nous, which appears in the 14th position. It is the first journal in the list to consider only one subject area. It also happened for #4 (Business, Management & Accounting), since the first journals are also related to subject area #11 (Economics, Econometrics & Finance). Therefore, we considered the journal Academy of Management Annals, the 4th in the list but the first that is only related to #4. Table 2 presents several known journals, such as Nature (established in 1869), Science (1880), and Quarterly Journal of Economics (1886). For each journal, the table also indicates the initial year considered in this study and the total number of articles for DS1. All the metadata related to the papers published in that year or afterwards were downloaded from the Scopus engine.

Tables 3 and 4 show the journals that were considered for DS2. All the selected journals are listed in SCImago in just one subject area. The adopted process for journal inclusion was iterative. We went through the list of SCImago journals for a given subject area, ranked according to the SJR criterion, and searched for a journal that fits exclusively in that area. We then searched in Scopus for all papers published in the selected journal ISSN within the 2018–2020 period. If the returned number of papers was smaller than 1,000, then we selected the next highest ranked journal for the same subject area, until at least 1,000 papers were reached. For instance, the final Scopus search query for the subject area #7, which covers three journals, was:

```
PUBYEAR AFT 2017 AND (ISSN(1935-8237) OR ISSN(2352-7110) OR ISSN(2162-237X))
```

This query searches the articles published after 2017 in the three journals with the indicated ISSNs. Whenever the number of articles by Scopus was higher than 2,000, only the metadata of the first 2,000 were considered. It should be noted that for the journals that have two ISSNs, as indicated in SCImago, the query includes both ISSNs, just to make sure that all articles of that journal were considered.

The entries in the files retrieved from Scopus and related with the papers metadata were “cleaned” using a computer program written in the Python language. Firstly, only the list of authors for each paper was considered and thus the other fields were discarded. Then, data errors, inconsistencies, lack of data, wrong spellings, etc. were eliminated/corrected. Only the 26 letters of the Latin/Roman alphabet (A to Z) were considered. Diacritics were removed, thus many non-Latin letters, such as â, ã, á, à, ä, ă, æ, ç, č, ğ, ł, ń, ø, ş, ţ, ż, were replaced by the most similar Latin letter (e.g., ‘á’ by ‘a’). In total, datasets DS1 and DS2 include respectively around 83 and 32 thousand papers that have two or more authors and that are analyzed in Section 4.

3.3 Ordering indicators

In this study, we adopt two main alphabetic author order indicators:

1. the measurement of fully ordered author papers; and
2. a degree of alphabetic ordering, measured using the Inversion criterion.

These indicators are detailed in the next subsections.

3.3.1 Percentage of fully ordered articles (PFOA)

Regarding the first indicator, we have implemented a Python program to analyze, for each paper, how many authors it contains and if the list of authors is either ordered or not ordered.

It is important to notice that there are lines of authors that are accidentally in alphabetical order, i.e., the authors are ordered, but that was not the criterion used to place them in the authors list. If, for example, 2-author papers are considered, around 50% of them would be ordered. With six co-authors, there are $6! = 720$ different combinations to place them and only one is alphabetically ordered. Thus, the probability to find an alphabetically ordered 6-author line (that was not specifically arranged in alphabetical order) is smaller than 0.0014%. The higher the number of authors, the smaller is this value. To cope with this issue, we define a baseline for the first indicator, which is defined as 1/N!, with N representing the number of authors. In this work, the baseline is compared with the Percentage of Fully Ordered Articles (PFOA), defined as:

$$\mathrm{PFOA}(D) = \frac{\mathrm{FO}(D)}{\#D} \tag{1}$$

D is the dataset with a total of $\#D$ author lists and $\mathrm{FO}(D)$ denotes the number of fully ordered lists in D. The dataset D is defined according to an analysis criterion. For example, it can include all papers from DS1 that have only N = 4 authors. The D members are author sequence lists $l =< a_1, a_2, ..., a_N >$, where a_i denotes the i-th author of the paper. A list is fully ordered if the alphabetic condition $a_i < a_j$ is true for all i and j where $i < j$.

3.3.2 Degree of alphabetic ordering (DAO)

In the field of Computer Science, the efficiency of sorting algorithms has been well studied. Thus, there are several methods to measure the sorting degree (more precisely, its inverse, i.e., the degree of disorder), such as the eleven metrics proposed by Estivill-Castro and Wood (1992). The most common metric is the number of inversions that exist in a list or sequence. Let $\mathrm{Inv}(l)$ denote the number of inversions in list $l =< a_1, a_2, ..., a_N >$, where (i, j) is an inversion if $i < j$ and $a_i > a_j$. The maximum number of inversions in a list with N elements is thus $\frac{N \times (N-1)}{2}$, which occurs for the inversely ordered list. For example, the lists <A, B, C, Z, D> and <Z, A, B, C, D> have one and four inversions, respectively. The pair (Z,D) is the only inversion in the first list. There are four inversions in the second list, the pairs (Z,A), (Z,B), (Z,C), and

(Z,D). In this paper, we adapt the inversion metric to measure the Degree of Alphabetic Ordering (DAO) of a list l, within the $[0, 1]$ range, by using:

$$\mathrm{DAO}(l) = 1 - \frac{\mathrm{Inv}(l) \times 2}{\mathrm{N} \times (\mathrm{N}-1)} \tag{2}$$

A DAO value of zero indicates a fully inverted ordered list. The higher the value, the more ordered is the list. When $\mathrm{DAO}(l) = 100\%$, it corresponds to the fully ordered list. As an example, Table 5 shows several $\mathrm{DAO}(l)$ values for 5-author papers. In this paper, we perform the distinct ordering analyses, by considering a dataset with several lists ($D = \{l_1, l_2, ..., l_{\#D}\}$). In such cases, we measure the overall $\mathrm{DAO}(D)$ value as the average of all $\mathrm{DAO}(l)$ values.

The DAO measure has some interesting properties. Firstly, it provides a numeric score that is more informative than the binary fully ordered measurement (as shown in Table 5). For example, it can provide a high order value (e.g., 90%) for the papers that are almost ordered or that are actually ordered but that is not correctly detected by our Python program (e.g., usage of the Danish alphabetic ordering). Secondly, for any fixed number of paper authors (N), a random list of authors (r) tends to produce a $\mathrm{DAO}(r) = 50\%$, which is the baseline value considered for the DAO indicator. In effect, the 50% limit was confirmed experimentally by generating 10 000 random author sequences, each sequence with a number of authors that ranged from $N = 2$ to $N = 30$. The obtained average DAO values were very close to the 50% limit (49.99667%).

In this manuscript, we calculate for each dataset and for all papers with N-authors (dataset D_N), the mean of the DAO values:

$$\mathrm{DAO}(N) = \frac{\sum_{l \in D_N} \mathrm{DAO}(l)}{\#D_N} \tag{3}$$

4 Results

Using our Python program, for each dataset (DS1 and DS1) we computed the PFOA and DAO ordering indicators. Since we want to check if the adoption of alphabetic ordering is more likely to occur for papers with more authors, the indicator overall percentages were computed for different number of authors, namely from $\mathrm{N} = 2$ to $\mathrm{N} = 9$, also including a special value of $\mathrm{N} > 9$, which denotes all papers with at least 10 authors.

Tables 6 and 7 present the fully ordered (PFOA) results for DS1 and DS2. Similarly, Tables 8 and 9 show the degree of alphabetic ordering (DAO) values for the two datasets. The first row of each table presents the baseline values. All PFOA and DAO overall values that are 10 percentage points higher than the baseline are highlighted with a **boldface** font. For each subject area, we also track the maximum number of authors (column max.) and the average percentage of intentionally alphabetical publications. This last indicator was proposed by Waltman (2012) and it can be used to estimate the probability that the authors of a paper intentionally listed their names alphabetically. In the tables, we use the criterion adopted in Waltman (2012) and highlight the values higher than 15%.

We first analyze the maximum number of authors, which clearly shows differences among the subject areas. Some scientific areas have a smaller maximum number of authors, such as: Mathematics (5 for DS1, 6 for DS2); and Business, Management & Accounting (5 for DS1 and 8 for DS2). Other scientific subjects have a much higher number of authors, including Arts & Humanities (2932 for DS1), Multidisciplinary (2422 for DS1) and Medicine (506 for DS2). As a demonstrative example, Figure 6 shows two histograms of the numbers of paper authors ($N \in \{2, \dots, max\}$) for the Economics (#11, left plot) and Environmental Science (#14, right graph) areas. The figure reveals two distinct patterns for the typical number of authors that appear in each area. Economics papers tend to have just two authors, while most Environmental Science articles have from four to eight authors.

Regarding the alphabetic ordering, there is an overall consistency in the obtained results for both datasets and for both indicators (PFOA and DAO). For instance, highlighted results (when compared with baseline values) tend to appear in similar cases for all ordering results (Tables 6–9). Moreover, Table 10 presents several Pearson correlations that were computed when varying the number of paper authors from $N = 2$ to $N = 5$ (range that appears in all 27 scientific subjects). The correlations show a very positive alignment between the two ordering indicators (PFOA and DAO), with just one 0.76 correlation and several values above 0.90. Also, there is a positive relationship in the alphabetic ordering measurements obtained for both datasets, with the correlations ranging from 0.73 to 0.95. Turning to the comparison among the different scientific areas, there are a few areas that present a consistent alphabetic pattern. In particular, we can identify two main patterns:

- **strong alphabetic degree**:
  - DS1 and DS2 – Economics (#11) and Mathematics (#18);
  - DS1 – Business, Management & Accounting (#4);
- **moderate alphabetic degree**:
  - DS1 and DS2 – Social Sciences (#26)
  - DS1 – Physics & Astronomy (#24), Biochemistry, Genetics & Molecular Biology (#3, e.g., $N \in \{7, 8, 9, > 9\}$) and Earth & Planetary Sciences (#10, e.g., $N \in \{3, 4, 5\}$));
  - DS2 – Arts & Humanities (#2, e.g., $N \in \{2, 3, 4\}$).

The average percentage of intentionally alphabetical publications indicator tends to follow these two main patterns (e.g., for DS1, the value is 95% for Mathematics and 36% for Physics & Astronomy).

The DAO differences for DS1 and D2 can be visualized in Figure 6, which includes the eight subject areas previously listed as having an interesting alphabetic ordering degree pattern. Figure 6 allows to visually confirm that the strong alphabetic ordering areas (#11 and #18) maintain the same level of indicator values for DS1 (top graph) and DS2 (bottom plot). The obtained ordering results are aligned with the ones made by Waltman (2012), which highlights an alphabetic author degree usage in the fields of Economics, Mathematics and Physics.

Finally, it should be noted that in some cases, there is a slight alphabetic ordering increase for papers that have a larger number of authors. For example, the average DAO values rise from $N = 9$ to $N > 9$ for several subject areas. This occurs in six areas (#2, #3, #6, #9, #10 and #13) for DS1 (Table 8) and five areas (#1, #9, #10, #14 and #25) for DS2 (Table 9). After noticing this behavior, we have performed an additional analysis to further check with statistical confidence if the ordering degree increases when there is a high number of authors ($N > 9$). Using the Wilcoxon signed-rank non parametric statistical test (Hollander et al. 2013), we have computed the 95% confidence intervals for the median DAO values of two sets: I - containing all DAO scores related with papers with $N \in \{2, 3, ..., N_* - 1\}$; and II - with all DAO values from papers with $N \in \{N_*, ..., max\}$. Using computer code written in the R tool, we performed a cycle to find the lowest N_* value (if any) such that: $N_* > 9$, II contains at least 10 papers (#II≥10) and the estimated Wilcoxon median for II is higher than 50%. The obtained results are shown in the last column of Tables 8 and 9. Each cell contains the values: N_* (estimated median, cardinality of II). The N_* analysis confirms that for most of the highlighted $N > 9$ column DAO scientific areas, there is an increase in the ordering degree, namely: DS1 - #1, #2, #20, #23; and DS2 - #1, #9, #10, #14, #20. For example, for the Pharmacy (#23) subject and DS1, there are 32 papers with 17 or more authors with a Wilcoxon estimated DAO median of 76% that is statistically higher when compared with the 67% median related with the $N \in \{2, ..., 16\}$ papers.

5 Limitations

While interesting results were achieved, some study limitations are here discussed. Some of the subject areas considered in this study are too generic (e.g., Engineering, Medicine, Arts & Humanities, and Social Sciences). Each generic area includes subareas (e.g., Engineering can be subdivided into Electrical, Civil, Mechanical Engineering and others) that may have different publication patterns and practices. Moreover, the adopted Scopus database querying system is limited and thus we employed a semi-automated data collection that considered a relatively small number of journals for each dataset (in some cases, just one). While a large number of papers was retrieved (e.g., more than 1 000 for DS2), it can be case that the patterns of publication in those journals are not totally representative of those that are observable in the corresponding areas or subareas. This issue is more relevant to DS1, since this dataset only considers one journal for each subject area (with one exception). Therefore, the results for DS1 should be seen with some caution, as they may not adequately represent the typical patterns of publication of the respective field. Furthermore, we do not distinguish the type of analyzed research papers. These include review papers that might have a significant difference in the number of authors. However, we believe that the pattern of placement of authors follows the same practices. For instance, we did not found substantial authorship order differences two journals that were selected for area #3 ("Nature Reviews Genetics" from DS1 and "Cell" from DS2). In addition, we note that the review papers were also included in (Waltman 2012). Finally, survey papers tend to produce a high number of citations, thus it makes sense that several review journals are highly ranked and thus appear in DS1.

Another aspect is related with the number of papers considered for the different subject areas, which differs quite substantially in some subject areas. For DS1, the number of papers for Multidisciplinary is almost 71 000 and for Health Professions is it just 138. These differences, namely the small number of papers for some areas, can be problematic if the sample of papers is not representative. We notice however that this issue was mitigated for DS2, as the process used to collect the metadata assures that the number of papers fits in the range {1000, . . . , 2000}.

The detection of an alphabetically ordered author sequence may fail in a few cases. In fact, analyzing the degree of alphabetical order of a list of authors is not a completely objective process. One reason lies in the fact that small words with only lowercase letters (e.g., da, de, del, den, der, di, do, dos, du, la, le, te, ten, ter, thi, van, van, von), that usually precede surnames, are ignored in our study. If however the preceding small word has an initial capital letter, it is considered to be part of the surname. This may differ from the rules followed by the authors. The Python program considers "Almeida A., De Barbosa B., Carvalho C." to be disordered, but "Almeida A., de Barbosa B., Carvalho C." to be ordered. The surname of the second author is "De Barbosa" in the former case and "Barbosa" in the latter. Another reason is the use of non-Latin letters that can affect the analysis whether a list of authors is ordered. For example, the Danish/Norwegian alphabet includes three letters (æ, ø, å) that are considered to be the last ones (i.e.,they appear after z). However, when they are transformed into Latin letters (to ae, oe, aa) the ordering analysis is likely to change. For example, the list of authors "Bratbak G., Tsagaraki T.M., Øvreås L." is ordered according to the danish/norwegian alphabet, but when latinised ("Bratbak G., Tsagaraki T.M., Oevreaas L.") it becomes disordered. Nevertheless, the number of papers where such situations occur is residual and seems to not affect the overall results presented in this manuscript. Furthermore, it should be noted that the proposed degree order indicador (DAO) can handle most of these exceptions, as it can still provide a high ordering degree score (e.g., 90%) when a non perfect alphabetic match was detected.

6 Conclusions

This paper provides a systematic bibliometric study of the tendencies and patterns related to the alphabetic approaches taken to position the author names in scholarly publications. This study addresses 27 scientific fields and is supported by two different datasets composed of articles published in journals of those fields. The analysis uses two ordering indicators that measure the degree of order of the authors list of a set of articles. The main results show that three fields show a strong tendency to have their authors ordered alphabetically: Economics; Mathematics; and Business, Management & Accounting. The fields of Social Sciences; Physics & Astronomy; Biochemistry, Genetics & Molecular Biology; Earth & Planetary Sciences; and Arts & Humanities have a moderate tendency. According to our study, the other 19 fields do not generically follow the approach of alphabetically ordering the authors. Nevertheless, we have detected a slight alphabetic degree increase when the total number of papers is high (with 10 or more papers) in a few scientific domains (e.g., Agriculture & Biological Sciences; and Multidisciplinary).

In future work, we intend to perform a similar author alphabetic degree analysis that includes more journals and scientific subareas (e.g., electronic engineering). It would also be interesting to calculate the two ordering indicators (PFOA and DAO) on sublists of authors related with research institutions (to check alphabetic order usage of authors from the same institution or laboratory).

Author contributions

J. M. Fernandes performed the conceptualization, methodology, software, investigation, resources, formal analysis, writing – original draft, writing - review and editing. P. Cortez contributed with methodology, validation, formal analysis, writing – review and editing, and visualization.

Acknowledgements This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020. We would like to thank Rui Mendes (from U. Minho) for the initial version of the Python program that was used to automatically calculate the various metrics for our datasets. We also acknowledge Jorge Sousa Pinto and José Nuno Oliveira (from U. Minho) for discussions on how to measure the order degree of a list.

References

Baerlocher MO, Newton M, Gautam T, Tomlinson G, Detsky AS (2007) The meaning of author order in medical research. Journal of Investigative Medicine 55(4):174–180, DOI 10.2310/6650.2007.06044

Basili VR (1992) Software modeling and measurement: The goal/question/metric paradigm. Tech. rep., University of Maryland

Bornmann L, Osório A (2019) The value and credits of n-authors publications. J Informetrics 13(2):540–554, DOI 10.1016/j.joi.2019.03.001

Broad WJ (1981) The publishing game: Getting more for less. Science 211:1137–1139, DOI 10.1126/science.7008199

Estivill-Castro V, Wood D (1992) A survey of adaptive sorting algorithms. ACM Computing Surveys 24(4):441–476, DOI 10.1145/146370.146381

Falagas ME, Kouranos VD, Arencibia-Jorge R, Karageorgopoulos DE (2008) Comparison of scimago journal rank indicator with journal impact factor. The FASEB journal 22(8):2623–2628

Fernandes JM (2014) Authorship trends in software engineering. Scientometrics 101(1):257–271, DOI 10.1007/s11192-014-1331-6

Fernandes JM, Monteiro MP (2017) Evolution in the number of authors of computer science publications. Scientometrics 110(2):529–539, DOI 10.1007/s11192-016-2214-9

Frandsen TF, Nicolaisen J (2010) What is in a name? credit assignment practices in different disciplines. J Informetrics 4(4):608–617, DOI 10.1016/j.joi.2010.06.010

Glänzel W, Schlemmer B, Schubert A, Thijs B (2006) Proceedings literature as additional data source for bibliometric analysis. Scientometrics 68(3):457–473, DOI 10.1007/s11192-006-0124-y

Grant I (1989) Multiple authorship. BMJ: British Medical Journal 298(6670):386

Greene M (2007) The demise of the lone author. Nature 450(7173):1165–1165

Hagen NT (2014) Counting and comparing publication output with and without equalizing and inflationary bias. J Informetrics 8(2):310–317, DOI 10.1016/j.joi.2014.01.003

Henriksen D (2016) The rise in co-authorship in the social sciences (1980-2013). Scientometrics 107(2):455–476, DOI 10.1007/s11192-016-1849-x

Hollander M, Wolfe DA, Chicken E (2013) Nonparametric statistical methods. John Wiley & Sons

Hu X (2009) Loads of special authorship functions: Linear growth in the percentage of "equal first authors" and corresponding authors. J Assoc Inf Sci Technol 60(11):2378–2381, DOI 10.1002/asi.21164

Lisée C, Larivière V, Archambault É (2008) Conference proceedings as a source of scientific information: A bibliometric analysis. J Assoc Inf Sci Technol 59(11):1776–1784, DOI 10.1002/asi.20888

Maciejovsky B, Budescu DV, Ariely D (2009) The researcher as a consumer of scientific publications: How do name-ordering conventions affect inferences about contribution credits? Marketing Science 28(3):589–598, DOI 10.1287/mksc.1080.0406

Marušić A, Bošnjak L, Jerončić A (2011) A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. Plos one 6(9):e23477, DOI 10.1371/journal.pone.0023477

Onwude JL, Staines A, Lilford RJ (1993) Multiple author trend worst in medicine. British Medical Journal 306(6888):1345, DOI 10.1136/bmj.306.6888

Peffers K, Hui W (2003) Collaboration and author order: Changing patterns in is research. Communications of the Association for Information Systems 11(1):10, DOI 10.17705/1CAIS.01110

Peidu C (2019) Can authors' position in the ascription be a measure of dominance? Scientometrics 121(3):1527–1547, DOI 10.1007/s11192-019-03254-1

Persson O, Glänzel W, Danell R (2004) Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. Scientometrics 60(3):421–432, DOI 10.1023/B:SCIE.0000034384.35498.7d

Sauermann H, Haeussler C (2017) Authorship and contribution disclosures. Science Advances 3(11):e1700404, DOI 10.1126/sciadv.1700404

Strange K (2008) Authorship: Why not just toss a coin? American Journal of Physiology-Cell Physiology 295(3):C567–C575, DOI 10.1152/ajpcell.00208.2008

Trueba FJ, Guerrero H (2004) A robust formula to credit authors for their publications. Scientometrics 60(2):181–204, DOI 10.1023/B:SCIE.0000027792.09362.3f

Vardi MY (2009) Conferences vs. journals in computing research. Commun ACM 52(5):5, DOI 10.1145/1506409.1506410

Vavryčuk V (2018) Fair ranking of researchers and research teams. PLoS ONE 13(4):e0195509, DOI 10.1371/journal.pone.0195509

Vrettas G, Sanderson M (2015) Conferences versus journals in computer science. J Assoc Inf Sci Technol 66(12):2674–2684, DOI 10.1002/asi.23349

Waltman L (2012) An empirical analysis of the use of alphabetical authorship in scientific publishing. J Informetrics 6(4):700–711, DOI 10.1016/j.joi.2012.07.008

Weber M (2018) The effects of listing authors in alphabetical order: a review of the empirical evidence. Research Evaluation 27(3):238–245, DOI 10.1093/reseval/rvy008
Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in production of knowledge. Science 316(5827):1036–1039, DOI 10.1126/science.1136099

#	subject area	#	subject area
1	Agricultural & Biological Sciences	15	Health Professions
2	Arts & Humanities	16	Immunology & Microbiology
3	Biochemistry, Genetics & Molecular Biology	17	Materials Science
4	Business, Management & Accounting	18	Mathematics
5	Chemical Engineering	19	Medicine
6	Chemistry	20	Multidisciplinary
7	Computer Science	21	Neuroscience
8	Decision Sciences	22	Nursing
9	Dentistry	23	Pharmacology, Toxicology & Pharmaceutics
10	Earth & Planetary Sciences	24	Physics & Astronomy
11	Economics, Econometrics & Finance	25	Psychology
12	Energy	26	Social Sciences
13	Engineering	27	Veterinary
14	Environmental Science		

Table 1 The 27 subject areas addressed in this study

#	journal	ISSN	initial year	number articles
1	Genome Biology	1474-760X	2000	4 689
2	Nous	1468-0068	1989	1 088
3	Nature Reviews Genetics	1471-0056	2000	3 263
4	Academy of Management Annals	1941-6520	2010	182
5	Nature Reviews Chemistry	2397-3358	2017	255
6	Chemical Reviews	1520-6890	1924	5 054
7	SoftwareX	2352-7110	2015	345
8	Journal of Operations Management	0272-6963	1980	1 336
9	Periodontology 2000	0906-6713	1993	1 013
10	Annual Review of Astronomy and Astrophysics	1545-4282	1990	466
11	Quarterly Journal of Economics	0033-5533	1886	4 437
12	Nature Energy	2058-7546	2016	806
13	Advanced Materials	0935-9648	1989	19 191
14	Energy and Environmental Science	1754-5692	2008	3 783
15	Vital and Health Statistics [Series 2]	0083-2057	1965	138
16	Nature Reviews Immunology	1474-1733	2001	3 283
17	Nature Reviews Materials	2058-8437	2016	404
18	Journal of the American Mathematical Society	1088-6834	1988	966
19	CA Cancer Journal for Clinicians	1542-4863	1950	2 190
20a	Nature	1476-4687	1992	33 907
20b	Science	0036-8075	2004	37 946
21	Nature Reviews Neuroscience	1471-0048	2000	3 425
22	World Psychiatry	2051-5545	2011	656
23	Nature Reviews Drug Discovery	1474-1776	2002	3 264
24	Reviews of Modern Physics	0034-6861	1929	3 326
25	Annual Review of Psychology	0066-4308	1950	1 222
26	Administrative Science Quarterly	0001-8392	1975	605
27	Annual Review of Animal Biosciences	2165-8110	2013	163

Table 2 Selected journals for DS1.

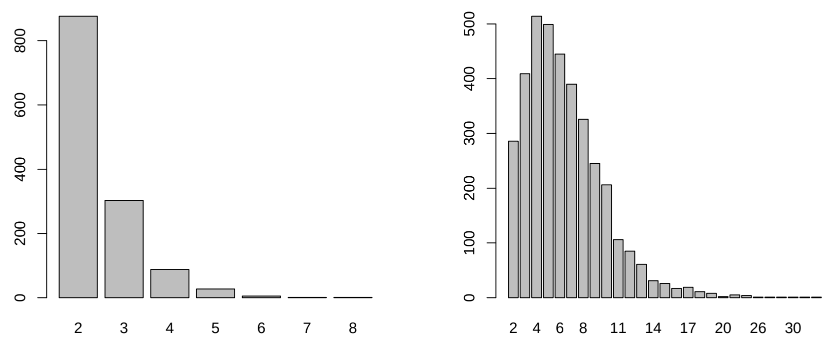

Fig. 1 Histograms of the number of authors (x-axis presents N; $y-$axis shows the number of papers) for two example scientific areas (Economics – left plot; Environmental Science – right plot).

#	journal	ISSN
1	Trends in Ecology and Evolution	0169-5347
	Ecology Letters	1461-023X
	Annual Review of Entomology	0066-4170
	Studies in Mycology	0166-0616
	Ecological Monographs	0012-9615
2	Nous	1468-0068
	The Philosophical Review	0031-8108
	Ethics	1539-297X
	Nous-Supplement: Philosophical Issues	1533-6077
	British Journal for the Philosophy of Science	1464-3537
	Philosophy and Phenomenological Research	0031-8205
	Philosophical Quarterly	0031-8094
	Mind	0026-4423
	Philosophical Studies	0031-8116
3	Nature Reviews Molecular Cell Biology	1471-0072
	Cell	0092-8674
4	Academy of Management Annals	1941-6520
	Academy of Management Journal	0001-4273
	Academy of Management Review	0363-7425
	Strategic Management Journal	1097-0266
	Organization Science	1526-5455
	Journal of Business Venturing	0883-9026
	Journal of Retailing	0022-4359
5	Catalysis Science and Technology	2044-4761
6	Chemical Reviews	1520-6890
	Chemical Society Reviews	0306-0012
7	Foundations and Trends in Machine Learning	1935-8237
	SoftwareX	2352-7110
	IEEE Transactions on Neural Networks and Learning Systems	2162-237X
8	OR Spektrum	0171-6468
	Annals of Operations Research	0254-5330
9	Periodontology 2000	0906-6713
	Journal of Clinical Periodontology	1600-051X
	Clinical Oral Implants Research	1600-0501
	International Endodontic Journal	1365-2591
10	Reviews of Geophysics	8755-1209
	Annual Review of Marine Science	1941-0611
	Nature Geoscience	1752-0908
	Earth System Science Data	1866-3516
	Bulletin of the American Meteorological Society	1520-0477
11	Quarterly Journal of Economics	0033-5533
	Journal of Political Economy	0022-3808
	Econometrica	0012-9682
	Review of Economic Studies	0034-6527
	American Economic Journal: Macroeconomics	1945-7707
	American Economic Review	0002-8282
	American Economic Journal: Applied Economics	1945-7790
	Journal of Economic Literature	0022-0515
12	Joule	2542-4351
	Renewable and Sustainable Energy Reviews	1364-0321
13	IEEE Communications Surveys and Tutorials	1553-877X
	Automatica	0005-1098
14	Global Change Biology	1365-2486

Table 3 Selected journals for DS2.

#	journal	ISSN
15	Journal of Physiotherapy	1836-9553
	Physical Therapy	0031-9023
	Physiotherapy Research International	1358-2267
	Musculoskeletal Science and Practice	2468-7812
	Journal of Chiropractic Medicine	1556-3715
16	Annual Review of Microbiology	0066-4227
	Annual Review of Virology	2327-0578
	mBio	2161-2129
17	Progress in Materials Science	0079-6425
	Annual Review of Materials Research	1531-7331
	Acta Materialia	1359-6454
18	Journal of the American Mathematical Society	1088-6834
	Inventiones Mathematicae	0020-9910
	Publications Mathématiques	0073-8301
	Duke Mathematical Journal	0012-7094
	Communications on Pure and Applied Mathematics	0010-3640
	Acta Numerica	0962-4929
	Annales Scientifiques de l'Ecole Normale Superieure	0012-9593
	Acta Mathematica	0001-5962
	Geometric and Functional Analysis	1420-8970
	Annales de l'Institut Henri Poincare	0294-1449
	Memoirs of the American Mathematical Society	0065-9266
19	CA - A Cancer Journal for Clinicians	1542-4863
	New England Journal of Medicine	0028-4793
20	Nature	1476-4687
21	Nature Reviews Neuroscience	1471-0048
	Nature Neuroscience	1097-6256
	Neuron	0896-6273
22	Clinical and Translational Immunology	2050-0068
	International Journal of Nursing Studies	0020-7489
	NursingPlus Open	2352-9008
	Journal of Nursing Scholarship	1547-5069
	Journal of Nursing Management	0966-0429
23	Annual Review of Pharmacology and Toxicology	1545-4304
	Trends in Pharmacological Sciences	0165-6147
	Advanced Drug Delivery Reviews	0169-409X
	International Journal of Pharmaceutical Investigation	2230-9713
	British Journal of Pharmacology	0007-1188
24	Reviews of Modern Physics	0034-6861
	Advances in Physics	1460-6976
	Nature Physics	1745-2473
	Physics Reports	0370-1573
25	Annual Review of Psychology	0066-4308
	Personality and Social Psychology Review	1088-8683
	Psychological Inquiry	1532-7965
	Psychological Science in the Public Interest & Supplement	1529-1006
	Journal of Applied Psychology	0021-9010
	Perspectives on Psychological Science	1745-6916
	Psychological Review	0033-295X
	Educational Psychologist	1532-6985
	Psychological Science	0956-7976
26	National Vital Statistics Reports	1551-8922
	American Journal of Political Science	0092-5853
	Quarterly Journal of Political Science	1554-0634
	American Political Science Review	1537-5943
	Political Analysis	1047-1987
	American Sociological Review	0003-1224
	Review of Educational Research	0034-6543
	Journal of Politics	1468-2508
27	Veterinary and Comparative Oncology	1476-5829
	Veterinary Research	0928-4249
	Journal of Veterinary Emergency and Critical Care	1479-3261
	Veterinary Pathology	1544-2217

Table 4 Selected journals for DS2 (cont.).

author list	FO	DAO
<A,B,C,D,E>	1	100%
<A,B,C,E,D>	0	90%
<B,A,C,E,D>	0	80%
<B,D,A,C,E>	0	70%
<B,E,A,C,D>	0	60%
<B,E,A,D,C>	0	50%
<B,E,D,A,C>	0	40%
<E,B,D,A,C>	0	30%
<E,D,B,A,C>	0	20%
<E,D,C,A,B>	0	10%
<E,D,C,B,A>	0	0%

Table 5 Examples of FO and DAO values for 5-author lists.

		N-authors										
#	**subject area**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**>9**	**max.**	**%int.**
-	baseline	50	17	4	1	0	0	0	0	0	-	
1	Agricult. & Biolog. Sc.	52	17	3	1	1	1	1	1	0	386	1
2	Arts & Humanities	**70**	**80**	0	-	-	-	-	0	0	46	**41**
3	Bioch., Genet. & Molec.	53	16	7	10	**11**	**14**	**20**	**20**	0	144	3
4	Business, Manag. & Acc.	**65**	**28**	8	0	**33**	0	0	-	-	8	**17**
5	Chemical Eng.	50	24	9	0	**14**	0	0	0	0	24	4
6	Chemistry	55	24	9	7	3	1	2	0	3	32	7
7	Computer Science	49	14	6	4	6	10	13	0	0	19	1
8	Decision Sciences	58	24	13	6	0	0	-	-	-	7	12
9	Dentistry	51	21	4	0	0	0	0	0	0	16	2
10	Earth & Planetary Sc.	55	**31**	**20**	**33**	-	-	-	-	-	5	13
11	Economics	**89**	**87**	**88**	**85**	**60**	0	0	-	-	8	**80**
12	Energy	55	25	**18**	6	6	0	0	0	0	59	5
13	Engineering	53	18	6	2	1	0	1	0	0	32	1
14	Environmental Science	51	19	3	1	0	0	0	0	0	45	0
15	Health Professions	56	9	**20**	0	0	0	0	0	0	18	4
16	Immun. & Microbiology	48	19	6	5	0	0	0	0	**10**	90	0
17	Materials Science	42	12	7	0	0	0	0	-	0	15	-4
18	Mathematics	**98**	**96**	**89**	**86**	-	-	-	-	-	5	**95**
19	Medicine	50	15	6	0	0	0	4	0	0	42	0
20a	Multidisciplinary	51	20	8	3	1	1	0	0	0	2 422	2
20b		51	22	7	3	2	1	1	1	1	2 932	2
21	Neuroscience	50	19	9	2	5	0	0	0	0	42	1
22	Nursing	46	23	4	0	0	0	0	0	0	101	-2
23	Pharmacy	59	19	6	6	2	3	0	0	3	65	8
24	Physics & Astronomy	**70**	**49**	**35**	**32**	**26**	6	**24**	**31**	**15**	48	**36**
25	Psychology	57	**27**	**17**	0	0	-	-	-	-	6	14
26	Social Sciences	**60**	20	**22**	**11**	0	-	-	-	-	6	15
27	Veterinary	53	19	8	0	0	0	0	-	-	8	4

Table 6 Percentage of fully ordered articles per number of authors (from 2 to >9) for DS1 (values 10 percentage points higher than the baseline are in **boldface**). Column max. indicates the highest number of authors in the subject area. Column %int. shows the percentage of intentionally alphabetical publications (values higher than 15% are highlighted in bold)

		N-authors										
#	subject area	2	3	4	5	6	7	8	9	>9	max.	%int.
-	Baseline	50	17	4	1	0	0	0	0	0	-	
1	Agricult. & Biolog. Sc.	51	20	5	2	1	0	0	0	0	73	2
2	Arts & Humanities	43	**67**	0	0	-	-	0	-	0	46	**49**
3	Bioch., Genet. & Molec.	47	19	9	0	0	0	0	0	0	72	-1
4	Business, Manag. & Acc.	**62**	**31**	12	2	**20**	0	0	-	-	8	**16.9**
5	Chemical Eng.	44	19	5	1	0	0	0	1	1	15	0
6	Chemistry	52	19	7	1	2	0	0	0	0	25	2
7	Computer Science	52	20	5	1	1	3	0	0	0	24	2
8	Decision Sciences	58	**39**	**24**	**14**	**11**	0	**50**	-	-	8	**21**
9	Dentistry	53	17	6	1	0	0	0	0	0	42	1
10	Earth & Planetary Sc.	49	18	6	1	2	0	0	0	0	501	0
11	Economics	**98**	**93**	**88**	**85**	**50**	**100**	-	-	-	7	**93**
12	Energy	55	20	6	5	0	0	3	0	1	32	3
13	Engineering	53	22	6	3	0	0	0	0	0	11	4
14	Environmental Science	51	17	3	1	0	0	0	2	1	104	0
15	Health Professions	51	20	3	1	0	0	0	0	0	21	1
16	Immun. & Microbiology	46	17	3	1	0	0	0	0	0	104	-1
17	Materials Science	52	12	5	1	0	0	0	0	0	20	0
18	Mathematics	**97**	**94**	**96**	**100**	**75**	-	-	-	-	6	**94**
19	Medicine	49	19	3	0	0	0	5	0	0	506	0
20	Multidisciplinary	54	19	10	2	2	0	4	3	0	559	2
21	Neuroscience	53	15	4	0	0	1	0	0	0	95	1
22	Nursing	53	16	5	1	0	0	0	0	0	25	1
23	Pharmacy	55	17	3	0	0	0	0	0	0	30	2
24	Physics & Astronomy	**67**	**48**	**24**	6	8	2	4	3	1	99	14
25	Psychology	52	15	5	2	2	0	0	0	0	124	1
26	Social Sciences	**81**	**55**	**42**	**21**	0	**20**	0	0	-	9	**53**
27	Veterinary	40	23	3	2	0	0	0	0	0	22	0

Table 7 Percentage of fully ordered articles per number of authors (from 2 to >9) for DS2 (values 10 percentage points higher than the baseline are in **boldface**). Column max. indicates the highest number of authors in the subject area. Column %int. shows the percentage of intentionally alphabetical publications (values higher than 15% are highlighted in bold)

		N-authors										
#	**subject area**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**>9**	**max.**	**N_***
-	Baseline	50	50	50	50	50	50	50	50	50	-	-
1	Agricult. & Biolog. Sc.	52	49	48	49	50	50	53	51	57	386	40 (62%,50)
2	Arts & Humanities	70	87	50	-	-	-	-	56	45	46	
3	Bioch., Genet. & Molec.	53	48	53	56	48	**63**	**70**	**62**	**68**	144	
4	Business, Manag. & Acc.	65	51	50	60	71	71	57	-	-	8	
5	Chemical Eng.	50	52	43	48	51	**67**	57	51	40	24	
6	Chemistry	55	54	53	53	50	50	55	56	**61**	32	
7	Computer Science	49	50	46	57	51	**60**	56	57	50	19	
8	Decision Sciences	58	55	54	49	**63**	**90**	-	-	-	7	
9	Dentistry	51	53	51	54	52	52	52	51	42	16	
10	Earth & Planetary Sc.	55	**63**	**67**	**80**	-	-	-	-	-	5	
11	Economics	**89**	**94**	**95**	**91**	**89**	**90**	32	-	-	8	
12	Energy	55	51	**61**	50	54	47	45	50	51	59	
13	Engineering	53	51	51	50	50	49	50	50	47	32	
14	Environmental Science	51	49	48	49	50	49	49	48	54	45	
15	Health Professions	56	33	56	46	50	33	**70**	**61**	40	18	
16	Immun. & Microbiology	48	53	50	56	**62**	46	57	**78**	58	90	
17	Materials Science	42	46	50	49	46	49	42	0	41	15	
18	Mathematics	**98**	**97**	**96**	**94**	-	-	-	-	-	5	
19	Medicine	50	46	52	51	53	50	55	52	**63**	42	
20a	Multidisciplinary	51	52	52	51	51	51	50	50	**65**	2 422	31 (60%,948)
20b		51	52	51	52	51	52	50	51	**71**	2 932	21 (62%,1309)
21	Neuroscience	50	49	48	53	52	**62**	54	**71**	**69**	42	
22	Nursing	46	54	58	56	51	56	57	48	57	101	
23	Pharmacy	59	53	54	53	55	**61**	55	**60**	**72**	65	17 (76%,32)
24	Physics & Astronomy	**70**	**70**	**68**	**73**	**71**	**65**	**68**	**81**	**74**	48	
25	Psychology	57	55	54	54	58	-	-	-	-	6	
26	Social Sciences	**60**	50	**65**	**72**	**70**	-	-	-	-	6	
27	Veterinary	53	53	57	45	55	**70**	**63**	-	-	8	

Table 8 Ordered degree mean percentage per number of authors (from 2 to >9) for DS1 (values 10 percentage points higher than the baseline are in **boldface**). Column max indicates the highest number of authors in the subject area. Column N_* shows that value and the estimated median and cardinality of II.

		N-authors										
#	**subject area**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**>9**	**max.**	**N$_*$**
-	Baseline	50	50	50	50	50	50	50	50	50	-	-
1	Agricult. & Biolog. Sc.	51	54	53	55	53	55	53	**62**	**72**	73	11 (69%,113)
2	Arts & Humanities	**76**	**67**	**94**	40	-	-	54	-	45	46	
3	Bioch., Genet. & Molec.	47	52	49	47	46	51	47	49	50	72	
4	Business, Manag. & Acc.	**62**	58	55	49	**64**	54	**62**	-	-	8	
5	Chemical Eng.	44	55	50	49	50	51	51	51	41	15	
6	Chemistry	52	52	50	49	51	48	50	50	51	25	
7	Computer Science	52	52	47	50	54	52	51	58	52	24	
8	Decision Sciences	58	**64**	**61**	57	56	53	**61**	-	-	8	
9	Dentistry	53	48	50	46	50	51	51	53	**68**	42	10 (59%,83)
10	Earth & Planetary Sc.	49	48	51	52	54	55	55	54	**73**	501	10 (63%,370)
11	Economics	**98**	**96**	**96**	**93**	**90**	**100**	-	-	-	7	
12	Energy	55	53	52	51	51	51	50	51	51	32	
13	Engineering	53	51	51	51	50	51	50	49	30	11	
14	Environmental Science	51	50	50	48	50	53	52	55	**69**	104	10 (62%,265)
15	Health Professions	51	56	50	50	52	53	50	52	55	21	
16	Immun. & Microbiology	46	49	50	49	51	50	48	47	48	38	
17	Materials Science	52	47	50	51	50	50	49	50	46	20	
18	Mathematics	**97**	**97**	**98**	**100**	**97**	-	-	-	-	6	
19	Medicine	49	50	50	53	55	54	51	52	53	506	
20	Multidisciplinary	54	52	53	57	54	51	54	51	58	559	58 (67%,16)
21	Neuroscience	53	51	50	47	49	52	50	48	54	95	
22	Nursing	53	49	49	50	49	49	54	53	51	25	
23	Pharmacy	55	51	50	49	51	53	50	50	49	30	
24	Physics & Astronomy	**67**	**70**	**64**	53	55	52	49	55	56	99	
25	Psychology	52	51	48	52	49	52	51	42	**60**	124	
26	Social Sciences	**81**	**72**	**73**	58	33	**76**	25	51	-	9	
27	Veterinary	40	49	47	47	49	52	53	52	50	22	

Table 9 Ordered degree mean percentage per number of authors (from 2 to >9) for DS2 (values 10 percentage points higher than the baseline are in **boldface**). Column max indicates the highest number of authors in the subject area. Column N_* shows that value and the estimated median and cardinality of II.

		N-authors			
Variable 1	**Variable 2**	**2**	**3**	**4**	**5**
PFOA (DS1)	PFOA (DS2)	0.82	0.91	0.95	0.93
DAO (DS1)	DAO (DS2)	0.89	0.81	0.73	0.73
PFOA (DS1)	DAO (DS1)	1.00	0.98	0.95	0.76
PFOA (DS2)	DAO (DS2)	0.90	0.97	0.80	0.97

Table 10 Pearson correlation values for combinations of the ordering indicators (PFOA, DAO) and for the two datasets.

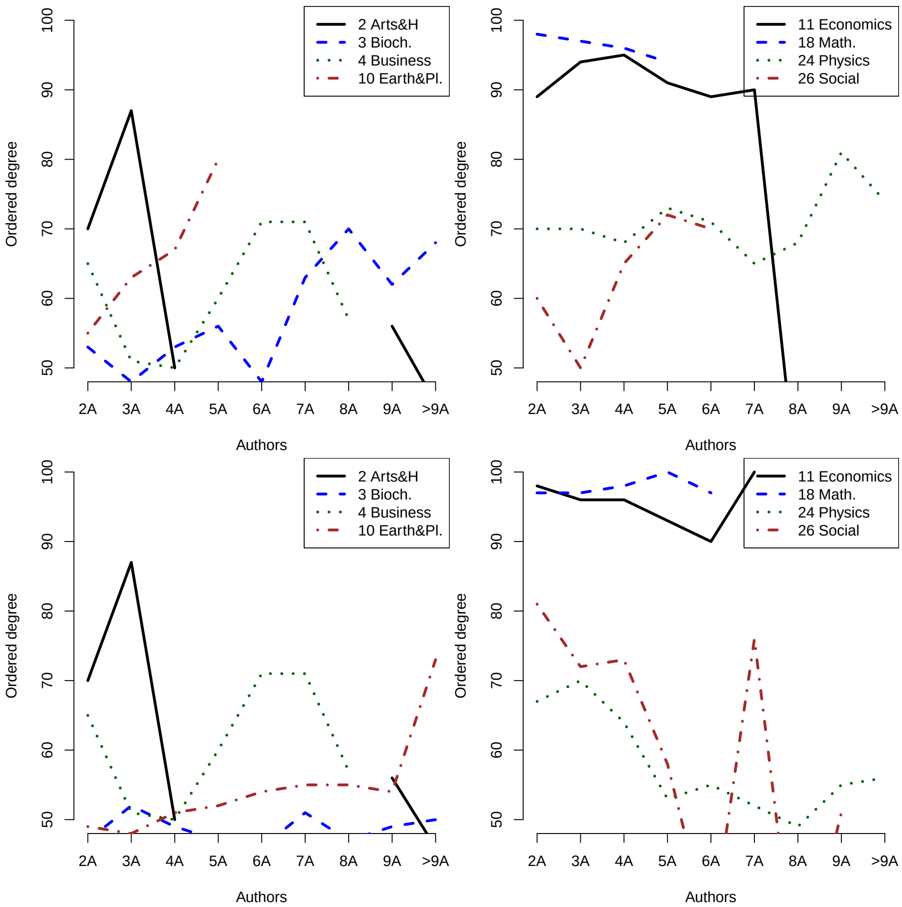

Fig. 2 Ordered degree by number of authors for selected scientific areas (DS1 – top graphs; DS2 – bottom graphs).