Creating a collection of publications categorized by their research guarantors into the *Scopus ASJC* scheme

Jesús M. Álvarez-Llorente; Vicente P. Guerrero-Bote; Félix De-Moya-Anegón

Nota: Este artículo se puede leer en español en:
https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87486

Recommended citation:

Álvarez-Llorente, Jesús M.; **Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2023). "Creating a collection of publications categorized by their research guarantors into the *Scopus ASJC* scheme". *Profesional de la información*, v. 32, n. 7, e320704.

https://doi.org/10.3145/epi.2023.dic.04

Manuscript received on 12th September 2023
Accepted on 10th October 2023

Jesús M. Álvarez-Llorente ✉
https://orcid.org/0000-0002-4901-3457

Universidad de Extremadura
Departamento de Ingeniería de Sistemas Informáticos y Telemáticos
Plazuela Ibn Marwan
06001 Badajoz, Spain
llorente@unex.es

Vicente P. Guerrero-Bote
https://orcid.org/0000-0003-4821-9768

Universidad de Extremadura
Departamento de Información y Comunicación
Plazuela Ibn Marwan
06001 Badajoz, Spain
guerrero@unex.es

Félix De-Moya-Anegón
https://orcid.org/0000-0002-0255-8628

SCImago Research Group
18220 Granada, Spain
felix.moya@scimago.es

Abstract

Given the need in Scientometrics to get beyond merely classifying scientific production based on the classification of the journals in which it is published, there have been many attempts to classify papers directly. Little has been done, however, to check how reliable the results are. In this work, a collection of publications was generated which we call an Author's Assignation Collection (AAC) comprising 13449 papers referenced in the *Scopus* database and classified by their research guarantor with fractional weighting in terms of *Scopus*'s own *ASJC* scheme. The methodological approach taken is described, and the collection's representativeness is evaluated and compared with the journal-based classification. There stand out both the great number of papers assigned by their research guarantors to more than one category (at times with even the same weight) and how frequently authors assigned categories which were not assigned to the journals in which their paper was published.

Keywords

Scientometry; Scientific classification; Scientific production; Paper-by-paper classification; Author's Assignation Collection; Authors; Research guarantors; Representation of papers; 2020; Test documentary collection; *Scopus*; *ASJC*.

Funding

Grant project PID2020-115798RB-I00 funded by *Ministerio de Ciencia e Innovación* of Spain (*Micin*), *Agencia Estatal de Investigación* (*AEI*) / 10.13039/501100011033.

1. Introduction

Scientometric studies require a reliable bibliographic database that covers the research papers published in the main scientific journals. The papers also need to be classified by discipline so that the progress of each discipline can be quantified. This classification is needed not only to quantify the research in each discipline but also to normalize the impact in that discipline, since publication and citation habits vary from one discipline to another (**Althouse** *et al.*, 2009; **Lancho-Barrantes**; **Guerrero-Bote**; **Moya-Anegón**, 2010; **Opthof**; **Leydesdorff**, 2010; **Bornmann**; **Leydesdorff**, 2017; **Bornmann**; **Tekles**; **Leydesdorff**, 2019).

What have until now been most used (**Gómez-Crisóstomo**, 2011; **Wang**; **Waltman**, 2016) are the classifications of the bibliographic databases themselves, *Scopus*'s *ASJC* (*All Science Journal Classification*) (*Elsevier*, 2023), and the *JCR* categories. Using scientific journals as scientometric units is quite common, and they have also been used to visualize the structure of science (**Leydesdorff**; **Moya-Anegón**; **Guerrero-Bote**, 2010; 2015; **Hassan-Montero**; **Guerrero-Bote**; **Moya-Anegón**, 2014). These journals' classifications are extended to the research papers they publish. The classification systems need to include some multidisciplinary categories, and many journals are assigned to various categories because they publish work corresponding to more than one. Nevertheless, not all the work a journal publishes are from all the categories to which it is assigned, indeed, quite the contrary is the case. All of this leads to great imprecision in both quantifying and normalizing the impact.

> "Not all the work a journal publishes are from all the categories to which it is assigned, indeed, quite the contrary is the case"

Numerous attempts have been made to improve these classification systems, and have generally aimed at classifying individual papers according to their own characteristics rather than those of the journal they belong to. Among these characteristics there stand out those based either on citation networks (direct citation, co-citation, bibliographic coupling, etc.) or text analysis (frequency of terms, etc.). **Šubelj**, **Van-Eck** and **Waltman** (2016) provide a discussion of these methods.

Some of these approaches are based on using automatic clustering systems to generate a new category classification scheme in which to distribute publications (**Klavans**; **Boyack**, 2005, 2006; **Waltman**; **Van-Eck**, 2012; **Janssens**; **Glänzel**; **De-Moor**, 2008; **Janssens** *et al.*, 2009). The results produced by these systems tend to change greatly as new literature is introduced into the classification, and have a randomness factor that can lead to disparate outcomes each time the procedure is restarted, even with the same sets of starting publications. Unfortunately, many bibliometric studies require classifications that are persistent and stable over time, even after the addition of new publications as they arise, so this type of classification is not usually widely accepted by the scientific community.

Other systems try to reorganize publications maintaining the category scheme of the journals, but also considering each publication's reference network to estimate the most precise category in which to assign it. This is done by **Glänzel**, **Schubert** & **Czerwon** (1999) and **Glänzel**, **Thijs** & **Huang** (2021), for example, to categorize articles from *WoS* multidisciplinary journals, and by **Milojević** (2020) to uniquely assign *WoS* publications. However, all of these works assume to be valid only certain assignations of papers to the category of their journal, and these are then used as trivial cases (starting points) with which to recursively solve the path of the citation network. This means that they are based not on the total number of references but on a smaller set. The classifications obtained through these methods have either not been evaluated or have been evaluated only in a very basic way.

In this paper, we describe our generation of a collection of documents from among those indexed by *Scopus*, and which their corresponding authors have classified using *Scopus*'s own *ASJC* scheme. The aim is for it to serve, with its limitations, as a possible further way to evaluate classification algorithms and the *Scopus* journal-based classification. Throughout this work and in future work, we shall use the term Author's Assignation Collection (AAC) to refer to this collection comprising the set of documents plus their classification.

For this, thousands of corresponding authors as research guarantors were surveyed (**De-Moya-Anegón** *et al.*, 2013) for them to determine the most appropriate category or categories in which to classify their works. We shall select a sufficiently large and representative sample from *Scopus*, and we shall have to answer some research questions about the responses obtained, such as:

- Has the response obtained been homogeneous by country, proportionately distributed by subject, and adequately representative of all journals?
- How many categories do the authors assign?
- How do they distribute the weight among the different categories?
- To what extent do these distributions coincide with the assignations that arise from the database's journal classification?

> "The classifications obtained through these methods have either not been evaluated or have been evaluated only in a very basic way"

2. Method and data

For this work, we used an April 2022 snapshot of the *Scopus* database (to which *SCImago* has access by agreement with its owner, the company *Elsevier*).

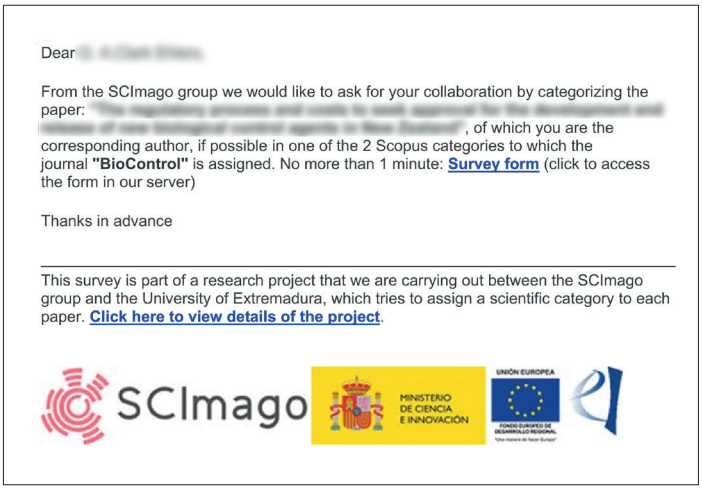
Scopus is known as the world's largest scientific database. It appeared in 2004 (**Hane**, 2004; **Pickering**, 2004) as an alternative to the *Thomson Reuters Web of Science* (*WoS*), covering most of the journals included in *WoS* and more (**Guerrero-Bote**; **Moya-Anegón**, 2012), and providing metadata on scientific documents and on citation links between these documents (**Guerrero-Bote** *et al.*, 2021).

The *Scopus* database uses the *ASJC* classification. This classifies journals into 27 subject areas, one of which is Multidisciplinary, which is where clearly multidisciplinary journals such as *Science* or *Nature* are classified. The other 26 subject areas are subdivided into 311 specific subject areas or categories, but each of those 26 subject areas has a miscellaneous category: Agricultural and Biological Sciences (miscellaneous), Arts and Humanities (miscellaneous), Biochemistry, Genetics and Molecular Biology (miscellaneous), etc.

In order to specify this classification a little more precisely, we made a "fractional assignation" of the journals to the categories. This fractional assignation consists of the fact that, if a journal is ascribed to 5 categories, each of those 5 affiliations is weighted by 1/5. We also eliminated both the Multidisciplinary subject area and the miscellaneous categories, distributing the weight among the corresponding categories. We were left with 26 subject areas and 285 categories or specific subject areas. The weight assigned to an affiliation to the Multidisciplinary subject area is divided among the 285, and the weight assigned to the miscellaneous categories is divided among the rest of the categories of the same subject area. In this way, there are journals that have the same weight for all the categories to which they belong and others that have different weights. This is a consequence of direct assignation, of possible assignation to a miscellaneous category, and of possible assignation to the Multidisciplinary subject area.

Although some classifications have forced each work to be assigned to a single category (**Milojević**, 2020; **Waltman**; **Van Eck**, 2012), there are currently many studies on multidisciplinarity (**Zhang**; **Rousseau**; **Glänzel**, 2016; **Huang** *et al.*, 2021; **Thijs**; **Huang**; **Glänzel**, 2021), so that we have considered allowing authors to assign more than one category. To do so, the research-guarantor corresponding authors (**De-Moya-Anegón** *et al.*, 2013) are asked to assign up to 5 categories for each work, indicating the percentage for which the work would belong to each category. They are asked to assign as few categories as possible and that, in so far as possible, the assigned categories be from the categories that *Scopus* assigns to the journal in which the work was published. One must bear in mind that, when authors submit a paper, they do so knowing the scope of the journal and the categories assigned to it, and the review process that the papers follow is oriented to the said scope and category.

The survey is done by email (Figure 1 shows an example).


Dear [redacted]

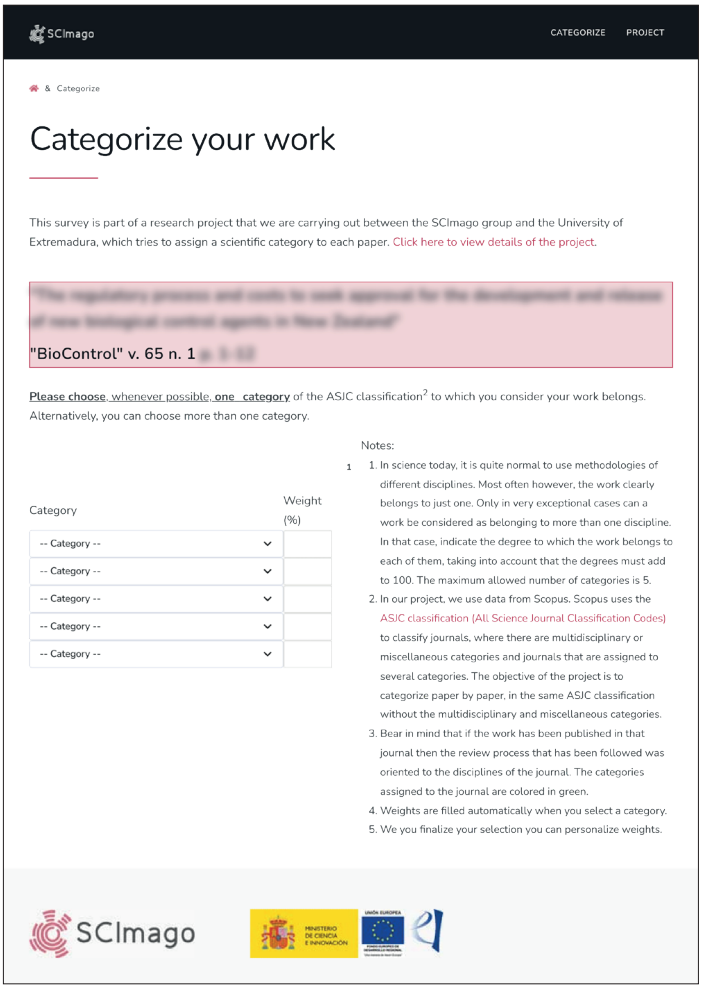
From the SCImago group we would like to ask for your collaboration by categorizing the paper: "[redacted]", of which you are the corresponding author, if possible in one of the 2 Scopus categories to which the journal **"BioControl"** is assigned. No more than 1 minute: **Survey form** (click to access the form in our server)

Thanks in advance

This survey is part of a research project that we are carrying out between the SCImago group and the University of Extremadura, which tries to assign a scientific category to each paper. **Click here to view details of the project**.

SCImago · MINISTERIO DE CIENCIA E INNOVACIÓN · UNIÓN EUROPEA

Figure 1. Example of an email sent.


SCImago · CATEGORIZE · PROJECT

& Categorize

Categorize your work

This survey is part of a research project that we are carrying out between the SCImago group and the University of Extremadura, which tries to assign a scientific category to each paper. Click here to view details of the project.

"[redacted]"
"BioControl" v. 65 n. 1 [redacted]

Please choose, whenever possible, one category of the ASJC classification[2] to which you consider your work belongs. Alternatively, you can choose more than one category.

Category	Weight (%)
-- Category --	
-- Category --	
-- Category --	
-- Category --	
-- Category --	

Notes:

1. In science today, it is quite normal to use methodologies of different disciplines. Most often however, the work clearly belongs to just one. Only in very exceptional cases can a work be considered as belonging to more than one discipline. In that case, indicate the degree to which the work belongs to each of them, taking into account that the degrees must add to 100. The maximum allowed number of categories is 5.
2. In our project, we use data from Scopus. Scopus uses the ASJC classification (All Science Journal Classification Codes) to classify journals, where there are multidisciplinary or miscellaneous categories and journals that are assigned to several categories. The objective of the project is to categorize paper by paper, in the same ASJC classification without the multidisciplinary and miscellaneous categories.
3. Bear in mind that if the work has been published in that journal then the review process that has been followed was oriented to the disciplines of the journal. The categories assigned to the journal are colored in green.
4. Weights are filled automatically when you select a category.
5. We you finalize your selection you can personalize weights.

SCImago · MINISTERIO DE CIENCIA E INNOVACIÓN · UNIÓN EUROPEA

Figure 2. Example of the information collection form.

The survey form can be accessed in the links. It is as shown in Figure 2. Note that in both the email and the form the author is asked to assign a single category whenever possible, and better if it is one of the journal's categories.

To generate the sample, we start from the list of active *Scopus* journals for which the SJR is calculated (**Guerrero-Bote**; **Moya-Anegón**, 2012). We decided to focus on a recent year even if it were not the most recent, which is why 2020 was chosen. In that year, the SJR is calculated for 34 169 journals. It is considered that a sample of 15 000 works could be both feasible and more than sufficient.

For all the journals to be represented based on their size, it was decided that the sample would include one paper for every 200 papers a journal published, in particular, the integer part of dividing by 200 the number of papers published. This led to a total of 8751 papers representing 3271 journals. The rest of the journals publish fewer than 200 papers. For these journals, one paper is taken from those which are most prestigious, as measured by having an SJR greater than 0.6. Of these, there were 6338, making a total of 15 089 papers.

Once the selection of the journals had been made and the corresponding number of their papers to be taken established, all the papers published in them in 2020 that included the corresponding author's email in the database were taken. To avoid bombarding an author with several emails, we kept only one work per author.

We then assigned a random order number to each paper from each journal. Taking those papers with a random order number less than or equal to the number of papers assigned to each journal, we sent the first wave of emails on 17-10-2022.

We received 1123 responses for this first wave of 15 089 emails. This led us to launch a second wave with 13 966 emails corresponding to the following papers from each incomplete journal by its order number. For this second wave we obtained 1017 responses, so that there were still 12 949 left, proceeding in the same way with a third wave, and so on.

We ended the survey on 14-01-2023, having received 13 449 responses, which represents 89.13% of our objective.

3. Results

To check how robust the sample of responses received was, we compared it with the total number of publications in 2020 from different perspectives. Throughout this section, we shall refer to three sets of publications as follows:

a) "Citables 2020": The set of citable *Scopus* papers of 2020.

b) "Sent": The selection of papers to whose authors the invitation to participate in the survey had been sent (a subset of Citables 2020).

c) "Received": The set of papers for which we received a response from the authors (a subset of Sent).

We chose as a first verification that of the country of affiliation of the corresponding authors of the works in the sample. Table 1 lists the percentages by country of the corresponding authors of the set of Citables 2020, of Sent, and of Received, for countries with a greater than 1% Citables 2020 percentage. Figure 3 shows a plot of these data is shown. The complete table with all the countries is given in Annex 1. In the calculation of these percentages, it had to be taken into account that there are works with multiple affiliations that may cause them to be added to more than one country.

One observes that the choice of works for the survey (Sent) has the same proportional distribution by country as in the total of Citables 2020. The response obtained, being also quite proportional, shows some striking data, such as the low response of authors affiliated to the countries with the highest proportion of scientific output (China especially, and the United States to a lesser extent), compared with the high response rate of such countries as Italy, Spain, and Brazil.

There also stands out the difference between the percentages of Sent works and the total Citables 2020 in the cases of China and the United States. This small imbalance is probably due to the fact that, as explained above, in the case of journals with fewer than 200 articles, only those with an SJR greater than 0.6 are considered, which today is commoner in journals of the United States than in those of China.

Table 1. Percentages of affiliation by the country of the authors.

Country	% Citables 2020	% Sent	% Received
China	20.63	18.53	5.97
United States	16.25	19.19	12.59
India	5.57	4.64	5.47
United Kingdom	4.25	5.10	3.10
Germany	3.98	4.45	4.90
Russian Federation	3.29	2.21	3.61
Japan	3.28	3.18	2.23
Italy	3.17	3.02	6.56
Spain	2.41	2.74	7.32
South Korea	2.36	2.33	1.21
Canada	2.36	2.65	2.09
France	2.35	2.63	2.95
Brazil	2.33	2.12	4.11
Australia	2.25	2.56	2.85
Iran	1.84	1.31	1.78
Turkey	1.37	1.16	2.06
Indonesia	1.36	1.18	1.89
Poland	1.33	1.48	1.56
Netherlands	1.23	1.72	1.17

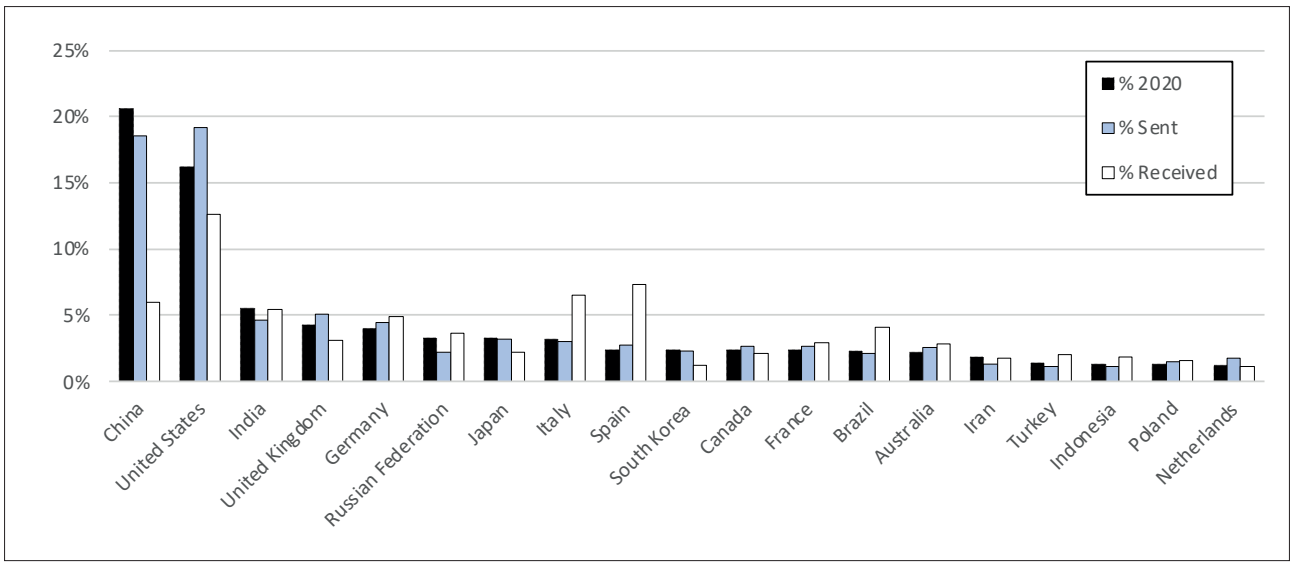


Figure 3. Plot of the affiliation percentages by country of the authors.

To check the thematic distribution of the sample, let us also compare the percentages of subject areas and categories.

Table 2 lists the percentages by subject area of the set of Citables 2020, of Sent, and of Received, for the 26 subject areas. Figure 4 shows a plot of these data.

Table 2. Assignation percentages by area.

ASJC	Description	% 2020	% Sent	% Received
1100	Agricultural and Biological Sciences	4.71	4.90	5.23
1200	Arts and Humanities	2.24	1.19	1.21
1300	Biochemistry, Genetics and Molecular Biology	5.32	6.28	6.45
1400	Business, Management and Accounting	1.51	2.31	2.37
1500	Chemical Engineering	2.05	1.94	1.72
1600	Chemistry	3.98	3.88	3.45
1700	Computer Science	8.10	6.11	6.80
1800	Decision Sciences	0.68	0.57	0.77
1900	Earth and Planetary Sciences	3.38	3.57	3.14
2000	Economics, Econometrics and Finance	0.97	1.62	2.18
2100	Energy	2.38	1.97	2.24
2200	Engineering	10.64	8.43	7.60
2300	Environmental Science	3.95	4.45	3.94
2400	Immunology and Microbiology	1.18	1.42	2.21
2500	Materials Science	5.52	5.01	5.36
2600	Mathematics	3.95	4.49	5.60
2700	Medicine	19.81	19.78	17.33
2800	Neuroscience	1.24	1.57	1.73
2900	Nursing	1.00	0.91	1.44
3000	Pharmacology, Toxicology and Pharmaceutics	1.83	1.91	1.65
3100	Physics and Astronomy	6.51	6.43	4.37
3200	Psychology	1.32	2.25	2.51
3300	Social Sciences	6.02	7.31	8.99
3400	Veterinary	0.53	0.43	0.26
3500	Dentistry	0.42	0.45	0.34
3600	Health Professions	0.75	0.83	1.10

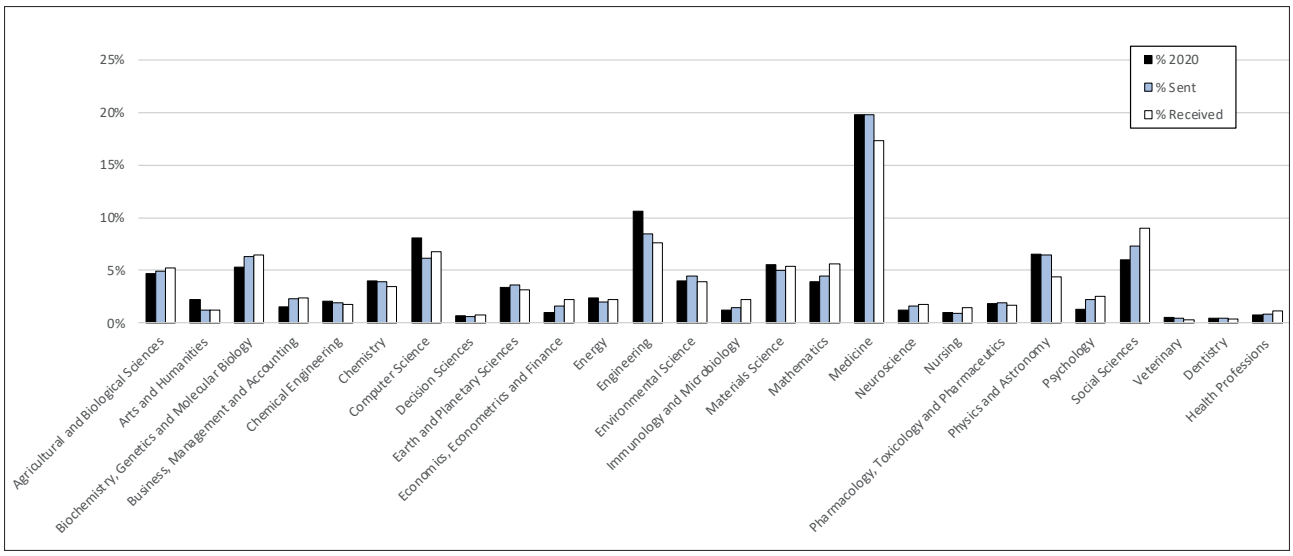


Figure 4. Graphical representation of assignation percentages by area.

As can be seen, there is very little variation in the percentages of assignation to subject areas between the sets of Citables 2020, Sent, and Received. At a finer grain, for the 285 specific subject areas, the distribution again shows little variation between Citables 2020, Sent, and Received. In this case, due to its length, we relegate the complete percentage data table to Annex 2.

Once verified that the sample used is robust in terms of its thematic variety and affiliation by country, we shall analyse the data collected, i.e., the categorizations that the authors made of their papers into the different specific subject areas. Table 3 compares the categorization percentages by number of categories, i.e., the percentage of papers that have a single category assigned, the percentage of those that have two categories assigned, etc., where:

a) Citables 2020 refers again to the set of 2020 citable papers from *Scopus*, classified with the "fractional assignation" of the journals.

b) Sample refers to the set of papers completed in the survey (Received), classified with the "fractional assignation" of the journals.

c) Survey refers to the set of papers completed in the survey (Received), but classified with the new assignation made by the authors.

The Items column indicates the number of papers in the set, and the Assignations column indicates the total number of assigned categories that accumulate those papers. As can be seen, while there is little variation between the categorization percentages of the total citable in 2020 and the sample, which serves to strengthen the validity of the sample, there is great variation from the categorization obtained from the survey.

Firstly, it called our attention that there is no greater percentage of authors who assign their works to a single category, despite the indications given in the questionnaire and in the email, considering the effort made in works such as **Milojević** (2020) or in **Waltman** & **Van-Eck** (2012) for achieving a categorization system that classifies papers into unique categories. Also striking is the high percentage of papers with 4 or more categories in the set of citable from *Scopus* and in the set of papers from the survey in the "fractional assignation". Let us recall that, as explained in the Introduction, the "fractional assignation" consists of eliminating the subject area Multidisciplinary, so that the papers assigned to that subject area are assigned to all categories (with a weight of 1/285), and miscellaneous categories are also removed from each subject area, reassigning their papers to all other categories of the subject area (with a weight of 1 divided by the number of remaining categories). Therefore, any paper originally belonging to the subject area Multidisciplinary or to any of the miscellaneous categories will, in the "fractional assignation," be assigned to a large number of categories (although with little weight in each of them).

Table 3. Categorization percentages by number of assigned categories.

Source	Items	Assignations	%1	%2	%3	%4	%>4
Citables 2020	3246022	56360548	15.57	17.81	11.79	7.77	47.05
Sample	13449	248621	15.01	18.83	11.83	7.09	47.25
Survey	13449	26141	44.85	30.70	14.40	5.39	4.68

For this reason, it may be pertinent to also analyse these percentages from the perspective of the original *Scopus* assignation, but excluding from the statistics the papers in the Multidisciplinary subject area and in any of the miscellaneous categories (both for the percentages of the original assignation and for those of the authors' assignation). Table 4 lists

these values. Again, one observes little variation between the percentages of original categorization of the total of 2020 citable and of the survey. It does seem interesting that the elimination of the statistics of all these presumably multidisciplinary papers (which represent approximately 43% of those received in the survey and 42% of the total citable papers) does not seem to have substantially modified the authors' assignation percentages. It is true that by including them in the statistics, the percentage of unique assignations significantly decreases in favour of multiple ones, but with a very homogeneous distribution among those of 2, 3, 4, and more categories.

> "It called our attention that there is no greater percentage of authors who assign their works to a single category, despite the indications given in the questionnaire and in the email"

Table 4. Categorization percentages by number of assigned categories excluding the Multidisciplinary and miscellaneous categories.

Source	Items	Assignations	%1	%2	%3	%4	%>4
Citables 2020	1895436	4900726	26.67	30.3	19.6	12.35	11.1
Sample	7613	18798	26.52	33.0	20.4	11.6	8.5
Survey	7613	14356	47.432	29.7	13.9	4.926	4.09

Table 5 presents data similar to those of Table 3, but instead of with the number of assigned categories, with the number of winning assigned categories understood as being those in which the categories with the greatest weight have exactly the same weight.

Table 5. Categorization percentages by number of winning categories.

Source	Items	Winners	%1 w.	%2 w.	%3 w.	%4 w.	%5 w.
Citables 2020	3246022	30600572	27.95	23.72	15.39	9.67	23.28
Sample	13449	131659	28.49	25.10	15.70	9.15	21.55
Survey	13449	21037	65.58	21.25	7.26	2.97	2.93

As in Table 4, one sees that the sample reflects the total 2020 set fairly accurately. Likewise, there is a major increase in the results obtained from the survey with respect to the number of papers with a winning category, although, as in Table 4, the large number of categories the authors assign with equal weight is still striking.

Table 6 presents the range of weights the authors assign to the areas of their works, as well as the percentage of categories they assign to their works that are included among those the database assigns to the journal, which we denote by "coincidence". For example, for the first band (Bin 1), in 306 papers the authors assigned a total of 411 areas with weights between 0% and 10% (>=0 and <10). In 190 of these 411 assignations, this assignation was also found among the journal's areas, representing 46.23% coincidence.

Table 6. Assignations classified by weight, and percentages of coincidences with the respective journal.

Bin	Min wt	Max wt	Items (papers)	Assignations	Coincidences	%
1	0	9.34	306	411	190	46.23
2	10	18	998	1400	655	46.79
3	20	29	2411	5653	2706	47.87
4	30	39	2001	4176	2315	55.44
5	40	48.94	635	785	480	61.15
6	50	55	3124	5743	3833	66.74
7	60	68	444	444	336	75.68
8	70	75	682	682	531	77.86
9	80	89	555	555	452	81.44
10	90	99.99	198	198	161	81.31
11	100	100	6094	6094	5317	87.25
Total	-	-	13449	26141	16976	64.94

Figure 5 shows the number of assignations within each weight band of Table 6. One observes that most of the assignations are around the values of 20%, 30%, 50%, and 100%, which, contrasted with the data in Table 5 which indicates a very low percentage of works with many winning categories, leads us to think that authors tend to use round numbers to distribute the weight of the different categories of their works.

It seems logical to think that those author-made assignations with greater weight should have a greater likelihood of coincidence (i.e., of being among those assigned to the respective journals), and this is indeed reflected in the table. However, we find the percentage of assignations in which this is not the case to be very high, also considering that the papers published in journals in the Multidisciplinary subject area will always coincide with the journal since these journals are assigned to all categories due to the "fractional assignation." The same occurs, although with a lower probability, with the journals in the miscellaneous categories, and, furthermore, these coincidences can be understood as very weak coincidences.

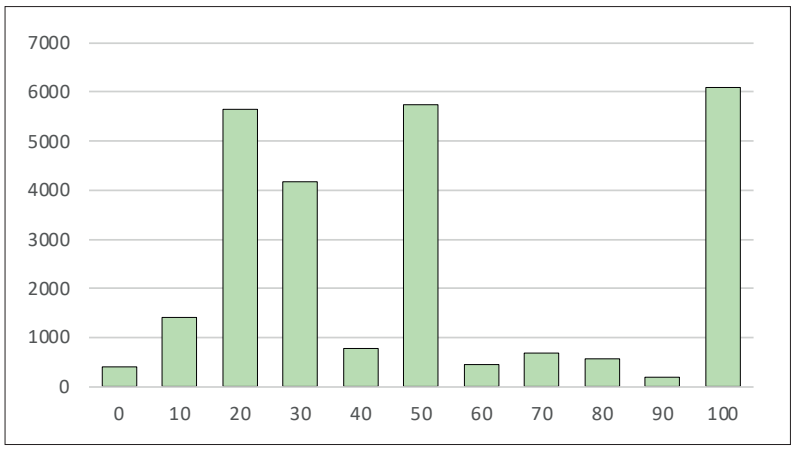


Figure 5. Number of assignations per range of assignation weights.

These cases of weak coincidence make us think that, to better estimate the coincidence, rather than treating it as a logical value, it should be weighted taking into account the concordance between the assignation percentage that the author establishes in a category (author weight) and the weight that the journal has in the category with the "fractional assignation" (journal weight). We have defined a coincidence weight that is calculated as the sum, for each coincident assignation made by the author, of the lesser of the weights – that of the author or that of the journal. Thus, if all the author assignations coincide with those of the journal and with the same weight, the final weight of the coincidence will be 100%.

Table 7 lists the percentages of coincidences and their weights according to the order of assignation. In the second row for example, for 2788 papers the authors made assignations to categories with a weight that made them come in second place (non-winning categories), with a total of 3411 assignations (i.e., in some cases there were assignations to two or more categories with the same weight, remaining in second place because there was another assignation with greater weight). On 53.00% of the occasions there was coincidence (those assignations were included among those of the journal). The average weight of the assignations made by the authors in those cases was 23.69%, although the average weight of the coincidences by item was 7.67%.

For the first row, in the 13449 cases (total), categories (sometimes divided into several) with winning weight were assigned, a total of 21041 assignations, among which in 68.33% of the cases there was coincidence. The average weight of the authors' assignations was 59.06% and the coincidence per item was 20.91%.

Table 7. Percentages of coincidences and their weights by order of assignation.

Order	Items	Assignations	Coincidences	Average author wt	Average coincidence wt
1	13449	21037	68.33	59.06	20.91
2	2788	3411	53.00	23.69	7.67
3	984	1187	47.85	14.40	4.60
4	314	400	45.00	9.65	2.79
5	106	106	42.45	7.41	1.96

The data in Table 7 indicate that the coincidence rate, as well as its weight, is greater in the first-order assignations, and that they decrease noticeably as the order number increases. Furthermore, the low average percentage of the weight of the coincidences is striking, even for the first-order assignations. Here one must take into account the effect produced by the "fractional assignation" since, in the publications of journals in the Multidisciplinary subject area, the weight of the coincidence will be extremely low –in the most favourable of cases (when its author assigns it to 5 subject areas), it will have a maximum coincidence percentage of 5×1/285 = 1.75%. Although to a lesser extent, the same is the case with the papers of each subject area's miscellaneous detailed subject area. For example, papers published in category 3301 Social Sciences (miscellaneous) of the subject area Social Sciences would be limited to 5×1/22 ≈ 22.72%.

Table 8 presents just the winning assignations according to the number of winners. For example, a single winning category was assigned in 8818 papers, with an average weight of 89.10%, which resulted in coincidence with the journal in 83.19% of the occasions with an average weight in the coincidence of 32.93%. There were two winning categories in 2858 papers (therefore, 2858×2=5716 winning assignations) with an average weight of 48.92%, in which 66.10% were coincidences with an average weight of 17.16%.

> "The low average percentage of the weight of the coincidences is striking, even for the first-order assignations"

Table 8. Coincidence percentages and their weights according to the number of winning categories.

Winners	Items	Assignations	Coincidences	Average author wt	Average coincidence wt
1	8820	8820	83.19	89.10	32.93
2	2858	5716	66.10	48.92	17.16
3	977	2931	55.00	33.25	10.30
4	400	1600	50.50	25.00	7.12
5	394	1970	42.64	20.00	5.02

One observes that the coincidences are higher the fewer the winning categories.

> In some particular case that we verified manually, we saw that the classification made by the authors is inconsistent with the references used as intellectual bases

We also investigated possible correlations of the percentage of categories assigned to journals and the percentage of winners assigned to the journal with the paper's number of references, citation, normalized citation, number of authors, with the corresponding author's prestige as measured by their number of papers or by their brute force (number of papers × average normalized citation), with the average of the authors, finding no significant correlations at 1%.

The only minimally significant correlation found was with the journal's SJR (0.067 and 0.068) in the sense that the greater the journal's SJR, the greater the probability that the categories assigned in the survey are among those assigned to the journals, which could be interpreted as that high-impact journals contain papers on subjects which (according to their authors) are more closely linked to the journal's subject area.

A significant negative correlation was also found between the percentage of assignations included in the journal and the percentage of winners included in the journal with the number of categories assigned and with the number of winners (of the order of -0.30). This reveals a certain tendency that the more categories the author assigns, the lower will be the probability that they coincide with those of the journal.

4. Conclusions

In this work, a collection of papers has been generated representative of those indexed in *Scopus* in 2020, categorized by the corresponding authors themselves as research guarantors (**De-Moya-Anegón** *et al.*, 2013) using the same *ASJC* scheme in a fractional way with up to a maximum of 5 categories, which we have named Author's Assignation Collection. The publications in the collection closely represent the thematic variety by area and category, as well as by country of affiliation of their authors, of the complete set of *Scopus* publications. However, as we have shown, authors of all nationalities did not respond equally.

The most important thing is that, despite having been explicitly urged to use few assignations, and to match them in so far as possible with those assigned by the journal, the author's responses show what is, in our opinion, a high proportion of multiple assignations that do not coincide with the journals. This deviation from the journal's theme is more notable the greater the number of assignations made by the author.

In some particular case that we verified manually, we saw that the classification made by the authors is inconsistent with the references used as intellectual bases. For example, there are some cases in which the authors assigned a paper to the Library and Information Science area without including a single reference to a paper that can be considered as from that area.

However, given the importance of the human factor involved in a survey-based methodology, we cannot state determinatively that the authors' classification is unquestionably more accurate, or that it presents better scientometric characteristics. What is certain is that there is a striking deviation between the journal-based classifications established in *Scopus* and the classification that the authors of the works themselves believe to be most appropriate, and to which, in some way, any classification should converge.

> The Author's Assignation Collection (AAC) that we have created can be used as a further classification of reference for evaluating other classification systems of scientific documents collected in *Scopus*

With all this, we consider that the Author's Assignation Collection (AAC) that we have created can be used as a further classification of reference for evaluating other classification systems of scientific documents collected in *Scopus* that also use the *ASJC* scheme.

5. References

Althouse, Benjamin M.; **West, Jevin D.**; **Bergstrom, Carl T.**; **Bergstrom, Theodore** (2009). "Differences in impact factor across fields and over time". *Journal of the Association for Information Science and Technology*, v. 60, n. 1, pp. 27-34. *https://doi.org/10.1002/asi.20936*

Bornmann, Lutz; **Leydesdorff, Loet** (2017). "Skewness of citation impact data and covariates of citation distributions: A large-scale empirical analysis based on Web of Science data". *Journal of informetrics*, v. 11, n. 1, pp. 164-175.
https://doi.org/10.1016/j.joi.2016.12.001

Bornmann, Lutz; **Tekles, Alexander**; **Leydesdorff, Loet** (2019). "How well does I3 perform for impact measurement compared to other bibliometric indicators? The convergent validity of several (field-normalized) indicators". *Scientometrics*, v. 119, n. 2, pp. 1187-1205.
https://doi.org/10.1007/s11192-019-03071-6

De-Moya-Anegón, Félix; **Guerrero-Bote, Vicente P.**; **Bornmann, Lutz**; **Moed, Henk F.** (2013). "The research guarantors of scientific papers and the output counting: a promising new approach". *Scientometrics*, v. 97, 421-434.
https://doi.org/10.1007/s11192-013-1046-0

Elsevier (2023). *What is the complete list of Scopus*, February 02. *Subject Areas and All Science Journal Classification Codes (ASJC)? - Scopus: Access and use Support Center*.
https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus

Glänzel, Wolfgang; **Schubert, András**; **Czerwon, Hans-Jürgen** (1999). "An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis". *Scientometrics*, v. 44, n. 3, pp. 427–439.
https://doi.org/10.1007/bf02458488

Glänzel, Wolfgang; **Thijs, Bart**; **Huang, Ying** (2021). "Improving the precision of subject assignment for disparity measurement in studies of interdisciplinary research". In: W. Glänzel, S. Heeffer, P.S. Chi, R. Rousseau, *Proceedings of the 18th International Conference of the International Society of Scientometrics and Informetrics (ISSI 2021)*, Leuven University Press, pp. 453–464.
https://kuleuven.limo.libis.be/discovery/fulldisplay?docid=lirias3394551&context=SearchWebhook&vid=32KUL_KUL:Lirias

Gómez-Crisóstomo, María-Rocío (2011). *Estudio y comparación de la web of Science y Scopus (1996-2007)*. Tesis de doctorado. Universidad de Extremadura.
https://portalcientifico.uah.es/documentos/5da82cb42999523a1dc44829?lang=en

Guerrero-Bote, Vicente P.; **Chinchilla-Rodríguez, Zaida**; **Mendoza, Abraham**; **De-Moya-Anegón, Félix** (2021). "Comparative analysis of the bibliographic data sources dimensions and scopus: An approach at the country and institutional levels". *Frontiers in research metrics and analytics*, v. 5.
https://doi.org/10.3389/frma.2020.593494

Guerrero-Bote, Vicente P.; **De-Moya-Anegón, Félix** (2012). "A further step forward in measuring journals' scientific prestige: The SJR2 indicator". *Journal of informetrics*, v. 6, n. 4, pp. 674-688.
https://doi.org/10.1016/j.joi.2012.07.001

Hane, Paula J. (2004). *Elsevier announces Scopus service*, March 15.
https://newsbreaks.infotoday.com/NewsBreaks/Elsevier-Announces-Scopus-Service-16494.asp

Hassan-Montero, Yusef; **Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2014). "Graphical interface of the Scimago Journal and Country Rank: an interactive approach to accessing bibliometric information". *El profesional de la información*, v. 23, n. 3.
https://doi.org/10.3145/epi.2014.may.07

Huang, Ying; **Glänzel, Wolfgang**; **Thijs, Bart**; **Porter, Alan L.**; **Zhang, Lin** (2021). *The comparison of various similarity measurement approaches on interdisciplinary indicators*. Feb - KU Leuven.

Janssens, Frizo; **Glänzel, Wolfgang**; **De-Moor, Bart** (2008). "A hybrid mapping of information science". *Scientometrics*, v. 75, n. 3, pp. 607–631.
https://doi.org/10.1007/s11192-007-2002-7

Janssens, Frizo; **Zhang, Lin**; **De-Moor, Bart**; **Glänzel, Wolfgang** (2009). "Hybrid clustering for validation and improvement of subject-classification schemes". *Information processing & management*, v. 45, n. 6, pp. 683–702.
https://doi.org/10.1016/j.ipm.2009.06.003

Klavans, Richard; **Boyack, Kevin W.** (2005). "Identifying a better measure of relatedness for mapping science". *Journal of the Association for Information Science and Technology*, v. 57, n. 2, pp. 251-263.
https://doi.org/10.1002/asi.20274

Klavans, Richard; **Boyack, Kevin W.** (2006). "Quantitative evaluation of large maps of science". *Scientometrics*, v. 68, n. 3, pp. 475-499.
https://doi.org/10.1007/s11192-006-0125-x

Lancho-Barrantes, Barbara S.; **Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2010). "What lies behind the averages and significance of citation indicators in different disciplines?". *Journal of information science*, v. 36, n. 3, pp. 371-382.
https://doi.org/10.1177/0165551510366077

Leydesdorff, Loet; **De-Moya-Anegón, Félix**; **Guerrero-Bote, Vicente P.** (2010). "Journal maps on the basis of Scopus data: A comparison with the Journal Citation Reports of the ISI". *Journal of the American Society for Information Science and Technology*, v. 61, n. 2, pp. 352-369.
https://doi.org/10.1002/asi.21250

Leydesdorff, Loet; **De-Moya-Anegón, Félix**; **Guerrero-Bote, Vicente P.** (2015). "Journal maps, interactive overlays, and the measurement of interdisciplinarity on the basis of scopus data (1996–2012)". *Journal of the Association for Information Science and Technology*, v. 66, n. 5, pp. 1001-1016.
https://doi.org/10.1002/asi.23243

Milojević, Staša (2020). "Practical method to reclassify Web of Science articles into unique subject categories and broad disciplines". *Quantitative science studies*, v. 1, n. 1, pp. 183-206.
https://doi.org/10.1162/qss_a_00014

Opthof, Tobias; **Leydesdorff, Loet** (2010). "Caveats for the journal and field normalizations in the CWTS ('Leiden') evaluations of research performance". *Journal of informetrics*, v. 4, n. 3, pp. 423-430.
https://doi.org/10.1016/j.joi.2010.02.003

Pickering, B. (2004). "Elsevier prepares Scopus to rival ISI Web of science". *Information world review*, 8.

Šubelj, Lovro; **Van-Eck, Nees-Jan**; **Waltman, Ludo** (2016). "Clustering scientific publications based on citation relations: A systematic comparison of different methods". *PLoS one*, v. 11, n. 4, e0154404.
https://doi.org/10.1371/journal.pone.0154404

Thijs, Bart; **Huang, Ying**; **Glänzel, Wolfgang** (2021). *Comparing different implementations of similarity for disparity and variety measures in studies on interdisciplinarity*. FEB Research Report MSI_2103, Report No. MSI_2103.
https://lirias.kuleuven.be/retrieve/610314

Waltman, Ludo; **Van-Eck, Nees-Jan** (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the Association for Information Science and Technology*, v. 63, n. 12, pp. 2378-2392.
https://doi.org/10.1002/asi.22748

Wang, Qi; **Waltman, Ludo** (2016). "Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus". *Journal of informetrics*, v. 10, n. 2, pp. 347-364.
https://doi.org/10.1016/j.joi.2016.02.003

Zhang, Lin; **Rousseau, Ronald**; **Glänzel, Wolfgang** (2016). Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. *Journal of the Association for Information Science and Technology*, v. 67, n. 5, pp. 1257-1265.
https://doi.org/10.1002/asi.23487

6. Annexes

Annex 1. Complete list of affiliation percentages by country of the authors

Country	Citables 2020 %	Sent %	Received %	Country	Citables 2020 %	Sent %	Received %
China	20.63015	18.52546	5.96702	Indonesia	1.35858	1.17807	1.88906
United States	16.24934	19.19230	12.58621	Poland	1.33076	1.48093	1.55922
India	5.57112	4.64144	5.47226	Netherlands	1.22525	1.71849	1.16942
United Kingdom	4.25447	5.09919	3.10345	Taiwan	0.96073	1.05235	1.03448
Germany	3.98272	4.44903	4.89505	Malaysia	0.88135	0.65086	1.13193
Russian Federation	3.29436	2.20958	3.60570	Switzerland	0.87462	1.03220	1.19940
Japan	3.27829	3.17927	2.23388	Sweden	0.80177	1.16071	1.05697
Italy	3.16538	3.01603	6.55922	Egypt	0.68262	0.64183	0.79460
Spain	2.40927	2.74096	7.31634	Mexico	0.67758	0.62793	1.24438
South Korea	2.36397	2.32767	1.21439	Portugal	0.64170	0.68906	2.00900
Canada	2.35498	2.65483	2.09145	South Africa	0.64124	0.59390	0.71214
France	2.34842	2.62843	2.95352	Saudi Arabia	0.63869	0.50846	0.67466
Brazil	2.33221	2.11511	4.10795	Belgium	0.62897	0.76269	0.55472
Australia	2.25274	2.56106	2.84858	Pakistan	0.56870	0.39524	0.44978
Iran	1.83511	1.30866	1.78411	Denmark	0.54454	0.75922	0.43478
Turkey	1.36696	1.15932	2.06147	Czech Republic	0.53950	0.49943	0.62969

Country	Citables 2020 %	Sent %	Received %
Austria	0.52935	0.54041	0.77961
Israel	0.50487	0.63696	0.87706
Ukraine	0.49631	0.25145	0.56222
Norway	0.49356	0.68559	0.53223
Thailand	0.49269	0.41121	0.88456
Iraq	0.47436	0.44664	0.76462
Hong Kong	0.46572	0.43969	0.39730
Greece	0.46264	0.47720	1.00450
Singapore	0.44408	0.40635	0.18741
Finland	0.42363	0.58348	0.35982
Romania	0.38178	0.34036	0.60720
Argentina	0.36531	0.40079	1.36432
Viet Nam	0.35986	0.29035	0.46477
Colombia	0.34956	0.30355	0.77211
Chile	0.34562	0.34661	0.84708
New Zealand	0.32949	0.38968	0.38231
Ireland	0.31905	0.38760	0.26237
Nigeria	0.30580	0.22506	0.59970
Hungary	0.26290	0.25284	0.59970
Morocco	0.24156	0.21116	0.34483
Slovakia	0.19884	0.16254	0.23238
Algeria	0.19669	0.14865	0.33733
Tunisia	0.18645	0.17991	0.30735
United Arab Emirates	0.18486	0.15143	0.20240
Croatia	0.18324	0.15768	0.31484
Bangladesh	0.17848	0.11600	0.21739
Serbia	0.17149	0.15907	0.34483
Bulgaria	0.15296	0.10489	0.29235
Slovenia	0.14246	0.15559	0.20990
Ethiopia	0.14049	0.18546	0.32234
Jordan	0.13938	0.09794	0.14243
Philippines	0.12230	0.09308	0.17991
Peru	0.11911	0.07224	0.16492
Kazakhstan	0.11813	0.08058	0.11994
Ecuador	0.11778	0.09655	0.18741
Lithuania	0.10099	0.10141	0.08996
Qatar	0.08538	0.06946	0.08246
Uzbekistan	0.08506	0.07224	0.12744
Lebanon	0.08295	0.08822	0.08246
Cyprus	0.07952	0.07780	0.13493
Ghana	0.07793	0.08058	0.12744
Estonia	0.06850	0.07710	0.11994
Sri Lanka	0.06427	0.05696	0.11244
Kenya	0.06264	0.07780	0.14993
Macao	0.05365	0.05418	0.01499
Latvia	0.05325	0.03612	0.03748
Nepal	0.05043	0.04446	0.07496
Cuba	0.04736	0.02084	0.05997
Kuwait	0.04521	0.04515	0.06747
Oman	0.04513	0.04098	0.11244

Country	Citables 2020 %	Sent %	Received %
Belarus	0.04356	0.02709	0.07496
Cameroon	0.03965	0.05279	0.14993
Bosnia and Herzegovina	0.03741	0.03959	0.07496
Azerbaijan	0.03680	0.01875	0.03748
Luxembourg	0.03666	0.05071	0.08996
Uruguay	0.03350	0.03543	0.10495
Uganda	0.03213	0.04723	0.04498
Tanzania	0.03092	0.03890	0.08246
Iceland	0.02700	0.03265	0.00000
Venezuela	0.02517	0.01737	0.02999
Georgia	0.02381	0.01875	0.02999
Costa Rica	0.02320	0.01806	0.02999
Bahrain	0.02297	0.01181	0.00750
Armenia	0.02253	0.01250	0.05247
Palestine	0.02216	0.01598	0.05997
Malta	0.02062	0.02084	0.05997
Zimbabwe	0.01926	0.01945	0.02249
Puerto Rico	0.01845	0.02640	0.00750
Macedonia	0.01830	0.01181	0.01499
Jamaica	0.01700	0.02362	0.02249
Sudan	0.01627	0.01806	0.02999
Syrian Arab Republic	0.01566	0.01042	0.01499
Vatican City State	0.01552	0.01042	0.01499
Senegal	0.01517	0.01598	0.03748
Yemen	0.01369	0.01528	0.03748
Botswana	0.01331	0.01598	0.03748
Brunei Darussalam	0.01262	0.00972	0.02249
Myanmar	0.01198	0.01181	0.00000
Benin	0.01114	0.01181	0.00750
Montenegro	0.01108	0.01042	0.01499
Panama	0.01096	0.01945	0.02999
Malawi	0.01047	0.00903	0.01499
Moldova	0.01038	0.00903	0.00750
Trinidad and Tobago	0.01003	0.01598	0.01499
Libya	0.00954	0.00834	0.04498
Albania	0.00940	0.01042	0.04498
Zambia	0.00934	0.01320	0.01499
Rwanda	0.00922	0.00764	0.02249
Côte d'Ivoire	0.00902	0.01598	0.02249
Burkina Faso	0.00815	0.01042	0.02249
Namibia	0.00769	0.00834	0.00000
Kyrgyzstan	0.00754	0.01111	0.01499
Mauritius	0.00751	0.00556	0.00000
Fiji	0.00737	0.00764	0.00000
Paraguay	0.00658	0.00278	0.00750
Mozambique	0.00655	0.00834	0.03748
Bolivia	0.00650	0.00556	0.01499
Democratic Republic Congo	0.00621	0.00903	0.02999
Reunion	0.00516	0.00556	0.00750
Mongolia	0.00510	0.00556	0.00750

Country	Citables 2020 %	Sent %	Received %
North Korea	0.00510	0.00556	0.00000
Cambodia	0.00505	0.00347	0.00750
Madagascar	0.00493	0.00695	0.02249
Kosovo (UNMIK)	0.00473	0.00417	0.00000
Afghanistan	0.00386	0.00556	0.00750
Honduras	0.00360	0.00347	0.01499
Tajikistan	0.00357	0.00278	0.00000
Dominican Republic	0.00351	0.00417	0.00750
Congo	0.00345	0.00417	0.01499
Bhutan	0.00336	0.00556	0.01499
Togo	0.00334	0.00347	0.00000
Guatemala	0.00328	0.00417	0.00000
Laos	0.00319	0.00139	0.00000
Guadeloupe	0.00313	0.00556	0.00750
French Guiana	0.00313	0.00417	0.00000
Mali	0.00270	0.00139	0.00000
Gambia	0.00261	0.00347	0.00000
Gabon	0.00261	0.00347	0.00000
Monaco	0.00247	0.00139	0.00000
Barbados	0.00244	0.00347	0.00000
Papua New Guinea	0.00215	0.00278	0.00750
New Caledonia	0.00212	0.00208	0.00750
Niger	0.00203	0.00208	0.00750
Grenada	0.00200	0.00208	0.00000
Liechtenstein	0.00197	0.00139	0.00000
Sierra Leone	0.00191	0.00208	0.01499
Saint Kitts and Nevis	0.00180	0.00000	0.00000
French Polynesia	0.00177	0.00069	0.00000
Angola	0.00177	0.00278	0.00750
Swaziland	0.00174	0.00000	0.00000
Martinique	0.00162	0.00139	0.00000
Eritrea	0.00142	0.00208	0.00000
Burundi	0.00136	0.00278	0.00750
Lesotho	0.00122	0.00000	0.00000
Guinea	0.00116	0.00069	0.00000
Mauritania	0.00113	0.00208	0.00000
Guam	0.00107	0.00139	0.00000
Somalia	0.00102	0.00278	0.01499
Greenland	0.00099	0.00139	0.00000
Bahamas	0.00096	0.00069	0.00000
El Salvador	0.00093	0.00000	0.00000
Faroe Islands	0.00087	0.00208	0.00750
Belize	0.00084	0.00069	0.00750
Haïti	0.00081	0.00139	0.00000
Maldives	0.00078	0.00069	0.00000
San Marino	0.00078	0.00000	0.00000
Seychelles	0.00078	0.00069	0.00000

Country	Citables 2020 %	Sent %	Received %
Suriname	0.00073	0.00208	0.00000
Guyana	0.00070	0.00069	0.00000
Nicaragua	0.00067	0.00139	0.00000
Cape Verde	0.00064	0.00069	0.00000
Guinea-Bissau	0.00061	0.00069	0.00750
Samoa	0.00058	0.00000	0.00000
Chad	0.00046	0.00000	0.00000
Liberia	0.00046	0.00000	0.00000
Central African Republic	0.00046	0.00139	0.00000
Bermuda	0.00038	0.00139	0.00000
Antigua and Barbuda	0.00038	0.00069	0.00000
Djibouti	0.00035	0.00069	0.00000
Republic of South Sudan	0.00035	0.00069	0.00000
Aruba	0.00035	0.00000	0.00000
Gibraltar	0.00029	0.00000	0.00000
Falkland Islands (Malvinas)	0.00026	0.00000	0.00000
Turkmenistan	0.00026	0.00139	0.00750
Vanuatu	0.00026	0.00000	0.00000
Virgin Islands (U.S.)	0.00026	0.00000	0.00000
Solomon Islands	0.00023	0.00000	0.00000
Equatorial Guinea	0.00023	0.00000	0.00000
Timor-Leste	0.00020	0.00000	0.00000
Curaçao	0.00020	0.00000	0.00000
Andorra	0.00020	0.00069	0.00000
Cayman Islands	0.00017	0.00139	0.00000
Federated States of Micronesia	0.00015	0.00000	0.00000
Mayotte	0.00015	0.00000	0.00000
Palau	0.00015	0.00000	0.00000
Dominica	0.00012	0.00000	0.00000
Saint Vincent and the Grenadines	0.00012	0.00000	0.00000
Virgin Islands (British)	0.00009	0.00000	0.00000
Anguilla	0.00009	0.00000	0.00000
Turks and Caicos Islands	0.00009	0.00069	0.00000
Nauru	0.00009	0.00000	0.00000
Svalbard and Jan Mayen	0.00006	0.00000	0.00000
Tonga	0.00006	0.00000	0.00000
Saint Helena	0.00006	0.00000	0.00000
Saint Lucia	0.00003	0.00000	0.00000
Comoros	0.00003	0.00000	0.00000
Norfolk Island	0.00003	0.00000	0.00000
American Samoa	0.00003	0.00000	0.00000
Northern Mariana Islands	0.00003	0.00000	0.00000
Tuvalu	0.00003	0.00000	0.00000
South Georgia and the South Sandwich Islands	0.00003	0.00000	0.00000

Annex 2. Assignation percentages by specific subject areas

ASJC	Description	% 2020	% Sent	% Received
1102	Agronomy and Crop Science	0.57	0.51	0.68
1103	Animal Science and Zoology	0.57	0.58	0.62
1104	Aquatic Science	0.41	0.44	0.39
1105	Ecology, Evolution, Behavior and Systematics	0.78	1.01	1.05
1106	Food Science	0.74	0.69	0.79
1107	Forestry	0.23	0.21	0.23
1108	Horticulture	0.23	0.19	0.16
1109	Insect Science	0.24	0.23	0.23
1110	Plant Science	0.71	0.81	0.86
1111	Soil Science	0.24	0.23	0.22
1202	History	0.48	0.18	0.19
1203	Language and Linguistics	0.03	0.03	0.13
1204	Archeology (arts and humanities)	0.12	0.14	0.08
1205	Classics	0.06	0.03	0.00
1206	Conservation	0.06	0.04	0.15
1207	History and Philosophy of Science	0.14	0.17	0.15
1208	Literature and Literary Theory	0.38	0.05	0.04
1209	Museology	0.05	0.04	0.00
1210	Music	0.09	0.06	0.02
1211	Philosophy	0.39	0.28	0.34
1212	Religious Studies	0.24	0.10	0.05
1213	Visual Arts and Performing Arts	0.22	0.06	0.05
1302	Aging	0.13	0.13	0.26
1303	Biochemistry	0.83	0.94	0.76
1304	Biophysics	0.25	0.25	0.28
1305	Biotechnology	0.40	0.40	0.50
1306	Cancer Research	0.555	0.59	1.05
1307	Cell Biology	0.54	0.67	0.77
1308	Clinical Biochemistry	0.24	0.27	0.10
1309	Developmental Biology	0.16	0.22	0.19
1310	Endocrinology	0.19	0.24	0.17
1311	Genetics	0.60	0.78	0.80
1312	Molecular Biology	0.68	0.88	0.82
1313	Molecular Medicine	0.31	0.36	0.23
1314	Physiology	0.31	0.38	0.36
1315	Structural Biology	0.13	0.16	0.15
1402	Accounting	0.11	0.21	0.27
1403	Business and International Management	0.24	0.35	0.38
1404	Management Information Systems	0.10	0.12	0.08
1405	Management of Technology and Innovation	0.22	0.30	0.27
1406	Marketing	0.18	0.32	0.43
1407	Organizational Behavior and Human Resource Management	0.13	0.26	0.32
1408	Strategy and Management	0.34	0.47	0.37
1409	Tourism, Leisure and Hospitality Management	0.14	0.21	0.22
1410	Industrial Relations	0.06	0.08	0.03
1502	Bioengineering	0.38	0.35	0.29
1503	Catalysis	0.47	0.47	0.43

ASJC	Description	% 2020	% Sent	% Received
1504	Chemical Health and Safety	0.15	0.13	0.05
1505	Colloid and Surface Chemistry	0.19	0.18	0.20
1506	Filtration and Separation	0.19	0.17	0.11
1507	Fluid Flow and Transfer Processes	0.34	0.32	0.44
1508	Process Chemistry and Technology	0.33	0.31	0.21
1602	Analytical Chemistry	0.70	0.71	0.70
1603	Electrochemistry	0.46	0.43	0.36
1604	Inorganic Chemistry	0.58	0.56	0.42
1605	Organic Chemistry	0.87	0.82	0.92
1606	Physical and Theoretical Chemistry	0.88	0.88	0.77
1607	Spectroscopy	0.50	0.48	0.29
1702	Artificial Intelligence	1.07	0.71	1.66
1703	Computational Theory and Mathematics	0.33	0.29	0.32
1704	Computer Graphics and Computer-Aided Design	0.30	0.25	0.12
1705	Computer Networks and Communications	1.20	0.76	0.81
1706	Computer Science Applications	1.38	1.09	1.51
1707	Computer Vision and Pattern Recognition	0.53	0.43	0.69
1708	Hardware and Architecture	0.53	0.33	0.10
1709	Human-Computer Interaction	0.42	0.37	0.39
1710	Information Systems	0.75	0.60	0.51
1711	Signal Processing	0.54	0.41	0.37
1712	Software	1.06	0.86	0.33
1802	Information Systems and Management	0.35	0.14	0.16
1803	Management Science and Operations Research	0.19	0.21	0.37
1804	Statistics, Probability and Uncertainty	0.14	0.22	0.24
1902	Atmospheric Science	0.39	0.45	0.42
1903	Computers in Earth Sciences	0.13	0.15	0.20
1904	Earth-Surface Processes	0.27	0.30	0.33
1905	Economic Geology	0.11	0.13	0.05
1906	Geochemistry and Petrology	0.32	0.31	0.24
1907	Geology	0.43	0.46	0.42
1908	Geophysics	0.35	0.31	0.37
1909	Geotechnical Engineering and Engineering Geology	0.49	0.41	0.47
1910	Oceanography	0.23	0.27	0.25
1911	Paleontology	0.15	0.21	0.22
1912	Space and Planetary Science	0.40	0.44	0.14
1913	Stratigraphy	0.11	0.13	0.03
2002	Economics and Econometrics	0.65	1.16	1.78
2003	Finance	0.32	0.46	0.40
2102	Energy Engineering and Power Technology	0.89	0.62	0.82
2103	Fuel Technology	0.37	0.33	0.14
2104	Nuclear Energy and Engineering	0.29	0.24	0.16
2105	Renewable Energy, Sustainability and the Environment	0.83	0.78	1.13
2202	Aerospace Engineering	0.52	0.33	0.36
2203	Automotive Engineering	0.30	0.24	0.16
2204	Biomedical Engineering	0.48	0.44	0.46
2205	Civil and Structural Engineering	0.73	0.69	0.81
2206	Computational Mechanics	0.20	0.17	0.24

ASJC	Description	% 2020	% Sent	% Received
2207	Control and Systems Engineering	1.03	0.96	0.62
2208	Electrical and Electronic Engineering	2.54	1.85	1.93
2209	Industrial and Manufacturing Engineering	0.94	0.90	0.62
2210	Mechanical Engineering	1.33	0.99	0.99
2211	Mechanics of Materials	0.86	0.67	0.42
2212	Ocean Engineering	0.27	0.26	0.19
2213	Safety, Risk, Reliability and Quality	0.56	0.27	0.22
2214	Media Technology	0.23	0.16	0.09
2215	Building and Construction	0.43	0.34	0.37
2216	Architecture	0.22	0.15	0.12
2302	Ecological Modeling	0.12	0.15	0.12
2303	Ecology	0.49	0.55	0.79
2304	Environmental Chemistry	0.43	0.47	0.44
2305	Environmental Engineering	0.42	0.40	0.59
2306	Global and Planetary Change	0.16	0.22	0.13
2307	Health, Toxicology and Mutagenesis	0.40	0.46	0.17
2308	Management, Monitoring, Policy and Law	0.40	0.54	0.39
2309	Nature and Landscape Conservation	0.23	0.28	0.16
2310	Pollution	0.51	0.57	0.38
2311	Waste Management and Disposal	0.33	0.33	0.23
2312	Water Science and Technology	0.46	0.50	0.54
2402	Applied Microbiology and Biotechnology	0.16	0.16	0.36
2403	Immunology	0.36	0.48	0.64
2404	Microbiology	0.35	0.42	0.70
2405	Parasitology	0.13	0.14	0.18
2406	Virology	0.18	0.22	0.32
2502	Biomaterials	0.47	0.48	0.55
2503	Ceramics and Composites	0.53	0.47	0.37
2504	Electronic, Optical and Magnetic Materials	1.26	1.06	0.58
2505	Materials Chemistry	0.89	0.81	1.08
2506	Metals and Alloys	0.60	0.52	0.64
2507	Polymers and Plastics	0.60	0.56	0.63
2508	Surfaces, Coatings and Films	0.59	0.54	0.45
2509	Nanoscience and Nanotechnology	0.58	0.58	1.05
2602	Algebra and Number Theory	0.23	0.31	0.48
2603	Analysis	0.25	0.32	0.73
2604	Applied Mathematics	0.78	0.92	0.86
2605	Computational Mathematics	0.25	0.25	0.29
2606	Control and Optimization	0.45	0.22	0.26
2607	Discrete Mathematics and Combinatorics	0.16	0.17	0.22
2608	Geometry and Topology	0.16	0.24	0.38
2609	Logic	0.12	0.15	0.10
2610	Mathematical Physics	0.17	0.21	0.38
2611	Modeling and Simulation	0.45	0.39	0.63
2612	Numerical Analysis	0.12	0.16	0.20
2613	Statistics and Probability	0.32	0.50	0.76
2614	Theoretical Computer Science	0.51	0.64	0.31
2702	Anatomy	0.17	0.16	0.14

ASJC	Description	% 2020	% Sent	% Received
2703	Anesthesiology and Pain Medicine	0.35	0.35	0.28
2704	Biochemistry (medical)	0.19	0.20	0.27
2705	Cardiology and Cardiovascular Medicine	0.93	0.81	0.98
2706	Critical Care and Intensive Care Medicine	0.29	0.30	0.23
2707	Complementary and Alternative Medicine	0.27	0.21	0.12
2708	Dermatology	0.40	0.44	0.40
2709	Drug Guides	0.13	0.12	0.03
2710	Embryology	0.14	0.13	0.04
2711	Emergency Medicine	0.26	0.24	0.13
2712	Endocrinology, Diabetes and Metabolism	0.45	0.48	0.46
2713	Epidemiology	0.26	0.31	0.61
2714	Family Practice	0.21	0.17	0.08
2715	Gastroenterology	0.43	0.43	0.37
2716	Genetics (clinical)	0.28	0.34	0.29
2717	Geriatrics and Gerontology	0.26	0.27	0.20
2718	Health Informatics	0.37	0.34	0.27
2719	Health Policy	0.42	0.46	0.55
2720	Hematology	0.36	0.36	0.33
2721	Hepatology	0.26	0.27	0.17
2722	Histology	0.19	0.20	0.06
2723	Immunology and Allergy	0.45	0.49	0.25
2724	Internal Medicine	0.36	0.35	0.20
2725	Infectious Diseases	0.62	0.66	0.85
2726	Microbiology (medical)	0.38	0.41	0.31
2727	Nephrology	0.25	0.25	0.24
2728	Neurology (clinical)	0.70	0.79	0.62
2729	Obstetrics and Gynecology	0.50	0.47	0.45
2730	Oncology	0.98	0.94	0.96
2731	Ophthalmology	0.47	0.45	0.45
2732	Orthopedics and Sports Medicine	0.59	0.57	0.52
2733	Otorhinolaryngology	0.35	0.33	0.20
2734	Pathology and Forensic Medicine	0.38	0.35	0.20
2735	Pediatrics, Perinatology and Child Health	0.74	0.63	0.49
2736	Pharmacology (medical)	0.57	0.55	0.29
2737	Physiology (medical)	0.29	0.29	0.14
2738	Psychiatry and Mental Health	0.74	0.95	0.88
2739	Public Health, Environmental and Occupational Health	1.05	1.10	1.23
2740	Pulmonary and Respiratory Medicine	0.42	0.45	0.37
2741	Radiology, Nuclear Medicine and Imaging	0.71	0.63	0.53
2742	Rehabilitation	0.25	0.25	0.23
2743	Reproductive Medicine	0.21	0.21	0.19
2744	Reviews and References (medical)	0.13	0.12	0.07
2745	Rheumatology	0.26	0.26	0.26
2746	Surgery	1.19	1.07	1.02
2747	Transplantation	0.22	0.21	0.11
2748	Urology	0.37	0.38	0.26
2802	Behavioral Neuroscience	0.15	0.20	0.24
2803	Biological Psychiatry	0.12	0.14	0.08

ASJC	Description	% 2020	% Sent	% Received
2804	Cellular and Molecular Neuroscience	0.20	0.27	0.29
2805	Cognitive Neuroscience	0.18	0.24	0.41
2806	Developmental Neuroscience	0.09	0.12	0.08
2807	Endocrine and Autonomic Systems	0.08	0.09	0.05
2808	Neurology	0.31	0.35	0.42
2809	Sensory Systems	0.11	0.15	0.16
2902	Advanced and Specialized Nursing	0.07	0.06	0.05
2903	Assessment and Diagnosis	0.02	0.02	0.11
2904	Care Planning	0.02	0.02	0.01
2905	Community and Home Care	0.04	0.04	0.01
2906	Critical Care Nursing	0.04	0.04	0.02
2907	Emergency Nursing	0.04	0.04	0.01
2908	Fundamentals and Skills	0.04	0.03	0.02
2909	Gerontology	0.05	0.05	0.05
2910	Issues, Ethics and Legal Aspects	0.04	0.04	0.08
2911	Leadership and Management	0.06	0.04	0.10
2912	LPN and LVN	0.03	0.03	
2913	Maternity and Midwifery	0.04	0.04	0.05
2914	Medical and Surgical Nursing	0.04	0.03	0.05
2915	Nurse Assisting	0.02	0.02	0.02
2916	Nutrition and Dietetics	0.25	0.27	0.40
2917	Oncology (nursing)	0.04	0.04	0.03
2918	Pathophysiology	0.02	0.02	0.06
2919	Pediatrics	0.04	0.04	0.14
2920	Pharmacology (nursing)	0.02	0.02	0.00
2921	Psychiatric Mental Health	0.02	0.02	0.12
2922	Research and Theory	0.03	0.02	0.11
2923	Review and Exam Preparation	0.02	0.02	
3002	Drug Discovery	0.35	0.35	0.42
3003	Pharmaceutical Science	0.58	0.56	0.56
3004	Pharmacology	0.61	0.67	0.39
3005	Toxicology	0.29	0.33	0.28
3102	Acoustics and Ultrasonics	0.45	0.42	0.22
3103	Astronomy and Astrophysics	0.61	0.67	0.68
3104	Condensed Matter Physics	1.54	1.57	0.93
3105	Instrumentation	0.90	0.70	0.27
3106	Nuclear and High Energy Physics	0.65	0.69	0.50
3107	Atomic and Molecular Physics, and Optics	1.09	1.04	0.91
3108	Radiation	0.42	0.42	0.20
3109	Statistical and Nonlinear Physics	0.42	0.47	0.34
3110	Surfaces and Interfaces	0.43	0.45	0.32
3202	Applied Psychology	0.22	0.42	0.42
3203	Clinical Psychology	0.30	0.46	0.46
3204	Developmental and Educational Psychology	0.30	0.52	0.52
3205	Experimental and Cognitive Psychology	0.16	0.28	0.37
3206	Neuropsychology and Physiological Psychology	0.11	0.17	0.17
3207	Social Psychology	0.23	0.40	0.57
3302	Archeology	0.10	0.14	0.18

ASJC	Description	% 2020	% Sent	% Received
3303	Development	0.20	0.28	0.23
3304	Education	1.23	1.79	2.22
3305	Geography, Planning and Development	0.53	0.70	0.66
3306	Health (social science)	0.22	0.24	0.41
3307	Human Factors and Ergonomics	0.05	0.06	0.11
3308	Law	0.48	0.36	0.29
3309	Library and Information Sciences	0.19	0.17	0.19
3310	Linguistics and Language	0.49	0.50	0.54
3311	Safety Research	0.10	0.06	0.07
3312	Sociology and Political Science	0.64	0.98	1.21
3313	Transportation	0.18	0.19	0.34
3314	Anthropology	0.16	0.16	0.25
3315	Communication	0.25	0.32	0.54
3316	Cultural Studies	0.33	0.17	0.17
3317	Demography	0.07	0.12	0.07
3318	Gender Studies	0.09	0.12	0.17
3319	Life-span and Life-course Studies	0.06	0.06	0.03
3320	Political Science and International Relations	0.29	0.39	0.59
3321	Public Administration	0.11	0.17	0.16
3322	Urban Studies	0.12	0.18	0.30
3323	Social Work	0.06	0.07	0.11
3399	E-learning	0.06	0.08	0.13
3402	Equine	0.17	0.13	0.04
3403	Food Animals	0.19	0.15	0.13
3404	Small Animals	0.17	0.14	0.09
3502	Dental Assisting	0.06	0.06	0.01
3503	Dental Hygiene	0.06	0.06	0.03
3504	Oral Surgery	0.12	0.13	0.15
3505	Orthodontics	0.09	0.09	0.06
3506	Periodontics	0.08	0.10	0.09
3602	Chiropractics	0.02	0.01	
3603	Complementary and Manual Therapy	0.02	0.01	0.01
3604	Emergency Medical Services	0.01	0.01	0.02
3605	Health Information Management	0.05	0.06	0.07
3606	Medical Assisting and Transcription	0.01	0.01	0.01
3607	Medical Laboratory Technology	0.04	0.03	0.06
3608	Medical Terminology	0.01	0.01	0.00
3609	Occupational Therapy	0.02	0.02	0.03
3610	Optometry	0.02	0.03	0.07
3611	Pharmacy	0.05	0.04	0.09
3612	Physical Therapy, Sports Therapy and Rehabilitation	0.21	0.23	0.21
3613	Podiatry	0.01	0.01	0.01
3614	Radiological and Ultrasound Technology	0.08	0.06	0.05
3615	Respiratory Care	0.01	0.01	0.03
3616	Speech and Hearing	0.05	0.08	0.09
3699	Sports Science	0.15	0.21	0.34