

A Comparison of Citation Disciplinary Structure in Science between the G7 Countries and the BRICS Countries

Ting Yue^{1,2}, Liying Yang^{1†}, Per Ahlgren^{1,3}, Jielan Ding^{1,2},
Shuangqing Shi⁴, Rainer Frietsch⁵

Citation: Ting Yue, Liying Yang, Per Ahlgren, Jielan Ding, Shuangqing Shi, Rainer Frietsch (2018). A comparison of citation disciplinary structure in science between the G7 countries and the BRICS countries.

Vol. 3 No. 3, 2018

pp 14–31

DOI: 10.2478/jdis-2018-0012

Received: Jul. 4, 2018

Revised: Aug. 17, 2018

Accepted: Aug. 24, 2018

¹National Science Library of the Chinese Academy of Sciences, Beijing, 100190, China

²University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China

³KTH Library, KTH Royal Institute of Technology, Osquars backe 25, Stockholm, 10044, Sweden

⁴Library of Peking University Health Science Center, Beijing, 100083, China

⁵Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany

Abstract

Purpose: This study aims to compare the characteristics of citation disciplinary structure between the G7 countries and the BRICS countries.

Design/Methodology/Approach: In this contribution, which uses about 1 million Web of Science publications and two publications years (1993 and 2013), we compare the G7 countries and the BRICS countries with regard to this type of structure. For the publication year 2013, cosine similarity values regarding the citation disciplinary structures of these countries (and of nine other countries) were used as input to cluster analysis. We also obtained cosine similarity values for a given country and its citation disciplinary structures across the two publication years. Moreover, for the publication year 2013, the within-country Jeffreys-Matusita distance between publication and citation disciplinary structure was measured.

Research limitations: First, the citation disciplinary structures of countries depend on multiple and complex factors. It is therefore difficult to completely explain the formation and change of the citation disciplinary structure of a country. This study suggests some possible causes, whereas detailed explanations might be given by future research. Second, the length of the citation window used in this study is three years. However, scientific disciplines differ in their citation practices. Comparison between citations across disciplines using the same citation window length may affect the citation discipline structure results for some countries.

Practical limitations: First, the results of this study are based on the WoS database. However, in this database some fields are covered to a greater extent than others, which may affect the results for the citation discipline structure for some studied countries. In future research, we might repeat this study using another database (like Scopus) and, in that case, we would like to make comparisons between the two outcomes. Second, the use of a constant journal set



yielded that a large share of the journals covered by WoS year 2013 is ignored in the study. Thus, disciplinary structure is studied based on a quite restricted set of publications. The three mentioned limitations should be kept in mind when the results of this study are interpreted.

Originality/value: Disciplinary structure on country level is a highlighted topic for the S&T policy makers, especially for those come from developing countries. This study observes the disciplinary structure in the view of academic impact, and the result will provide some evidence to make decision for the discipline strategy and funding allocation. Besides, Jeffreys-Matusita distance is introduced to measure the similarity of citation disciplinary structure and publication disciplinary structure. By applying this measure, some new observations were drawn, for example, “Based on the comparison of publication disciplinary structure and citation disciplinary structure, the paper finds most BRICS countries have less impact with more publications”.

Findings: The outcome of the cluster analysis indicates that the G7 countries and BRICS countries are quite heterogeneous regarding their citation disciplinary structure. For a majority of the G7 countries, the citation disciplinary structure tend to be more stable compared to BRICS countries with regard to the years 1993 and 2013. Most G7 countries, with United States as an exception, turned out to have lower values on the Jeffreys-Matusita distance than BRICS countries, indicating a higher degree of heterogeneity between the publication and the citation disciplinary structure for the latter countries. In other words, BRICS countries still receive much less citations in most disciplines than their publication output would suggest. G7 countries can still expect more citations than is to be expected based on their publication output, thereby generating relatively more impact than BRICS countries.

Keywords Citation disciplinary structure; Country; Bibliometrics

1 Introduction

The concept of revealed comparative advantage (Balassa, 1965) suggests a specialization of countries in terms of their export portfolio. National economies tend to focus on what they do—relatively, not necessarily absolutely—best. Transferring this idea to science systems (Grupp, 1998; Schmoch et al., 2000; van Raan, 2003) profiles of relative advantage and disadvantage of the disciplinary profile of countries gets into focus. Independent of both disciplinary and country size effects, areas of intensive activity in the profile, one could also say in the disciplinary structure, of each country can be identified.

For individual countries, scientific disciplinary structure could reflect the distribution of scientific funds over disciplines and the characteristics of Science & Technology (S&T) strategy. S&T input, as well as output, could be used to analyze disciplinary structure of countries. The analysis of disciplinary structure based on S&T input might reveal how resources and funds are distributed over disciplines (Julio, 1998; Shibayama, 2011; Zhu, 2006; Wu, 2015). For S&T output, research



publications are typically used to analyze disciplinary structure. In this study, we deal with two types of publication-based disciplinary structures: publication volume disciplinary structure and citation disciplinary structure. For a given country, publication volume disciplinary structure could mirror the features of S&T strategy policy for different disciplines, while citation disciplinary structure could reflect the general composition of the academic impact within individual disciplines.

Publication volume and citation disciplinary structure have been addressed in earlier research. Glänzel et al. (2006) and Yue (2008) described the local characteristics of disciplinary structure, whereas a number of studies have compared differences in disciplinary profiles between countries (e.g. Liang et al., 2006; Kozłowski et al., 1999; Yang et al., 2012; Schulz & Manganote, 2012), or treated the evolution of such profiles over time (e.g. Bongioanni et al., 2014; Yang et al., 2012; Li N., 2017). Ding et al. (2018) compared the publication volume disciplinary structure in three top journals (Nature, Science, Proceedings of the National Academy of Sciences of the United States of America) between countries. Aksnes et al. (2014) investigated how the increasing publication output of China influences Netherland's disciplinary profile. A model to explain a country's research profile and competitiveness in certain disciplines was developed by Harzing & Giroud (2014). Daraio et al. (2017) used the pseudo-likelihood approach (Aurell & Ekeberg, 2012), introduced in the physics of complex systems, to assess interdependencies between scientific disciplinary profiles at the country level. Other studies have focused on the citation impact of disciplines at the country level, or compared citation impact and publication volume regarding individual disciplines (e.g. Kozłowski et al., 1999; King, 2004; Yue, 2008; Radosevic & Yoruk, 2014).

Yang et al. (2012) investigated the disciplinary structures of the publication volume of the G7 countries (Canada, France, Germany, Italy, Japan, United Kingdom and United States) and the BRICS countries (Brazil, China, India, Russia and South Africa). The results showed that the two groups had different characteristics in terms of publication volume disciplinary structure. Our study, however, focuses on citation disciplinary structure, and compares the G7 countries and the BRICS countries with regard to this type of structure. Since the publication volume and the citation disciplinary structure for a given country might be more or less similar, we also analyze the relation between publication volume and citation disciplinary structure for the countries of the study. We thereby address the question how the publication output and the impact—reflected in citations—of science systems of different countries differ.

The remainder of this paper is structured as follows. In the next section, the data, indicators and methods of the study are described. The third section reports the results and discusses them, whereas the final section put forward conclusions.



2 Data and methods

2.1 Data

The data source of the study is Bibmet (2017 Quarter 1 version), the bibliometric version of the Web of Science (WoS) at KTH Royal Institute of Technology (Sweden). Bibmet is a relational database, based on WoS raw data purchased from Clarivate Analytics. Bibmet, which is updated quarterly, contains about 51 million WoS publications from the publication year 1980 onwards. The database also includes citation relations between WoS source publications.

To compare the disciplinary structure over time, two publication years were used: 1993 and 2013. The year 2013 was selected in order to have a fairly recent end year, and at the same time ensuring that all (or almost all) citations from WoS publications to publications published in the end year were present in the Bibmet version used in the study. Further, we first used 2003 as start year, but since very small differences in disciplinary structure, compared to year 2013, were observed, we decided to use 1993 as start year instead. For citations, we used 3-year windows. For example, for the publication year 1993, the citation window is 1993–1995. The data were collected in the following way:

1. We used a list of journals obtained from the 2013 Journal Citation Reports Science Edition. This list was named JCR_list. JCR_list contains 8,534 journals, together with their ISSNs.
2. Each journal in JCR_list has been mapped to exactly one Essential Science Indicators (ESI) class. These classes—except “Economics & Business”, “Multidisciplinary” and “Social Sciences, General”—were used in the study. However, classes related to biology and medicine were manually merged by the authors, which resulted in 10 broad science fields. Thus, each journal in JCR_list is mapped to exactly one of these 10 fields. In Table 1, the mapping from ESI classes to fields is displayed.
3. Standardization of the ISSNs in the JCR_list was performed. For instance, long hyphens were replaced by short hyphens.
4. For the two considered publication years, 1993 and 2013, we retrieved from Bibmet all publications published in one of these two years. However, we only used publications of the WoS document types “Article” and “Review” and appearing in the journal indexes of WoS. The ISSNs for (the journals of) the publications were standardized (cf. step 3 above). Publications such that their journals lacked an ISSN were excluded from the study.



Table 1. The mapping from ESI classes to science fields.

Field(full)	Field(abb.)	ESI class
Agricultural Sciences	AGR	Agricultural Sciences
Environment/Ecology	ENV	Environment/Ecology
Geosciences	GEO	Geosciences
Engineering	ENG	Computer Science Engineering Materials Science
Biology	BIO	Biology & Biochemistry Microbiology Molecular Biology & Genetics Plant & Animal Science
Medicine	MED	Clinical Medicine Immunology Neuroscience & Behavior Pharmacology & Toxicology Psychiatry/Psychology
Mathematics	MAT	Mathematics
Chemistry	CHE	Chemistry
Physics	PHY	Physics
Space Science	SPA	Space Science

5. For both publication years, we extracted the unique ISSNs from the two corresponding publication sets of step 4. This gave rise to two lists of ISSNs, one per year.
6. In order to have a constant journal set—with respect to the two involved publication years—JCR_list was reduced to those ISSNs that are present in both ISSN lists of step 5. Our final list includes 3,221 journals. If a constant journal set would not be used, the interpretation of the results of the study would be difficult. For instance, an increase in relative publication volume for a given country in a given science field between 1993 and 2013 might be an effect of an increase in the number of journals (in the field) in WoS for the country from the former to the latter year.
7. For each of the two publication years and for each ISSN in our final list of journals, the publications published in the year, satisfying the conditions of step 4, and appearing in the journal corresponding to the ISSN, were extracted and assigned to the science field of the journal. This gave rise to 1,085,849 publications in total. However, 14,415 publications turned out to lack addresses and were therefore excluded. Thus, the number of publications included in the study is equal to 1,071,434.
8. For the 1,071,434 publications, address data was extracted in order to obtain the contribution of countries to a publication.
9. Finally, citation counts were gathered from Bibmet.



Table 2. Journal and publication volume by publication year.

Publication year	# journals	# publications
1993	3,221	436,440
2013	3,221	634,994
Total	3,221	1,071,434

With reference to point 8 above, a publication might involve more than one country because of co-authorship. Due to this, we applied fractional counting of publications and citations. Since there are no associations in WoS between author names and addresses before the publication year 2008, address fractionalization was used. For instance, if a publication has two China addresses and three United States addresses, and if it has received 10 citations, China is assigned 2/5 and United States 3/5 of the publication, while China is assigned 2/5 x 10 = 4 and United States 3/5 x 10 = 6 citations.

In Table 2, journal and publication volume statistics are reported.

2.2 Indicators and methods

As we use fractionalization, when we refer to the number of publications and the number of citations for countries, we actually mean magnitude of publication fractions and magnitude of fractions of citations counts, respectively.

For country i , we define P_{ijy} as the number of publications of i in field j in publication year y and C_{ijy} as the number of citations to i in j during the three-year window starting with y . Let P_{jy} (C_{jy}) be the total number of publications (citations) in j in y (during the three-year window starting with y). We then define R_{Pijy} and R_{Cijy} as

$$R_{Pijy} = \frac{P_{ijy}}{P_{jy}} \times 100 \tag{1}$$

$$R_{Cijy} = \frac{C_{ijy}}{C_{jy}} \times 100 \tag{2}$$

Thus, R_{Pijy} gives the percentage of the publications of country i in field j in publication year y relative to the total number of publications in j in y , whereas R_{Cijy} gives the percentage of the citations to country i in field j during the three-year window starting with y , relative to the total number of citations in field j during the three-year window starting with y .

The disciplinary structure of the publication volume for a country i in publication year y can now be operationalized as the vector V_{Piy} , whereas the disciplinary



Research Paper

structure of the citations for country i for the three-year window starting with year y can be operationalized as the vector V_{Ciy} :

$$V_{Piy} = (R_{Pi1y}, R_{Pi2y}, \dots, R_{Piny}) \quad (3)$$

$$V_{Ciy} = (R_{Ci1y}, R_{Ci2y}, \dots, R_{Ciny}) \quad (4)$$

where $n = 10$.

The cosine measure (Salton & McGill, 1983) was applied to measure the similarity between the citation disciplinary structures for one country in different publication years (1993, 2013), and between such structures for different countries in the same publication year (2013):

$$Sim(V_{Ciy}V_{Ci'y'}) = \frac{\sum_{j=1}^{10} R_{Cijy} \times R_{Ci'jy'}}{\sqrt{\sum_{j=1}^{10} (R_{Cijy})^2} \times \sqrt{\sum_{j=1}^{10} (R_{Ci'jy'})^2}} \quad (5)$$

where i' and c' stand for a country and a year, respectively.

A cluster analysis was used to group countries based on cosine similarity values for pairs of countries. We used the average linkage, a well-known hierarchical method, for the clustering (Everitt et al., 2001). The Silhouette measure (Kaufman & Rousseeuw, 1990) was used in order to obtain the best number of clusters. This measure contrasts coherence to separation by comparing within-cluster dissimilarity to between-cluster dissimilarity. Since the measure is defined in terms of dissimilarities, we converted the cosine similarity values to corresponding dissimilarity values by subtracting a given similarity value from 1. The cluster solution with the highest overall Silhouette value turned out to have two clusters for publication volume disciplinary structure and five clusters for citation disciplinary structure.

It might be the case that there are G7 and/or BRICS countries that are more similar, with respect to the citation disciplinary structure, to countries outside the G7-BRICS group than to countries within this group. In order to have a more comprehensive view of the citation disciplinary structure patterns, we extended the set of countries used in the cluster analysis with the nine countries (apart from the G7-BRICS countries) with the largest publication volumes in 2013. Thus, 21 countries were used in the cluster analysis.

With regard to the comparison of the disciplinary structure of the publication volume and of the citations within a given country i , we are interested in the distance between the publication and citation percentages R_{Pijy} and R_{Cijy} ($j = 1, \dots, 10$). More precisely, we are interested in the distance between the vectors V_{Piy} and V_{Ciy} . The



Jeffreys-Matusita distance (Bruzzone L. et al., 1995) was applied to measure the distance for a given country i in year 2013. This measure (with respect to publication year 2013 and for country i) is defined as:

$$Dist(V_{Pi2013}, V_{Ci2013}) = \left[\sum_{j=1}^{10} \left(\sqrt{R_{Pij2013}} - \sqrt{R_{Cij2013}} \right)^2 \right]^{1/2} \quad (6)$$

Clearly, the Jeffreys-Matusita distance is similar to the well-known Euclidian distance. However, the latter is more dependent on the magnitudes of the components of the involved vectors. For instance, the Euclidian distance between two vectors with very small components, and thus with small norms, might underrate the distance between the vectors. In view of this, and in our context, we consider the Jeffreys-Matusita distance to be a better choice than Euclidian distance.

Moreover, in order to have a summary value for country i in publication year 2013 across all 10 fields, we calculated means of the elements of the vectors V_{Pi2013} and V_{Ci2013} :

$$Mean(V_{Pi2013}) = \frac{1}{10} \sum_{j=1}^{10} R_{Pij2013} \quad (7)$$

$$Mean(V_{Ci2013}) = \frac{1}{10} \sum_{j=1}^{10} R_{Cij2013} \quad (8)$$

3 Results

This section reports citation disciplinary structure results for countries in 2013 and within countries across 1993 and 2013. The section further reports the outcome of the comparison of publication volume and citation disciplinary structure within countries in 2013. Note that 21 countries, and not only the 12 G7/BRICS countries, are taken into account with regard to the first subsection (cf. the section “Data and methods”).

3.1 Citation disciplinary structure patterns for the G7 and BRICS countries in 2013

Figure 1 visualizes the outcome of the cluster analysis, which is based on the citation disciplinary structures of the G7 and BRICS countries, as well as on the corresponding structures of the nine countries, apart from the G7/BRICS countries, with the largest publication volumes in 2013. The clusters and countries are represented horizontally, whereas the 10 fields are represented vertically, with the field names located in the bottom of the figure (see Table 1 for the full names of the fields). The dendrogram for the analysis occurs to the right of the last column of the figure. Table 3 gives the $R_{Cij2013}$ values (Equation 2) across the 10 fields for the 21 considered countries for year 2013.



Research Paper

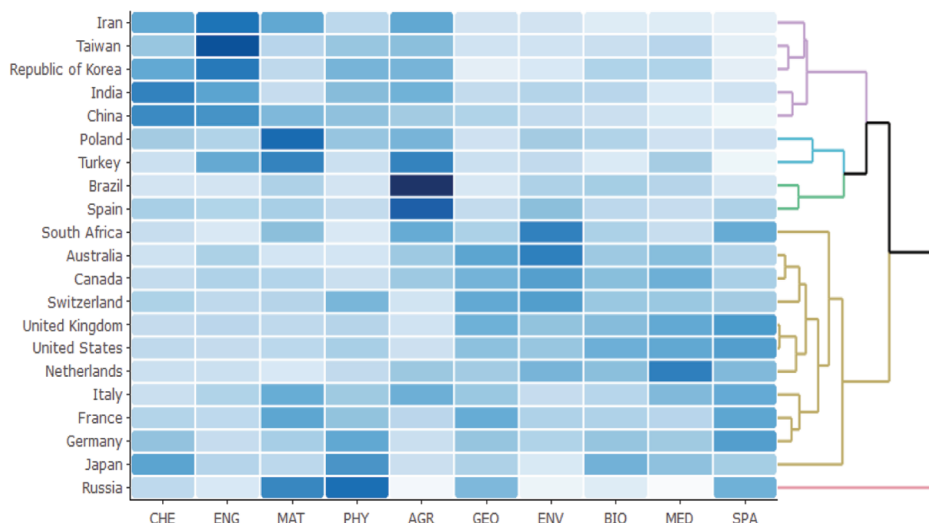


Figure 1. Cluster analysis outcome of citation disciplinary structure of 21 countries (2013).

Table 3. $R_{Cij2013}$ values for 21 countries (2013).

Field	CHE	ENG	MAT	PHY	AGR	GEO	ENV	BIO	MED	SPA
Australia	1.7	2.6	1.5	1.5	2.9	4.1	5.2	2.9	3.3	2.4
Brazil	0.9	0.9	1.5	0.9	4.4	0.8	1.5	1.6	1.4	0.8
Canada	2.2	2.7	2.6	2.0	3.1	3.9	4.6	3.5	4.0	2.8
China	24.8	23.7	17.5	16.0	14.1	12.6	10.7	9.3	6.4	3.0
France	3.7	3.3	6.1	4.7	3.4	5.8	3.8	3.8	3.5	6.1
Germany	6.2	3.9	5.3	8.0	3.7	6.1	5.0	6.1	5.7	8.6
India	4.5	3.6	1.7	2.9	3.2	1.8	2.1	2.0	1.1	1.4
Iran	1.8	2.5	1.8	1.0	1.8	0.7	0.7	0.5	0.5	0.4
Italy	2.4	3.2	4.9	3.7	4.8	3.8	2.5	3.0	4.4	5.0
Japan	5.2	3.0	2.8	5.8	2.3	3.2	1.6	4.6	4.0	3.4
Republic of Korea	3.3	4.5	1.8	3.0	3.0	0.8	1.1	2.1	2.1	0.8
Netherlands	1.1	1.1	0.8	1.3	1.8	1.7	2.2	2.0	3.2	2.1
Poland	1.0	0.9	2.1	1.1	1.3	0.6	1.0	0.9	0.6	0.6
Russia	1.0	0.6	2.3	2.6	0.2	1.6	0.3	0.5	0.1	1.7
South Africa	0.3	0.2	0.5	0.2	0.6	0.4	0.8	0.4	0.3	0.6
Spain	3.1	2.9	3.1	2.4	7.2	2.4	3.8	2.6	2.3	3.0
Switzerland	1.3	1.1	1.2	1.8	0.8	2.0	2.2	1.5	1.5	1.4
Taiwan	1.3	2.8	1.0	1.3	1.4	0.7	0.7	0.8	1.0	0.4
Turkey	0.6	1.3	1.7	0.6	1.7	0.6	0.7	0.4	0.9	0.2
United Kingdom	3.7	4.2	4.0	4.4	3.0	6.9	5.8	6.2	7.3	8.2
United States	18.9	17.3	19.6	23.1	15.0	27.9	26.4	32.5	34.9	37.5

*For each country, values in bold face are the highest two $R_{Cij2013}$ values among the 10 fields

For each country i and field j represented in Figure 1, the darkness of the corresponding grid is related to the value of $R_{Cij2013}$: the darker the color is, the larger the value is. Countries belonging to the same cluster tend to be represented by similar rows with respect to darkness levels across the fields. Thus, such countries tend to have similar citation disciplinary structures, given our operationalization of such structures.

For the 5-cluster solution, i.e. the solution with the highest overall Silhouette value, the G7 countries belong to the same cluster, together with Australia, Netherlands, South Africa and Switzerland, whereas the five BRICS countries are distributed over four groups. We observe that South Africa is the only BRICS country that belongs to the same cluster as the G7 countries. Thus, all the cluster neighbors of the other four BRICS countries are non-G7 countries. These results indicate that the G7 countries and BRICS countries are quite heterogeneous regarding citation disciplinary structure.

United States and United Kingdom are the two countries that are grouped first in the hierarchical clustering process (Figure 1), which indicates that they are similar in citation disciplinary structure. Because the countries are grouped based on cosine similarity values, and cosine-based measure is independent of size effects and only focuses on the disciplinary profiles of the citations. The data in Table 3 suggest, in more detail, why they are similar: Both countries have their highest $R_{Cij2013}$ values in Space Science and Medicine, and their smallest value in Agriculture. As is visible in Figure 1, Germany, France and Italy are also grouped early in the process. Space Science is the strongest field for those three countries in terms of citation impact. For example, for Germany, the $R_{Cij2013}$ value of Space Science is 8.6%, while the $R_{Cij2013}$ value of other fields are lower than 8% (Table 3). Apart from Space Science, Italy and France are comparatively strong in Mathematics from a citation point of view. Table 3 shows that for France and for a majority of the fields, the $R_{Cij2013}$ value is lower than 4%, while the $R_{Cij2013}$ value for Mathematics exceeds 6%. Like Germany, France and Italy, Canada Australia and Switzerland are grouped early. For these three countries, Life Sciences (Medicine, Biology) and Earth Sciences (Environment/Ecology, Geosciences) are prominent fields regarding citation impact.

For China and India, both grouped together early in the clustering process, the citation disciplinary structure is dominated by Chemistry and Engineering. For China, Chemistry and Engineering have $R_{Cij2013}$ values of about 25% and 24%, respectively, whereas the percentages in the other fields are less than 18%. Agriculture Sciences dominates for both Brazil and South Africa, even if the two countries do not belong to the same cluster until the next last step in the process has been executed, when the number of clusters is equal to 2 (Figure 2). For Brazil, the world share of Agriculture Science is 4.41%, compared to less than 2% in the other fields.



Also South Africa has a better performance in Agriculture Sciences compared to the other fields.

Note that Russia is not grouped with another country until the last step of the clustering process has been executed, when the number of clusters is equal to 1. This indicates Russia has unique citation disciplinary structure characteristics, relative to the other considered countries. Russia's academic achievements in Mathematics and Physics are globally acknowledged. This is reflected by the citation disciplinary structure of Russia, where the $R_{Cij2013}$ values of these two fields are much higher compared to the other fields.

3.2 Citation disciplinary structure within G7 and BRICS countries across 1993 and 2013

Table 4 gives similarity values, which are based on the citation disciplinary structures in 1993 and 2013, for the G7 countries and BRICS countries. A given value indicates the similarity between the 1993 and the 2013 citation disciplinary structures of the corresponding country. The closer the similarity value is to 1, the more similar are the citation disciplinary structures for a country with respect to the years 1993 and 2013.

Table 4. Similarity values based on citation disciplinary structures in 1993 and 2013.

Country	$Sim(V_{C1993}, V_{C2013})$	Country	$Sim(V_{C1993}, V_{C2013})$
Brazil	0.739	Italy	0.942
Canada	0.971	Japan	0.950
China	0.907	Russia	0.951
France	0.978	South Africa	0.963
Germany	0.969	United Kingdom	0.968
India	0.974	United States	0.982

It is clear from Table 4 that all cosine values, except the one for Brazil, are greater than 0.9. Thus, in general the citation disciplinary structures within countries changed to a small extent across the two considered years (the similarity value is equal to or greater than 0.95 for nine of the 12 countries). We can see that the similarity values of most G7 countries are slightly higher compared to a majority of BRICS countries. For most G7 countries, then, the citation disciplinary structure tend to be more stable compared to BRICS countries with regard to the periods 1993 and 2013.

China has a lower similarity value than the other countries, apart from Brazil (0.907, Table 4). Compared to the other G7 and BRICS countries, the citation disciplinary structures of Brazil and China are more dissimilar regarding the two years. In view of this, we studied the citation disciplinary structures of China and



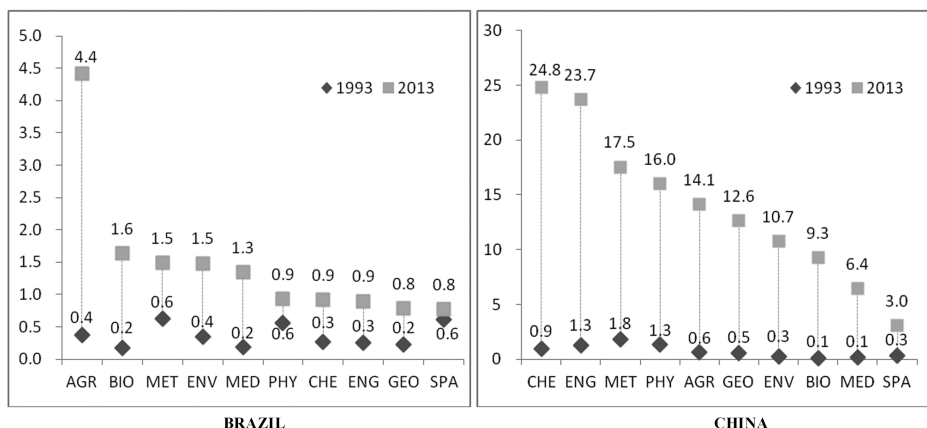


Figure 2. $R_{Cij1993}$ and $R_{Cij2013}$ values across fields for Brazil and China.

Brazil in more detail. Figure 2 displays the $R_{Cij1993}$ and $R_{Cij2013}$ values across fields for Brazil and China.

For 1993 publications, Brazil’s citation impact in Mathematics, Space Science and Physics are relatively strong, whereas the performance in Biology and Medicine is relatively poor. For 2013 publications, however, Agricultural Sciences exhibits the (by far) strongest performance, with a large increase in world share compared to 1993. Besides, the corresponding shares for Biology, Environment/Ecology and Medicine have also increased a lot. For instance, the values for Environment/Ecology are 0.4% (1993) and 1.5% (2013).

Regarding China, several fields show quite remarkable increases in world shares, especially Chemistry and Engineering. The world shares of Chemistry and Engineering increased with more than 20 percentage units from 1993 to 2013, while the shares for Biology and Medicine only increased with less than 10 percentage points. The findings indicate that the enormous upsurge of publications in the 20 years of our observation period hardly had an impact on the citation structure of the countries under analysis. The upsurge resulted in an absolute upgrade of the numbers of citations, but the disciplinary profiles are rather stable, even in transition countries like Brazil or China. However, there are changes in the profiles as the example of China shows. While the basic research fields of mathematics and physics were the largest in terms of citations in the Chinese profile at the beginning of the 1990s, they shifted towards the more application oriented fields of engineering and also chemistry (including material sciences). These changes, however, have a measurable but limited impact on the overall citation profile. There is more stability than change in the system.



3.3 Comparison of publication volume and citation disciplinary structure within G7 and BRICS countries in 2013

Jeffreys-Matusita distances between publication and citation disciplinary structure within G7 and BRICS countries and in year 2013 are reported in Table 5. Also reported are the means of the elements of corresponding vectors, which are the world shares of publication volume and citations.

Table 5. Jeffreys-Matusita distances between publication and citation disciplinary structure and vector means within G7 and BRICS countries.

country	$Dist(V_{Pi2013}, V_{Ci2013})$	$Mean(V_{Pi2013})$	$Mean(V_{Ci2013})$
United States	2.1	19.2	25.3
Russia	1.9	2.6	1.1
Brazil	1.3	2.5	1.5
India	1.2	3.6	2.4
China	1.0	14.8	13.8
United Kingdom	0.8	4.3	5.4
Japan	0.8	4.6	3.6
Germany	0.7	5.0	5.9
France	0.5	3.9	4.4
Italy	0.4	3.5	3.8
South Africa	0.4	0.6	0.4
Canada	0.3	2.9	3.2

Most G7 countries, with United States as an exception, have lower values on the Jeffreys-Matusita distance than BRICS countries, indicating a higher degree of heterogeneity between the publication and the citation disciplinary structure for the latter countries. The distance for United States is much higher compared to the other G7 countries, which indicates a substantial difference between the two types of disciplinary structures.

For the G7 countries except Japan, $Mean(V_{Ci2013})$ is larger than $Mean(V_{Pi2013})$, which indicates that on average across the fields, the citation impact performance is stronger than publication volume performance. This result is particularly striking for USA, with $Mean(V_{Ci2013})$ equal to 25.3, much higher than $Mean(V_{Pi2013})$ (19.2), and with the largest Jeffreys-Matusita distance of 2.1, indicating that USA gets much more citation impact with fewer publications. For the BRICS countries and Japan, $Mean(V_{Ci2013})$ is less than $Mean(V_{Pi2013})$: The citation impact of these countries lags behind publication volume. This is the case especially for Russia. $Mean(V_{Ci2013})$ (1.09) is considerably smaller than $Mean(V_{Pi2013})$ (2.58), which indicates that Russia has less impact with more publications.



4 Discussion and conclusions

In this study, which uses about 1 million publications and two publications years (1993 and 2013), we have dealt with citation disciplinary structure at the country level. We focused on the G7 and the BRICS countries. For the publication year 2013, cosine similarity values regarding the citation disciplinary structures of these countries (and of nine other countries) were used as input to cluster analysis. For the same publication year, the within-country Jeffreys-Matusita distance between publication and citation disciplinary structure was measured. We also obtained cosine similarity values for a given country and its citation disciplinary structures across the two publication years.

The citation disciplinary structure patterns for the G7 and BRICS countries, regarding year 2013, are quite different, which is indicated by the outcome of the cluster analysis. For most G7 countries, the academic impact of Space Sciences exceeds the impact of the other fields. Space Sciences has the highest citation world share for these countries, with Canada and Japan as exceptions. Among the BRICS countries, the citation disciplinary structure patterns are heterogeneous. China and India perform well in Chemistry and Engineering compared to the other fields, Brazil and South Africa in Agriculture Sciences, whereas Russia performs comparatively well in Mathematics and Physics.

For a given country, the pattern of the publication disciplinary structure can be assumed to be related to the S&T level of the country (Kozłowski et al., 1999; King, 2004; Kostoff et al., 2007). Based on the results of this study, citation disciplinary structure is also relevant to the S&T level of a country. USA and United Kingdom represent high S&T level countries, and they have higher world share citation impact in Space science and Medicine. Some countries with lower S&T levels, like China and India, perform better in Chemistry, Physics and Engineering, while the citation impact in Medicine and Biology lags behind. Chemistry, Physics and Engineering can be seen as the material-related fields, focusing on the studies of non-living systems, whereas Medicine and Biology can be regarded as life-related fields, which are helpful in improving the quality of life. Countries with relatively low S&T level usually give priority to the development of material-related fields, and high S&T level countries often shift their development priorities to the life-related fields.

S&T policy or strategy might also make contributions to the evolution of disciplinary structure. About one decade ago, Brazil created a new economic model named “Natural Knowledge Economy” (DEMOS, 2008). The field changes for Brazil in world share citation impact from 1993 to 2013 (from best relative performance in Mathematics, Space Science and Physics in 1993 to best relative



performance in Agricultural Sciences in 2013) might partly be explained by the establishment and maintenance of this economic model.

The citation disciplinary structure is related to a country's national conditions (apart from S&T level and policy), such as history and natural resources. Both Mathematics and Space Science harvested more citations compared to other fields for France and Italy. In these two countries, research in astronomy and mathematics started many centuries ago, leading to the high impact of contemporary research in these fields. For Canada and South Africa, countries rich in natural resources, the citation impact in Environment/Ecology is relatively high.

Based on the analysis of the citation disciplinary structure within countries across publication years, we can see, comparing 1993 to 2013, that for most G7 and BRICS countries, their citation disciplinary structures did not change to any greater extent, which reveals that the evolution of the citation disciplinary structure can be seen as a drawn-out process. This is related to the fact that it might take quite a long time to raise the S&T level of a country. The drawn-out process is also related to the circumstance that the unique national conditions of a country typically do not change rapidly.

For the comparison of publication volume and citation disciplinary structure in the year 2013, the G7 countries get more academic impact with less publications, whereas the academic impact lags behind publication volume for BRICS countries. If we regard publication volume as input, and citation impact of the publications as output, then a comparison of input and output could tell us something about scientific output efficiency or at least productivity. In this study, the G7 countries tend to have higher output efficiency than the BRICS countries. The scientific output efficiency of a country is affected by its S&T level. On the other hand, the improvement of scientific output efficiency could promote the enhancement of the S&T level. Low S&T level countries (such as the BRICS countries) usually have relatively low output efficiency. To focus on publishing publications with high scientific quality, rather than on publication volume, might be a fruitful strategy for such countries regarding S&T level enhancement.

The citation disciplinary structures of countries depend on multiple and complex factors. It is therefore difficult to completely explain the formation and change of the citation disciplinary structure of a country. This study suggests some possible causes, whereas detailed explanations might be given by future research.

Finally, we have to mention three limitations. First, the results of this study are based on the WoS database. However, in this database some fields are covered to a greater extent than others, which may affect the results for the citation discipline structure for some studied countries. In future research, we might repeat this study using another database (like Scopus) and, in that case, we would like to make



comparisons between the two outcomes. Second, the length of the citation window used in this study is three years. However, scientific disciplines differ in their citation practices. Comparison between citations across disciplines using the same citation window length may affect the citation discipline structure results for some countries. Third, the use of a constant journal set yielded that a large share of the journals covered by WoS year 2013 is ignored in the study. Thus, disciplinary structure is studied based on a quite restricted set of publications. The three mentioned limitations should be kept in mind when the results of this study are interpreted.

Acknowledgements

The authors would like to thank Zhesi Shen for his contribution to the cluster analysis of the paper.

Author Contributions

Liying Yang (yangly@mail.las.ac.cn, corresponding author) proposed the research idea, designed the research framework, and revised the manuscript. Ting Yue (yuet@mail.las.ac.cn) did the data analysis, drafted and revised the manuscript. Per Ahlgren (perahl@kth.se) collected the data, did the data statistics, drafted and revised the manuscript. Jielan Ding (dingjielan@mail.las.ac.cn) searched the literature and revised the manuscript. Shuangqing Shi (shisq@mail.las.ac.cn) processed the data, made the tables and drew the figures. Rainer Frietsch (rainer.frietsch@isi.fraunhofer.de) performed the research and revised the manuscript.

References

- Aksnes, D. W., Leeuwen, T. N. V., & Sivertsen, G. (2014). The effect of booming countries on changes in the relative specialization index (RSI) on country level. *Scientometrics*, 101(2), 1391–1401.
- Aurell, E., & Ekeberg, M. (2012). Inverse Ising inference using all the data. *Physical Review Letters*, 108(9), 090201.
- Bongioanni, I., Daraio, C., & Ruocco G. (2014). A quantitative measure to compare the disciplinary profiles of research systems and their evolution over time. *Journal of Informetrics*, 8(3), 710–727.
- Balassa, B. (2010). Trade liberalisation and “revealed” comparative advantage, *Manchester School*, 33(2), 99–123.
- Bruzzone, L., Roli F., & Serpico S. B. (1995). An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection. *IEEE Transactions on Geoscience and Remote Sensing*, 33(6), 1318–1321.
- Bound, K. (2008). Brazil: The natural knowledge-economy. DEMOS. https://www.demos.co.uk/files/Brazil_NKE_web.pdf.



- Daraio, C., Fabbri F., Gavazzi, G., Izzo, M. G., Leuzzi, L., Quaglia, G., & Ruocco, G. (2017). Assessing the interdependencies between scientific disciplinary profiles at the country level: A pseudo-likelihood approach. 2017 ISSI Conference Proceedings. China: Wuhan University, 2017, 1448–1459.
- Ding, J., Ahlgren, P., Yang, L., & Yue, T. (2018). Disciplinary structures in Nature, Science, and PNAS: journal and country levels. *Scientometrics* (3), 1–36.
- Everitt, B., Landau, S. & Leese, M. (2001). *Cluster Analysis* (Forth Edition.). London: Arnold.
- Glänzel, W., Debackere, K., & Meyer, M. (2008). “Triad” or “tetrad”? On global changes in a dynamic world. *Scientometrics*, 74(1), 71–88.
- Glänzel, W., Leta, J., & Thijs, B. (2006). Science in Brazil. Part 1: A macro-level comparative study, *Scientometrics*, 67(1), 67–86.
- Glänzel, W., & Schlemmer, B. (2007). National research profiles in a changing Europe (1983–2003): An exploratory study of sectoral characteristics in the Triple Helix. *Scientometrics*, 70(2), 267–275.
- Grupp, H. (1998). Foundations of the Economics of Innovation: Theory, measurement, and practice. *Books*, 30(4), 379–379.
- Harzing, A. W., & Giroud, A. (2014). The competitive advantage of nations: An application to academia. *Journal of Informetrics*, 8(1), 29–42.
- Julio, C. R. P., Maria, M. L. E., & Dirce, M. T. Z. (1998). Brazilian sciences and government funding at the state of Sao Paulo. *Scientometrics*, 43(2), 177–188.
- Kaufman, L. & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. DBLP, New York: Wiley.
- King, D. A. (2004). The scientific impact of nations. *Nature*, 430(6997), 311–6.
- Kozłowski, J., Radosevic, S., & Ircha, D. (1999). History matters: the inherited disciplinary structure of the post-communist science in countries of central and Eastern Europe and its restructuring. *Scientometrics*, 45(1), 137–166.
- Li, N. (2017). Evolutionary patterns of national disciplinary profiles in research: 1996–2015. *Scientometrics*, 111(1), 1–28.
- Liang, L. M., Havemann, F., Heinz, M., & Wagner-Dobler, R. (2006). Structural similarities between science growth dynamics in China and in western countries. *Scientometrics*, 66(2), 311–325.
- Radosevic, S., & Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. *Scientometrics*, 101(3), 1897–1924.
- Schulz P. A., & Manganote E. J. T. (2012). Revisiting country research profiles: learning about the scientific cultures. *Scientometrics*, 93(2), 517–531.
- Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.
- Schmoch, U., Licht, G., & Reinhard, M. (eds.) (2000). *Wissens-und technologietransfer in Deutschland*. Stuttgart: Fraunhofer IRB Verlag.
- Shibayama, S. (2011). Distribution of academic research funds: a case of Japanese national research grant. *Scientometrics*, 88(1), 43–60.
- Raan, A. F. J. V. (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung - Theorie und Praxis*(1), 20–29.



- Wu, J. (2015). Distributions of scientific funding across universities and research disciplines. *Journal of Informetrics*, 9(1), 183–196.
- Yang, L. Y., Yue, T., Ding, J. L., & Han, T. (2012). A comparison of disciplinary structure in science between the G7 and the BRIC countries by bibliometric methods. *Scientometrics*, 93(2), 497–516.
- Yue, H. J. (2008). Prediction and relative dynamic performance on field of research of China. *Studies in Science of Science*, 26(3), 530–538.
- Zhu, W. T. (2006). A study on NSFC's interdisciplinary research funding system. *Bulletin of National Natural Science Foundation of China*. 20(3), 184–189.



This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

