InfoBrief
NCSES National Center for Science and Engineering Statistics

October 2017 ■ NSF 18-300

Science and Engineering Publication Output Trends: 2014 Shows Rise of Developing Country Output while Developed Countries Dominate Highly Cited Publications

by Karen E. White, Carol Robbins, Beethika Khan, and Christina Freyman[1]

Publication output in peer-reviewed science and engineering (S&E) journals, books, and conference proceedings serves as an indicator of scientific research activity. New data show that S&E publication output has continued to grow, reaching 2.3 million globally in 2014, with the United States and China being the two largest producers in 2014 (19% and 17% of the world total). When counted together the European Union countries produced more S&E publications than the United States or China. Globally, S&E publications output grew at an average annual rate of 6% between 2004 and 2014, the most recent 10 years for which data are available. The growth rate varied substantially across world regions. For example, in Iran, China, and India, growth rates were 22%, 14%, and 14%, respectively, compared with 3% in the United States.

This InfoBrief updates information on the international bibliometric trends presented in *Science and Engineering Indicators: 2016*. The bibliometric data summarizes the output volume of the world's peer-reviewed journals, books, and conference proceedings (publications), providing insights into the development of scientific and technological capabilities around the globe. Data on publications and citations were obtained by the National Center for Science and Engineering Statistics within the National Science Foundation from Elsevier's Scopus bibliometric database, which provides article-level information on peer-reviewed documents and conference proceedings.[2]

Comparisons of Overall Scientific Publication Output

Publication Output, by Country

In 2014, researchers in the United States and China were the world's largest producers of S&E publications of any single country, producing 19% and 17%, respectively, of the 2.3 million world total (table 1). The next closest countries by count of research publications were Germany, India, Japan, and the United Kingdom, who each produced 4% to 5% of world output (approximately 100,000 publications each). The nine other countries that make up the top 15 countries by output (France, Italy, South Korea, Canada, Spain, Brazil, Australia, Russia and Iran) each produced between 2% and 3% (between 37,000 and 74,000 publications).

Among the 15 countries with the highest S&E publication output in 2014, the growth rates varied over the past decade and smaller producers typically experienced large growth rates from small bases. The average U.S. annual growth of 3% was below the world's average annual growth (6%) and lower than other top producers, except for Japan (1%) (table 1). Over the past decade, developed countries that produce over 50,000 papers or more annually experienced relatively low average annual growth rates, for example, Germany (4%), the United Kingdom (3%), and France (3%).

Publication Output, by Field

The distribution of S&E publication output by field provides an indication of the priority and emphasis of scientific research in different geographic locations (table 2). On a global scale, almost 40% of publications were in journals classified as covering the biological sciences, medical sciences, or other life sciences.

TABLE 1. Science and engineering articles in all fields, by region, country, or economy: 2004 and 2014

Rank	Region, country, or economy	2004	2014	Average annual growth rate (%)	2014 world total (%)	2014 cumulative world total (%)
na	World	1,272,362	2,290,294	6.1	100.0	na
1	United States	336,194	431,623	2.5	18.8	18.8
2	China	110,388	395,588	13.6	17.3	36.1
3	Germany	72,177	107,747	4.1	4.7	40.8
4	India	28,752	106,574	14.0	4.7	45.5
5	Japan	95,999	103,793	0.8	4.5	50.0
6	United Kingdom	75,119	101,536	3.1	4.4	54.4
7	France	53,375	74,269	3.4	3.2	57.7
8	Italy	42,647	70,453	5.1	3.1	60.8
9	South Korea	27,029	63,748	9.0	2.8	63.5
10	Canada	40,624	60,916	4.1	2.7	66.2
11	Spain	30,977	56,604	6.2	2.5	68.7
12	Brazil	18,814	53,152	10.9	2.3	71.0
13	Australia	26,277	52,269	7.1	2.3	73.3
14	Russia	26,869	43,487	4.9	1.9	75.2
15	Iran	4,952	36,539	22.1	1.6	76.8

na = not applicable.

NOTES: Articles refer to publications from a selection of peer-reviewed journals, books, and conference proceedings in science and engineering fields from Scopus. Articles are classified by their year of publication and are assigned to a region, country, or economy on the basis of the institutional addresses listed in the article. The region, country, or economy shown each produced 36,000 publications or more in 2014. Rankings are based on the 2014 total. Articles are credited on a fractional-count basis (i.e., for articles from multiple countries, each country receives fractional credit on the basis of the proportion of its participating authors). Data are not directly comparable to *Science and Engineering Indicators: 2016;* see data sources and methodology section on data filters. Supporting tables available upon request.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics; SRI International; Science-Metrix; Elsevier, Scopus abstract and citation database (www.scopus.com), accessed December 2016.

Engineering publications made up another 17% of the global output.

The S&E publication portfolios of five major producers—the United States, the European Union, China, Japan, and India—displayed distinct differences by field. In 2014, almost half (49%) of the United States' publications were focused on biological sciences, medical sciences, or other life sciences, compared with 39% for the world at large. Researchers in the United States also produced a higher proportion of publications in psychology (4%) and social sciences (7%) than did the world.

As in the United States, most of the S&E publications in the European Union and Japan were focused on the fields of biological sciences, medical sciences, and other life sciences; these three fields together accounted for 41% of the European Union's publications and 42% of Japan's publications. However, compared with the United States and the world, Japan had larger publication shares in the fields of chemistry (10%) and physics (14%). China's research portfolio showed a different pattern than the one for the United States and the world, with a greater focus on engineering (28%) and chemistry (13%).

Another notable exception to the global averages was India's portfolio, which had a high concentration of publications in computer sciences (14%). India's proportion of research by field was also above the world average in chemistry (13%), biological sciences (18%), and engineering (18%).

International Collaboration

S&E research has steadily become more global over the past decade. The percentage of worldwide publications produced with international collaboration—that is, by authors with institutional addresses from at least two countries—rose from 30% to 35% between 2004 and 2014 (figure 1). The worldwide trend is seen in U.S. authors' international collaborations. Thirty-four percent of authors affiliated with U.S. institutions participated in international collaborations in 2014, with the U.S. community increasing its international collaboration rate over the 2004–14 period by 10 percentage points.

U.S. authors collaborated most frequently with authors from China, the second-largest producer of S&E

TABLE 2. Science and engineering research portfolios of selected region, country, or economy, by field: 2014

(Percent)

Field	World	United States	EU	China	Japan	India
All articles (number)	2,290,294	431,623	638,834	395,588	103,793	106,574
Agricultural sciences	2.2	1.2	1.9	2.1	1.6	3.1
Astronomy	0.6	0.8	0.9	0.2	0.5	0.4
Biological sciences	15.8	18.6	15.2	13.9	14.6	18.1
Chemistry	8.1	5.4	6.7	12.5	9.5	12.5
Computer sciences	8.9	6.5	9.5	9.7	7.9	14.2
Engineering	17.1	11.7	14.0	27.8	17.2	18.4
Geosciences	5.6	5.1	5.3	6.9	3.7	4.5
Mathematics	2.6	2.1	2.7	2.7	1.8	2.0
Medical sciences	22.1	28.0	24.1	12.4	27.5	15.9
Other life sciences	1.2	2.3	1.2	0.2	0.3	0.2
Physics	9.0	7.5	8.8	10.6	13.6	8.5
Psychology	1.7	3.7	2.1	0.2	0.5	0.2
Social sciences	5.1	7.2	7.5	0.8	1.2	1.9

EU = European Union.

NOTES: Articles refer to publications from a selection of peer-reviewed journals, books, and conference proceedings in science and engineering fields from Scopus. Articles are classified by their year of publication and are assigned to a region, country, or economy on the basis of the institutional addresses listed in the article. Articles are credited on a fractional-count basis (i.e., for articles from multiple countries, each country receives fractional credit on the basis of the proportion of its participating authors). Percentages may not add to 100% because of rounding. Data are not directly comparable to *Science and Engineering Indicators: 2016;* see data sources and methodology section on data filters. Supporting tables available upon request.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics; SRI International; Science-Metrix; Elsevier, Scopus abstract and citation database (www.scopus.com), accessed December 2016.

publications (table 3). Researchers in China accounted for about 20% of U.S. internationally coauthored publications in 2014. Other substantial sources of coauthors for U.S. authors include the United Kingdom, Germany, and Canada, each accounting for between 11% and 13% of U.S. internationally coauthored publications.

Data on international collaboration of other countries indicate that publications from authors from South Korea, China, and Canada are notable for having high collaboration rates with U.S. authors (49%, 46%, and 44%, respectively) (table 3).

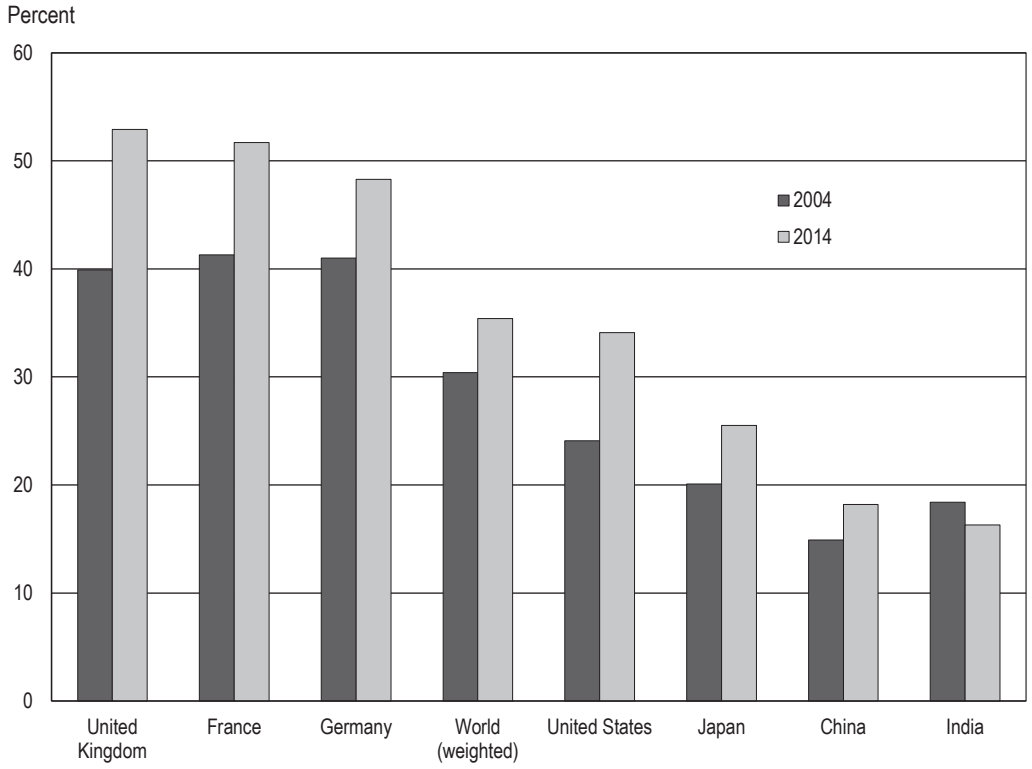
Assessments of Scientific Impact

Publication data can be used to indicate scientific impact by counting how many times an article is cited in another journal article, conference proceeding or book. Those with more citations can be said to be more impactful to the scientific discipline. Within the cited publications there is a small subset that receives a high number of citations—highly cited publications (HCP).

HCPs are publications that are most frequently cited within other researchers' papers, conference proceedings and books; for this study, the top 1% most cited publications were selected. This indicator is frequently used to examine the scientific influence of a country's research entities by measuring how many high-impact publications are produced by a country's research entities relative to their expected contribution to the world's leading research.

To create the HCP score we first created a dataset of the top 1% most cited publications in the world for each year. If each country's publication output within the top 1% pool was cited in an equal proportion to their output in the overall pool, then each country would have an HCP of 1. That is, a country with an HCP near 1 contributes close to their expected share of highly cited publications, a country with an HCP greater than 1 contributes more than their expected share, and a country with an HCP below 1 contributes less than their expected share. For example, assume the world output was 10,000 articles and there were two countries, with country *x* producing 7,000 articles and country y producing 3,000 articles. If both countries had the same impact in the citation records, then country *x* would have 70 highly cited articles and country *y* would have 30 highly cited articles in the top 100 most cited articles in the world. Each country would have an HCP of 1. The scores would be different if one of the countries produced a higher proportion of the highly cited articles. For example, if country *y* produced 50 of the most

FIGURE 1. Science and engineering articles internationally coauthored, by selected region, country, or economy: 2004 and 2014



NOTES: Articles refer to publications from a selection of peer-reviewed journals, books, and conference proceedings in science and engineering fields from Scopus. Articles are classified by their year of publication and are assigned to a region, country, or economy on the basis of the institutional addresses listed in the article. Articles are credited on a whole-count basis (i.e., one count for each collaborating institution within the region, country, or economy). Data are not directly comparable to *Science and Engineering Indicators: 2016;* see data sources and methodology section on data filters. Supporting tables available upon request.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics; SRI International; Science-Metrix; Elsevier, Scopus abstract and citation database (www.scopus.com), accessed December 2016.

highly cited articles, then their HCP score would be 1.7 (score is calculated by taking the country's number of highly cited articles (50) divided by the country's expected share (30)).

Authors associated with U.S. institutions contributed nearly twice the expected volume of highly cited publications in 2013 (HCP score of 1.9) (figure 2). China's publication citations met its expected share, and Japan and India were each below their expected levels in 2013 (HCP scores of 0.8 and 0.6, respectively).

HCP scores have changed over time with some exceptions. United States authors slightly increased their HCP score from 1.8 in 2003 to 1.9 in 2013. Contributions from China to the top 1% most cited publications rose in the same decade, with HCP scores of 0.6 in 2003 to 1.0 in 2013. In 2003, Indian publications were in a similar position to China's; yet India's HCP has remained relatively flat while China's rose. HCP scores for Japanese publications have increased over the past 10 years, from 0.6 in 2003 to 0.8 in 2013. The HCP score for Chinese publications roughly equaled that of the Japanese in 2011; however, China's score surpassed Japan's in 2012.

Several European countries have been particularly notable in leading the European Union's increase from 1.0 in 2003 to 1.3 in 2013. For publications released in 2013, those by Swiss authors were within the global highly cited research nearly three times more often than would be expected (HCP score of 2.8). Publications from the Netherlands and Sweden also made very strong contributions (HCP scores of 2.6 and 2.2, respectively).

TABLE 3. International coauthorship of science and engineering articles with the United States, by selected region, country, or economy: 2014

(Percent)

Region, country, or economy	U.S. share of region's, country's, or economy's international articles	Region's, country's, or economy's share of U.S. international articles
World	39.3	na
China	45.9	19.8
United Kingdom	29.5	13.0
Germany	28.7	11.4
Canada	44.3	10.5
France	25.1	7.7
Italy	29.0	6.6
Australia	28.9	5.9
Japan	32.7	5.5
South Korea	49.2	5.2
Spain	24.8	4.9
Netherlands	29.9	4.7
Switzerland	30.1	4.2
India	32.8	3.4
Brazil	34.8	3.4
Sweden	27.6	3.0

na = not applicable.

NOTES: Articles refer to publications from a selection of peer-reviewed journals, books, and conference proceedings in science and engineering fields from Scopus. Articles are classified by their year of publication and are assigned to a region, country, or economy on the basis of the institutional addresses listed in the article. Articles are credited on a whole-count basis (i.e., one count for each collaborating institution within the region, country, or economy). Data are not directly comparable to *Science and Engineering Indicators: 2016;* see data sources and methodology section on data filters. Supporting tables available upon request.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics; SRI International; Science-Metrix; Elsevier, Scopus abstract and citation database (www.scopus.com), accessed December 2016.

Data Sources and Limitations

The counts, coauthorships, and citations discussed in this section are derived from information about research materials in S&E fields published in peer-reviewed scientific and technical journals, books, and conference proceedings (publications). This information, termed metadata, includes title, publication and journal information, and author name collected in Elsevier's Scopus database.

In preparing the dataset as a basis for bibliometric evaluation, a set of filters has been applied to the S&E titles covered in the Scopus database. The goal of the bibliometric data analysis presented in this report is to measure only publications containing robust and novel research. Bibliometric experts have increasingly noted that there is a tide of low-quality publications that lack substantive peer review.[3] The following two publication sets were removed from the Scopus database to exclude low-quality publications from the bibliometric data included in this report.

- Free online journals and proceedings (called open access) flagged by the Directory of Open Access Journals (DOAJ)[4] for failing to adhere to their list of best practices or as being suspected of editorial misconduct.[5]
- Elsevier's own analysis of publication quality produced a list of journals and conferences for removal beginning in 2014. To create a time series from Scopus, the titles were removed retroactively for all publication years.[6]

Use of these two data filters means that, while broadly consistent, the data presented in this InfoBrief are not directly comparable to the data presented in *Science and Engineering Indicators: 2016.* In addition, Scopus is a dynamic database where publications, especially conference proceedings, may enter the database with some delay. The upcoming *Science and Engineering Indicators: 2018* will update the bibliometric data through 2016. Additional data and supporting tables are available from the authors upon request.

Methodology

Number of Publications, Using Full Counting and Fractional Counting

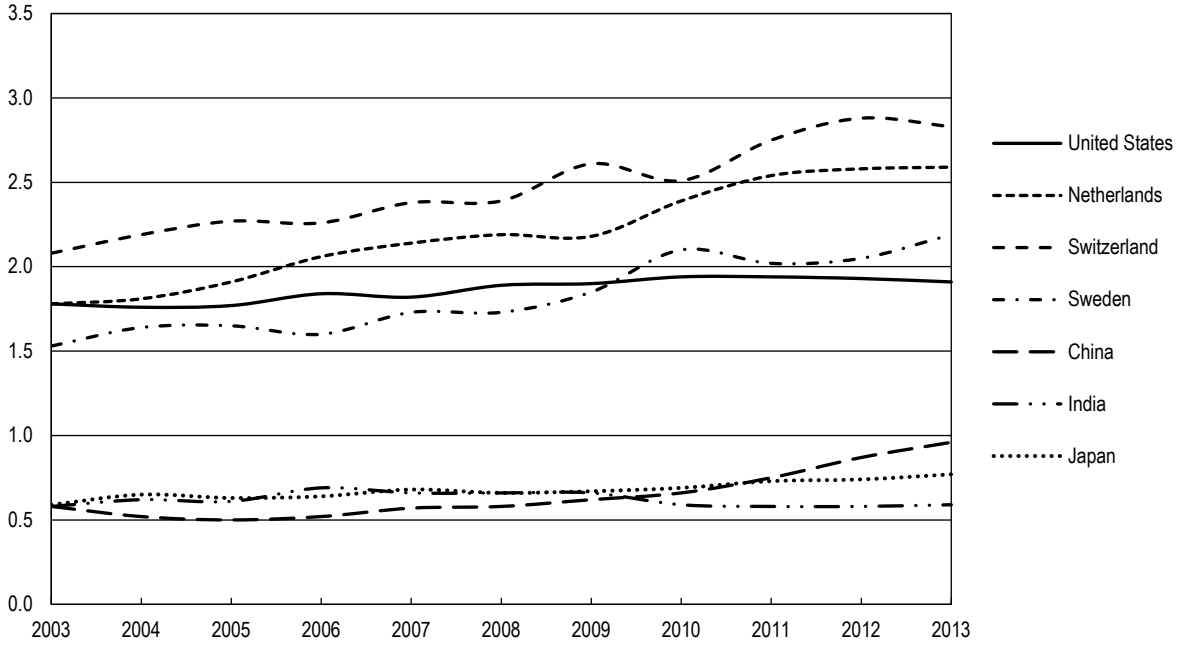
Publication counts are the number of peer-reviewed publications produced by a given country, region, economy, or institutional sector. Publications coauthored by multiple countries or institutional sectors are counted in two ways. Fractional counting divides the publication count by the proportion of each of the countries or institutional coauthors named on the publication. Fractional counting enables the counts to sum up to the number of total publications (tables 1 and 2). Whole counting (also called full or integer counting) assigns one count to each country or institutional sector involved in coauthoring the publication, irrespective of their proportionate involvement in authorship (table 3 and figures 1 and 2).

Timeliness of the Citation-Based Measures

For all citation-based measures, a certain amount of time must be allowed for the published work to have an impact on subsequent research because of the delay between the appearance of a publication and its being read, understood, and taken up in subse-

FIGURE 2. Science and engineering articles that are in the top 1% of cited articles, by selected region, country, or economy: 2003–13

Percent share



EU = European Union.

NOTES: Highly cited publications are those that are most frequently cited by other researchers in their publications; for this study, the top 1% most cited publications were selected, relative to all the country's publication in that period and field. The share of articles for a country in the world's top 1% of cited articles is also termed the highly cited publication (HCP) score. It is computed as follows: $Sx = HCPx/Px$, where Sx is the share of output from country x in the top 1% most cited articles; $HCPx$ is the number of articles from country x that are among the top 1% most cited articles in the world; and Px is the total number of papers from country x in the database that were published in 2013 or earlier. Citations are presented for the year of publication, showing the counts of subsequent citations from peer-reviewed literature. At least 3 years of data following publication are needed for a meaningful measure. Articles that cannot be classified by country or field are excluded. Articles are classified by their year of publication and are assigned to a region, country, or economy on the basis of the institutional addresses listed in the article. The world average stands at 1.00% for each period and field. Data are not directly comparable to *Science and Engineering Indicators: 2016;* see data sources and methodology section on data filters. Supporting tables available upon request.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics; SRI International; Science-Metrix; Elsevier, Scopus abstract and citation database (www.scopus.com), accessed December 2016.

quent research. Normally, a window of at least 2 years is allowed, although allowing a window of 3 years generally facilitates measurements that more robustly reflect long-term trends; the 3-year window has been applied here. Accordingly, impact assessments for this InfoBrief cover publications appearing in 2013 or earlier. However, because of the need for timeliness, citation data for the most recent year (i.e., 2013) are based on individual citation windows ranging from 24 to 36 months, depending on the month in which each publication was released.

Notes

1. Karen E. White (kewhite@nsf.gov; 703-292-4344), Carol Robbins (crobbins@nsf.gov; 703-292-7801), and Beethika Khan (bkhan@nsf.gov; 703-292-4669) are with the Science and Engineering Indicators Program, National Center for Science and Engineering Statistics, National Science Foundation, 2415 Eisenhower Avenue, Suite W14200, Alexandria, VA 22314. Christina Freyman is with SRI International, Arlington, VA.

2. More information on the selection of documents can be found at http://www.elsevier.com/online-tools/scopus/content-overview and https://www.elsevier.com/solutions/scopus/content/content-policy-and-selection.

3. For discussions of journal requirements for robust and novel submissions see http://www.nature.com/authors/policies/peer_review.html. Articles on low-quality publications include https://www.nytimes.com/2016/12/29/upshot/fake-academe-looking-much-like-the-real-thing.html?_r=0, http://www.nytimes.com/2013/04/08/health/for-scientists-an-exploding-world-of-pseudo-academia.html, http://science.sciencemag.org/content/342/6154/60.full, and http://www.nature.com/news/predatory-publishers-are-corrupting-open-access-1.11385.

4. The DOAJ list of excluded journals is available at https://docs.google.com/spreadsheets/d/183mRBRqs2jOyP0qZWXN8dUd02D4vL0Mov_kgYF-8HORM/edit.

5. Note that DOAJ also flags serials that are no longer available in open access (OA); although an important and evolving phenomenon in the research landscape, OA status is not associated here with any specific demarcation of quality—low or high—and thus the titles flagged by DOAJ for OA-related reasons alone will not be filtered out of the database for this InfoBrief.

6. Elsevier's principles of quality can be found at https://www.elsevier.com/solutions/scopus/content/content-policy-and-selection and https://doaj.org/bestpractice. In 2014, during its periodic re-evaluation of items flagged for follow-up, Elsevier's Content Selection and Advisory Board elected to remove 42 titles as of 2014. In the InfoBrief database these 42 titles are retroactively removed to create a valid time series for bibliometric analysis, even though Elsevier does not claim that these titles were necessarily of low quality before 2014.