

November 2021

Research Impact, Research Output, and the Role of International Collaboration

CSET Data Brief



AUTHORS
Autumn Toney
Melissa Flagg

Introduction

Over the last two decades, global policy and national strategy have encouraged international research collaborations as a way to tackle massive research challenges surrounding emerging technologies.¹ However, due to increasing political concerns over research collaborations with China, recent studies have analyzed how these collaborations affect research output and impact, with a particular focus on collaborations between the United States and China.²

Studies to clarify which countries lead in scientific research output and which countries contribute the highest impact work typically find that while both China and the United States are leaders in research publication output, the United States leads in high-impact research.³ However, most studies focus on one specific country's collaborations in a narrow area of research during a short period of time.⁴ This data brief addresses the question of how international collaboration relates to research publication impact and output across all fields of research for the 10 countries with the highest publication output from 2010 to 2019, giving a much more comprehensive analysis of output and impact across the major science and technology producing nations.

Takeaways:

- Analyzing collaborative research impact and output together provides a more nuanced perspective into who is “leading” in scientific research.
- Countries collaborating with China experience a disproportionate decline in research publication output relative to China when collaborations with China are removed. However, China experiences a significant decrease in highly cited research publication output if international collaborations are removed.
- Generally, publications with four or more country collaborators are higher impact, meaning more highly cited, than publications with two or three country collaborators.

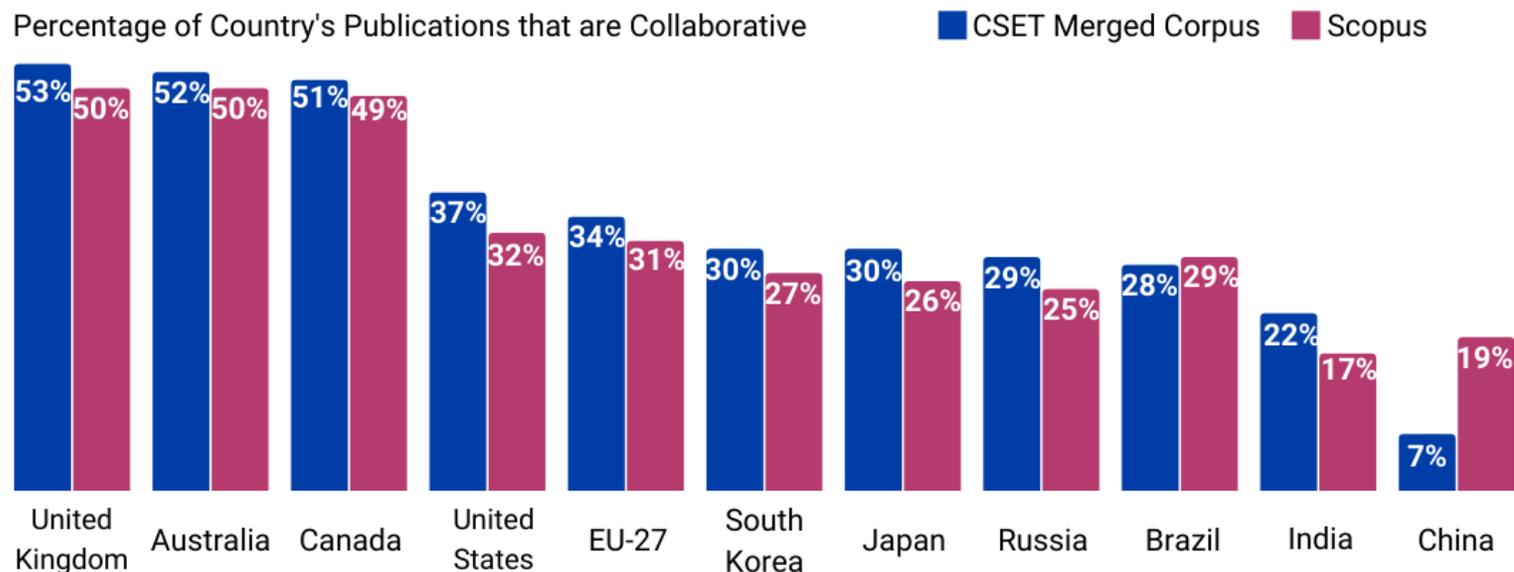
Publication Data

In a recent CSET policy brief, *Research Security, Collaboration, and the Changing Map of Global R&D*, we analyzed research publications from the Scopus database published between 2010 and 2019, and used this analysis to create a global map of such research.⁵ We ensure comparability in this data brief by using research publications published in the same time period, 2010–2019, that as of February 2021 were linked in CSET’s Map of Science, which is generated from the CSET Merged Corpus of scholarly literature from Dimensions, Web of Science, Microsoft Academic Graph, China National Knowledge Infrastructure (CNKI), arXiv, and Papers With Code.⁶

Because our analysis focuses on the effects of international collaboration, we filter for research publications that have country affiliation data. Country affiliations are assigned to a publication based on the author’s organization address. If a publication has multiple authors from different countries it will be assigned multiple countries. In this way, we can identify publications with international collaborations by checking if there are more than one distinct country listed. As with our earlier policy brief, we treat the 27 countries in the European Union as one entity (EU-27), because the EU has significant funding programs, like the Framework Programmes for Research and Innovation, and countries in the EU collaborate at high rates with one another.⁷ We consider a publication with more than one distinct country assigned to it as a collaborative publication.⁸

When we compared international collaboration rates for the top 10 countries by publication output in the Scopus dataset and CSET’s Merged Corpus, we found that all countries have approximately equivalent international collaboration rates except for China (shown in Figure 1). While China has a low international collaboration rate in the Scopus dataset, with 19 percent, it drops to 7 percent in CSET’s Merged Corpus. Because CSET’s Merged Corpus contains publication data from CNKI, we observe more Chinese publications solely with Chinese authors (more details in the Appendix).

Figure 1. Comparison of the percentages of collaborative publications from the top 11* countries by research publication output (2010–2019).



*Brazil has the 10th highest publication count in the CSET merged corpus and the 11th highest publication count in Scopus. Russia has the 11th highest publication count in CSET merged corpus and the 10th highest publication count in Scopus. Therefore, both Brazil and Russia are included in the figure.

Source: CSET Merged Corpus and Scopus

In addition to analyzing international collaboration publication rates, we explore how collaboration correlates with high-impact research. While high-impact research can be measured in varying ways, we use citations as a standard indicator of impact. Specifically, we compare the citation percentiles of publications by research field and year to account for variance in citation patterns from different areas of research, as well as to ensure that publications are only being assigned a citation ranking based on other publication citation values from the same publication year.⁹ In this analysis, we consider publications in and above the 90th percentile to be highly cited.¹⁰

Findings

We explore global trends in collaborative publications from the top 10 countries with the highest publication output using CSET's Merged Corpus. Table 1 provides the total number of publications that each country produced during 2010–2019. For direct country comparisons in this study, we provide the rest of our results in percentages. However, it is important to note the scale difference in publication output between China and the rest of the top 10 countries. Out of all the publications studied in this paper, China was listed on 43 percent of them, with the EU-27 and the United States in second and third place respectively, both being listed on 18 percent of the publications.

Table 1. Number of total publications from 2010–2019 by country

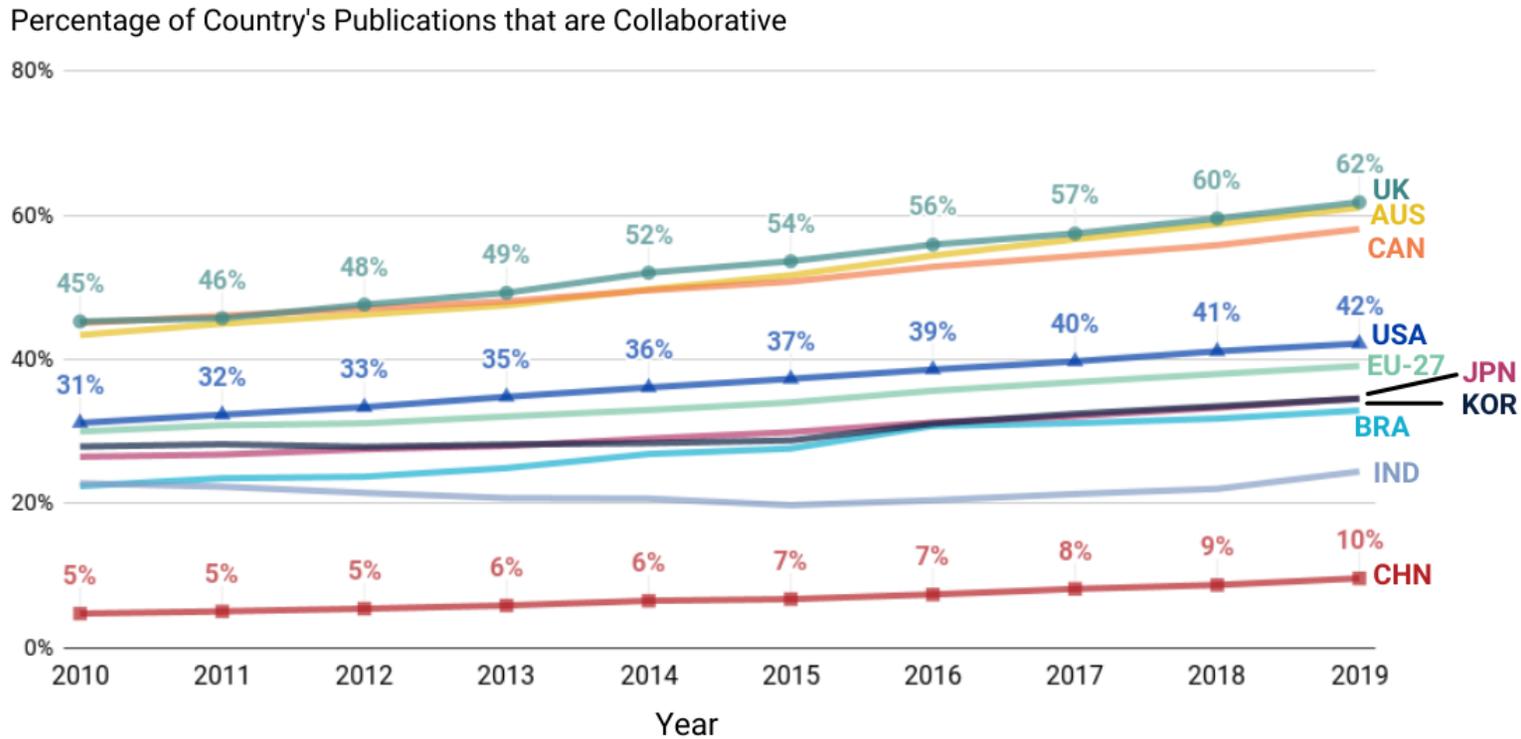
Country	Number of Publications	Percentage of Total Publications	Number of Collaborative Publications	Percentage of Total Collaborative Publications	Percentage of Publications with Collaboration
China	17,845,080	43%	1,227,244	12%	7%
EU-27	7,709,514	18%	2,637,848	26%	34%
United States	7,452,698	18%	2,743,435	27%	37%
United Kingdom	2,202,946	5%	1,173,332	12%	53%
Japan	1,406,559	3%	421,239	4%	30%
India	1,332,748	3%	287,456	3%	22%
Canada	1,141,781	3%	583,236	6%	51%
Australia	1,023,147	2%	534,824	5%	52%
South Korea	851,045	2%	257,026	3%	30%
Brazil	835,499	2%	234,884	2%	28%
Total	41,801,017		10,100,524		

Source: CSET Merged Corpus.

Which countries are the most collaborative?

First, we analyzed the trends in international collaboration rates from research published between 2010 and 2019, as shown in Figure 2. We observed that China maintains the lowest rate and the United Kingdom, Australia, and Canada maintain the highest rates of international collaboration throughout the 10-year period of analysis. Overall, the English-speaking countries (United Kingdom, Australia, Canada, and United States) lead with the highest rates of international collaboration, while all countries have an increasing rate of collaboration.

Figure 2. Research Publication International Collaboration Rates by Country (2010–2019).



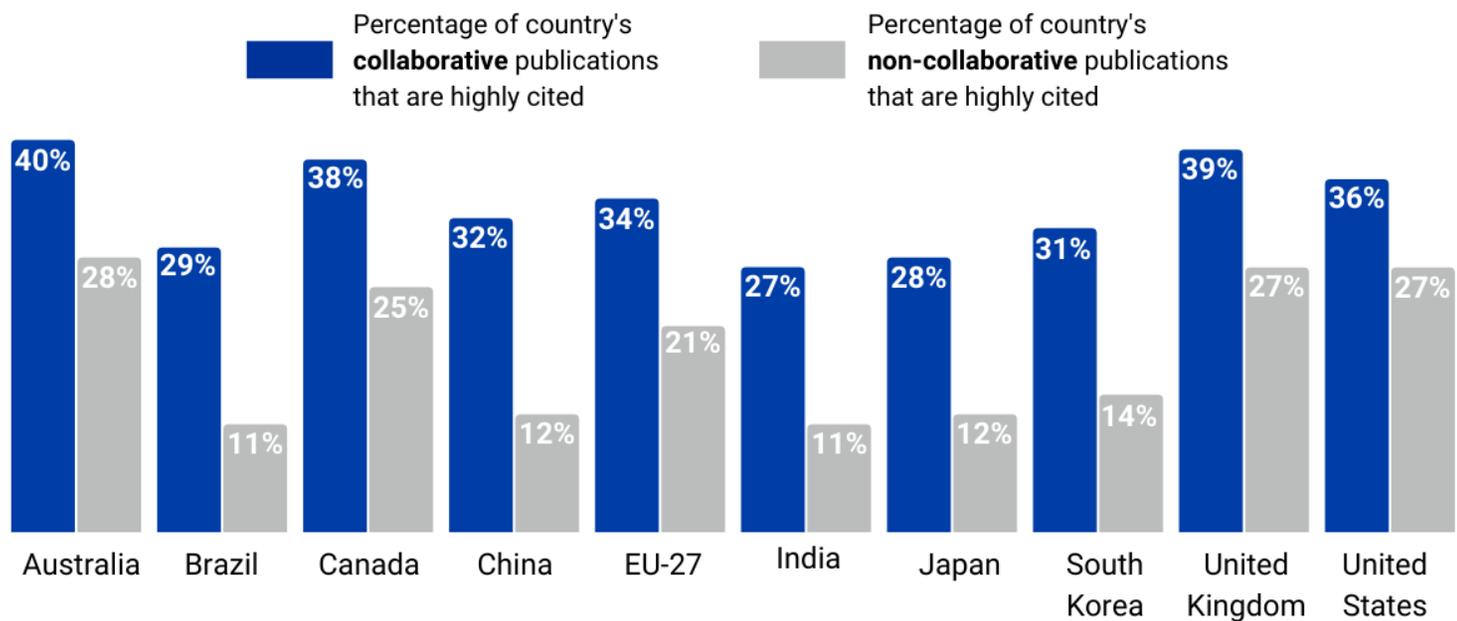
Source: CSET Merged Corpus.

In general, the collaboration rate ordering (from highest to lowest) of the United Kingdom, Australia, Canada, the United States, EU-27, Japan, South Korea, Brazil, India, and China remains consistent over time. The only order shift we observe is between South Korea and Japan. In 2010, Japan's collaboration rate was slightly below South Korea's, and by 2019 Japan's collaboration rate slightly overtook South Korea's.

How does international collaboration correlate with research impact?

Here, we take a closer look at the difference in research publication impact between publications with international collaboration ("collaborative") and those without ("non-collaborative").¹¹ In Figure 3 we see that for all countries a higher percentage of collaborative publications are highly cited compared to non-collaborative publications. Australia, Canada, and the United Kingdom lead in having both the highest percentage of collaborative and non-collaborative publications that are highly cited. While Brazil, China, India, Japan, and South Korea have comparable percentages of highly cited publications (both collaborative and non-collaborative), India falls in last place for highly cited papers in both collaborative and non-collaborative research with 27 percent and 11 percent, respectively.

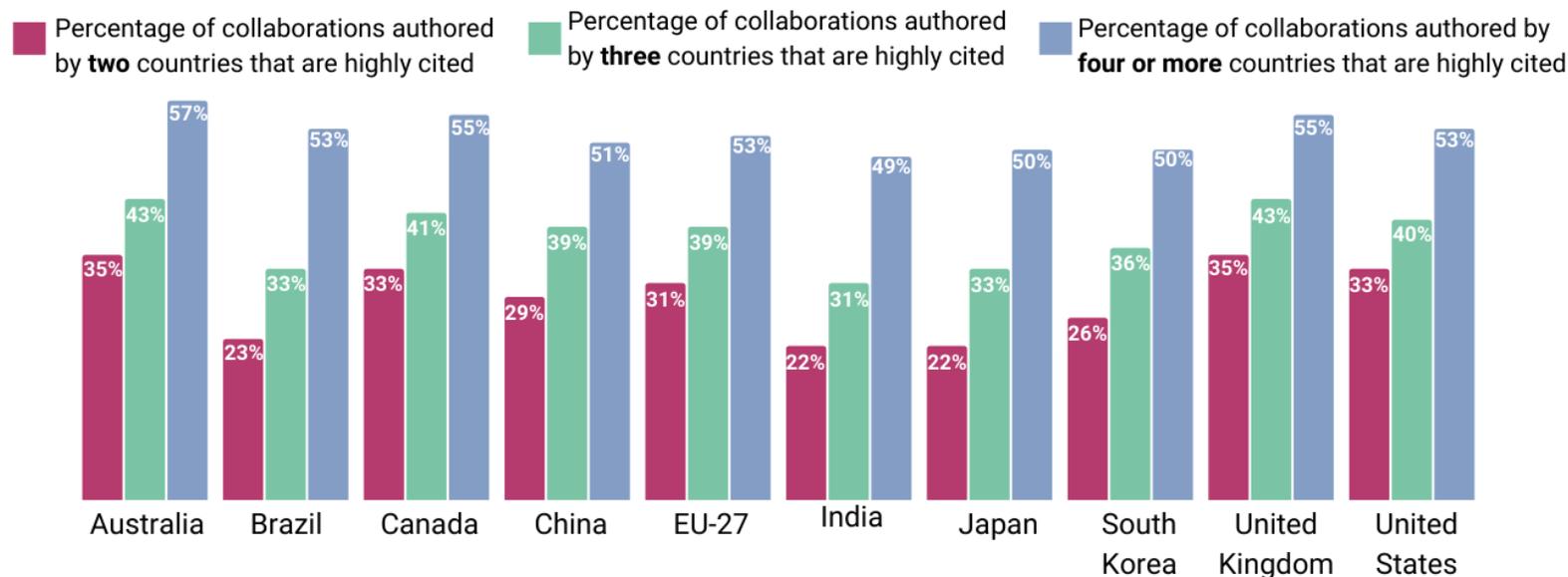
Figure 3. Percentage of highly cited collaborative and non-collaborative publications by country (2010–2019).



Source: CSET Merged Corpus.

We further investigate the effect of international collaboration on highly cited research by breaking down the collaboration size—the number of different countries on each collaborative publication. Figure 4 displays the percentage of publications with different collaboration sizes (two countries, three countries, and four or more countries) that are highly cited. Again, we see that Australia, Canada, and the United Kingdom lead in percentage of collaborative publications that are highly cited, regardless of the collaboration size.

Figure 4. Percentages of highly cited publications from different collaboration sizes by country.

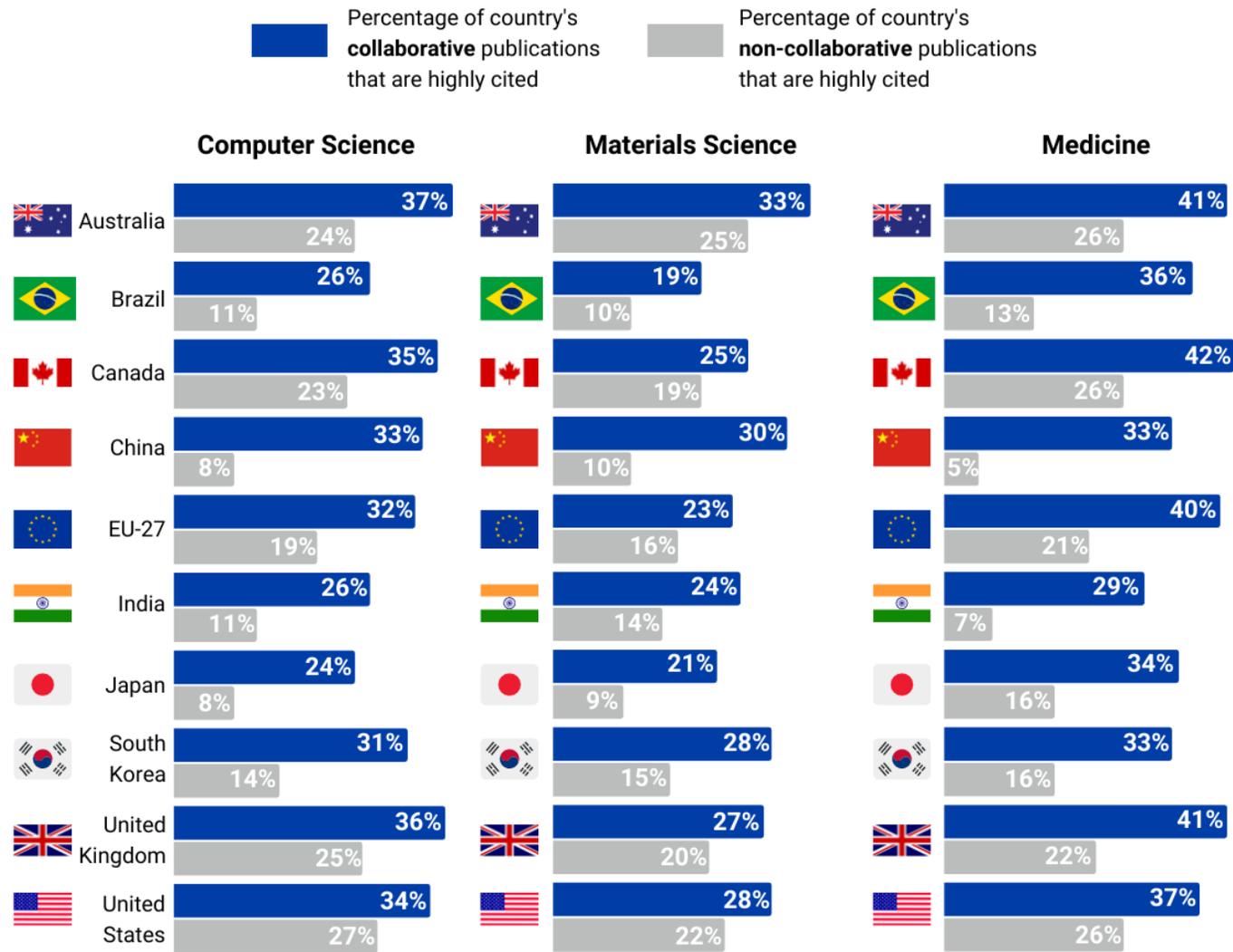


Source: CSET Merged Corpus.

There is a clear trend for all countries: publications with four or more country collaborators have the highest rate of high-impact publications, with three-country collaborations and two-country collaborations falling in second and third place respectively. Brazil has the biggest difference in percentages between two-country and four-or-more country collaborations, with a 30 percentage point increase. These results highlight the importance of international collaboration for highly cited research across all of science.

While Figures 3 and 4 above show results for collaborative and non-collaborative publications that are highly cited from all research fields, we investigate if there are differences in these results if we analyze research fields individually. Figure 5 displays the percentage of collaborative and non-collaborative publications that are highly cited from computer science, materials science, and medicine.

Figure 5. Percentages of highly cited publications from collaborative and non-collaborative publications by country for computer science, materials science, and medicine.



Source: CSET Merged Corpus

Out of the three fields selected, we find that for each country, medicine has the highest percentages of collaborative publications that are highly cited. With the exception of China, India, the United Kingdom, and the United States, medicine also has the highest percentages of non-collaborative publications that are highly cited. For China and India, materials science has the highest percentage of non-collaborative publications. However, for the United States and the United Kingdom, computer science has the highest percentage of non-collaborative publications.

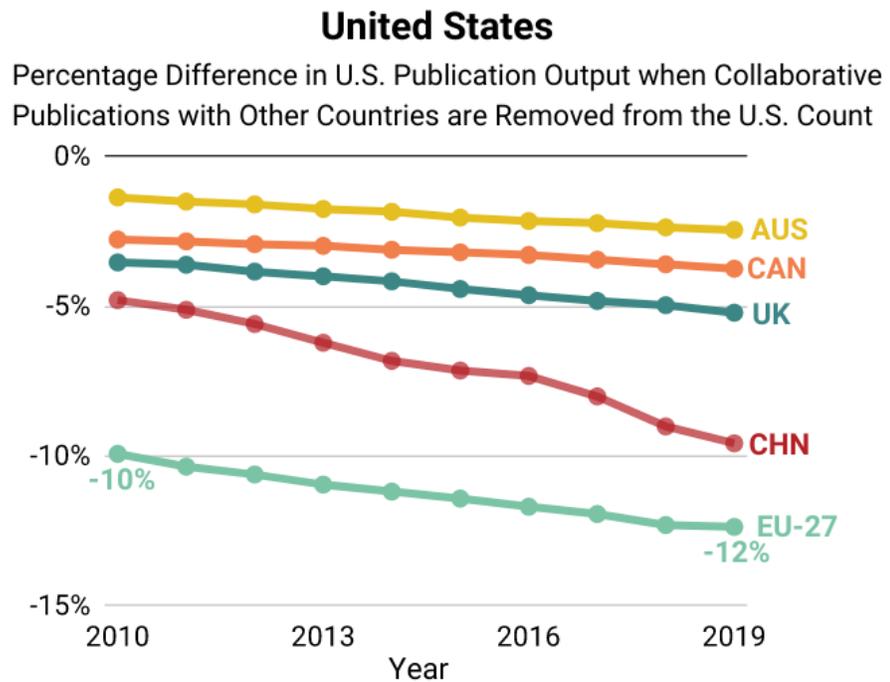
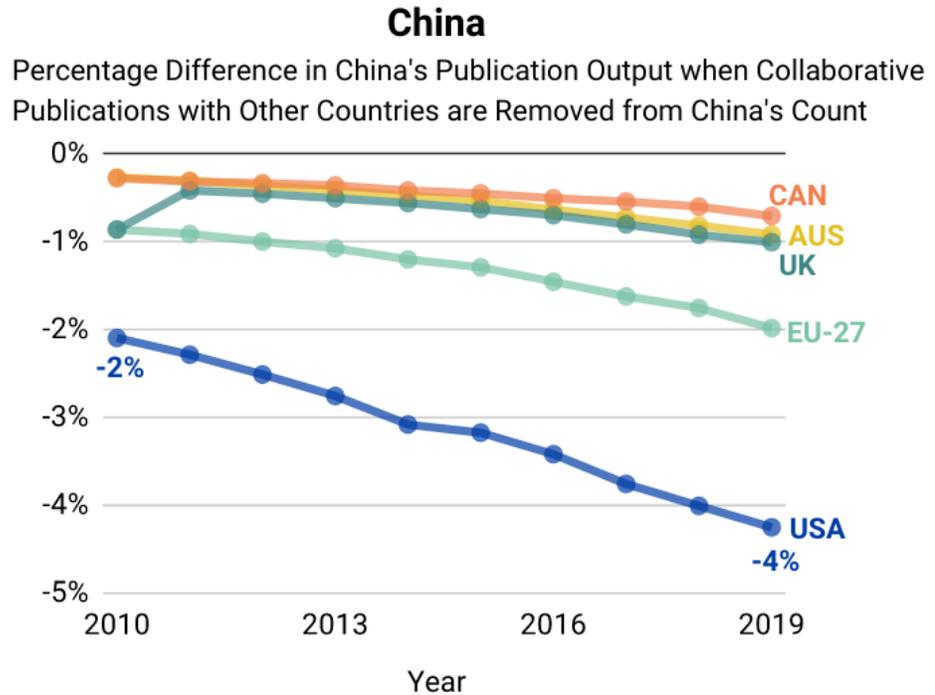
Generally, publications with international collaboration are cited more than publications with no international collaboration.

What effect does international collaboration have on research output?

Because China and the United States are recognized as global leaders in research publication output, we investigate the effect that international collaborations with China and the United States have on other countries' research outputs.

We compute the percentage difference in research publication output for China if its top five collaborators—Australia, Canada, EU-27, the United Kingdom, and the United States—are removed. We did the same comparison for collaborations with the United States, with China among the top five. Figure 6 displays the percentage difference in China's and the United States' research publication outputs if collaborations with the other five countries are removed.

Figure 6. China's and the United States' percentage differences in publication output when collaborative publications with other countries are removed from their publication counts (2010–2019).

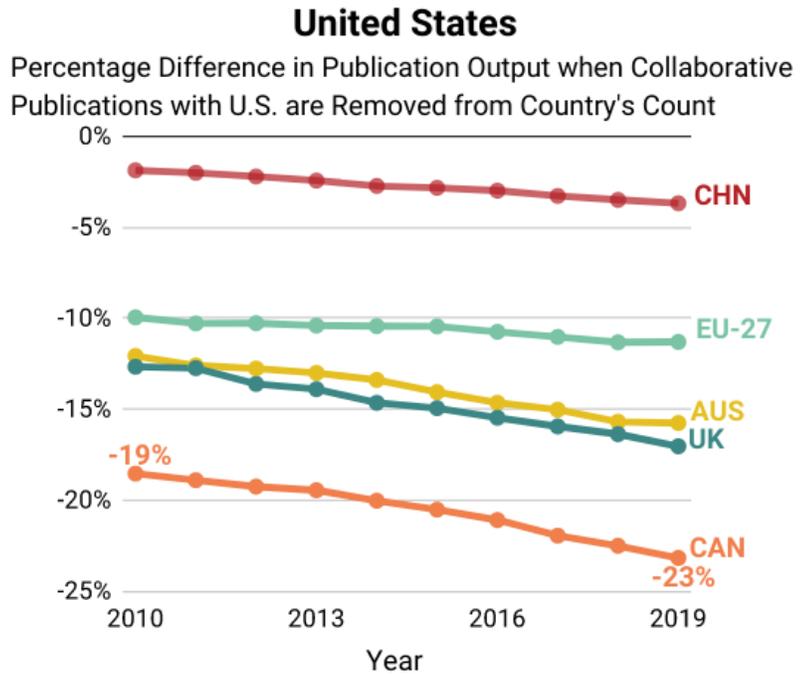
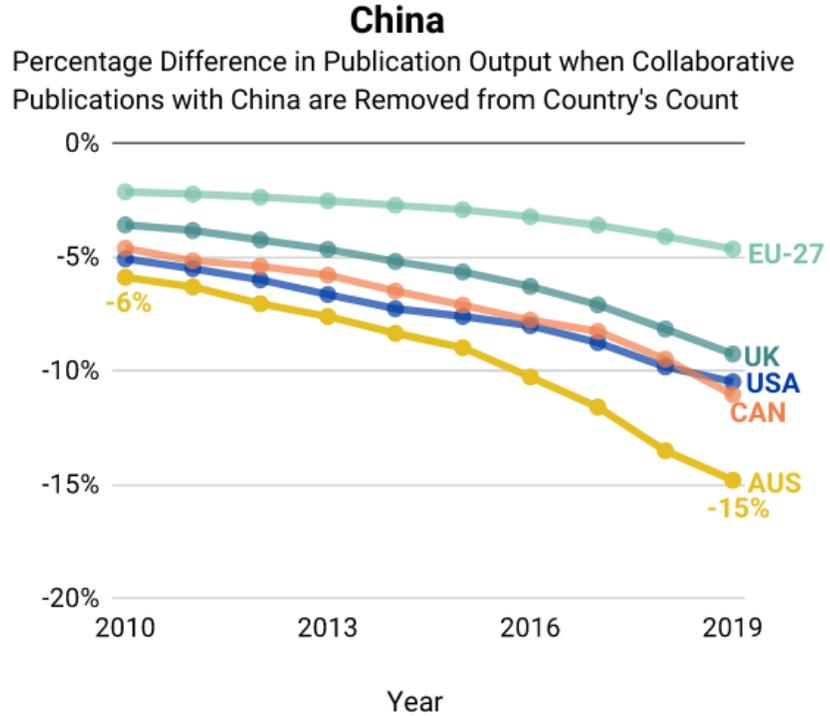


Source: CSET Merged Corpus.

China loses a noticeably smaller percentage of research publication output than the United States overall. Yet, removing collaborations with the United States has the largest impact on China's publication output, resulting in a 4 percent decrease in research output in 2019, a total of 84,299 research publications. For the United States, removing collaborations with the EU-27 results in the largest impact on publication output, for a 12 percent decrease in U.S. research output in 2019, a total of 93,792 publications.

Figure 7 displays the opposite of Figure 6: the percentage difference in the five countries' research output if collaborations with China and the United States are removed.

Figure 7. Percentage difference in a given country’s publication output when collaborative publications with China and the United States are removed from that country’s counts (2010–2019).



Source: CSET Merged Corpus.

For both China and the United States, removing their collaborations with the other countries has a significantly larger percentage impact on those nations' outputs than on China's and the United States' research publication outputs. When collaborations with China are removed, Australia loses approximately 15 percent of its research output, a total of 18,309 research publications, in 2019. When collaborations with the United States are removed, Canada loses approximately 23 percent of its research output, a total of 28,455 research publications, in 2019.

In terms of relative country output, countries collaborating with China experience a more significant percentage decrease to their publication outputs when their collaborations with China are removed than China experiences when collaboration publications are removed. However, in terms of publication counts, the values are on a slightly different scale. For example, in 2019 Australia loses 15 percent of its publications (18,309 publications), whereas China loses 4 percent of its publications (84,299 publications) when collaborations with the United States are removed.

Conclusion

Because scientific research benefits from international collaborations, it is important for researchers, policymakers, and national leaders to understand the implications of collaborative research. This data brief explores the role of international collaboration in research publications in terms of research output and impact. While several studies have analyzed specific countries or areas of research, we provide a full overview of all of science for the top 10 countries with the highest research output during 2010–2019. This analysis provides a more nuanced view into which countries are leading in scientific research publications and the important role of international collaboration in research output and impact.

We find that international collaboration does not result in equal outcomes for all countries. Specifically, we find that if collaborations with China were restricted, China's research publication output would decrease far less than that of

collaborating countries (e.g., Australia). However, a restriction on collaborating with China would significantly decrease China's highly cited publications. Additionally, publications with more countries contributing have higher percentages of high-impact work. Our results suggest that countries should analyze their international collaborator relationships to understand how investing or divesting from specific partnerships will affect their research impact and output.

Authors

Autumn Toney is a data research analyst at CSET. Melissa Flagg was formerly a senior research fellow at CSET and is currently the founder of Flagg Consulting, as well as a visiting fellow at the Perry World House.

Acknowledgments

For feedback and assistance, we would like to thank Catherine Aiken, Kevin Boyack, Paul Harris, Igor Mikolic-Torreira, Shelton Fitch, and Lynne Weil.



© 2021 by the Center for Security and Emerging Technology. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc/4.0/>.

Document Identifier: doi: 10.51593/20210050

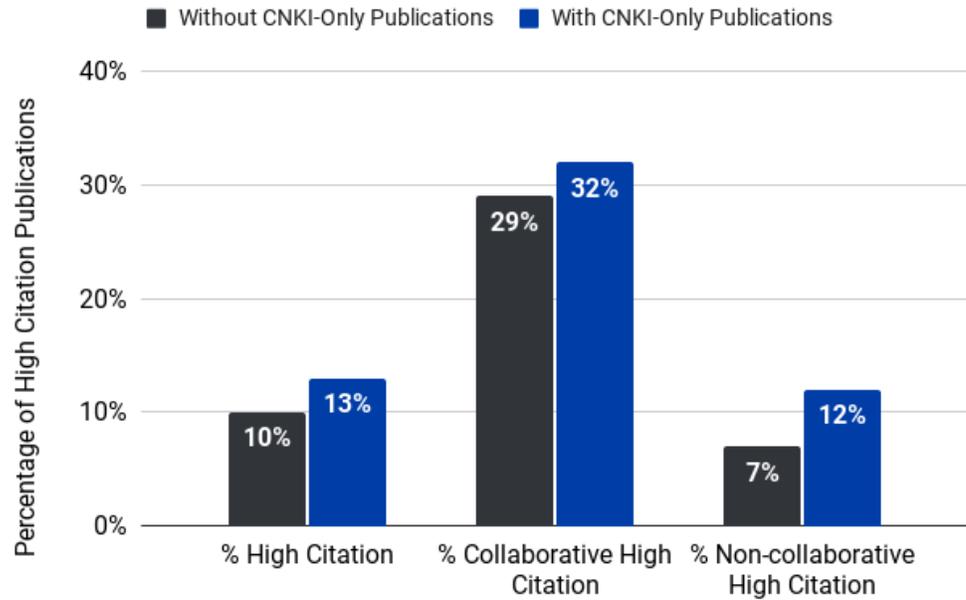
Appendix

Comparing Collaboration Rates in CSET's Merged Corpus to Scopus

In Figure 1 in the main text, we compare the collaboration rates of the top 11 countries by research publication output between CSET's Merged Corpus and Scopus. For both datasets we analyze publications that are linked in the Map of Science and we filter for research publications that are labeled as books, book chapters, conference publications, journal publications, or repositories. We find that all countries except for China have similar collaboration rates.

In order to understand the discrepancy between CSET's Merged Corpus and Scopus in regards to China's collaboration rates, we compute the percentages of highly cited publications, collaborative highly cited publications, and non-collaborative highly cited publications for CSET's Merged Corpus (with CNKI only publications) and CSET's Merged Corpus (without CNKI only publications). CNKI-only publications refer to publications that are only found in the CNKI database, and were not linked in any other database used to generate the Merged Corpus.

Figure A. Comparing Highly Cited Chinese Publication results with CNKI-Only Publications and Without CNKI-Only publications



Source: CSET Merged Corpus and Scopus.

The results in Figure 1 show that using CNKI-only publications does not cause China to appear less highly cited.

European Union Collaboration Rates

In the CSET policy brief *Research Security, Collaboration, and the Changing Map of Global R&D*, the European Union countries are treated as one entity. Our research also considers the 27 EU countries as one group, and in Table 1A we provide details of collaboration rates for each of the EU countries. Here we treat each EU-27 country individually, therefore if two or more EU-27 countries are on a publication, we consider that publication collaborative. The countries are sorted in descending order by number of publications produced between 2010–2019.

Table 1A. European Union Country Details on Publication Output and Collaboration Rates

Country	Number of Publications	% Collaborative Publications	% Highly Cited Publications	% Highly Cited Collaborative Publications
Germany	1,827,772	50%	30%	37%
France	1,216,310	56%	29%	35%
Italy	1,118,828	48%	31%	38%
Spain	965,472	47%	28%	36%
Netherlands	645,037	60%	39%	44%
Poland	477,645	31%	16%	29%
Sweden	415,404	62%	36%	41%
Belgium	353,073	65%	35%	40%
Denmark	277,433	61%	38%	44%
Austria	266,737	63%	31%	37%
Portugal	251,028	52%	27%	35%
Czech Republic	226,440	45%	19%	30%
Finland	215,840	60%	34%	40%
Greece	202,177	50%	28%	36%
Romania	169,662	33%	13%	27%
Ireland	156,115	58%	32%	38%
Hungary	114,417	53%	23%	33%
Slovakia	77,162	44%	15%	25%
Croatia	63,787	43%	19%	32%
Slovenia	59,988	51%	24%	33%

Bulgaria	49,686	53%	16%	26%
Lithuania	42,454	39%	18%	33%
Estonia	35,369	61%	29%	38%
Latvia	29,075	49%	20%	35%
Cyprus	28,721	69%	31%	36%
Luxembourg	22,156	74%	33%	36%
Malta	7,214	57%	27%	37%

Endnotes

¹ Ribeiro, Leonardo Costa, et al. "Growth patterns of the network of international collaboration in science." *Scientometrics* 114.1 (2018): 159-179.

² See examples: Sooho Lee and Barry Bozeman, "The Impact of Research Collaboration on Scientific Productivity." *Social Studies of Science* 35, no. 5 (2005): 673-702; Jonathan Adams, "The Fourth Age of Research," *Nature* 497 (2013): 557-560; Jenny J. Lee and John P. Haupt, "Winners and Losers in US-China Scientific Research Collaborations," *Higher Education* 80, no. 1 (2020): 57-74.

³ See examples: Jeff Tollefson, "China Declared Largest Source of Research Articles," *Nature* 553, no. 7689 (2018): 390; Karen White, "Publications Output: U.S. Trends and International Comparisons," National Science Foundation and National Science Board, December 17, 2019.

⁴ See examples: Tuan V. Nguyen, Thao P. Ho-Le, and Ut V. Le, "International Collaboration in Scientific Research in Vietnam: An Analysis of Patterns and Impact," *Scientometrics* 110, no. 2 (2017): 1035-1051; Maki Kato and Asao Ando, "The Relationship between Research Performance and International Collaboration in Chemistry," *Scientometrics* 97, no. 3 (2013): 535-553; András Schubert and Tibor Braun, "International Collaboration in the Sciences 1981-1985," *Scientometrics* 19, no. 1 (1990): 3-10.

⁵ Melissa Flagg, Autumn Toney, and Paul Harris, "Research Security, Collaboration, and the Changing Map of Global R&D" (Center for Security and Emerging Technology, June 2021), <https://cset.georgetown.edu/publication/research-security-collaboration-and-the-changing-map-of-global-rd/>.

⁶ Dimensions is provided by Digital Science, Web of Science is provided by Clarivate Analytics, and China National Knowledge Infrastructure is furnished for use in the United States by East View Information Services, Minneapolis, MN, USA.

⁷ EU-27 does not include the United Kingdom, full country list in the Appendix.

⁸ Because we are treating the EU-27 countries as one entity, if an EU-27 country collaborates with another EU-27 country on a publication we do not consider that publication to be collaborative.

⁹ There are 19 research fields: Art, Biology, Business, Chemistry, Computer Science, Economics, Engineering, Environmental Science, Geography, Geology, History, Materials Science, Mathematics, Medicine, Philosophy, Physics, Political Science, Psychology, and Sociology.

¹⁰ We also ran our analysis treating the 99th percentile as highly cited, but the collaboration rate trends were similar across all countries, so we chose to use the 90th in order to include more publications.

¹¹ Research publications with one distinct country are considered non-collaborative, despite the number of authors.