https://doi.org/10.51408/issi2025_033

Self Citations in Academic Excellence: Analysis of the Top 1% Highly Cited India-Affiliated Research Papers

Kiran Sharma[1], Parul Khurana[2]

[1]*kiran.sharma@bmu.edu.in*
School of Engineering & Technology, BML Munjal University, Gurugram, Haryana-122413 (India)
Center for Advanced Data and Computational Science, BML Munjal University, Gurugram, Haryana-122413 (India)

[2]*parul.khurana@lpu.co.in*
School of Computer Applications, Lovely Professional University, Jalandhar - Delhi G.T. Road, Phagwara, Punjab – 144411 (India)

Abstract

Citations demonstrate the credibility, impact, and connection of a paper with the academic community. Self citations support research continuity, but, if excessive, may inflate metrics and raise bias concerns. The aim of the study is to examine the role of self citations towards the research impact of India. To study this, 3.58 million papers affiliated with India from 1947 to 2024 in the Scopus database were downloaded, and 2.96 million were filtered according to document type and publication year up to 2023. Further filtering based on high citation counts identified the top 1% of highly cited papers, totaling 29,556. The results indicate that the impact of Indian research, measured by highly cited papers, has grown exponentially since 2000, reaching a peak during the 2011–2020 decade. Among the citations received by these 29,556 papers, 6% are self citations. Papers with a high proportion of self citations (>90%) are predominantly from recent decades and are associated with smaller team sizes. The findings also reveal that smaller teams are primarily domestic, whereas larger teams are more likely to involve international collaborations. Domestic collaborations dominate smaller team sizes in terms of both self citations and publications, whereas international collaborations gain prominence as team sizes increase. The results indicate that while domestic collaborations produce a higher number of highly cited papers, international collaborations are more likely to generate self citations. The top international collaborators in highly cited papers are the USA, followed by UK, and Germany.

Introduction

Citations are essential in academic research, acknowledging previous work, demonstrating integrity, and situating new studies within a broader scientific context. They serve as a key metric for assessing the impact of research, with high citation counts reflecting significant contributions to the field. Furthermore, citations facilitate knowledge dissemination, foster collaboration, and link studies between disciplines (Bornmann and Daniel, 2008). However, self citations, where authors cite their own work, provide continuity by linking new findings to prior contributions, especially in cumulative research. They also increase the visibility of newly published papers, which can attract external citations by highlighting related work (Hyland, 2003). However, excessive self citations can artificially inflate the citation metrics, misrepresenting the true influence of the paper, and raising concerns about bias (Fowler and Aksnes, 2007). Thus, while citations and self citations are vital

tools for measuring academic impact, their appropriate use is essential to maintain credibility and transparency in research.

Moreover, when self citations are used in an excessive or strategic manner to inflate citation metrics, it distorts the author's as well as the organization's academic influence (Moed, 2006). Today, citations in the form of scientific influence are used by various academic and government organizations for hiring, promotions, institutional prestige, bridging knowledge across different fields, fostering interdisciplinary research and funding decisions (Van Leeuwen, 2013). In such cases, self citations may create citation loops to potentially skew critical bibliometric indicators (Taham- tan and Bornmann, 2019). The existence of groups in the form of "citation cartels" also engage in reciprocal citation practices, which further compounds the issue (Hillman and Baydoun, 2019). This trend underscores the urgent need for a nuanced understanding of self citations across different academic backgrounds as high citations indicate that the particular study has substantial contribution in the field of research (Hirsch, 2005).

As the scientific community understood the elevation, narrative, and opportunistic power of self citations, concerns arose about the ethical implications of artificial citations (Van Noorden and Chawla, 2019). Some argued in favor of self citations, stating it as a reflection of specialization, while others presented them as manipulations (Costas et al., 2010). The tipping point came when researchers unravel the self citation patterns at the level of authors, country exhibitions, and academic organizations (Hellsten et al., 2007).

Citations in the academic world work as the thread that weaves the vast fabric of human knowledge. If utilized properly and for the advancement of the community, they are more than just numbers (Hodge, 2025; Szomszor et al., 2020). They enhance human knowledge by guaranteeing coherence and continuity. Institutions and financial agencies need to take into account the caliber of contributions rather than just the quantity of citations (Hussein et al., 2024). Most significantly, the scientific community needs to keep improving its evaluation methods so that a researcher's effect is determined by academic merit rather than metric manipulation (Martin, 2013).

Research objectives

The study examines the top 1% of highly cited papers and aims to achieve the following objectives:

1. Evaluate the concentration of self citations in the highly cited papers over year and decades.
2. Investigate the influence of team size on self citation patterns.
3. Explore the impact of domestic and international collaborations on highly cited papers and the associated concentration of self citations.

Methodology

In figure 1, the flow chart outlines the process of selecting the top 1% highly cited research papers affiliated with India, based on data retrieved from Scopus. A total of 3.58 million papers from 1947 to 2024 were downloaded from Scopus, searching for

the affiliation country as "India". The dataset was then refined to include only articles, conference proceedings, and reviews, focusing on publications up to 2023, resulting in 2.96 million papers from 1947 to 2023. Of these, articles accounted for 2.28 million (76.29%), conference proceedings for 0.56 million (18.94%) and reviews for 0.14 million (4.78%). The papers were organized in descending order based on the number of citations received. Further filtering identified the top 1% of highly cited papers, totalling 29,556. Within this group, articles comprised 21,645 papers (73.23%), reviews 7,100 (24%) and conference papers 811 (2.74%). Finally, the filtering process systematically narrowed down a massive dataset of more than 3.5 million papers to a smaller set of highly impactful publications. The top 1% highly cited papers (29,556) represent the most influential research outputs affiliated with India, highlighting the country's global academic and scientific contributions.

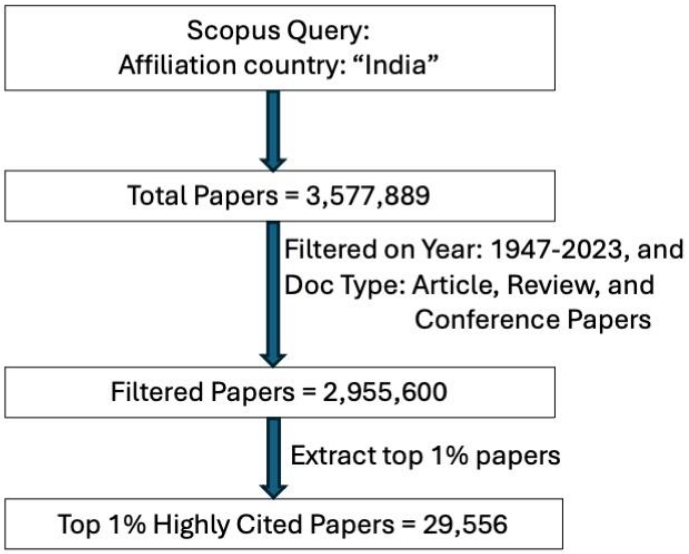

Figure 1. Search strategies undertaken to identify top 1% highly cited papers affiliated to India.

Results and Discussion

Citations vs. self citations

Citations and self citations are indispensable tools for academic research, helping to recognize prior work, measure impact, and foster scholarly communication. Striking the right balance between self-referencing and engaging with the broader academic community is essential to maintain the integrity and quality of research. Figure 2 represents the trends in total citations and number of papers (logarithmic scale) affiliated with India over time (from 1947 to 2023). Very few highly cited papers were published during 1947–1980 (early stage), as reflected by the flat portion of the blue star line. Citations are also minimal, but the red dotted line shows occasional spikes (possibly due to a few influential papers published during this period). The

graph demonstrates India's remarkable progress in producing highly influential research papers, particularly post-2000.

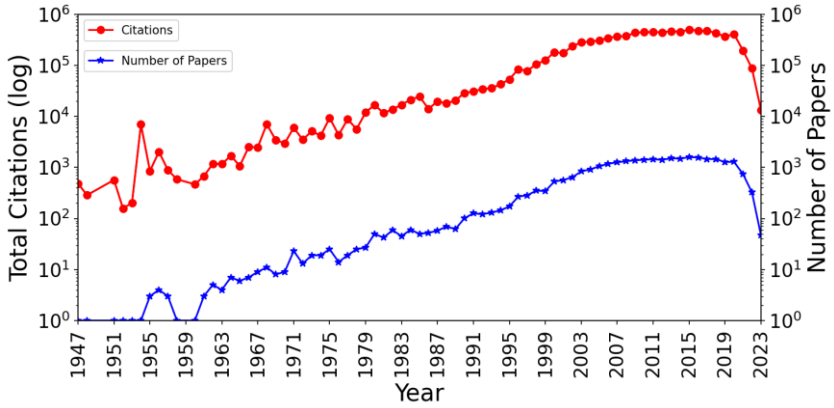

Figure 2. Year-wise trend of number of publications and total citations received.

A total of 29,556 highly cited papers received 97,53,620 citations in total, as shown in Table 1. Of these 9.75 million citations, 0.59 million (593,321) are self-cited, that is, 6% citations are self citations. Figure 3 represents the trend line of the average self citations over years received by the highly cited papers. The data exhibit fluctuations, with periods of increase and occasional declines, but the overall trend reveals a steady growth in self citations over time. The positive slope of the red trend line confirms this upward trajectory, suggesting that self citations have generally become more frequent in recent years.

Table 1. Search strategies undertaken to identify top 1% highly cited papers affiliated to India.

Decades	TP		Total Citations	Self Citations	Open Access	
	Count	%Count			Yes	No
1947-1950	2	0.01	801	0	0	2
1951-1960	15	0.05	13,083	26	1	14
1961-1970	61	0.21	24,585	364	3	58
1971-1980	193	0.65	76,997	2,325	8	185
1981-1990	547	1.85	1,92,674	8,658	34	513
1991-2000	2,042	6.91	7,88,617	45,757	175	1,867
2001-2010	9,680	32.75	33,79,239	1,85,869	1,204	8,476
2011-2020	14,614	49.45	48,98,374	3,25,447	3,929	10,685
2021-2023	2,402	8.13	3,79,250	24,875	973	1,429
Total	**29,556**	**100**	**97,53,620**	**5,93,321**	**6,327**	**23,229**

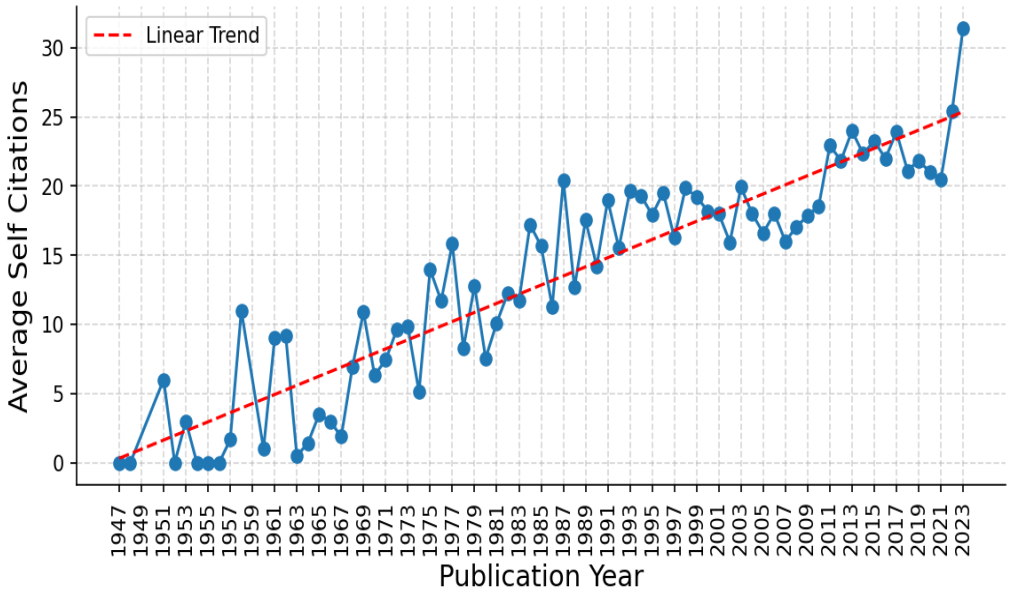

Figure 3. Average self citations over the years. Red dashed line represents the trend line with slope 0.33.

Table 1 provides an analysis of the top 1% highly cited papers affiliated with India, organized by decades. The table includes the count of papers, their percentage contribution to the total, the total citations, share of self citations and whether the papers were published under Open Access (Yes) or not (No). Among the 29,556 highly cited papers, 21.4% (6,327) were published as open access, while 78.59% (23,229) were not open access. The transition to Open Access is evident, with nearly 21% of recent highly cited papers being openly accessible. The production of highly cited papers increased dramatically after 2000, with nearly 90% of the top 1% papers produced between 2001 and 2023. The most significant contribution came from the 2011–2020 decade (49.45%). Citations reflect the growing impact of Indian research globally. Papers from 2011–2020 have received the highest citations (48,98,374), nearly half of the total. This trend demonstrates India's rising contribution to global research impact, particularly in recent decades, through both high citation counts and improved accessibility. In contrast, an analysis of the proportion of self citations to total citations reveals that 6.5% of self citations refer to papers published during 2021–2023, 6.6% to those from 2011–2020 and 5.5% to papers from 2002–2010. Table 2 presents a list of papers in which self citations account for more than 90% of their total citations. These papers are primarily from recent decades and are associated with smaller team sizes.

Table 2. List of papers with more than 90% of self citations.

S. no.	Paper title	Pub year	Team size	Total citations	Self citations	% Self citations
1	Analysis, adaptive control and synchronization of a seven-term novel 3-D chaotic system	2013	2	176	173	98.3
2	Global chaos synchronization of a family of n-scroll hyperchaotic chua circuits using backstepping control with recursive feedback	2013	2	148	145	98.0
3	Sliding controller design of hybrid synchronization of Four-Wing Chaotic systems	2011	2	182	178	97.8
4	Sliding mode control based global chaos control of Liu-Liu-Liu-Su chaotic system	2012	1	165	161	97.6
5	Adaptive anti-synchronization of Uncertain Tigan and Li Systems	2012	2	162	158	97.5
6	Active controller design for generalized projective synchronization of four-scroll chaotic systems	2011	2	164	159	97.0
7	A new eight-term 3-D polynomial chaotic system with three quadratic nonlinearities	2014	1	181	175	96.7
8	Global chaos control of hyperchaotic Liu system via sliding control method	2012	1	163	157	96.3
9	Anti-synchronization of Lu and Pan chaotic systems by adaptive nonlinear control	2011	2	161	155	96.3
10	Global chaos synchronization of hyperchaotic Pang and hyperchaotic Wang systems via adaptive control	2012	2	152	146	96.1
11	A new six-term 3-D chaotic system with an exponential nonlinearity	2013	1	197	189	95.9
12	Generalized Projective Synchronization of Two-Scroll Systems via Adaptive Control	2012	2	162	155	95.7
13	Anti-synchronization of hyperchaotic lorenz and hyperchaotic chen systems by adaptive control	2011	2	158	151	95.6
14	The generalized projective synchronization of hyperchaotic lorenz and hyperchaotic Qi systems via active control	2011	2	165	157	95.2
15	Hybrid synchronization of n-scroll chaotic chua circuits using adaptive backstepping control design with recursive feedback	2013	2	161	152	94.4
16	Analysis, properties and control of an eight-term 3-D chaotic system with an exponential nonlinearity	2015	1	154	145	94.2
17	Generalised projective synchronisation of novel 3-D chaotic systems with an	2014	1	166	155	93.4

	exponential non-linearity via active and adaptive control					
18	Adaptive synchronization of chemical chaotic reactors	2015	1	149	136	91.3
19	Global chaos synchronization of WINDMI and Coullet chaotic systems using adaptive backstepping control design	2014	2	160	146	91.3
20	Analysis, control and synchronisation of a six-term novel chaotic system with three quadratic nonlinearities	2014	1	192	175	91.1
21	Analysis and anti-Synchronization of a novel chaotic system via active and adaptive controllers	2013	1	184	167	90.8
22	Analysis, adaptive control and anti-synchronization of a six-term novel jerk chaotic system with two exponential nonlinearities and its circuit simulation	2015	5	150	136	90.7
23	Analysis and adaptive synchronization of two novel chaotic systems with hyperbolic sinusoidal and cosinusoidal nonlinearity and unknown parameters	2013	1	191	173	90.6

Team size vs. self citations

According to Wuchty et al. (2007), larger teams tend to dominate in producing highly cited research due to their ability to combine various expertise, tackle complex problems, and leverage collaborative networks. In contrast, smaller teams are more likely to focus on niche and innovative topics, which may gain recognition more gradually. This high- lights the role of team size in shaping research impact and citation patterns. Figure 4(a) illustrates the authors with multiple highly cited papers. 73.47% of the authors have a single paper appearing in the top 1% highly cited papers. 13.55% authors have two papers in the highly cited list followed by 5.2% having 3 papers, etc. Figure 4(b) represents the distribution of teams appearing in the paper and the number of publications. A negative correlation exists between the number of papers and the number of authors, suggesting that smaller teams (fewer authors) are more prolific in producing papers. In contrast, larger teams contribute fewer papers, likely due to the increased complexity and coordination involved in collaborative research. As team size grows, the number of papers decreases significantly, with some exceptions for very large teams.

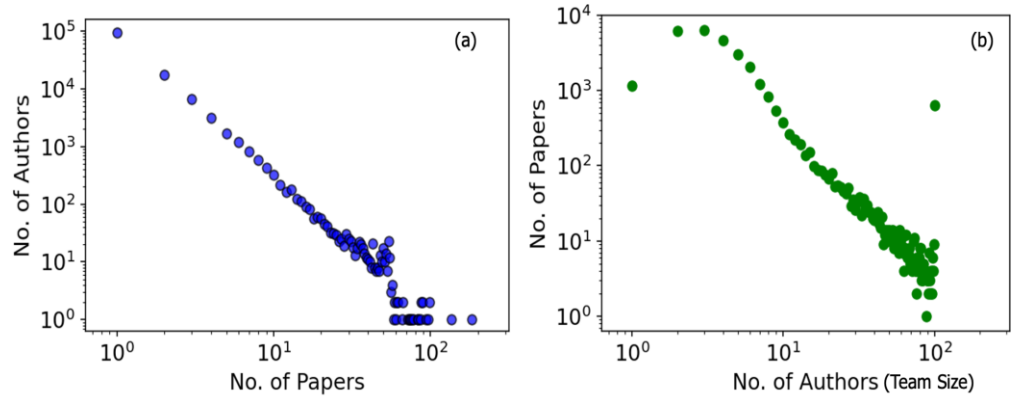

Figure 4. (a) Number of papers vs. count of authors. (b) Team size vs. number of papers.

In addition, Figure 5 represents the average self citations based on team size. The average self citations generally increase as the team size grows, particularly in smaller teams. There are notable peaks and troughs, indicating variability in self citation practices as team sizes change. Teams with fewer members (1–25) exhibit a more consistent, gradual increase in average self citations, indicating relatively steady behavior. The red dashed line illustrates the overall increasing trend, suggesting a positive correlation between team size and average self citations, although the data show variability.

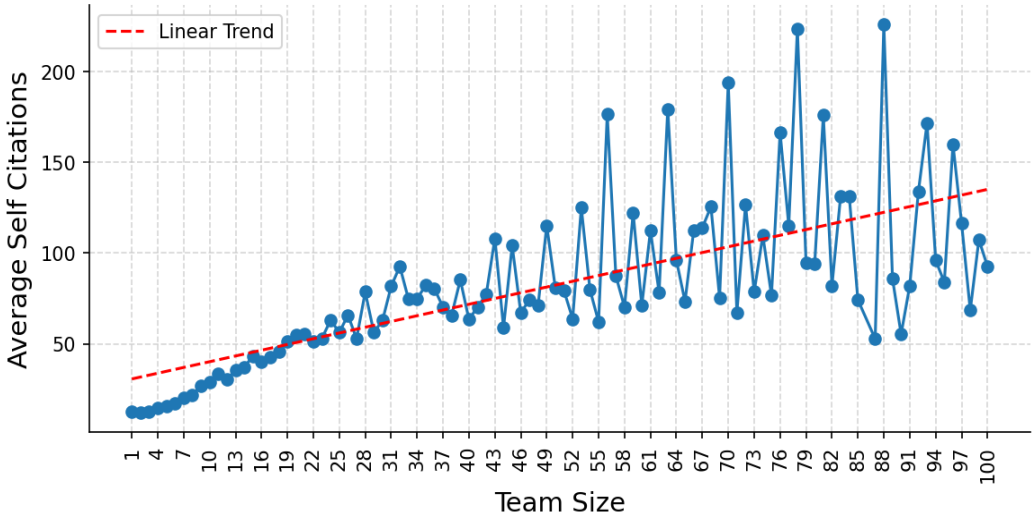

Figure 5. Team size vs average self citations.

Collaboration pattern vs. self citations

An influential and highly cited paper often the result of interdisciplinary teams and collaborative efforts. According to Uzzi et al. (2013), interdisciplinary collaboration and diversity in research teams are significant factors contributing to innovative and impactful research, as they bring together varied perspectives and expertise, which

increase the likelihood of producing groundbreaking work. The table 3 provides an overview of the distribution of highly cited papers and self citations between domestic and international collaborations. Out of the total 29,556 highly cited papers, domestic collaborations contribute the majority, accounting for 56.9%, while international collaborations contribute 43.1%. The total citation distribution indicates that 49.7% of total citations belong to domestic papers, while 50.3% of the total citations belong to international papers. However, when examining self citations, international collaborations dominate with 67% self citations, compared to 33% self citations from domestic collaborations. This indicates that while domestic collaborations produce a higher number of highly cited papers, international collaborations are more likely to generate self citations. This could reflect the broader scope, larger teams, and multidisciplinary nature of international projects, leading to higher interconnectedness and frequent self referencing. In contrast, domestic collaborations, often involving smaller teams, focus on national-level research with relatively fewer self citations.

Table 3. Distribution of number of papers, corresponding total and self citations.

Category	Domestic		International		Total
	Count	In %	Count	In %	
No. of papers	16,823	56.9%	12,733	43.1%	29,556
No. of total citations	48,43,057	49.7%	49,10,563	50.3%	97,53,620
No. of self citations	1,95,752	33%	3,97,569	67%	5,93,321

Figure 6 represents different aspects of the relationship between team size and the nature of collaborations (domes. vs. international). The figure on the left represents the number of publications versus team size where publications domestic collaborations peak at smaller team sizes and gradually decline as team size increases, showing minimal activity for larger teams. International collaborations demonstrate a consistent trend, maintaining a higher level of publications for medium-to-large team sizes compared to domestic collaborations. The spike in international collaborations for the largest team size is likely driven by highly collaborative or global scale projects. Similarly, the right figure represents the number of self citations vs. team size where self citations for domestic collaborations are higher for smaller team sizes and decline sharply as the team size increases. International collaborations exhibit a steadier decline in self citations, with smaller team sizes showing relatively lower self citations compared to domestic ones. There is a noticeable peak for international collaborations at the larger team size (likely an outlier). In general, domestic collaborations dominate smaller team sizes in terms of both self citations and publications, whereas international collaborations gain prominence as team sizes increase. In addition, the USA is the topmost collaborator followed by the UK and Germany.

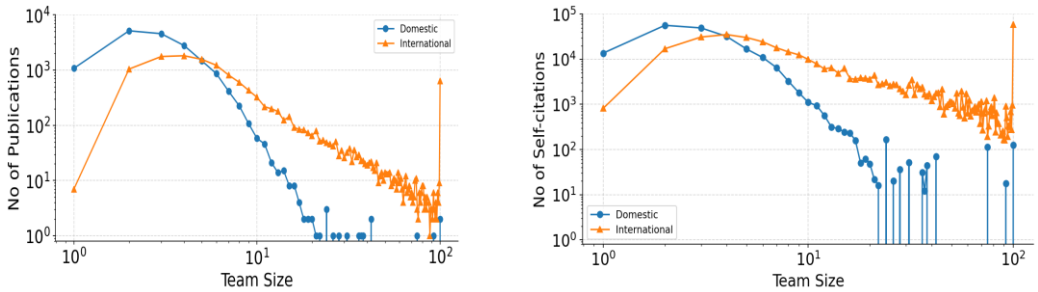

Figure 6. Domestic vs. international collaboration. (left) Team size vs. number of publications. (b) Team size vs. self citations.

Conclusion

In conclusion, citations and self citations play a crucial role in academic research by measuring impact, ensuring research continuity, and fostering collaboration. The study on highly cited Indian research papers highlights that domestic collaborations contribute a greater number of highly cited papers, while international collaborations generate more self citations, likely due to broader networks and multidisciplinary projects. Team size also influences citation patterns, with smaller teams producing more papers but relying more on self citations, whereas larger teams are linked to international collaborations and higher citation visibility. Striking a balance between self-referencing and engaging with the broader academic community is essential to maintain credibility and ensure meaningful research impact.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, citation practices vary across disciplines, making direct comparisons challenging, especially in fields where self citation rates are naturally higher. Second, limitation lies in the classification of team sizes, as it does not account for variations in individual author contributions, which can influence citation impact. Third, the study differentiates between domestic and international collaborations but does not fully capture the complexity of multi-country partnerships that may affect citation trends. Lastly certain fields, such as biomedical research, have higher citation frequencies than others, making direct comparisons across disciplines challenging.

Data availability

The datasets used in the study will be available from the corresponding author on request.

Acknowledgment

Authors gratefully acknowledges the Research and Development Cell, BML Munjal University for their financial support through the seed grant (No: BMU/RDC/SG/2024-06), which made this research possible.

Conflict of interest

The authors declare no conflict of interest.

References

Bornmann, L., Daniel, H.D., 2008. What do citation counts measure? a review of studies on citing behavior. Journal of documentation 64, 45–80.

Costas, R., van Leeuwen, T., Bordons, M., 2010. Self citations at the meso and individual levels: effects of different calculation methods. Scientometrics 82, 517–537.

Fowler, J., Aksnes, D., 2007. Does self citation pay? Scientometrics 72, 427–437.

Hellsten, I., Lambiotte, R., Scharnhorst, A., Ausloos, M., 2007. Self citations, co-authorships and keywords: A new approach to scientists' field mobility? Scientometrics 72, 469–486.

Hillman, J.R., Baydoun, E., 2019. Quality assurance and relevance in academia: a review. Major challenges facing higher education in the Arab world: Quality assurance and relevance, 13–68.

Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences 102, 16569– 16572.

Hodge, D.R., 2025. Assessing scholarly impact: Conducting bibliometric analyses, in: Handbook of Research Methods in Social Work. Edward Elgar Publishing, pp. 76–83.

Hussein, T.M., Ateeq, A., Ateeq, R.A., Agarwal, S., 2024. Self citation in scholarly work: Balancing self-reference with scientific integrity, in: Business Sustainability with Artificial Intelligence (AI): Challenges and Opportunities: Volume 2. Springer, pp. 361–369.

Hyland, K., 2003. Self citation and self-reference: Credibility and promotion in academic publication. Journal of the American Society for Information Science and technology 54, 251–259.

Martin, B.R., 2013. Whither research integrity? plagiarism, self-plagiarism and coercive citation in an age of research assessment.

Moed, H.F., 2006. Citation analysis in research evaluation. volume 9. Springer Science & Business Media.

Szomszor, M., Pendlebury, D.A., Adams, J., 2020. How much is too much? the difference between research influence and self citation excess. Scientometrics 123, 1119–1147.

Tahamtan, I., Bornmann, L., 2019. What do citation counts measure? an updated review of studies on citations in scientific documents published between 2006 and 2018. Scientometrics 121, 1635–1684.

Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical combinations and scientific impact. Science 342, 468–472.

Van Leeuwen, T., 2013. Bibliometric research evaluations, web of science and the social sciences and humanities: a problematic relationship? Bibliometrie-Praxis und Forschung 2.

Van Noorden, R., Chawla, D.S., 2019. Hundreds of extreme self-citing scientists revealed in new database. Nature 572, 578–580.

Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316, 1036–1039.