

A Preliminary Study of Self-Citation Behavior Among Library and Information Science Authors in China: Postprint

Authors: Li Xinyu, Lei Jiaqi, Step 1

Date: 2022-11-27T00:00:00+00:00

Abstract

[Purpose/Significance] Author self-citation is a common phenomenon in academia and an important factor considered in citation analysis and scientific evaluation, yet how to handle self-citation in scientific evaluation has remained controversial. Exploring the patterns of author self-citation can provide recommendations for appropriately addressing author self-citation. [Method/Process] First, we systematically reviewed research findings related to author self-citation, then analyzed author self-citation at both the author and document levels using data from journals in the field of library and information science in China, along with statistical methods such as regression analysis. [Results/Conclusion] Empirical research demonstrates that author self-citation behavior is generally infrequent in the field of library and information science in China, that author self-citation frequency has a strong positive correlation with the number of publications by the author, that author self-citation has a relatively small impact on the h-index, and that document self-citation has a positive effect on its citation frequency by others, among other findings.

Full Text

A Preliminary Study on Author Self-Citation Behaviors in Library and Information Science in China

Li Xinyu, Lei Jiaqi, Bu Yi

Department of Information Management, Peking University, Beijing 100871

Abstract:

[Purpose/Significance] Author self-citation is a common phenomenon in academia and an important factor to be considered in citation analysis and scientific evaluation. However, how to properly handle self-citation in scientific evaluation has long been controversial. Exploring the patterns of author

self-citation can provide recommendations for appropriately addressing this issue. [Method/Process] This paper first reviews relevant research on author self-citation, then analyzes author self-citation at both the author and document levels using data from Chinese library and information science journals and statistical methods such as regression analysis. [Result/Conclusion] Empirical evidence shows that author self-citation behaviors in Chinese library and information science are generally infrequent. The frequency of author self-citation exhibits a strong positive correlation with the number of published articles, while the impact of self-citation on the h-index is relatively small. Furthermore, a document being self-cited by its author has a positive effect on its citation frequency by others.

Keywords: author self-citation; citation analysis; library and information science; scientific evaluation

Classification Number: G253.1

Self-citation is a common citation phenomenon in academia. The definition of self-citation varies for different entities such as authors, journals, institutions, and countries [1-2]. Among these, journal self-citation and author self-citation are generally the focus of self-citation analysis. Citation analysis constitutes an important method in bibliometrics, with self-citation analysis being a crucial component. Self-citation behavior can, to some extent, reveal the citing intentions of entities and the continuity of their research, making self-citation research essential for both bibliometrics and scientific evaluation. Traditional metrics commonly used in scientific evaluation, such as journal impact factor and h-index, are based on citation counts. In most cases, self-citations are included in these counts without distinguishing them from citations by others, creating opportunities for entities to manipulate scientific evaluation through self-citation. Additionally, to make scientific evaluation more objective and effective, the Journal Citation Reports (JCR), as one of the scientific evaluation databases, began providing a metric that excludes self-citations—the impact factor without self-citations.

Self-citation is a controversial topic in academia. Journal self-citation has sometimes become a manipulation tool for journals seeking high citation counts and impact factors. JCR publishes an annual list of “suppressed” journals that are no longer indexed, primarily due to excessive self-citation and other anomalous citation behaviors. In 2021, JCR suppressed ten journals for these reasons. Some scholars argue that improper self-citation behavior has significantly impacted scientific evaluation [4], while others defend self-citation as mostly representing normal citation practices [5]. M. Schreiber identified three reasons for self-citation: (1) reasonable self-citation when previous experimental setups or theoretical models are needed for continuity but repetitive description is avoided; (2) somewhat controversial self-citation due to authors being most familiar with their own work, making it easier to cite; and (3) undisputed improper self-citation aimed at increasing personal citation counts and h-index

[6]. E. Garfield maintained that self-citation is neither inherently good nor bad [7]; the key lies in how it is viewed and utilized. Under current circumstances, the direction for resolving self-citation controversies involves finding ways to distinguish excessive from appropriate self-citation and exploring methods to properly handle self-citation in scientific evaluation. This paper analyzes these issues from two perspectives: (1) summarizing important research related to self-citation, and (2) using Chinese library and information science literature data to examine author self-citation, analyzing bibliometric factors correlated with self-citation metrics and discussing the impact of author self-citation on scientific evaluation in Chinese library and information science.

2 Research Status

2.1 Definitions of Self-Citation Metrics

The term “self-citation rate” has been used inconsistently in self-citation research [9]. A more precise approach distinguishes between “self-citing rate” and “self-cited rate.” In the 1975 introduction to JCR, Garfield explained that journals have two self-citation rates [10]: the self-citing rate (the proportion of self-citations among all references) and the self-cited rate (the proportion of self-citations among all citations received). The self-citing rate is a synchronous indicator, while the self-cited rate is a diachronic indicator. Broadly speaking, the formulas for these rates are:

Formula (1) Self-citing rate = Number of self-citations by an entity / Total number of citations by the entity

Formula (2) Self-cited rate = Number of self-citations received by an entity / Total number of citations received by the entity

For author self-citation, previous research has lacked a unified, standardized definition. Since a document may have multiple authors with different authorship orders, various definitions of author self-citation are possible. Most current studies define author self-citation as an author citing their own previously published work, which is the definition adopted in this paper. Under this definition, the formulas for author self-citing rate and author self-cited rate are:

Formula (3) Author self-citing rate = Number of an author’ s self-citations / Total number of an author’ s citations

Formula (4) Author self-cited rate = Number of an author’ s self-citations received / Total number of an author’ s citations received

It is worth noting that this paper’ s definition of author self-citation considers all authors and does not require the self-citing author to be the first author of either the citing or cited paper. However, some studies, due to data limitations, treat cases where the first authors of two cited documents are the same person as author self-citation [11-12]. The author self-citing rate reflects the importance of self-citation among knowledge sources, while the author self-cited rate reflects its importance in academic influence.

2.2 Factors Related to Self-Citation

Literature review reveals that author self-citation and self-cited behaviors correlate with several factors:

- (1) **Time since publication:** Studies show that the longer the citation time window after publication, the lower the share of author self-citations in all citations received [13-15].
- (2) **Discipline:** Self-citation frequencies vary across disciplines. For instance, H. Snyder and S. Bonzi found that author self-citation rates were approximately 15% in physics, 6% in social sciences, and 3% in humanities [16].
- (3) **Gender:** Some studies found that male authors self-cite more frequently than female authors, though this difference may stem from productivity disparities [17-18].
- (4) **Author productivity/journal article count:** Research indicates that author productivity positively correlates with author self-citation volume [14], and journals with significantly increased publication volumes have higher self-cited rates [19].
- (5) **Co-authorship:** M.R. Davarpanah and F. Amel found that the number of authors per paper positively correlates with author self-citing count [14]. Jiang Ying et al. discovered that the first author's self-citation percentage is much higher than other authors, with self-citing counts generally decreasing by author order [20].
- (6) **Citation count:** D.W. Aksnes found that an author's total citation count negatively correlates with their self-cited rate—the higher the total citations, the lower the self-cited rate [15].

These correlations between author self-citation and the above factors require further verification, and other factors remain to be explored.

2.3 Impact of Self-Citation on Scientific Evaluation

Self-citation phenomena may affect journal impact factors, scholar h-index, and citations by others. Previous studies have reached varying conclusions due to differences in data and methods.

Journal self-citation may influence journal impact factor. Some scholars argue it has a positive effect [21-22]. For example, A. Fassoulaki et al. demonstrated a significant positive correlation between self-citing rate and journal impact factor using data from six anesthesiology journals [22]. However, other studies found no significant effect or even a negative impact [23-26]. For instance, J.M. Campanario and A. Molina identified 123 journals with declining impact factors from 1998-2006 JCR data and concluded that self-citation did not significantly increase impact factors, suggesting that impact factor manipulation through

self-citation is ineffective [26]. Some research suggests the effect depends on journal influence or publication year [27-29].

Author self-citation may influence document impact. Studies show that self-citation can increase citations by others. J.H. Fowler et al. used Norwegian citation data and Poisson regression to find that each additional self-citation increased citations by others by approximately 1 in the following year, 2.83 cumulatively by year four, and 3.65 by year ten [30].

Author self-citation may also affect h-index and rankings. Scholars hold different views on this impact. Some argue the effect is small [31-32]. For example, Cha Ying examined the top 10 most-cited scholars from a library and information science impact report (2000-2004) and found that removing self-citations had minimal impact on h-index and rankings, with an average h-index decrease of 5.6% [31]. However, other research suggests self-citation significantly affects h-index, particularly for early-career researchers with lower h-indices, where self-citations should be excluded [6].

Based on these studies, this paper conducts empirical research on self-citation in Chinese library and information science. Section 4 will examine relationships between self-citation and factors such as time and productivity, the impact of author self-citation on citations by others, and the effect of self-citation on h-index.

3 Data and Methods

This study obtained bibliographic data for 98,948 articles published between 1955-2017 from 17 library and information science journals in the CNKI database, along with their references (only those indexed by CNKI) and citation data up to 2018. These 17 journals were identified through the intersection of the 7th edition of the *Chinese Core Journals Overview* [33] and the 2017-2018 CSSCI source journal directory, covering most important journals and literature in the field. During data cleaning, 25,782 articles with erroneous or missing publication year information, missing author information, or zero references were removed, leaving 73,160 articles published between 1957-2017 by 38,556 authors. shows the number of articles per journal in the dataset, with *Library and Information Service* publishing the most. [Figure 1: see original paper] displays the publication timeline; due to the small number of articles from 1957-1990, these years are shown separately on the left side of the figure.

This study aims to develop a deeper understanding of self-citation by exploring its correlates and impact on paper influence in the information science field. The research questions are: (1) What are the characteristics of self-citation levels, self-citing rates, and self-cited rates in Chinese library and information science? (2) Are the frequency and rate of self-citation correlated with time since publication, author productivity, and collaboration patterns, and what are these relationships? (3) How does author self-citation affect authors' h-

index and rankings? (4) Does author self-citation in a document influence its citation frequency by others?

Data processing and analysis were conducted primarily using Python, SPSS, and Stata. To examine the impact of self-citation on document influence, this study employed negative binomial regression models. When the dependent variable is a count variable (i.e., the number of events), count models should be used. These models are appropriate for discrete integer values with many zeros. Count models include Poisson and negative binomial regression. However, when data exhibit overdispersion, negative binomial regression is preferred over Poisson regression. To study the effect of self-citation on citations by others, this paper used panel data with citations by others (total citations minus self-citations) as the dependent variable. The independent variables focused on self-citing count and self-cited count, with control variables including reference count, publication year, average author h-index, journal impact factor in the publication year, and number of authors per document.

4 Results Analysis

4.1 Overall Self-Citation Status

[Figure 2: see original paper] shows the average self-citing rate of articles published from 1980-2017. During 2000-2017, the average self-citing rate remained relatively stable at approximately 4%-6%. Although the average self-citing count increased overall during 2000-2017 (from 0.13 for 2000 publications to 0.59 for 2017 publications), the self-citing rate did not change significantly because reference counts also increased overall.

The 73,160 articles in the dataset cited 430,505 times, with 4.50% being author self-citations. These articles received 811,301 citations, with 2.97% being author self-citations. A total of 12,782 articles (17.47%) contained author self-citations, and 13,698 articles (18.72%) were cited by their authors. The 38,556 authors had an average self-citing rate of 1.82% and an average self-cited rate of 2.05%. The distributions of author self-citing and self-cited rates were similar: most authors had rates in [0, 10%), some in [10%, 20%), with only 2.17% of authors exceeding 20% self-citing rate and 2.52% exceeding 20% self-cited rate.

4.2 Analysis of Factors Related to Author Self-Citation

4.2.1 Relationship Between Author Self-Citation and Time Since Publication Author self-citations occurred on average 2.65 years after publication (self-citations in the publication year counted as year 0). In contrast, citations by others occurred on average 4.47 years after publication. As shown in Figure 3: see original paper, self-citations peaked in the first year after publication and decreased gradually in years 2-20. Figure 3: see original paper, (c), and (d) demonstrate that for documents with varying citation levels, self-citations generally occur earlier than citations by others.

4.2.2 Relationship Between Author Order and Self-Citation in Co-Authored Papers shows the average proportion of self-citations by authors in different positions among papers with self-citing behavior. This proportion refers to the ratio of self-citations by authors in a given position to the total self-citations in that document. Among papers with self-citation behavior, the first author's self-citations accounted for an average of 73.99% of all self-citations in the document. Overall, authors in later positions had lower average self-citation percentages. For papers with various collaboration sizes, the first author's average self-citation proportion was always the highest among all authors. However, as the number of authors increased from 1 to 8 or more, the first author's average self-citation proportion continuously decreased.

4.2.3 Relationship Between Author Self-Citation and Productivity, Career Length, and Collaboration As shown in , authors' self-citing counts have varying degrees of positive correlation with publication count, career length, and collaboration frequency. Similarly, authors' self-cited counts show weak to moderate correlations with these factors. Therefore, under normal circumstances, authors with more publications, higher total collaboration frequency, longer academic careers, and higher h-indices tend to have higher self-citing and self-cited counts. This positive correlation is conditional: when authors are grouped by publication count or h-index, the positive correlations between self-citing/self-cited counts and collaboration, citation, and citation-received metrics are more easily observed for productive or influential author groups, as shown in and .

4.3 Impact of Author Self-Citation on h-Index

The data show that the average h-index including self-citations was 2.074, decreasing to 2.028 after excluding self-citations. Removing self-citations caused h-indices to decrease by 0 to 4 points, with an average decrease of 0.045 and an average reduction rate of 1.83%. Wilcoxon signed-rank tests revealed significant differences between h-indices with and without self-citations ($p = 0.000$).

To examine the impact of self-citation on h-indices and rankings of influential, well-known authors in library and information science, this study calculated the h-indices and citation counts of the top 20 most-cited authors after removing self-citations. The h-index decrease ranged from 0 to 6.06%, with 11 authors experiencing no change. The average h-index decreased by 1.77% after removing self-citations.

Thus, for both the broader group of all authors and the more specific group of highly-cited authors in Chinese library and information science, self-citation has a limited impact on h-index.

4.4 Impact of Author Self-Citation on Citations by Others

Since citation counts by others are count data with overdispersion, this study employed Stata's NB2 negative binomial regression with robust standard errors. The regression results show that the 95% confidence interval for the alpha parameter is (0.90, 0.92), rejecting the null hypothesis that $\alpha = 0$ and confirming significant overdispersion, making negative binomial regression more appropriate than Poisson regression.

The regression results are presented in . Except for "number of authors," all independent variables are significant. The results are interpreted using incidence rate ratios (IRR), which represent the ratio of new to old incidence rates when an explanatory variable increases by 1. As shown in , when the number of self-citations increases by 1, the number of citations by others decreases by an average of 9.86%, holding other variables constant. When the number of self-citations received increases by 1, citations by others increase by an average of 12.82%. Thus, when a document is cited by its author, its visibility increases, raising the likelihood of being cited by others. Conversely, increased author self-citing may indicate lower document quality or bias, resulting in fewer citations by others.

5 Conclusions

The analysis yields the following conclusions:

5.1 Overall Low Level of Self-Citation in Chinese Library and Information Science

Analysis of 73,160 articles published between 1955-2017 reveals that Chinese library and information science researchers have an average self-citing rate of 1.82% and an average self-cited rate of 2.05%. Overall, 4.50% of all citations are author self-citations, and 2.97% of all citations received are author self-citations. Among all documents, 17.47% contain author self-citations, and 18.72% are cited by their authors. Compared with previous research [10, 14-15], self-citation levels in Chinese library and information science are relatively low. For example, H. Snyder and S. Bonzi's study of journal literature from multiple fields in humanities, social sciences, and natural sciences (1980-1989) found that 9% of all citations were self-citations, with approximately 6% in social sciences [16]. Although self-citation levels vary by discipline and time period, Chinese library and information science appears to maintain generally reasonable self-citation levels.

5.2 Positive Correlation Between Self-Citation Frequency and Publication Count, Collaboration Frequency, and Career Length

These phenomena likely result from the strong positive correlation between self-citation frequency and author productivity. Higher productivity creates more

opportunities for self-citation and may strengthen self-citation due to research continuity and relevance. Therefore, absolute self-citation frequency should not be used to judge whether authors engage in excessive or appropriate self-citation.

5.3 Author Order Correlates with Self-Citation in Co-Authored Papers

For papers of various collaboration sizes, the first author's self-citations consistently represent the highest proportion of all self-citations in the document, suggesting that first authors may have greater cumulative research advantages and play more important roles [20]. As collaboration size increases, the first author's proportion of self-citations decreases, likely due to the distribution of research work. Thus, self-citation proportions in co-authored papers may partially reflect individual authors' contributions.

5.4 Document Self-Citation Correlates with Time Since Publication

Author self-citations occur on average 2.65 years after publication, peaking in the second year and decreasing thereafter. This relatively short interval suggests that self-citation in Chinese library and information science is largely related to research inheritance and relevance.

5.5 Limited Impact of Self-Citation on h-Index Values and Rankings

For all authors in the dataset, removing self-citations decreased h-indices by an average of 1.83%. For the top 20 most-cited authors, the average decrease was 1.77%, with some scholars experiencing minor ranking changes. Thus, self-citation has a limited impact on h-index for both highly-cited and ordinary authors in this field.

5.6 Author Self-Citing and Self-Cited Status Affect Citations by Others

Under controlled conditions for author count, journal impact factor, average author h-index, and publication year, each additional self-citation decreases citations by others by 9.86% on average, while each additional self-citation received increases citations by others by 12.82% on average. This suggests that when authors self-cite, the diverse motivations may include improper self-citation, potentially reducing overall document quality and subsequent citations. When a document is self-cited, its visibility and influence increase, making it more likely to be noticed and cited by others.

In summary, author self-citation is not common in Chinese library and information science, and excessive self-citation is likely rare. The impact of author self-citation on h-index appears small. While the effect of author self-citation on citations by others is real, it is likely also related to document quality. Although

methods exist to identify excessive self-citation or appropriately incorporate self-citation into scientific evaluation [34-38], such approaches have not been widely adopted in practice, and their scientific fairness remains unverified. Given that self-citation correlates with many factors like publication count but has limited impact on h-index and rankings, simple removal of self-citations in scientific evaluation should be opposed. Developing scientific methods to handle self-citation requires further research and validation based on self-citation patterns. Future research should develop toward full-text analysis and functional assessment of citations, intelligently identifying the legitimacy of self-citation behavior through citation function and motivation analysis.

Due to time and technical constraints, this study has limitations, including incomplete data coverage and conclusions limited to the library and information science field, which should be addressed in future research.

References

- [1] Qiu Junping. Statistics and analysis of self-citation in scientific literature [J]. *Information Science*, 1989(6): 16-21.
- [2] Li Shaohong, Hou Jinchuan. Self-citation and self-citation analysis [J]. *Library*, 2001(6): 39-42.
- [3] CLARIVATE. HOPKINSON A. Journal citation reports help: title suppressions [EB/OL]. [2022-05-30]. <https://jcr.help.clarivate.com/Content/title-suppressions.htm>.
- [4] Li Yunjing, Hou Hanqing. The “regulation” of self-citation on impact factors in natural science journals [J]. *Journal of the China Society for Scientific and Technical Information*, 2006, 25(2): 7.
- [5] ANDRADE A, GONZÁLEZ-JONTER, CAMPANARIO J. Journals that increase their impact factor at least fourfold in a few years: the role of journal self-citations [J]. *Scientometrics*, 2009, 80(2): 515-528.
- [6] SCHREIBER M. Self-citation corrections for the Hirsch index [J]. *Europhysics Letters*, 2007, 78(3): 30002.
- [7] GARFIELD E. Journal citation studies XVII: journal self-citation rates—there’s a difference [J]. *Essays of an Information Scientist*, 1974, 52(2): 192-194.
- [8] GARFIELD E. Citation indexing—its theory and application in science, technology, and humanities [M]. New York: John Wiley, 1979: 245.
- [9] Jin Tiecheng. Misuse of academic journal self-citation rates and countermeasures [J]. *Science-Technology & Publication*, 2016(11): 96-98.
- [10] GARFIELD E. Journal citation reports: a bibliometric analysis of reference processed for the 1974 [R/OL]. (1975-01-01) [2022-09-10]. <http://garfield.library.upenn.edu/papers/jcr1975introduction.pdf>.
- [11] GARFIELD E. Is citation analysis a legitimate evaluation tool? [J]. *Scientometrics*, 1979, 1(4): 359-375.
- [12] Pan Yuntao, Wu Yishan. Self-citation and other-citation: endless stories [J]. *Science & Technology Review*, 2007(24): 85.

- [13] WOLFGANG G, BART T, BALÁZS S. A bibliometric approach to the role of author self-citations in scientific communication [J]. *Scientometrics*, 2004, 59(1): 63-77.
- [14] DAVARPANAH M R, AMEL F. Author self-citation pattern in scientific communication [J]. *Scientometrics*, 2011, 37(3): 171-173.
- [15] AKSNESS D W. A macro study of self-citation [J]. *Scientometrics*, 2003, 56(2): 235-246.
- [16] SNYDER H, BONZI S. Patterns of self-citation across disciplines (1980-1989) [J]. *Journal of Information Science*, 1998, 24(6): 431-435.
- [17] KING M M, BERGSTROM C T, CORRELL S J, et al. Men set their own cites high: gender and self-citation across fields and over time [J]. *Socius: Sociological Research for a Dynamic World*, 2017(3).
- [18] MALINIAK D, POWERS R, WALTER B F. The gender citation gap in international relations [J]. *International Organization*, 2013, 67(4): 889-922.
- [19] Liu Xiaomin. Analysis of the relationship between journal article count and journal self-citation [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2010, 21(2): 148-150.
- [20] Jiang Ying, Jin Bihui, Liu Xiaomin. Analysis of author collaboration and self-citation of co-authors in journal articles [J]. *Library and Information Service*, 2000, 44(12): 23-28.
- [21] Li Jianhui, Wang Zhikui, Xu Hong, et al. Quantitative study on the effect of self-citation on impact factors of scientific journals [J]. *Acta Editologica*, 2007(2): 154-157.
- [22] FASSOULAKI A, PARASKEVA A, PAPILAS K, et al. Self-citations in six anesthesia journals and their significance in determining the impact factor [J]. *British Journal of Anaesthesia*, 2000, 84(2): 266-269.
- [23] MIMOUNI M, RATMANSKY M, SACHER Y, et al. Self-citation rate and impact factor in pediatrics [J]. *Scientometrics*, 2016, 108(3): 1455-1460.
- [24] CAMPANARIO J M, GONZÁLEZ L. Journal self-citations that contribute to the impact factor: documents labeled “editorial material” in journals covered by the Science Citation Index [J]. *Scientometrics*, 2006, 69(2): 365-386.
- [25] Wen Fangfang. Zonal comparison and diachronic analysis of the relationship between journal self-citation and impact factor: taking management journals indexed by JCR as an example [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2020, 31(8): 110-117.
- [26] CAMPANARIO J M, MOLINA A. Surviving bad times: the role of citations, self-citations and numbers of citable items in recovery of the journal impact factor after at least four years of continuous decreases [J]. *Scientometrics*, 2009, 81(3): 859-864.
- [27] CAMPANARIO J M. The journal citation reports (SCI edition) with and without journal self-citation [J]. *Scientometrics*, 2011, 88(2): 1699-2407.
- [28] GIRI R. Influence of selected factors in journals' citations [J]. *Aslib Journal of Information Management*, 2019, 71(1): 90-104.
- [29] XIA X D, WU Y W. Journal self-citation analysis of some Chinese sci-tech periodicals [J]. *Serials Review*, 2011, 37(3): 171-173.
- [30] FOWLER J H, AKSNESS D W. Does self-citation pay? [J]. *Scientometrics*,

2007, 72(3): 427-437.

[31] Cha Ying. H-index and paper self-citation: taking Chinese scholars in library and information science as an example [J]. *Library Theory and Practice*, 2008(6): 36-38.

[32] HUANG M H, LIN W Y C. Probing the effect of author self-citations on h-index: a case study of environmental engineering [J]. *Journal of Information Science*, 2011, 37(5): 453-461.

[33] Zhu Qiang, He Jun, Cai Ronghua. *Chinese Core Journals Overview* (7th edition) [M]. Beijing: Peking University Press, 2015.

[34] Yu Liping, Wan Xiaoyun, Ju Chunhua. Research on correcting excessive self-citation of academic journal impact factors: natural impact factor [J]. *Information Studies: Theory & Application*, 2019, 42(11): 62-68.

[35] Liu Xueli, Zhou Zhixin, Fang Hongling, et al. Definition of self-citation rate and excessive self-citation in Chinese medical journals from 2005-2007 [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2009, 20(4): 624-626.

[36] HUMPHREY C, KISELEVA O, SCHLEICHER T. A time-series analysis of the scale of coercive journal self-citation and its effect on impact factors and journal rankings [J]. *European Accounting Review*, 2019, 28(2): 335-369.

[37] Jin Tiecheng. Comparison and analysis of using self-cited rate and 2-year self-cited rate to detect excessive journal self-citation [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2016, 27(9): 949-952.

[38] SZOMSZNOR M, PENDLEBURY D A, ADAMS J. How much is too much? The difference between research influence and self-citation excess [J]. *Scientometrics*, 2020, 123(2): 1119-1147.

Author Contributions:

Li Xinyu: Experimental design, data analysis, manuscript writing;

Lei Jiaqi: Manuscript review and revision;

Bu Yi: Research conceptualization, data preparation, literature review and revision.

A Preliminary Study on Author Self-Citation Behaviors in Library and Information Science in China

Li Xinyu, Lei Jiaqi, Bu Yi

Department of Information Management, Peking University, Beijing 100871

Abstract:

[Purpose/Significance] Author self-citations are a common phenomenon and an important factor to be considered in citation analysis and scientific evaluation. However, how to deal with author self-citations in scientific evaluation has always been controversial. This paper expects to provide suggestions for proper treatment of author self-citations by exploring the principles of author self-citations. [Method/Process] Based on a thorough survey on previous works related to author self-citations, this paper adopted the journal data of library and information science (Chinese literature) and implemented author-

and paper-level analyses of author self-citations by using statistical methods such as regression analysis. [Result/Conclusion] Empirical results show that Chinese library and information science researchers tend to have a low rate of author self-citations and that there is a strong positive correlation between the author self-citation frequency and the number of published articles. The paper also observes that author self-citations have limited impact on h-index and that being self-cited may boost the possibility of citations (by others).

Keywords: author self-citation; citation analysis; library and information science; scientific evaluation

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.