Library Holdings and Scholarly Work Impact Evaluation Postprint

Authors: Xiaojuan Liu, Ma Liang, Wei Yu, Zhao Zhuojing

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] To explore the application of library holdings in evaluating academic works and provide references for academic assessment. [Method/Process] This paper discusses the applicability of library holdings in evaluating the impact of academic works from three aspects: the generation mechanism, acquisition, and application of indicator data, and conducts empirical analysis and discussion using WorldCat as an example. [Results/Conclusion] Library holdings data are typically sourced from national or international library union catalogs. When acquiring data, appropriate retrieval methods should be selected based on the chosen indicator source to accurately obtain library holdings information for works. When applying library holdings, attributes such as the type and geographical distribution of holding libraries should be considered to evaluate the impact of works more comprehensively and accurately.

Full Text

Libcitations and the Evaluation of Scholarly Book Impact

Liu Xiaojuan, Ma Liang, Wei Yu, Zhao Zhuojing School of Government, Beijing Normal University, Beijing 100875

Abstract

[**Purpose/Significance**] This paper explores the application of libcitations in the evaluation of scholarly books to provide references for academic evaluation. [**Method/Process**] The applicability of libcitations in evaluating scholarly book impact is discussed from three aspects: the generation mechanism of indicator data, data acquisition, and application, with empirical analysis and discussion using WorldCat as an example. [**Result/Conclusion**] Libcitation data are typically derived from national or international library union catalogs.

When obtaining data, appropriate retrieval methods should be selected based on the chosen indicator source to accurately acquire collection information for works. When applying libcitations, attributes such as library type and geographic location should be considered to evaluate the impact of works more comprehensively and accurately.

Keywords: scholarly books; libcitations; impact evaluation

As a form of academic literature, scholarly books play a crucial role in disseminating research findings and promoting academic exchange, particularly in the social sciences and arts and humanities where their impact is especially significant [1]. In recent years, as publishing channels have continuously expanded, the number of published scholarly books has grown rapidly, making accurate and effective evaluation of scholarly book impact an increasingly urgent problem to solve—one that is vital for the publication, dissemination, and utilization of scholarly books.

Book evaluation involves analyzing and assessing the intrinsic quality, usage patterns, and developmental characteristics of book literature according to certain standards and methods, with the goal of revealing the inherent objective laws of the literature as a whole or of specific characteristic portions to better realize their academic value and social utility [2]. Unlike journal articles, scholarly books represent scholars' more in-depth, comprehensive, and systematic treatment of a particular discipline or specialized topic, containing far greater amounts of information than journal articles and possessing higher levels of specialization and academic depth. However, while journal articles already have relatively mature impact evaluation systems, scholarly book impact evaluation remains in its infancy due to factors such as lengthy volumes, large quantities, and difficulties in digitalization.

Currently, the evaluation of scholarly book impact primarily relies on qualitative analysis methods such as peer review, which depends on experts' knowledge and experience to assess scholarly books. However, peer review suffers from high time and labor costs and excessive subjectivity, making comprehensive and objective evaluation difficult. As technology and the environment have changed, an increasing number of data sources have become available for scholarly book evaluation, providing data support for quantitative assessment—for example, BKCI and Elsevier's Book Citation Database, with many book review websites also offering rich data. Nevertheless, the substantial academic content and time-intensive nature of scholarly books mean traditional citation indicators are insufficient for evaluating book impact [3-5]. Although book reviews can directly reflect readers' evaluations, academic book reviews suffer from extremely low coverage, while online book reviews lack reliability. More indicators need to be explored to improve scholarly book impact evaluation.

For scholarly books, library collection statistics reflect a different type of impact, yet how to reasonably utilize collection data to evaluate scholarly book

impact remains under exploration. Based on this, we explore the application of libcitations in scholarly book impact evaluation and conduct empirical analysis with data to provide references for scholarly book evaluation work.

2. Origin and Development of the Libcitation Indicator

In 2009, H.D. White et al. [6] coined the term "libcitation" to reflect the impact of scholarly books by the number of libraries that hold a given book in their collections. The introduction of libcitations provided new ideas for the quantitative evaluation of scholarly book impact. G. Halevi et al. [7] demonstrated that books have the highest coverage in libcitations compared to citations and other altmetrics indicators.

A.J.M. Linmans [8] selected books authored by teaching and research staff in the humanities at Leiden University as research subjects and found a certain correlation between book libcitations and WoS citations (Pearson's r = 0.29), though the correlation was not strong, suggesting that libcitations and citations reflect different types of impact—necessitating exploration of the specific meaning behind libcitation values. The study also found that book language affected results, with higher correlations for English-language works but non-significant results for Dutch works. A larger-scale study [9] selected 59,000 history books and 42,000 literature books cited in Scopus-indexed journals, comparing their Scopus citations with libcitations in Association of Research Libraries (ARL) libraries, non-ARL libraries, and all libraries, finding low Spearman correlation coefficients that indicated citations and libcitations measure different dimensions of book impact. A. Zuccala [10] reached similar conclusions and further investigated disciplinary differences, finding higher correlations in social sciences and arts and humanities than in science fields.

Additionally, scholars have examined the relationship between libcitations and book reviews. D. Shaw [11] analyzed 200 novels and found a high correlation (r = 0.620) between the number of reviews in Book Review Index and libcitations in the OCLC (Online Computer Library Center) database. K. Kousha [12] found a low but significant positive correlation between academic book review ratings in *Choice: Current Reviews for Academic Libraries* and WorldCat libcitations in social sciences, while the correlation increased significantly in science fields.

Domestic research on libcitations is limited, with most studies using libcitations as a standard for measuring book impact. Zhang Yu et al. [13] categorized 27 indicators—including citation frequency, number of book reviews, and sales volume—into four dimensions (citations, funding and awards, online book reviews, and book utilization) to construct an evaluation system, using CALIS (China Academic Library & Information System) libcitations to validate evaluation results.

In summary, although libcitations were introduced nearly a decade ago, current research has primarily focused on correlation analysis between libcitations and other indicators such as citations and book reviews. How libcitations reflect

book impact and how to reasonably utilize them in impact evaluation remain unclear and require in-depth exploration.

3. Applicability Analysis of Libcitations in Scholarly Book Impact Evaluation

Scientific and reasonable evaluation of scholarly book impact must first clarify the composition of such impact. Research impact manifests at two levels: academic impact and social impact [3, 14]. As one form of research output, the knowledge and ideas carried by scholarly books are disseminated through various channels to audiences, who undergo changes in thinking or behavior according to their needs and purposes, ultimately contributing to disciplinary or social development and progress—manifesting as academic and social impact respectively, which together constitute scholarly book impact. The influence of scholarly books on people occurs mostly imperceptibly, with reading representing the dissemination of academic ideas at the cognitive level.

Libraries, as specialized institutions that collect, organize, preserve, and disseminate literature, are primary venues for accessing and reading books. Libcitations can measure the dissemination of works; similar to citations, they represent libraries' "citation" behavior toward works—the number of libraries holding a work is analogous to the number of scholars citing it [15]. However, the motivations behind libraries' and scholars' "citation" of works differ. Meeting reader needs is the primary mission of library collection development. Facing today's vast book publishing landscape, major libraries and library consortia have established standardized acquisition systems following important principles including subject classification, language, genre, and publisher reputation. In collection development, librarians draw on their knowledge and experience, fully considering recommendations from experts and readers to select books that meet user needs. Although libraries typically use bulk purchasing rather than analyzing each work's individual value, librarians consider collection development principles, library budgets, and actual and potential user groups when ordering works to meet user expectations. Therefore, libraries' "citation" of works primarily serves to meet user demands.

Consequently, the magnitude of a scholarly book's libcitations can reflect the number of libraries that recognize its value, indirectly indicating the size of its potential user base and demonstrating its impact scope. It should be emphasized that after a work is collected by libraries, whether and how often it is borrowed by readers is difficult to comprehensively capture. Borrowing statistics only reflect partial actual usage, as readers may consult books in-library without borrowing them—usage and impact that cannot be accurately measured. Given this, although libcitations cannot precisely reflect actual impact, they can indicate potential impact scope and thus hold certain significance. Therefore, it can be concluded that the greater the libcitation count, the broader the impact scope and the greater the impact of the scholarly book.

In terms of data availability, libcitation data can be obtained from national or international library union catalogs. Similar to citation data, libcitation values vary significantly depending on the data source. Different union catalogs exhibit biases in library selection regarding geography, language, and library type. Taking China Academic Library & Information System (CALIS) as an example, it only selects domestic university libraries as member institutions, thus limiting its reflected impact scope to among Chinese universities.

When acquiring libcitation data, the first step is identifying appropriate search fields. Generally, a work's ISBN serves as the retrieval entry point. However, the same work may exist in different versions due to varying publication formats and media, resulting in non-unique ISBNs. Different union catalogs lack unified approaches to handling version issues. In WorldCat, for instance, bibliographic records are contributed by OCLC members, with OCLC management integrating data from multiple sources. Although data quality specialists clean and review member-submitted data, searching with a single ISBN may still yield multiple results because different member libraries describe works differently, and WorldCat integration does not completely eliminate these differences, leading to version ambiguity. While our investigation has not fully clarified WorldCat's approach to book versions, analysis of search results confirms that libraries in different result records are distinct; therefore, libcitation counts should be aggregated across multiple search results. Similarly, when using data sources other than WorldCat, appropriate acquisition methods should be adopted based on their data description and retrieval interface characteristics.

Finally, regarding application of libcitations in scholarly book impact evaluation, some citation analysis frameworks can be directly adapted for libcitation analysis. For example, introducing disciplinary difference analysis from citation studies into libcitation research allows more objective assessment of a work's status and impact based on its disciplinary libcitation distribution.

Unlike citations, a libcitation record reflects a scholarly book's impact on library users, with libraries' basic attributes and collection development goals directly affecting user group composition—that is, the book's audience. For instance, regarding library type, the American Library Association (ALA) classifies libraries into four categories based on served groups and functions: academic libraries, public libraries, school libraries, and special libraries [16]. User needs are the fundamental determinant of library survival and development; all library operations must center on users' information needs, with satisfying these needs as the organization's primary purpose [17].

Analysis of library types can further identify audience needs and purposes, thereby clarifying how scholarly book impact manifests across different dimensions. Additionally, the geographic distribution of holding libraries can reflect a work's impact scope. In communication studies, cultural artifacts' reach in dissemination areas typically measures their spread and cultural influence. The number of library systems in a country or region that hold a given book represents that book's impact in that country or region. Comparing library charac-

teristics and distribution across countries can further reveal differences in how scholarly books generate impact internationally. OCLC libraries are distributed across 175 countries worldwide, with over 60% in the United States and Chinese libraries comprising only 1.96%.

Therefore, when applying libcitations to evaluate scholarly book impact, analyzing the characteristics of holding libraries can further clarify specific impact situations. Based on WorldCat libcitation data for scholarly books, we conducted empirical research to deeply explore its application in scholarly book impact evaluation, aiming to provide scientifically sound recommendations for applying this indicator.

4. Empirical Analysis and Discussion

4.1 Research Design

Libcitation values can only roughly reflect scholarly book impact, while the hidden content behind these values is also crucial for application—for example, when facing two works with identical libcitation counts, how to accurately discriminate between their impact differences. Based on the data generation mechanism of libcitations, we extracted library type and geographic location, representing audience needs and geographic characteristics respectively, to deeply analyze the impact revealed by libcitations, including impact composition and coverage scope. We discuss the application of libcitations in scholarly book evaluation, specifically including: (1) distribution of libcitation counts for scholarly books; (2) distribution of scholarly books across library types; and (3) differences in libcitations across countries/regions.

4.2 Dataset Description

We used BKCI as the source database for scholarly books, retrieving all academic works published between 2005-2017 in BKCI's two sub-databases, yielding 1,146,217 records, including 81,642 book-type records. After preliminary processing, we obtained valid data for 77,919 scholarly books and their citation data. To quickly and effectively determine the value of libcitations in revealing scholarly book impact, we selected highly-cited scholarly books as our experimental dataset, as these works are generally considered to have high academic value and impact in citation evaluation systems. According to Price's Law, we ultimately identified 9,737 highly-cited works as our dataset; disciplinary thresholds for highly-cited scholarly books and the number of highly-cited works are shown in Table 1 .

Among all 9,737 highly-cited scholarly books, 7,798 were held by at least one library, distributed across 9,729 libraries in 100 countries worldwide.

WorldCat is a large-scale online cataloging database organized by OCLC (Online Computer Library Center) with participation from nearly 20,000 libraries

worldwide, possessing the world's largest and most comprehensive online cataloging system, currently searchable for collection information from over 65,000 libraries in more than 170 countries. We used WorldCat as the source database for libcitation information, developing a crawler program that used ISBN as the retrieval entry point to obtain libcitation data for scholarly books.

4.3 Distribution of Libcitation Counts for Scholarly Books

Libcitation counts can reflect the overall impact magnitude of scholarly books worldwide. Figure 1 [Figure 1: see original paper] calculates the annual changes in average libcitation counts for highly-cited scholarly books across disciplines. The figure shows that arts and humanities and social sciences books exhibit similar trends, displaying cyclical patterns of increase followed by decrease, with peaks in 2007 and 2013. Publication year has less impact on natural sciences, applied sciences, and life sciences/biomedicine works, with libcitation counts slightly decreasing as publication year increases.

In terms of disciplinary differences, arts and humanities works have the highest libcitation counts, followed by social sciences, significantly higher than the other three disciplines. This may be because users in natural sciences, applied sciences, and life sciences/biomedicine prefer accessing the latest technologies or theories, where electronic resources update far faster than print library books, leading users to prefer electronic resources and indirectly causing lower libcitation counts. According to an EBSCO survey report based on 128 academic libraries [19], libraries have increasingly allocated procurement funds toward electronic resources in recent years, particularly in applied and natural sciences.

4.4 Distribution of Scholarly Books Across Library Types

Further analysis of holding library types can clarify the composition of scholarly book impact. OCLC subdivides libraries into 12 types based on ALA definitions; this study adopts ALA's library type classification. The comparison between OCLC and ALA library classifications is shown in Table 2 .

Table 3 calculates the proportion of scholarly books across different library types by discipline, where the data represent the percentage of works in a discipline held by at least one specific library type relative to the total works in that discipline. Larger proportions indicate more scholarly books generating impact among specific user groups in that discipline.

The statistical results show that nearly all scholarly books across disciplines are simultaneously held by at least one academic library and one public library. Over 99% of scholarly books are held by at least one academic library, while proportions in school libraries are uniformly low. In special libraries, disciplinary proportions are slightly lower than in academic and public libraries, with social sciences showing the highest proportion in special libraries at 94.32%. Arts and humanities works have slightly higher proportions in public libraries than other

disciplines, while proportions in special and school libraries are significantly lower than in other disciplines.

Figure 2 [Figure 2: see original paper] calculates the distribution of holding libraries across types for scholarly books by discipline, enabling further comparison of impact magnitude among different user groups. Overall, among highly-cited scholarly books' holding libraries, academic libraries account for the highest proportion, followed by special libraries, with school libraries lowest. This relates to the positioning and served populations of different library types: school libraries primarily serve K-12 students, for whom scholarly books clearly do not meet needs.

By definition, academic libraries primarily serve researchers with professional backgrounds, whose reading purposes focus on knowledge accumulation for research; the distribution of works in academic libraries thus reflects academic impact more strongly. The other three library types serve primarily K-12 students, government personnel, doctors, private enterprise employees, and other non-researchers whose needs are mostly unrelated to research, tending to reflect social impact dimensions.

Across disciplines, nearly all scholarly books generate impact in both academic and social dimensions. However, regarding impact magnitude, academic libraries account for 50.71% and 50.28% of holding libraries in social sciences and arts and humanities respectively, indicating roughly equal distribution across both dimensions. In natural sciences, applied sciences, and life sciences/biomedicine, academic impact significantly exceeds social impact among the general public.

4.5 Differences in Libcitations Across Countries/Regions

Figure 3 [Figure 3: see original paper] shows annual changes in the average number of countries covered by scholarly books across disciplines. Overall, the number of countries where works are distributed decreases year by year. In 2017, only one scholarly book was distributed across 11 countries in 58 libraries; this data point's 偶然性 (randomness) caused a sharp increase in the 2017 trend. Disciplinary curves follow the overall trend but differ significantly in magnitude. Arts and humanities and social sciences works are distributed across significantly more countries than the other three disciplines, with smaller annual fluctuations, averaging around 20 countries per work and demonstrating broader geographic impact scope.

Analysis reveals that the distribution of scholarly books across countries follows the 80/20 rule: most works are concentrated in a few countries. The magnitude and composition of impact vary significantly across countries.

Table 4 calculates total and average libcitation counts for scholarly books across countries. In the table, "work count" refers to the number of highly-cited works held by a specific country; "libcitation count" refers to the number of holding

libraries in that country; and "libcitation proportion" is the ratio of a country's libcitations to the number of libraries from that country in OCLC. Countries are sorted descending by work count, showing only the top 20.

In the United States, 6,034 libraries hold 7,196 highly-cited works, with an average of 107.24 libcitations per work—far greater impact than other countries. However, the libcitation proportion is 13.62%, indicating the relative impact scope is lower than in 60% of listed countries. In China, although the number of held works is similar to the U.S., libcitations are far smaller, with only 128 libraries holding relevant works and an average of merely 4.69 libcitations per work. The libcitation proportion is 10.01%, not far from the U.S., indicating that in China, scholarly books impact only a small portion of user groups, with far less impact than in the U.S.

In France, scholarly books have the largest impact scope, with a libcitation proportion exceeding 40% and average impact per work of 4.27, slightly less than China. Sweden shows the smallest average libcitations, with works held by an average of only one library and a libcitation proportion of 6.42%, far below other countries.

Additionally, the composition of scholarly book impact differs significantly across countries. Table 5 shows the distribution of holding library types for highly-cited works across countries (note: if multiple works are held by the same library, that library is counted only once). The table displays the top 20 countries by work count. Results show significant variation in library type distribution across countries. In New Zealand and Australia, special libraries account for far larger proportions than other types, while in the United Kingdom, public libraries constitute the largest number of holding libraries. In other countries, academic libraries represent the highest proportion of holding libraries.

Regarding impact composition, in New Zealand, Australia, the United Kingdom, the United States, South Africa, and the Netherlands, scholarly books' academic impact is significantly lower than their social impact. In Canada, Singapore, and Denmark, impact is roughly equal across both dimensions. In other countries, academic impact is higher, particularly in Italy and France where only about 5% of works are distributed in non-academic libraries.

In summary, analyzing holding libraries' countries can further determine the geographic scope of scholarly book impact. Comparing library type distributions across countries can reveal differences in how scholarly books generate impact internationally. Results show that social sciences and arts and humanities works have broader geographic impact, while distribution patterns across countries are similar across disciplines. Moreover, impact magnitude varies considerably across countries, especially in the United States where average libcitations far exceed other nations. Comparing impact composition across countries reveals that in most nations, scholarly books generate greater impact within academia.

We have explored the applicability of libcitations in scholarly book impact eval-

uation. Through empirical analysis of libcitation information for highly-cited scholarly books, we have clarified the application of libcitations in book evaluation and, combined with experimental data, summarized how to apply the libcitation indicator in scholarly book impact evaluation.

First, libcitation indicator data sources originate from library union catalogs across countries and regions. Data source selection and quality directly affect evaluation result validity and accuracy, with different indicator platforms reflecting impact differently based on their functions and positioning. Therefore, appropriate data sources should be selected according to specific usage scenarios and evaluation purposes. Additionally, although evaluation results derive from objective data, human intervention exists in data source selection, indicator determination, and data acquisition processes; thus, quantitative evaluation results should serve only as one reference for final evaluation.

Second, when acquiring libcitation data, appropriate retrieval entry points should be selected based on source characteristics. Since identical works exist in different versions and union catalogs lack unified approaches to version handling, data acquisition should be based on thorough investigation of source databases to ensure indicator data scientificity and accuracy.

Finally, when applying libcitations for scholarly book impact evaluation, disciplinary differences should be fully considered. Simultaneously, attention to holding libraries' geographic location, reader groups, collection development goals, and related characteristics helps clarify scholarly book impact composition, enabling more accurate and comprehensive evaluation.

References

[1] GARFIELD E, EMERITUS C. Citation indexes for retrieval and research evaluation [EB/OL]. [2018-10-15]. http://www.garfield.library.upenn.edu/papers/cirtreseval-capri.html. [2] He Jun. Analysis of current status of book evaluation in China [J]. Journal of Academic Libraries, 2012(3): 106-110. [3] SMITH R. Measuring the social impact of research [J]. BMJ, 2001, 323(7312): 528. [4] SARLICC C, DUBINSKY EK, HOLMES KL. Beyond citation analysis: a model for assessment of research impact [J]. Journal of the Medical Library Association Jmla, 2010, 98(1): 17-23. [5] WILBERTZ J. Evaluating societal relevance of research [EB/OL]. [2018-10-15]. https://www.rug.nl/research/portal/files/14412466/Evaluating-societal-relevance.pdf. [6] WHITE HD, BOELLSK, YU H, et al. Libcitations: a measure for comparative assessment of book publications in the humanities and social sciences [J]. Journal of the American Society for Information Science & Technology, 2009, 60(6): 1083-1096. [7] HALEVI G, NICOLAS B, BAR-ILAN J. The complexity of measuring the impact of books [J]. Publishing research quarterly, 2016, 32(3): 1-14. [8] LINMANS AJM. Why with bibliometrics the humanities does not need to be the weakest link [J]. Scientometrics, 2010, 83(2): 337-354. [9] ZUCCALA A, GUNS R. Comparing

book citations in humanities journals to library holdings: Scholarly use versus 'perceived cultural benefit' (RIP) [C]//14th international society of scientometrics and informetrics conference. Wien: Facultas Verlags-und Buchhandels AG, 2013: 353-360. [10] ZUCCALA A. Correlating libcitations and citations in the humanities with worldcat and scopus data [EB/OL]. [2018-10-15]. https://www.researchgate.net/publication/303286873_{{{Correlating}}}{{libcitations}}{{{and}}}{{citations}
[11] SHAW D. An analysis of the relationship between book reviews and fiction holdings in OCLC [J]. Library & information science research, 1991, 13(2): 147-154. [12] KOUSHA K, THELWALL M. Alternative metrics for book impact assessment: can choice reviews be a useful source? [EB/OL]. [2018-10-15]. http://pdfs.semanticscholar.org/c69a/38d1ac5bafd75a3f54411452e8ac6d5f79d.pdf.
[13] Zhang Yu, Pan Yuntao, Yuan Junpeng, et al. On constructing a multidimensional evaluation system for Chinese scientific books' academic impact [J]. Library and Information Service, 2015, 59(7): 69-76. [14] Yu Houqiang, Qiu Junping. Theoretical research on stratification and aggregation of altmetrics indicators [J]. Library Journal, 2014, 33(10): 13-19. [15] WHITE HD, ZUCCALA AA. Libcitations, WorldCat, cultural impact, and fame [J]. Journal of the Association for Information Science and Technology, 2018, 69(12): 1502-1512. [16] American Library Association. Types of libraries [EB/OL]. [2018-10-15]. http://www.ala.org/educationcareers/libcareers/type. [17] Xian Peng. On the essential attributes of library type classification—also on the development of Party School libraries toward research-oriented special libraries [J]. Information and Documentation Services, 2002(6): 48-51. [18] Release of "Research Report on the Impact of Chinese Books in Overseas Library Collections (2016 Edition)" [J]. Journal of the National Library of China, 2016(5): 10. [19] A survey of collection development trends in academic libraries [EB/OL]. [2018-10-15]. https://www.ebsco.com/blog/article/collection-development-trends-in-academic-libraries.

Author Contributions: Liu Xiaojuan: Responsible for topic selection, research direction, and basic framework, guiding and revising the paper; Ma Liang: Responsible for data acquisition and analysis, drafting the paper; Wei Yu: Participated in data analysis and paper writing; Zhao Zhuojing: Participated in data analysis and paper writing.

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