Postprint: Adaptability Study of a Comprehensive Evaluation System for Chinese Books Incorporating Review Indicators

Authors: Zhu Shiqin, Qiu Yue, Chen Hongying

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] To address the deficiencies of current book evaluation systems in reflecting book content quality, this study incorporates book review metrics into the evaluation framework, analyzes the adaptability of the evaluation system across different disciplines, and explores Chinese book quality evaluation. [Method/Process] Sixteen indicators are selected from traditional bibliometric indicators, Altmetrics indicators, and book review indicators to construct a comprehensive Chinese book evaluation system weighted by the CRITIC method. Concurrently, empirical research is conducted on books from economics, computer science, and physics to analyze the utility of book review indicators in book quality evaluation and to examine the system' s adaptability across different disciplines. [Results/Conclusion] Empirical results indicate that the comprehensive Chinese book evaluation system demonstrates favorable adaptability for books across different disciplines, with evaluation results that can supplement the influence and quality of books as reflected by total citations and library holdings. Book review indicators can reflect readers' reading preferences for books in different disciplines, facilitating the identification of disciplinary books with high readability, scientific rigor, and knowledge content, thereby providing references for book quality evaluation activities such as reading promotion and bibliography recommendation.

Full Text

Research on the Adaptability of a Comprehensive Chinese Book Evaluation System Incorporating Review Indicators

Zhu Shiqin, Qiu Yue, Chen Hongying
Institute of Science and Technology Information, East China University of Science and Technology, Shanghai 200237

Abstract: *[Purpose/Significance]* To address the limitations of current book evaluation systems in reflecting book content quality, this study integrates book review indicators into the evaluation framework and applies it to books from different disciplines to analyze the system' s adaptability and explore quality evaluation for Chinese books. *[Method/Process]* Sixteen indicators were selected from traditional bibliometric metrics, Altmetrics, and book review indicators to construct a comprehensive Chinese book evaluation system weighted by the CRITIC method. Economics, computer science, and physics books were chosen for empirical research to analyze the utility of book review indicators in book quality evaluation and test the system' s adaptability across different disciplines. *[Results/Conclusions]* Empirical results demonstrate that the comprehensive Chinese book evaluation system exhibits good adaptability for books across different disciplines. The evaluation results can supplement book influence and quality as reflected by total citations and library holdings. Book review indicators can reflect reading preferences of readers from different disciplines, helping to identify books with high readability, scientific rigor, and knowledge value, thus providing references for book quality evaluation activities such as reading promotion and bibliographic recommendation.

Keywords: book evaluation; quality evaluation; book review; comprehensive evaluation system; adaptability

---

2. Research Status of Book Evaluation

Books are knowledge-intensive academic achievements that play a crucial role in cultural and knowledge inheritance. With the continuous growth in book publishing volume, how to evaluate books scientifically and effectively has become an urgent problem for publishing professionals and scholars. Unlike journals, book evaluation has lagged in development due to constraints such as carrier format, publication timing, and content length. While the internet has greatly enriched book evaluation indicators, existing research predominantly relies on external characteristics such as citation counts, borrowing volumes, and social media mentions, which only reflect book popularity and lack qualitative evaluation metrics that reveal content quality. This study constructs a comprehensive Chinese book evaluation system that incorporates book review indicators alongside traditional bibliometric and Altmetrics indicators. By processing reader reviews through sentiment analysis and attribute mining, the system reveals book quality through reader feedback on content, supplementing book influence evaluation while exposing content quality, thereby reducing the one-sidedness of single-type indicator evaluation and providing a comprehensive, feasible approach to Chinese book quality assessment.

2.1 Research on Book Evaluation Based on Traditional Bibliometric Indicators

Recognizing that citation indexes based solely on journals may be insufficient for evaluating citation impact across disciplines, K. Kousha et al. studied online citation data for books in social sciences, arts, and humanities, finding that such data could support citation impact assessment for these disciplines. Library holdings, representing the number of libraries holding a specified book, reflect librarians' selection of books and can reveal influence different from citations, compensating for limitations in citation analysis to some extent. Additionally, L. Cabezas-Clavijo et al. found no significant correlation between borrowing volume and citation counts, suggesting that borrowing volume can provide supplementary evaluation of citation impact.

2.2 Research on Book Evaluation Based on Altmetrics Indicators

The introduction of Altmetrics has attracted widespread attention from scholars, improving upon the time lag and Matthew effect inherent in traditional bibliometric indicators and providing new directions for book evaluation. A statistical analysis of Altmetrics coverage in humanities publications found that Twitter could serve as a reliable source of Altmetrics data, providing social media mention data for humanities books. Regarding the application value of Altmetrics indicators, A. A. Zuccala et al. found that Goodreads reader ratings could reveal the influence of history books. Moreover, significant correlations between downloads and citations across different disciplines (pure science, social science, engineering, and medicine) indicated that books with high download volumes have substantial reference value. X. Nan et al. studied citation indicators and Altmetrics from Douban Reading for 1,000 highly-cited books in CBKCI, demonstrating that Altmetrics and citation indicators reflect different aspects of book influence and complement each other in academic evaluation.

2.3 Research on Book Evaluation Based on Book Reviews

In research on book evaluation from the review perspective, K. Kousha et al. found that online book reviews tend to reflect the social impact of books in education and culture rather than academic impact. A study of Goodreads book indicators similarly showed that book reviews can reflect broader influence. Domestic scholars have conducted limited research based on book review content. Zhang Chengzhi et al. analyzed differences in book influence evaluation using review data from single versus multiple platforms, finding that integrating reader reviews from e-commerce and social media platforms enables more comprehensive book influence evaluation.

2.4 Research on Book Evaluation Systems

As book evaluation research has deepened, scholars have constructed book evaluation systems by combining indicators from different sources. Li Yingling

et al. proposed a five-dimensional book evaluation system and analysis model based on author, publisher, library, sales, and network public opinion. Liu Li et al. used expert surveys to obtain book quality evaluation indicators and weights, constructing a foreign academic book quality evaluation model based on core authors, core publishers, author affiliations, book editions, and review utilization. Yang Yuli et al. used the Harmonic method to calculate co-authors' contributions to books, comprehensively considering author academic reputation, book publication and circulation, and book reviews to construct a foreign book quality evaluation system. He Jun et al. used the analytic hierarchy process, expert surveys, and fuzzy membership analysis to establish a Chinese book comprehensive evaluation system from content quality and editorial quality perspectives, identifying core and extended books by discipline for reference. Xiao Aoxia et al. applied Bookmetrix indicator data to evaluate academic book influence, finding that Bookmetrix data has supplementary value but also noting that simple frequency statistics can distort academic impact and value, with significant disciplinary differences, suggesting that future research should incorporate more indicators.

2.5 Limitations of Existing Book Evaluation Research

In summary, existing book evaluation research primarily relies on readily available quantitative indicators such as citations, borrowing volumes, library holdings, and Altmetrics data, with only a few studies conducting fine-grained mining of content data. This limitation in research data means few studies can achieve book quality evaluation, particularly comprehensive evaluation reflecting book quality. This study adds review indicators to bibliometric metrics, using sentiment analysis and attribute mining to process book review data, revealing book quality through attributes and sentiments in reviews, and conducting empirical analysis with books from different disciplines to test the system' s adaptability.

---

3. Comprehensive Chinese Book Evaluation System

3.1 Indicator Selection Principles

In constructing the comprehensive Chinese book evaluation system, three principles guided indicator selection:

1. **Drawing on established indicators from existing research.** Considering that indicators from previous studies often have high research value and can inform selection, chosen indicators should be based on existing research findings to ensure scientific validity and reliability.
2. **Covering all aspects of book quality evaluation.** To ensure the system can identify books with high influence and quality, selected indicators should reflect book value from different levels and perspectives.

Simultaneously, indicators should complement each other to guarantee comprehensive book quality evaluation.

3. **Ensuring data accessibility.** Data availability is crucial for maintaining the evaluation system. The selection process should consider whether data sources are reliable and collection is convenient to ensure complete data acquisition.

3.2 System Construction

Based on these principles, sixteen secondary indicators were selected from traditional bibliometric indicators, Altmetrics, and book review indicators. These were categorized into three dimensions from a data source perspective: book utilization, book attention, and book reviews, forming the comprehensive Chinese book evaluation system shown in [Figure 1: see original paper].

Among the primary indicators, book utilization indicators consist of traditional bibliometric metrics that reflect academic influence, including frequently used total citations, library holdings, and borrowing counts, as well as average annual citations to reflect citation patterns for recently published books. These indicators represent usage data from academic utilization, collection, and reading processes, reflecting recognition in academic and library communities and popularity among readers. Data sources were selected as Duxiu website and Zhejiang Library to ensure data availability.

Book attention indicators reflect social influence, comprising Altmetrics such as Douban rating, Douban short review count, readers, want-to-read count, other editions, WeChat and Weibo mentions, Baidu inclusion count, average video play count, and average audio play count. These represent behavioral data generated when readers access book information in social network environments, sourced from Douban Reading, WeChat, Weibo, Baidu, Bilibili, and Ximalaya. As China' s largest reading sharing platform with high user activity and recognition, Douban Reading provides rating, review count, reader numbers, want-to-read numbers, and edition counts that present reader recommendation levels, dissemination depth and breadth, and market popularity. WeChat and Weibo mentions reflect popularity trends among general readers, compensating for the time lag in citation and borrowing indicators. Baidu inclusion count reflects the influence of electronic book versions mentioned in Baidu search engine. Video and audio average play counts are new indicators reflecting the impact of film and audiobook media on book content dissemination in new media reading contexts, enabling more comprehensive influence measurement.

Book review indicators reveal book quality. Reviews were segmented using sentence-ending punctuation marks (“。” , “.” , “!” , “?” , “...” , etc.). Sentiment orientation of segmented reviews was analyzed using the Baidu Sentiment Analysis API, and frequencies of content, author, and publication attribute words in segmented reviews were counted separately. When different attribute words for the same category appeared in a segmented sentence, they were not counted

repeatedly (examples shown in). Final indicator data represent the difference between positive and negative review frequencies for three attribute categories. Review attribute words were set based on existing research and high-frequency word distributions in current reviews, categorized into content, author, and publication attributes. Content attributes include words related to book content such as "part," "theory," "feeling," "content," "thinking," "viewpoint," etc. Author attributes consist of "author," "language," "writing style," "text," etc. Publication attributes cover editorial and publishing aspects including "editing," "publishing," "translation," "translator," "price," "packaging," "printing," "font," "cover," "paper," "edition," "illustration," etc., aiming to reflect reader opinions on content quality, author writing style, and publication quality.

3.3 CRITIC Method

This study employs the CRITIC method, an objective weighting approach that incorporates indicator data variability and inter-indicator correlation into weight calculation. This method is superior to entropy weight and standard deviation methods, producing weights that better reflect objective reality and reduce subjective interference in book quality evaluation.

When applying the CRITIC method to calculate indicator weights, conflict is first calculated based on inter-indicator correlation coefficients. Let C_j represent the conflict between indicator j and other indicators:

$$C_j = \sum_{k=1}^{n} (1 - r_{jk}) \quad \text{(Formula 1)}$$

The information quantity of the indicator is the product of conflict and standard deviation:

$$I_j = \delta_j C_j \quad \text{(Formula 2)}$$

The final indicator weight W_j is:

$$W_j = \frac{I_j}{\sum_{j=1}^{n} I_j} \quad \text{(Formula 3)}$$

Since the data in this study did not follow a normal distribution, Spearman correlation analysis was used to obtain correlation coefficients between indicators.

4. Empirical Analysis

4.1 Data Sources

To explore the evaluation effects of the Chinese book comprehensive evaluation system across different disciplines and test its adaptability, books from three research fields with distinct differences were selected: economics, computer science, and physics, representing social sciences, information sciences, and natural sciences respectively.

On February 1, 2020, Python programs were used to crawl books tagged "economics," "computer science," and "physics" on Douban Reading, selecting the top 20% of books in each discipline by comprehensive ranking, yielding 199 economics books, 200 computer science books, and 200 physics books.

By February 10, 2020, complete indicator data for the three disciplines were crawled. After removing duplicate books, books with missing data, and collections, 140 economics books, 137 computer science books, and 136 physics books were retained for analysis.

4.2 Indicator Weight Calculation

Secondary indicator weights for different disciplines were calculated using the CRITIC method:

Economics: $\omega_1 = (0.0325, 0.0327, 0.0785, 0.0591, 0.1158, 0.0442, 0.0449, 0.0598, 0.0439, 0.0651, 0.0438, 0.0546, 0.0$

Computer Science: $\omega_2 = (0.0414, 0.0332, 0.0967, 0.0851, 0.1047, 0.032, 0.0342, 0.0449, 0.1117, 0.0594, 0.0411, 0.05$

Physics: $\omega_3 = (0.0451, 0.0438, 0.0983, 0.0844, 0.087, 0.042, 0.0323, 0.0513, 0.0795, 0.0365, 0.0635, 0.0649, 0.0624, 0$

4.3 Evaluation Results Analysis

Indicator data were normalized using min-max standardization:

$$x'_{ij} = \frac{x_{ij} - \min_{1\le i\le m} x_{ij}}{\max_{1\le i\le m} x_{ij} - \min_{1\le i\le m} x_{ij}} \quad i = 1, 2, \dots, m;\ j = 1, 2, \dots, n \quad \text{(Formula 4)}$$

where n is the number of evaluation indicators and m is the number of evaluation objects.

4.3.1 Comprehensive Score Analysis Normalized book data were input into the evaluation system to obtain comprehensive score distributions for the three disciplines, shown in [Figure 2: see original paper].

The comprehensive score curves for all three disciplines exhibit long-tail distributions. Approximately 20% of books cluster in the 0.3-0.5 score range, where curves decline rapidly with significant inter-book score differences across disciplines. Most books score between 0.1-0.3, with economics books scoring higher

than computer science and physics books in this range, and score curves showing gentler slopes with smaller inter-book differences. Overall, the evaluation performance follows the market principle that "few books receive substantial attention while most perform averagely."

The Top 10 books across the three disciplines (see) show that most have high book attention scores, indicating that top-ranked books receive considerable attention from reading platforms, social media, and audio/video platforms. Book topics are relatively close to people's daily lives: economics books cover economic history, concepts, and financial management, showing readers' tendency to master economic patterns and obtain investment advice; computer science books involve hardware/software, mathematical thinking, and biographies, mostly introducing practical programming languages, algorithms, and operating systems, reflecting readers' primary goal of acquiring technical knowledge; physics books involve astrophysics, quantum theory, physics history, and basic physics, indicating readers' interest in non-common-sense physics knowledge and unknown fields like the universe.

4.3.2 Book Review Indicator Analysis To explore the application value of book review indicators, book review score distributions were plotted as shown in [Figure 3: see original paper]. Economics books show significantly higher review scores than computer science and physics books, indicating that review indicators have greater influence on economics books. The review indicator weights and short review quantities for this discipline rank among the top three. Meanwhile, review scores for the top 50% of books are relatively concentrated, with subsequent books showing decreasing trends, suggesting that books with higher review scores also have higher comprehensive scores, enabling the system to identify reader-approved books.

4.3.3 Impact of Book Review Indicators on Book Rankings When indicator data were input into an evaluation system without book review indicators, over 94% of books across all three disciplines experienced ranking changes after adding review indicators. The Top 10 books with the largest ranking differences are shown in .

Analysis of reviews for books with significantly changed rankings reveals that ranking declines show no disciplinary differences, with reader reviews indicating deficiencies in translation, content logic, and practicality. For ranking increases: economics books are mostly praised for rigorous logic, practicality, and inspiration; computer science books are mostly introductory texts with engaging writing that combines theory and practice, helping readers learn skills; physics books are popular science classics that combine scientific and popular elements with interesting content that stimulates thought. shows reviews for books with the largest ranking changes across the three disciplines.

Readers of economics books focus primarily on content, hoping to acquire theoretical knowledge, thus demanding strong logical frameworks, readability, and

knowledge value. Computer science books are technical, often involving software knowledge, and readers hope to master professional skills through reading, favoring foundational, readable books combining theory and practice. Readers of these two disciplines prefer practical content, while physics books, positioned as popular science with interesting content, are more likely to receive positive reviews.

4.3.4 Fine-Grained Analysis of Book Review Attributes To investigate how readers of different disciplines evaluate book attributes, content and publication review attributes were further subdivided (author attributes were not differentiated, summarized as "writing style"). The classification dimensions are shown in , with secondary attribute distributions across disciplines presented in [Figure 4: see original paper].

Readers across all disciplines show high attention to examples in books, indicating they hope to acquire knowledge through case studies. Additionally, readers of economics and physics books comment positively on writing style, while computer science book reviews involve more subjective feelings, reflecting that readers of social science and popular science books focus on reading fluency, while technical book readers care more about content comprehension. Comparing positive and negative review frequencies reveals numerous negative comments on translation, decreasing across economics, computer science, and physics. Given that over 75% of books in all three disciplines are foreign translations and translation quality significantly affects readability and logic, selecting high-quality translations is particularly important for publishers.

4.4 System Adaptability Analysis

Given the authority of total citations and library holdings in revealing book influence and quality, the evaluation results for the three disciplines were correlated with these two indicators to test the system' s evaluation effectiveness and adaptability. Before correlation analysis, Shapiro-Wilk tests confirmed that evaluation results, library holdings, and total citations did not follow normal distributions, so Spearman correlation tests were used. Sample sizes were $N_{economics} = 140$, $N_{computer\ science} = 137$, and $N_{physics} = 136$. Results are shown in .

Correlation analysis shows that evaluation results are positively correlated with library holdings and total citations at the 0.01 significance level, with moderate-to-weak correlation strength. This indicates that the evaluation results can supplement the influence and value reflected by library holdings and the academic impact reflected by total citations, providing reference value for book recommendation and quality evaluation activities.

By analyzing correlations between evaluation results and total citations/library holdings across disciplines, the comprehensive Chinese book evaluation system can be applied to quality evaluation of different disciplinary books, helping

readers and institutions like libraries and reading platforms identify books with high influence and content quality.

---

5. Conclusions

This study selected books from different research fields, analyzed differences in primary indicator scores and review indicator utility across disciplines, and examined correlations between evaluation results and total citations/library holdings to investigate system adaptability. Based on empirical results, the following conclusions can be drawn:

5.1 Good System Adaptability Across Disciplines

By inputting disciplinary book data into the evaluation system and conducting Spearman correlation analysis with total citations and library holdings, evaluation results show moderate-to-weak significant correlations with both indicators. This indicates the system can reflect influence and quality across disciplines, while incorporated review indicators integrate reader sentiments and evaluations, revealing books' value in readers' eyes. The system can be applied to evaluate books from different disciplines, identifying works with high reader approval, academic quality, and social media influence to support reading promotion and bibliographic recommendation.

5.2 Review Indicators Reflect Disciplinary Reading Preferences

Analysis of books with significantly changed rankings after adding review indicators shows that economics books emphasize theory and experience transmission, with readers focusing on logical frameworks and knowledge value. Computer science books are technical, and readers prefer foundational, readable works combining theory and practice to acquire skills. Readers of both disciplines favor practical content, while popular science physics books with interesting content receive better reviews. Review indicators can reveal reading preferences across disciplines, helping select books with high readability, scientific rigor, and knowledge value.

5.3 Readers Focus on Cases and Translation Quality

Statistical analysis of secondary review attributes shows that example attributes maintain high frequency across disciplines, indicating readers pay close attention to specific cases for knowledge acquisition. Meanwhile, negative comments on translation are numerous, decreasing across economics, computer science, and physics. Given that over 75% of books are foreign translations and translation quality significantly affects readability and logic, selecting high-quality translations is crucial for publishers.

Due to technical and time constraints, data sources for traditional bibliometric indicators, Altmetrics, and review indicators were relatively limited, with review data only using short reviews from Douban Reading. Future research could incorporate sentiment analysis and attribute mining of citation content and social media mentions to measure evaluation attitudes and attention attributes from different sources, making book quality evaluation more comprehensive.

References

[1] Kousha K, Thelwall M, Rezaie S. Assessing the citation impact of books: the role of Google Books, Google Scholar, and Scopus[J]. Journal of the American Society for Information Science and Technology, 2011, 62(11): 2147-2164.

[2] White H D, Boell S K, 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 and Technology, 2009, 60(6): 1083-1096.

[3] Linmans A J M. Why with bibliometrics the Humanities does not need to be the weakest link[J]. Scientometrics, 2010, 83(2): 337-354.

[4] Cabezas-Clavijo L, Robinson-García N, Torres-Salinas D, et al. Most borrowed is most cited? Library loan statistics as a proxy for monograph selection in citation indexes[EB/OL]. [2020-11-01]. https://arxiv.org/abs/1305.1488.

[5] Björn H. Using Altmetrics for assessing research impact in the humanities[J]. Scientometrics, 2014, 101(2): 1419-1430.

[6] Zuccala A A, Verleysen F T, Cornacchia R, et al. Altmetrics for the humanities: comparing Goodreads reader ratings with citations to history books[J]. Aslib journal of information management, 2015, 67(3): 320-336.

[7] Erfanmanesh M, Noorhidawati A, Abrizah A. What can Bookmetrix tell us about the impact of Springer Nature' s books[J]. Scientometrics, 2019, 121(1): 521-536.

[8] Nan X, Li M, Shi J. Using altmetrics for assessing impact of highly-cited books in Chinese Book Citation Index[J]. Scientometrics, 2020, 122(3): 1651-1669.

[9] Kousha K, Thelwall M. Can Amazon.com reviews help to assess the wider impacts of books?[J]. Journal of the Association for Information Science and Technology, 2016, 67(3): 566-581.

[10] Kousha K, Thelwall M, Abdoli M. Goodreads reviews to assess the wider impacts of books[J]. Journal of the Association for Information Science & Technology, 2017, 68(8): 2004-2016.

[11] Zhang Chengzhi, Tong Tiantian, Zhou Qingqing. Integrated book influence evaluation across different review platforms[J]. Journal of Intelligence, 2018, 37(9): 861-873.

[12] Li Yingling, Sun Xiaohui, Chen Jiubing. Construction of a five-dimensional book evaluation system and analysis model[J]. Information Science, 2013, 31(8): 77-80, 140.

[13] Liu Li, Yuan Xiling. Empirical research on foreign academic book quality evaluation[J]. Library and Information Service, 2011(21): 93-97, 102.

[14] Yang Yuli, Ding Yuan, Zhang Su. Empirical research on ESI single-book foreign language book quality evaluation system: taking computer science as an example[J]. Library Journal, 2015(7): 11-17.

[15] He Jun, Cai Ronghua. Research on Chinese book evaluation system[J]. Journal of Academic Libraries, 2016, 34(3): 51-58, 15.

[16] Xiao Aoxia, Yang Siluo. Application of Bookmetrix indicators in academic book influence evaluation[J]. Library Tribune, 2020, 40(9): 79-85.

[17] Zhou Qingqing, Zhang Chengzhi. Research progress and prospects of book influence evaluation[J]. Journal of Academic Libraries, 2020, 38(3): 24-33.

[18] Li Ming, Li Jiang, Chen Ming, et al. Exploratory analysis of Chinese academic book citations and Altmetrics indicators and their implications[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(6): 557-567.

Author Contributions

Qiu Yue: Data collection, processing, and analysis; paper writing.
Zhu Shiqin: Research design and methodology; paper guidance and revision.
Chen Hongying: Data collection and processing.

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