A Study on the Improvement of the z-index Evaluation by Scholars from the Perspective of Co-authorship

Xiaoqing He[1,2] and Xiaoyu Huang[2(✉)]

[1] School of Public Administration, Sichuan University, Chengdu, China
[2] Library, Sichuan University, Chengdu, China
hxy2007@scu.edu.cn

Abstract. The aim of this article is to propose an improved z-index based on the contribution of co-authors, in order to evaluate scholars' academic influence more accurately and precisely, as the z-index has a flaw in ignoring co-authorship. In this study, 20 representative scholars in the field of muscular dystrophy research from 2016 to 2020 were selected as examples, and their h-index, p-index, z-index, z_1-index, z_a-index and z_h-index, as well as the total number of publications, total citations, and average citation rate corresponding to each index, and the consistency index η reflecting citation distribution were calculated to empirically analyze the effectiveness of using different indices to evaluate scholars' academic influence. The results show that the z_h-index is superior to the z-index, z_1-index and z_a-index, not only in better reflecting the contribution of co-authors but also in helping to effectively curb academic "name-dropping" phenomena.

Keywords: z-index · z_h-index · co-author contribution rate · co-author

1 Introduction

The scientific and reasonable evaluation of scholars' academic influence is a prerequisite for ensuring the scientific and healthy development of the research environment and a driving force for promoting research and innovation. With the continuous emergence of research achievements in various fields, the objective and effective evaluation of scholars' academic influence has become a research hotspot in the current academic field. From Garfield's use of citation analysis to evaluate research output in 1955, to the h-index proposed by Hirsch in 2005 and the p-index proposed by Prathap in 2010, all of these indices aim to measure scholars' academic influence based on their publication and citation rates.

The z-index, proposed by Prathap in 2014 following the p-index, not only considers the quantity factor p (total publications) and the quality factor i (average citation rate) but also reflects the concentration of citations with the consistency index η [1]. Here, P represents the total number of publications, i represents the average citation rate, and η is the consistency index reflecting the distribution of citations, which is calculated as $\eta = X/E = (C^2/P)/\sum_{k=1}^{p} C^2{}_k$.

D. Kumar et al. (Eds.): IEIT 2023, AHSSEH 10, pp. 502–511, 2023.
https://doi.org/10.2991/978-94-6463-230-9_60

If C is the total number of citations, $i = C/P$, $X = iC = i^2P$, $C = \sum_{k=1}{}^{p} C_k$, E ($E = \sum_{k=1}{}^{p} C_k$), the z-index is calculated as:

$z = (\eta X)^{1/3} = (\eta^2 E)^{1/3} = [(C^4/P^2)/(\sum_{k=1}^{p} C^2{}_k)]^{1/3}$.

Prathap believes that C ($C = Pi$) can be seen as an indicator that reflects first-order performance, while X and E can be seen as indicators that reflect second-order performance. The simultaneous existence of X ($X = iC = i^2P$) and E ($E = \sum_{k=1}^{p} C^2{}_k$) provides a possibility for the proposal of a third-order performance index. Since the third index η is considered a good indicator to describe citation distribution, the z-index can be seen as a 3D performance evaluation index of quantity-quality-consistency or volume-velocity-variability. Although the z-index belongs to the third-order performance index, $z = Z^{1/3}$, and after taking the cube root, the z-index has the same dimension as the h-index and p-index [2]. Furthermore, Prathap also analogized the meaning of the z-index to the calculation principles of potential energy and kinetic energy in physics, suggesting that the z-index has the ability to serve as a new comprehensive academic evaluation index [3].

2 Problem Statement

As the z-index is a relatively new academic influence evaluation index, there is not much research on it both domestically and internationally [4]. Currently, research on the z_index abroad is mainly conducted by Prathap and focuses on providing a more in-depth interpretation of the various indicators in the z-index [5]. Additionally, Co-authoring has become a trend in scientific research, and a single paper usually involves multiple authors. However, the existing improved z indices do not consider the different research contributions of authors in co-authored papers [6]. Therefore, this article intends to take a sample of 20 representative scholars in the field of muscular dystrophy research and use four methods to assign weights to co-authors [7–9]. Mathematical expressions are used to explain the meaning of four z indices: (1) a z-index that completely ignores co-authorship and counts every participation equally; (2) a z_1-index that only considers sole and first authors; (3) a z_a-Index proposed by Prathap for the equal allocation of co-authorship weights based on the p index; and (4) a z_h-index that assigns weights based on the contribution rate of co-authors.

3 Method

3.1 Improved z-indexes from the Perspective of Co-author

Prathap did not explicitly define the calculation method for the variables C, P and E($E = \sum_{k=1}^{p} C^2{}_k$) involved in the formula for the z-index. Most scholars, including Prathap, use the following two methods to count publication records: one is to only count the publication records of single-authored or first-authored papers, completely ignoring co-authorship; the other is to count all papers in which an author participated, regardless of their authorship order.

Based on the three basic variables of the z-index: C (the total number of citations of an author's papers), P (the total number of papers published), and η (the consistency

index), and according to the co-authorship situation of the papers, this paper proposes four z-index calculation methods with different weightings. For example, if author K has published p papers in total, and the total number of authors for the k_{th} paper is w, and author K is ranked j_{th} in this paper, and the paper is cited C_k times, then C_k, $C = \sum_{k=1}^{p} C_k$.

(1) The z-index calculation method is to completely ignore co-authorship and count any participation as one paper [10].

If we completely ignore co-authorship and count any participation as one paper, regardless of whether it is a single-authored or co-authored paper, and we do not ignore the authorship order, then any participation will be counted as one paper. In this case, C, P, and η of author K are calculated based on all the papers in which the name of author K appears in the author list, regardless of their authorship order. The z-index calculation method is as follows:

$$z = (\eta^2 E)1/3 \left\{ \frac{\frac{C^4}{P^2}}{\sum_{k=1}^{p} C^2{}_k} \right\}^{1/3}$$

$$C = \sum_{k=1}^{p} C_k, E\ (E = \sum_{k=1}^{p} C^2{}_k)$$

$$\eta = (C^2/p)/E$$

(2) The z-index calculation method is to only count the papers where an author is the sole author or the first author. In this case, C, P, and η of an author are calculated based on all the papers where they are the sole author or the first author [11, 12].

If we only consider the papers where an author is the sole author or the first author, and we do not count the papers where they are not in these positions, then C, P, and E of author K are calculated based on the total number of papers published and the total citation counts of the papers where K is the sole author or the first author. So, when $d_k = \begin{cases} 0, j_k \neq 1 \\ 1, j_k = 1 \end{cases}$, the total number of articles published by the K author $P_1 = \sum_{k=1}^{p} d_k$, $C_1 = \sum_{k=1}^{p} d_k C_k$, $E_1 = \sum_{k=1}^{p} (d_k C_k)^2$, $\eta_1 = (C_1{}^2/P_1)/E_1$

Then, $z_1 = \left\{ \frac{\frac{(\sum_{k=1}^{p} d_k C_k)^4}{(\sum_{k=1}^{p} d_k)2}}{\sum_{k=1}^{p} (d_k C_k)^2} \right\}^{1/3}$

(3) The z_a-Index calculation method is to consider co-authorship, but not the authorship order, and to calculate the weighted contribution of each author in the co-authored papers.

For the papers that K author participated in, regardless of the number of co-authors and their authorship order, the weight of the paper is evenly distributed among all co-authors. That is, w_a author is assigned a contribution of $1/w_k$ [13], then,

$$P_a \sum_{k=1}^{p} \frac{1}{w_k},$$

$$C_a \sum_{k=1}^{p} \frac{1}{w_k} C_k$$

$$E_a = \sum_{k=1}^{p} \left(\frac{1}{w_k} C_k\right)^2$$

$$\eta_a = (C_a^2/P_a)/E_a$$

Then the formula for calculating z-index is adjusted to za:

$$z_a \left\{ \frac{\frac{\left(\sum_{k=1}^{p} \frac{1}{w_k}\right)^4}{\left(\sum_{k=1}^{p} \frac{1}{w_k} C_k\right)^2}}{\sum_{k=1}^{p} \left(\frac{1}{w_k} C_k\right)^2} \right\}^{1/3}$$

(4) The z_h-index is an academic indicator that takes into account the co-authorship factor and assigns weight based on each author's contribution rate to calculate their academic impact.

According to Hagen's proposed harmonic algorithm for author contribution rate levels, the weight of each participating author is assigned in a decreasing order based on the author's position of authorship in the co-authored paper. F_k represents the contribution of the k_{th} author to the paper, j_k represents the position of author k in the paper, and w_k represents the total number of authors of the paper. The formula for calculating the author's weight is as follows [14]:

$$F_k \frac{\frac{1}{j_k}}{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{w_k}} \tag{1}$$

Based on the aforementioned method of assigning weights according to the size of author contribution, this paper proposes to allocate weights to each participating author based on their different levels of contribution in the co-authored paper, i.e., their authorship order. Therefore,

$$P_h = \sum_{k=1}^{p} \frac{\frac{j_k}{w_k}}{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{w_k}}$$

$$C_h = \sum_{k=1}^{p} \frac{\frac{j_k}{w_k}}{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{w_k}} C_k$$

$$E_h = \sum_{k=1}^{p} \left(\frac{\frac{j_k}{w_k}}{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{w_k}} C_k \right)^2$$

$$\eta_h = (C_h^2 2/P_h)/E_h$$

Incorporating the author's co-authorship contribution value F_k into the calculation of the z_index yields the z_h-index, with the following formula for its calculation [15]:

$$Z_h \left\{ \frac{\left(\sum_{k=1}^{p} \frac{\frac{j_k}{w_k}}{1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{w_k}} \right)^4}{\left(\sum_{k=1}^{p} \frac{\frac{j_k}{w_k}}{1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{w_k}} C_k \right) 2} \right\}^{1/3}$$

$$\frac{\sum_{k=1}^{p} \left(\frac{\frac{j_k}{w_k}}{1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{w_k}} C_k \right)^2}{}$$

3.2 Data Collection and Analysis

This article uses the Science Citation Index Expanded (SCI) in Web of Science as the data source to analyze the relevant data of 20 representative scholars in the field of muscular dystrophy research from 2016 to 2020. This article calculates various indicators for 20 representative scholars in the field of muscular dystrophy research from 2016 to 2020, including the number of publications (including first author and other co-author positions), total citations, average citations, citations per paper, number and ranking of co-authors, and consistency index η. The article also calculates various indices, such as h-index, p-index, z-index, z_1-index, and z_h-index, corresponding to the total number of publications, total citations, and average citation rates, as well as consistency index η reflecting citation distribution. The results are shown in Table 1. Based on the statistical results, the rankings of z-index, z_1-index, z_a-index, z_h-index, h-index, and p-index are listed in Fig. 1.

4 Empirical Results

4.1 z-index, z_1-index, z_a-index, z_h-index Comparative Analysis of Index, h-index and p-index Ranking

As shown in Fig. 1, AARTSMA-RUS A ranks first according to z_1-index, z_a-Index, z_h-index, and p-index, but second according to z-index. Although AARTSMA-RUS A's number of publications is not outstanding among the 20 scholars, with only 44 publications, it has the highest number of publications as first author, and its co-authored papers have fewer co-authors and are ranked closer to the front. Its average citation frequency per paper and the weighted average citation frequency per paper based on

Table 1. 20 scholar's h-index, z-index and z_h-indicators corresponding to the index

Scholars	p	η	z-index	z_a-index	z_1-index	z_h-index
AARTSMA-RUS A	44	0.2283	15.043	15.79	20.44	18.6159
BUSHBY K	58	0.165	16.232	14.38	0.11	16.12923
GOEMANS N	51	0.0931	9.3014	6.37	8.80	6.392996
GORDISH-DRESSMAN H	35	0.1519	6.865	5.02	0	4.096068
GUGLIERI M	59	0.2657	9.8264	9.11	0	7.131816
KOMAKI H	48	0.2265	7.9282	7.71	4.96	7.813349
LOCHMULLER H	45	0.4387	12.895	11.45	0	10.50626
MCDONALD C	64	0.0827	1.8499	1.32	0.69	1.106337
MERCURI E	92	0.1348	13.003	10.12	6.87	8.970882
MUNTONI F	175	0.1119	9.0009	7.08	1.87	5.848026
NISHINO I	43	0.2032	4.8215	4.65	0.00	4.676052
PEGORARO E	39	0.1831	13.413	11.49	0.00	9.829436
RICOTTI V	41	0.164	7.1454	6.33	15.16	9.901801
SERVAIS L	43	0.1224	5.8183	4.80	0.00	3.964551
STRAUB V	114	0.2521	13.116	11.63	9.45	12.62843
TAKEDA S	53	0.2733	10.443	9.79	0.00	9.586707
TRIFILLIS P	44	0.0853	0.8244	0.55	0.00	0.329411
TULINIUS M	35	0.0992	6.7721	5.88	0.00	5.895729
VOIT T	46	0.1111	7.8677	5.70	0.00	5.788426
WONG B	51	0.0352	3.2191	2.85	0.50	2.038519

author contributions are also outstanding. The reason why it ranks second in terms of z_index is that its consistency index $\eta = 0.2283$ indicates that the concentration of its citation distribution is not low, leading to its ranking second in terms of z-index. However, because AARTSMA-RUS A's number of publications is not high, and the calculation of h-index is easily affected by the number of publications. For scholars with high citation frequency but few publications, their h-index will not be too high, resulting in AARTSMA-RUS A ranking only seventh according to h-index.

4.2 Typical Case Analysis Results

In order to better verify whether the z-index, z_1-index, z_a-index, and z_h-index can well reflect the different contribution rates of participating authors in co-authored papers, the publication situation of AARTSMA-RUS A was restored, as shown in Fig. 2. Situation 1: Paper 18 was published by AARTSMA-RUS A as the first author, and there were a total of two authors in this paper. If this paper were published solely by AARTSMA-RUS A, then

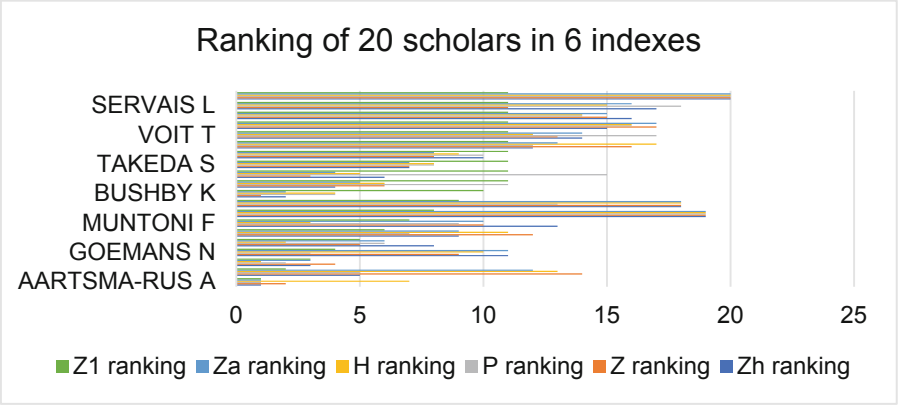

Fig. 1. 20 scholar's z-index, z_1-index, z_a-index, z_h-index, h-index and p-index ranking

the z_a-index would increase from 15.78915926 to 19.13194195, and the z_h-index would increase from 18.61589694 to 22.91319949. However, the z-index completely ignores the contribution rate of participating authors, so it cannot reflect this change. The z_1-index only counts papers published solely or as first author, and cannot reflect changes in the authorship of other authors who are not the first author, so the z_1-index remains unchanged. Situation 2: If the authorship position of AARTSMA-RUS A in paper 37 is changed from the second position to the first position, then the z_h-index will increase from 18.61589694 to 20.46123398, and the z_1-index will also show a corresponding increase, while the z-index and z_a-index do not capture this change. Situation 3: If an additional author is added among the two co-authors in paper 37, then both the z_a-index and z_h-index show different degrees of decline, while the z-index and z_1-index remain unchanged and do not reflect this change.

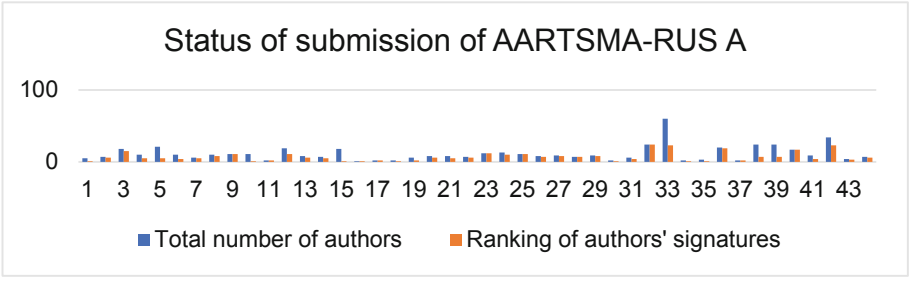

Fig. 2. Statuses of AARTSMA-RUS An issuance

5 Discussion

5.1 The z-index, z_1-index, z_a-index, and z_h-index All Reflect the Balance Between the Number and Quality of Scholars' Achievements and the Distribution of Citations

The process of proposing the z-index, z_1-index, z_a-Index, and z_h-index fully considered the fusion of P (reflecting the total number of quantity factors), i (reflecting the average citation rate of quality factors), and η (reflecting the degree of citation concentration) into a single comprehensive evaluation index, which realized the balance between quality, quantity, and concentration of citation distribution to varying degrees. Correlation analysis in Sect. 2.2.2 shows that the z-index, z_1-index, z_a-index, and z_h-index are significantly correlated with the total number of citations and the average citation rate at the 0.01 level (two-tailed), indicating that these four indices can have varying degrees of sensitivity to changes in the total number of citations and the average citation rate. When the number of citations of scholars' papers increases, these four indices can reflect this change in a timely manner. From the ranking analysis in Sect. 2.2.1, it can be seen that the ranking of z-index, z_1-index, z_a-index, and z_h-index is all affected by the comprehensive impact of the quantity and quality of scholars' papers and the distribution of citations.

5.2 The z_h-index Helps Curb the Phenomenon of "Publication Quantity and Citation Bubble" and "Gift Authorship"

Due to the calculation of the z-index, which counts every author as having equal credit for a publication, non-core authors and even gift authors can share the honor of a publication (publication quantity and citation frequency). This calculation method easily leads to the occurrence of the "publication quantity and citation bubble" phenomenon (i.e. equating the "publication quantity and citation frequency of a paper" with "each author participating in this paper publishing one paper and sharing all the citation frequency of this paper"), thereby exacerbating the occurrence of gift authorship. If a core author of a paper lists an unrelated person as an author, according to the calculation method of the z-index, their z-index remains unchanged, while the unrelated person listed as an author gains credit for one publication and shares all the citation frequency of the paper. This calculation method of the z-index allows gift authorship to occur without affecting the z-index of the listed authors, indirectly "encouraging" the emergence of gift authorship. However, the z_h-index considers the degree of co-authorship and author ranking to allocate credit for each paper, which effectively curbs the phenomenon of "publication quantity and citation bubble" and "gift authorship". If a core author of a paper lists an unrelated person as an author, their z_h-index will decrease correspondingly, and the credit that the unrelated person listed as an author can receive for this paper is limited.

6 Conclusion

Currently, there are not many studies on improving the z-index both domestically and internationally and the improvement of the z_index considering co-authorship is particularly lacking. However, this study shows that both the z-index that completely ignores

co-author contributions and the z_1-index that only considers the first author or sole authorship, as well as the z_a-index that evenly distributes credit to all authors for a paper, are all extreme in handling co-authorship. They either ignore or underestimate the contribution of authors listed higher in the authorship order, or exaggerate or completely ignore the contribution of authors listed lower in the authorship order. The z_h-index combines the advantages of the z-index and improves on its shortcomings in ignoring co-authorship, ensuring that the z_h-index not only considers quantity, quality, and citation distribution but also takes into account co-author contributions. The z_h- index is more comprehensive and reasonable than the z-index, z_1-index, and z_a-index and has great potential as an evaluation index. However, the z_h-index is still an evaluation index from a quantitative perspective. To avoid the limitations of a single index, the evaluation of scholars' academic influence should be comprehensive and include qualitative indicators to ensure a more reliable, fair, and scientific evaluation.

Declaration of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Gangan Prathap. A. (2014). Three-Class, Three-Dimensional Bibliometric Performance Indicator. *Journal of the Association for Information Science and Technology*, *65*(7), 1506-1508. https://doi.org/10.1002/asi.23120.
2. Gangan Prathap. A. (2015). Measures for Bibliometric Size, Impact, and Concentration. *Journal of the Association for Information Science and Technology*, 66(8), 1740-1741. https://doi.org/10.1002/asi.23364.
3. Gangan Prathap. A. (2014).The Zynergy-Index and the Formula for the h-index. *Journal of the Association for Information Science and Technology*, *65*(2), 426-427. https://doi.org/10.1002/asi.23046.
4. Gangan Prathap.A. (2014).Three-Dimensional Bibliometric Evaluation of Research in Polymer Solar Cells. *Scientometrics*, 101, 889-898. https://doi.org/10.1007/s11192-014-1346-z.
5. Gangan Prathap.A. (2017).Three-Dimensional Bibliometric Evaluation of Recent Research in India. *Scientometrics*, 110,1085-1097. https://doi.org/10.1007/s11192-016-2181-1.
6. Hagen N T. (2008). Harmonic Allocation of Authorship Credit: Source-Level Correction of Bibliometric Bias Assures Accurate Publication and Citation Analysis. *Plos One*, *3*(12), e4021. https://doi.org/10.1371/journal.pone.0004021.
7. He Xiaoqing, Wang Shengjie, Hu Lin. (2018). The Application of z-index in Evaluating Academic Influence of Researchers. *Information Studies: Theory & Application*, *41*(05), 50–54. https://doi.org/10.16353/j.cnki.1000-7490.2018.05.010.
8. He Xiaoqing, Wang Shengjie, Hu Lin. (2018). Application of z-index in journal evaluation. *Chinese Journal of Scientific and Technical Periodicals*, *29*(05), 509–514. https://doi.org/10.11946/cjstp.201709290831.
9. He Xiaoqing, Wang Shengjie, Hu Lin. (2018). Application of z-index to Institutions Evaluation. *Journal of Modern Information*, *38*(05), 82–86. 10.3969 /j.issn.1008–0821.2018.05.012.

10. Li Changling, Pai Yanxin, Xiangfu Zhong, Du Dehui. (2020). Discussion on the Method of Identifying Subjects Research Hot Spots of Highly Cited Using an Improved Z-Index. *Information Studies: Theory & Application*, *43*(06), 69–75 ≤ 96. 10.16353 /j.cnki.1000–7490. 2020. 06. 011.
11. Liu Yunmei, Li Changling, Liu Xiaohui. (2016). Discussion of P-index Based on the Co-author Contribution Division Value. *Library and Information Service*, *60*(21), 81–86. 10.13266 /j.issn. 0252–3116.2016.21.011.
12. Shen Junwei, Zheng Dejun, Zhang Zhenghui, Zhu Yufei and Kuipiqian. (2019). An empirical research on the characteristics of zynergy-index and its applicability in evaluating Chinese academic journals: Based on the analysis on the CSSCI source journals from three disciplines. *Chinese Journal of Scientific and Technical Periodicals*, *30*(10), 1113–1121.10.11946 /cjstp.201903180181.
13. Yang Wenxia, Liu Wenyun, Liu Jian, Huang Pingping. (2020). The Improvement of p-index of Author's Evaluation Based on Citation Distribution, Time and Co-authorship. *Journal of Intelligence*, *39*(11), 201–207. 10.3969 /j.issn.1002–1965.2020.11.030.
14. Yu Liping, Wang Zuogong. (2018). Applicability and Improvement of the z-index to Evaluate Academic Journals. *Journal of the China Society for Scientific and Technical Information*, *37*(11), 1132-1139. https://doi.org/10.3772/j.issn.1000-0135.2018.11.006.
15. Zhang Xuemei. (2020). Research on the Modification to Z Index Researcher Academic Evaluation Method Based on Representative Work System. *Library Science Research & Work*, *2020*(01), 22-28.