ATLANTIS
PRESS

Research on Online Course Quality Evaluation System Based on the Perspective of Users Comments

Fang GAO*, Yusi WEI

School of Business
Jianghan University
Wuhan, China

Abstract—**Nowadays students have a wider range of choice for knowledge acquisition benefit from the prevalence of online course construction in China. However, in the face of numerous online courses on various platforms that are dazzling and intermingled, how to guide students to make high-quality choice has become an urgent matter of the moment. Therefore, the quality evaluation of online courses will inevitably cause concern. This paper captures online users' reviews using the web crawler technology, combines with offline expert interviews and focus group interviews, focuses on users' evaluations on learning contents and effectives, experience and demands of online course platform, takes multidisciplinary integration into account, constructs a user-oriented online courses quality evaluation indicator system. The indicator system includes four first-level indicators which are teaching content, interactive experience, teacher's accomplishment and platform support, further includes fifteen second-level indicators.**

Keywords—online course; MOOC; user comment; quality evaluation

I. INTRODUCTION

The emergence of online courses such as MOOC stems from the original intention of colleges and universities to reduce the cost of education and improve the quality of education. It transmits knowledge and information on the basis of the network. At present, the main users (students) of MOOC just grow up in the network era. Therefore, compared with the traditional classroom, MOOC just complements its media habits.

With the rise of online courses such as MOOC, SPOC and micro courses, students have a wider choice of knowledge acquisition. However, in the face of many online courses on various platforms, which are dazzling and vary in good and bad, how to guide students to make high-quality choices has become an urgent task. Therefore, the quality evaluation of online courses is bound to attract attention. This paper finds that, unlike the traditional classroom quality evaluation research, the existing research on online course quality evaluation in China is relatively scarce, and the limited research results also have many deficiencies. For example, the evaluation indicators of traditional courses should be adapted appropriately to be used as the evaluation of online courses, or the evaluation system is too broad and not targeted, or only for one side of the course content. In addition, the existing research methods of online course quality evaluation in China are relatively single, focusing more on expert opinions and less on user (student) positions. In fact, students are the main users of online courses. Any evaluation system that ignores students' feelings is difficult to fully and truly reflect the quality of online courses and user satisfaction.

In view of this, this paper uses web crawler technology to capture online user comments, combining offline expert interviews and focus group interviews. From the perspective of users (students), by focusing on students' evaluation of online learning content and effect, learning experience and demands for online course platform, taking into account the integration of multi-disciplinary, a set of online course quality evaluation indicator system with high reliability and operability is constructed.

II. RESEARCH TRENDS AT HOME AND ABROAD

A. Relevant Domestic Research

The online course evaluation in China started from the "Online Course Evaluation Standard" issued by the Ministry of education in June 2002. Since then, the Ministry of Education issued the national quality curriculum evaluation indicators in 2010, including 5 first level indicators and 14 two level indicators and observation points. Obviously, these evaluation systems are relatively extensive, but they provide a basic framework for domestic related research, and lead to a lot of exploratory research.

At present, most of the research in this field in China adopts interview method, Delphi method, inductive summary method and so on. For example, Qiu and Ou(2015) designed the evaluation indicator system of MOOC, including 5 first level indicators and 13 second level indicators, by consulting experts and combining the characteristics of MOOC. Tong and Jia (2017) constructed the MOOC quality evaluation indicator system including three first level indicators and 26 second level indicators by using the methods of literature research and expert interview. Using Delphi method, Sun and Sun (2014)

designed an open curriculum quality evaluation indicator system including 5 first level indicators and 32 second level indicators. Zhang, Zhan and Li (2015) adopted the method of network metrology to design the evaluation system of network courses, which includes 4 first level indicators and 30 second level indicators. Based on the perspective of learning experience, Zhang (2017) constructed a MOOC quality evaluation system including 8 first level indicators and 22 second level indicators through questionnaire survey and Delphi method. On the basis of SWOT analysis of existing online resources, Mu (2015) proposed to evaluate online courses from three dimensions of technology, teaching and institutions.

B. Foreign Related Research

There are abundant research results on online open curriculum evaluation abroad, but the research on the construction of evaluation indicator system seems to be insufficient. The mainstream view is that there are four factors that affect online learning: the expected effect of the course and its accessibility, the flexibility for students and teachers, the emphasis on teacher-student interaction and the ability to achieve convenient interaction, and the experience of actively participating in the course. Butcher and Wilson-Strydom (2013) think that these four kinds of factors should be integrated into online curriculum evaluation. Therefore, the evaluation of online curriculum should pay attention to the following principles: student evaluation as the center, teacher evaluation as the basis, management evaluation as the reference, formative evaluation as the main, and summary evaluation as the auxiliary. Dayton and Vaughn (2007) designed a quality assurance system of online courses with 7 first level indicators and 25 second level indicators under the principle of educational practice objectives. Lin (2010) through summarizing the literature of online course quality research, designed an online course quality evaluation indicator system including 4 first level indicators and 16 second level indicators. The QM (Quality Matters) program in the United States has also created a national standard system for online courses, which involves higher education, basic education, continuing and vocational education and many other fields. Each field has relevant evaluation indicators for its online and hybrid courses. The indicator system has good scalability, can be applied to different disciplines, different stages of online course evaluation, and help to improve the quality of online education and learning effect, so it has been internationally recognized.

C. A Brief Review

As for online course quality evaluation, foreign related research started earlier, and the evaluation system established has good ductility, but due to many differences between Chinese and western education systems, most of them cannot be directly used by me. However, the development of online courses in China started late, and at the same time, there is a lack of a more perfect evaluation system. According to the current situation, more and more scholars have paid attention to the construction of online course evaluation system, but there are limitations in the research methods, most of which are Delphi method, case analysis method, inductive summary method, etc. Restricted by the traditional concept, the current evaluation of curriculum quality is more focused on the evaluation of teachers, and less attention is paid to the learning effect of students. However, the extremely low completion rate and high dropout rate not only waste a lot of network resources, but also have a negative impact on the enthusiasm of the construction of curriculum evaluation system.

III. DATA MINING AND THE GENERATION OF EVALUATION INDICATOR

In order to make the evaluation indicators reflect users' concerns comprehensively and objectively, this paper mainly develops online course quality evaluation system based on online user comments. First of all, we used the technology of web crawler to collect online user comments of iCourse website. Considering the great differences in course attributes of various disciplines, we set 6 courses to obtain comments, as far as possible, taking into account the history and philosophy, science and engineering, medicine, management and other disciplines.

We used the selenium test package and chrome webdriver headless browser in Python language to realize the whole process automatic grabbing. First, take the URL of the iCourse website as the crawler entry, control webdriver to open the website, and automatically enter the test account to log in. Then open the web address of the target course comment page, locate through the DOM tree of the web page, call the beautifulsoup library to parse HTML, and drive the click event to automatically click the next page, and automatically traverse the course comment one by one. Finally, we used the pandas library to store the captured comments as dateframe data types and export them to Excel files, so as to facilitate manual analysis. Finally, 6333 course comments were captured.

In view of the meaningless text noise in comments, we used regular expressions to match and delete them. In order to generate the initial evaluation indicator, we invited 6 professional teachers with senior titles and 12 students with relevant online course learning experience to participate in the discussion. Referring to the frequency of Chinese word segmentation, we manually divided online comments into multiple phenomenon summaries, such as video quality, teacher teaching, course quality, learning effect, content acceptance, etc., and then classified the content of each comment into different phenomenon summaries, and standardized to form the initial concept, that is, to judge which aspect of course quality users comment on, and finally got 31 evaluation indicators.

IV. QUESTIONNAIRE SURVEY AND INDICATOR SYSTEM CONSTRUCTION

A. Initial Investigation and Indicator Reduction

In order to avoid the adverse effects of too many indicators leading to lengthy questionnaires, we first simplified the indicators through small-scale surveys. Using Likert 5-point Scale, the 31 indicators generated were made into a paper-based questionnaire, and the effectiveness of each indicator for the evaluation of online course quality was measured by offline survey. The subjects of the survey were students who had

online learning experience in the past six months. A total of 60 questionnaires were sent out and 55 were recovered, of which 49 were valid, with an effective recovery rate of 81.67%. Based on the experience of Liu and Zhou [2011], the indicators with an average score of no more than 2 points are deleted directly, and finally 26 indicators are reserved.

B. Secondary Investigation and Formation of Indicator System

Once again, we use Likert 5-point Scale to make 26 indicators into online questionnaires, which were published on the Internet platform (https://www.wjx.cn/), to measure the effectiveness of each indicator in evaluating the quality of online courses. The following measures shall be taken: (1) the questionnaire shall be sent to the students who have online learning experience in the past half a year; (2) the IP address restriction shall be used to avoid repeated answers by the same respondents; (3) the missed answer reminder shall be set; (4) the questionnaire filled in less than one minute shall be deemed invalid. A total of 300 subjects were sent questionnaire links, 272 of which 261 were valid questionnaires, with an effective recovery rate of 87%.

1) Descriptive Statistical Analysis

Of the 261 samples, 45% were men and 55% were women. As the main users of online courses, 68% of the respondents at undergraduate level. From the perspective of geographical distribution, this survey has a wide geographical distribution, which has certain universal significance.

2) Extraction of Evaluation Indicators

Liu and Zhou (2011) research shows that: the evaluation indicator should be as refined as possible, because students are bored with the lengthy evaluation questionnaire and tend to avoid or perfunctory evaluation. Therefore, we used factor analysis and reliability test to simplify the indicators. KMO and Bartlett's test (Table I) show that KMO value is greater than 0.8, while Bartlett's test is significant at the level of 0.01, indicating that sampling is sufficient and data is suitable for factor analysis. The specific operations are as follows: (1) the orthogonal rotation method with the largest variance is used for factor analysis, with factor load of 0.8 as the intercept point, and the indicators lower than the standard are deleted. (2) In the reliability test, if Cronbach's α coefficient is significantly improved after deleting an indicator, the indicator will be deleted. Finally, 15 indicators are reserved.

TABLE I.　　RESULTS OF KMO AND BARTLETT'S TEST

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.851
Bartlett's Test of Sphericity	Approx. Chi-Square	1672.281
	df	105
	Sig.	0.000

3) Factor Extraction and Reliability and Validity Test

TABLE II.　　THE RESULTS OF FACTOR ANALYSIS AND CRONBACH'S α COEFFICIENT TEST

Indicator	Factor Loading				Cronbach's α coefficient	Variance interpretation rate
	Factor1	Factor2	Factor3	Factor4		
The usefulness of teaching content	0.825				0.887	31.779
Acceptability of teaching content	0.810					
The logic of teaching content	0.871					
Integrity of teaching content	0.803					
The forward-looking nature of teaching content	0.812					
Interaction of system		0.879			0.835	15.212
Accessibility of auxiliary materials		0.824				
Clarity of assessment methods		0.817				
Course guidance and evaluation support		0.889				
Teachers' teaching level			0.836		0.880	11.279
Teachers' teaching methods			0.877			
Teachers' teaching style			0.824			
Stability of platform performance				0.868	0.833	7.372
Technical support services of the platform				0.882		
The quality of curriculum management system				0.835		

The orthogonal rotation method with the largest variance was used for factor analysis. The results are shown in Table II. Four factors were extracted in total. After rotation, the variance interpretation rates were 31.779%, 15.212%, 11.279%, 7.372% respectively, and the cumulative variance interpretation rate was 65.642%. The maximum load of each indicator factor is 0.889, the minimum load is 0.803, and no cross load is found, which shows that the indicator system has a good

discrimination validity. The indicators of each factor are automatically aggregated into one factor by factor analysis alone, which shows that the indicator system has good aggregation validity.

The detection of Cronbach's α coefficient (Table II) shows that the overall α coefficient of 15 indicators is 0.874, the maximum value of α coefficient of each factor is 0.887, and the minimum value is 0.833, indicating that the internal consistency of indicator system is good.

C. Third Investigation and the Empowerment of Indicator System

After repeatedly considering and combining the opinions of the six professional teachers mentioned above, we named the factors of the indicator system, and finally named factors 1 to 4 as: teaching content, interactive experience, teacher literacy and platform support.

In order to reflect the difference of the importance degree of evaluation indicators, it is necessary to further empower

each indicator. For this reason, we made the developed indicator system into a questionnaire again (in the form of Excel table), and asked the subjects to measure the importance of online course quality according to each indicator. We assigned the total score of 100 points to these indicators. The more important an indicator is, the higher it will be. Otherwise, the lower it will be. We required that each indicator must be given a score, and there should be no vacancy. Moreover, the total score of 15 indicators should be exactly 100 points.

A total of 100 electronic questionnaires were distributed and 84 were recovered. Because Excel table can automatically help the subjects to calculate the total score and provide over score reminder, All the recovered questionnaires are effective with an effective recovery rate of 84%. We simply averaged the scores of each indicator, and adopted the principle of rounding nearby. We took the average scores of each indicator as its weight in the evaluation indicator system (Table III), and finally formed a complete online course quality evaluation system.

TABLE III.   THE WEIGHT OF INDICATOR SYSTEM

First level indicator	Second level indicator	Weight of second level indicator	Weight of first level indicator
Teaching content	The usefulness of teaching content	8	32
	Acceptability of teaching content	7	
	The logic of teaching content	7	
	Integrity of teaching content	5	
	The forward-looking nature of teaching content	5	
Interactive experience	Interaction of system	6	27
	Accessibility of auxiliary materials	8	
	Clarity of assessment methods	5	
	Course guidance and evaluation support	8	
Teacher literacy	Teachers' teaching level	9	25
	Teachers' teaching methods	8	
	Teachers' teaching style	8	
Platform support	Stability of platform performance	5	16
	Technical support services of the platform	6	
	The quality of curriculum management system	5	
Total		100	100

V.   CONCLUSIONS AND SUGGESTIONS

With the popularity of online course construction in China, the evaluation of online course quality is bound to attract attention. In this paper, we used the technology of web crawler to obtain online user comments, combined with offline expert interviews and focus group interviews, from the perspective of users, by focusing on users' evaluation of online learning content, effect, learning experience and demands for online course platform, and taking into account the integration of multi-disciplinary, built a user oriented online course quality evaluation indicator system. It includes four primary indicators, including teaching content, interactive experience, teacher literacy and platform support, and 15 secondary indicators.

Based on the principle of refining the indicator system as much as possible, we have adopted higher selection criteria in indicator refining. In order to test the effectiveness of this indicator system, we selected the "Advanced Mathematics" courses from three universities in the iCourse platform for verification survey, which are national level excellent courses, provincial-level excellent courses and school level excellent courses. According to the average scores of 20 students who were invited to test, the differences of scores of the three courses objectively reflect the differences of course grades. According to the feedback survey, students have a high degree of approval on the coverage and quantity of evaluation indicators.

From the perspective of the weight distribution of the indicator system, online course users pay more attention to the interaction experience besides the traditional classroom. Because the students can not directly contact the teachers themselves, so they put forward more demands for the guidance and evaluation support inside and outside the class.

This paper also notes that the current evaluation indicator system adopted by colleges and universities is mainly based on the general requirements of education laws and regulations on the profession of college teachers, so the evaluation of teachers' morality and style accounts for a considerable proportion, which is rarely involved in online comments. It can be seen that students pay more attention to the quality of the course itself than the formal content.

In this paper, the evaluation indicators were weighted, which provides operability for the institutions that usually adopt the percentage system evaluation mode. For each indicator, the weight we give is the upper limit of the teacher's score in the indicator, and students can score within the weight range. However, the indicator system of this paper is also applicable to the institutions that usually adopt the rating system.

REFERENCES

[1] X. Tong and X. Jia. "Construction of MOOC quality evaluation system". China Distance Education, 2017, (05), pp. 63-71; 80. (*In Chinese*).

[2] CELTS (Education Information Technology Standard Committee of the Ministry of Education). CELTS-22 online course evaluation. 2002-06-07. (*In Chinese*).

[3] X. Zhang. "Analysis of evaluation index system of excellent courses". China Audio Visual Education, 2012, (10), pp. 76-80. (*In Chinese*).

[4] Z. Zhang, N. Zhan, and Y. Li. "Research on evaluation methods of online courses based on Webometrics". Journal of Distance Education, 2015, (1), pp. 66-72. (*In Chinese*).

[5] K. Zhang. "Study on MOOC curriculum evaluation index system based on learning experience". Shaanxi Normal University, 2017. (*In Chinese*).

[6] D. Dayton and M. M. S. Vaughn. "Developing a quality assurance process to guide the design and assessment of online courses". Technical Communication, 2007, 54(4), pp. 475-489.

[7] H. F. Lin. "An application of fuzzy AHP for evaluating course website quality". Computers & Education, 2010, 54(4), pp. 877-888.

[8] Q. Liu and J. Zhou. "A study of student evaluation indicator system on teaching in Chinese university", Proceedings of the Conference on Information Technology and Education Teaching, 2011.

[9] L. Liu, Z. Liu, and Y. Luo. "Evaluation methods and enlightenment of MOOC education quality in Europe ". Open Education Research, 2015, 21 (05), pp. 57-65. (*In Chinese*).

[10] Z. Shi. "Restricting factors and countermeasures for the construction of online course quality evaluation system in colleges and universities". Industry and Technology Forum, 2018, 17 (20), pp. 278-279. (*In Chinese*).