Eng Int Syst (2024) 6: 659–669

Engineering
Intelligent Systems

Factors Affecting the Quality of Online Open Course Teaching in Universities Based on Big Data Analysis

Genchao Wei*

School of Economics and Management, Jiyuan Vocational and Technical College, Jiyuan 459000, Henan, China

By means of big data analysis, this study systematically explores and discusses the factors affecting the educational quality of college open online courses (MOOCs). The research on the MOOC platform data of A University in Chengdu, China, reveals the significant impact of teaching content, teacher quality, student engagement and platform technology on the quality of teaching and, subsequently, on students' learning outcomes. Results show that course content had the greatest influence on educational quality (regression coefficient 0.35, $p = 0.000$), followed by teacher quality (regression coefficient 0.30, $p = 0.000$), student engagement (regression coefficient 0.25, $p = 0.006$) and platform technology (regression coefficient 0.20, $p = 0.001$). The research shows that optimizing curriculum design, improving teachers' professional level, strengthening students' motivation to learn, and improving platform technology are key to improving the delivery of education via MOOCs. This study provides data support and a scientific basis for improving the quality of online education, which has important theoretical and practical significance, and can provide references for education administrators and curriculum designers who wish to devise more effective teaching strategies and policies.

Keywords: online open course; influencing factors; big data analysis; colleges and universities

1. INTRODUCTION

With the rapid development of information technology, open online courses (MOOCs) have gradually become an important part of higher education. Universities around the world have launched various types of open online courses to meet diverse learning needs. This new education model breaks the time and space constraints of traditional classrooms, and makes high-quality education resources more widely accessible. However, with the increasing number of MOOCs, the problem of evaluation and guarantee of educational quality has gradually emerged. Effective evaluation and improvement of the educational quality of MOOCs are essential to ensure their educational value. In this context, big data analysis can provide more dimensions of assessment data used to determine the quality of teaching. By mining and analyzing a large amount of data from online education platforms, various factors affecting the educational quality of MOOCs can be discussed in depth, with a view to providing a scientific basis and effective strategies for the improvement of online education. It is anticipated that this research will encourage universities and other researchers to apply big data technology to improve the educational quality of online open courses, so as to better serve the educational needs of students and society.

Globally, more research is being conducted on the educational quality of MOOCs, and the research methods and fields are becoming increasingly diversified. Yang and Loghej (2019) built and analyzed a model for the impact of emotional factors on the improvement of students' learning efficiency through big data analysis, and revealed the importance of

*Corresponding author's Email: genchao_wei@hotmail.com

emotional factors in online learning [1]. Yao and Lin (2023) adopted computational pedagogy to identify the key factors affecting the educational quality, providing a theoretical basis for improving the educational quality of MOOCs [2]. Chen et al. (2021) took the informatization construction of continuing pharmacology education as a case study and used big data analysis to discuss the etiology research methods of rheumatoid arthritis, providing a new perspective on online medical education [3].

During the COVID-19 pandemic, Liu et al. (2022) analyzed the factors influencing the evaluation of MOOCs in terms of effective teaching during the COVID-19 pandemic. The researchers conducted an exploratory study based on grounded theories, and pointed out the significant impact of teacher quality, teaching resources, interaction and technical support on educational quality evaluation [4]. Li et al. (2021) studied the reliability of peer assessment in MOOCs and its influencing factors, and found that course design, the clarity of assessment criteria and the professional background of evaluators have important influences on the reliability of assessment [5].

Le et al. (2023) conducted a study on the satisfaction of Vietnamese College of Education students in short-term online courses and found that course content, teaching methods, technical support and students' self-management ability are the main factors affecting students' satisfaction [6]. Zhang et al. (2023) analyzed online learning behavior and its influencing factors in STEM courses by studying the Open University Learning Analysis dataset (OULAD), and found that learning motivation, learning strategy and technology acceptance influenced the outcomes of online learning [7].

Chinese research on the educational quality of open online courses (MOOCs) is increasing and becoming more comprehensive due to the research methods constantly being refined, and the perspectives being diversified. Kong et al. (2018) designed a educational quality evaluation model based on fuzzy mathematics and a support vector machine (SVM) algorithm. They proposed a new evaluation method, and significantly improved the accuracy and scientific nature of educational quality evaluation [8]. Fang et al. (2019) studied the environment structure and practice of the Internet of Things through big data analysis, providing new technical support and implementation path for online education [9]. These studies provide important references for the improvement of the educational quality of MOOCs in China.

El-Sakran et al. (2022) found that the emergency distance learning mode had a significant impact on students' learning experience and academic outcomes, and put forward suggestions to improve the teaching mode [10]. Hsueh et al. (2022) used partial least squares structural equation model (PLS-SEM) to explore students' participation in online programming courses and found that the behavioral participation mode has an important impact on the learning outcomes [11]. Based on the technology acceptance model (TAM), Zhou et al. (2022) studied students' intention of using online education platforms and found that perceived usefulness, ease of use and attitude towards use are the key factors affecting students' intention to use this mode of learning [12].

These studies show that the application of big data analysis in MOOCs educational quality assessment is becoming increasingly widespread. Through the analysis of a large amount of data, various factors affecting educational quality can be comprehensively understood, and effective improvement strategies can be proposed [13, 14].

The purpose of this study is to systematically explore the factors affecting the educational quality of college MOOCs through big data analysis, and to provide a theoretical basis and practical guidelines for improving the educational quality of MOOCs. Specifically, through the construction and analysis of the model affecting the educational quality of MOOCs, key factors such as teaching content, teacher quality, student factors and platform technology and their interrelationships are identified and verified, and the specific influencing mechanism of these factors on the teaching and learning outcomes is revealed [15, 16]. The aim of the research is to improve the educational quality of online education in colleges and universities through scientific and comprehensive data analysis, so as to better meet the diversified learning needs of students and promote educational equity and resource sharing. Moreover, the research results can provide references for education administrators and curriculum designers to help them formulate more effective teaching strategies and policies to promote the sustainable development and innovation of online education.

2. THEORETICAL BASIS

2.1 Theory of Open Online Courses

2.1.1 Development Process and Definition

The evolution of open online courses (MOOCs) began in 2008, based on connectivism learning theory, which was proposed by Canadian education scholars George Siemens and Stephen Downes. This saw the emergence of the first MOOC. With the rapid development of Internet technology, MOOCs have rapidly gained worldwide attention and application. In 2011, Stanford University's "Artificial Intelligence" course attracted 160,000 student enrollments, indicating the rapid popularity and scale of MOOCs. Large MOOC platforms such as Coursera, edX, and Udacity have since been launched, providing a rich selection of courses for learners around the world. Typically, a MOOC has these characteristics: the course content is open and anyone can register and study for free; the course size is not limited and can accommodate a large number of students at any one time; online teaching and learning is conducted mainly through video lectures, interactive discussions and online quizzes. The core concept of MOOCs is to use Internet technology to achieve widespread sharing of high-quality educational resources, promote educational equity, and encourage lifelong learning.

2.1.2 Role of MOOCs in Education

MOOCs play an important role in education. By breaking the geographical and time constraints, MOOCs enable global learners to access high-quality educational resources anytime and anywhere, fostering educational equity and resource sharing. Secondly, MOOCs provide a powerful supplement and innovation to the traditional education model. Through the introduction of multimedia teaching, interactive discussion and online testing and other means, the teaching mode and content are enriched, and student participation and learning are improved.

MOOCs also promote changes in educational concepts and teaching methods. Through data analysis and learning behavior tracking, teachers can better understand students' learning needs and behavior patterns, and then optimize instructional design and strategies. MOOCs also offer new avenues for lifelong learning and career development, helping learners to constantly update their knowledge and skills to adapt to a rapidly changing social and professional environment. MOOCs not only provide a flexible and convenient way of learning, but also encourage the popularization, innovation and personalized development of education, becoming an indispensable part of the modern education system.

2.2 Educational Quality Assessment Theory

2.2.1 Definition and Connotation of Educational Quality

Educational quality refers to the degree to which teaching practices and their results meet the educational objectives and the needs of students, covering many aspects such as educational content, teaching methods, teacher quality, learning environment and student achievement.

The quality of education can be assessed according to: (1) the scientific and systematic course content, comprising a relevant and comprehensive curriculum design that is aligned with the Loge cognitive level and learning needs of students; (2) the effectiveness and innovation of teaching methods. Teachers can use a variety of teaching methods to stimulate students' interest in learning, encourage students' active participation, and ensure in-depth understanding; (3) the professional quality of teachers themselves, as teachers need to have not only an excellent knowledge of the subject matter, but also good teaching ability and communication skills. (4) the support and adaptability of the learning environment, including the physical environment and the psychological environment, can provide students with optimal conditions for learning; (5) the comprehensiveness and development of students' achievements, focusing not only on students' academic achievements, but also on a range of other skills and abilities. In general, the quality of education is a multi-dimensional and multi-level concept that requires continuous improvement through comprehensive evaluation and innovation.

2.2.2 Online Educational Quality Assessment

Online educational quality evaluation refers to the systematic evaluation of the teaching process and results of online courses, and involves scientific and proven methods and index systems to determine the educational value of online courses in terms of learning outcomes.

Online educational quality assessment examines: (1) course design and content quality, determining whether the course structure is reasonable, whether the content is scientific, systematic and practical; (2) teaching methods and the use of technology to determine whether the teaching methods and technological tools adopted by teachers for online teaching are effective, and whether they can stimulate students' interest in learning and participating; (3) teaching interaction and student engagement. By analyzing the frequency of teacher-student interaction, the quality of discussion and student participation, the interactivity of the course and the enthusiasm of the students are assessed; (4) student learning outcomes and satisfaction, assessing students' academic performance and comprehensive performance in the course, as well as students' satisfaction and feedback on the course; (5) the consistency of technical support and platform stability to ensure the reliability of the online teaching platform and the timeliness of technical support to ensure the smooth delivery of courses.

Through big data analysis, a massive amount of data on learning behaviors can be mined and analyzed to reveal key factors affecting educational quality, providing data support and a scientific basis for improving teaching and learning outcomes. Combining these evaluation dimensions, the quality of online education can be evaluated comprehensively and systematically, fostering continuous improvement and development of online education.

2.3 Big Data Analysis Theory

2.3.1 Definition and Characteristics of Big Data

Big data refers to data sets that are so large in volume, speed, and variety that they are beyond the capabilities of traditional data processing tools due to their size and complexity. The first feature of big data is its *volume*, and the amount of big data is usually in TB or even PB. The second feature is the *variety* of data, comprising structured data, semi-structured data and unstructured data, such as text, images, videos, social media data, etc. The third is the *velocity* of data generation. Big data requires the processing and analysis of vast amounts of data in real time. Fourth, the *value* of big data is poor in proportion to its size. Although big data contains a lot of information, the amount of useful information is low, and valuable information needs to be extracted through data mining and analysis technology. The fifth is data *veracity*. Big data is derived from a wide range of sources with uneven data quality, and effective technical means are needed to ensure the accuracy and credibility of data.

These characteristics mean that big data analysis requires advanced technologies and tools, such as distributed computing, machine learning and data mining, to process and analyze these complex data sets, mining useful information and knowledge from them to support decision making and innovation.

2.3.2 Big Data Analysis methods

(1) Data Mining: uses statistics, machine learning and database technology to discover potential patterns and relationships from big data. Common techniques include

Table 1 Applications of big data in education.

Application Area	Specific Applications
Personalized learning	Learning path recommendation, resource recommendation, personalized learning plans
Educational quality assessment	Real-time monitoring of teaching effectiveness, analysis of teaching process data
Education management	Analysis of student registration and attendance, optimization of resource allocation
Prediction and early warning	Prediction of learning outcomes, early warning of academic risks, analysis of behavioral patterns
Student support services	Counseling and tutoring services, monitoring of psychological well-being, learning advice
Teacher performance evaluation	Analysis of teaching effectiveness, identification of teacher development needs
Course development and improvement	Feedback on course design, optimization of teaching content, course evaluation

classification, clustering, association rules and sequential pattern mining.

(2) Machine Learning: automatically learns and predicts from data through algorithms and models, including supervised learning (such as regression analysis and support vector machines), unsupervised learning (such as cluster analysis) and reinforcement learning.

(3) Text Analytics: processes and analyzes unstructured text data through natural language processing (NLP) technology to extract useful information, such as sentiment analysis and topic modeling.

(4) Social Network Analysis: through graph theory and network analysis methods, studies the relationship and structure of nodes and edges in social networks, and reveals the patterns and trends in social networks.

(5) Stream Processing: real-time processing and analysis of continuously generated data streams, often used in areas such as financial transactions, sensor data and social media data.

(6) Data Visualization: graphically presents the results of data analysis to help understand and explain complex data relationships.

These big data analysis methods are widely used in various fields such as education, healthcare, finance, and transportation, providing strong support for data-driven decision making and innovation.

2.3.3 Application of Big Data in Education

The application of big data in education has become an important means of improving educational quality and education management. First of all, big data can be used for personalized learning, through the analysis of students' learning behavior and performance data, to develop personalized learning paths and recommended resources to improve learning results. Secondly, the application of big data in educational quality assessment, through data mining and analysis, can monitor and evaluate the teaching effect in real time, find problems in the teaching process and adjust them in time. In addition, the application of big data in education management, through the analysis of student registration, attendance, grades and other data, optimize the allocation of school resources and management decisions. Big data can also be used for prediction and early warning, by analyzing students' behavior and performance data, predicting students' learning outcomes and potential risks, and providing timely interventions. The various applications of big data in education are shown in Table 1 below.

3. RESEARCH METHODS

3.1 Research Design

3.1.1 Research Framework

The research framework adopted for this study aims to systematically explore the factors affecting the educational quality of MOOCs, and build a scientific and feasible research model combined with big data analysis methods. The research framework is depicted in Figure 1.

3.1.2 Model Construction

This paper analyzes the factors affecting the educational quality of MOOCs by constructing multiple models, explained below.

(1) Regression analysis model

Multiple linear regression model was used to analyze the contribution degree of each influencing factor to educational quality. Assuming that educational quality Y is determined by several independent variables $X_1, X_2, \ldots, X_n$, the model is shown in formula (1) below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon \tag{1}$$

where, β_0 is the constant term, β_i is the regression coefficient of each independent variable, and ϵ is the error term. The model is used to quantify the impact of teaching content,

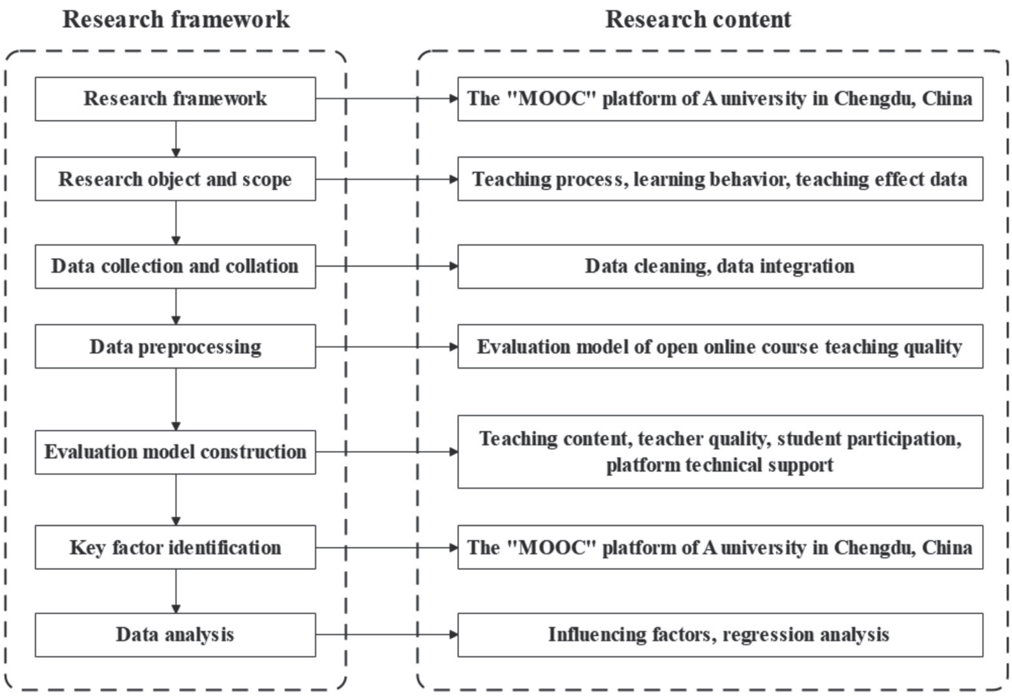

Figure 1 Research framework.

teacher quality, student engagement and platform technology on educational quality.

(2) Structural equation model

The relationships among potential variables were analyzed by structural equation model. The potential variables included teaching content quality (LCQ), teacher quality (TQ), student engagement (SE), platform technical support (PTS), etc. The model is shown in formula (2) and formula (3) below.

$$\begin{aligned} LCQ &= \lambda_1 X_1 + \lambda_2 X_2 + \epsilon_1 \\ TQ &= \lambda_3 X_3 + \lambda_4 X_4 + \epsilon_2 \\ SE &= \lambda_5 X_5 + \lambda_6 X_6 + \epsilon_3 \\ PTS &= \lambda_7 X_7 + \lambda_8 X_8 + \epsilon_4 \end{aligned} \tag{2}$$

$$Y = \beta_{LCQ} LCQ + \beta_{TQ} TQ + \beta_{SE} SE + \beta_{PTS} PTS + \epsilon_5 \tag{3}$$

where, λ is the load of the observed variable on the potential variable, β_i is the path coefficient of the potential variable on the result variable, and ϵ_i is the error term.

(3) Analytic hierarchy Process

The multi-level structure model of educational quality assessment is constructed by AHP, and the weight of each factor is determined. It is assumed that the objective layer A of educational quality assessment includes the criterion layer $B_1, B_2, \ldots, B_m$ and the specific index layer $C_1, C_2, \ldots, C_n$ under each criterion. The model is shown in formula (4) below.

$$\begin{aligned} A &= \sum_{i=1}^{m} w_i B_i \\ B_i &= \sum_{j=1}^{n} w_{ij} C_j \end{aligned} \tag{4}$$

where, w_i, w_{ij} is the weight of each criterion and indicator respectively.

3.1.3 Selection of Variables and Indicators

The selection of variables and indicators is a major step in the construction of a educational quality evaluation model, and involves multi-dimensional and multi-level factors. The research variables and their indicators are shown in Table 2 below.

3.2 Data source and collection

3.2.1 Data Source

The data source is mainly the MOOCs platform of A University in Chengdu, China. China Chengdu A University is a comprehensive university with rich educational resources and advanced teaching facilities, and has a leading position in the field of online education. The school's MOOC platform offers a large number of high-quality online courses across a wide range of disciplines, designed to meet the learning needs of different students. The data collected are as follows:

(1) Basic information about the course, such as course name, course category, course length and teaching teacher; Second,

(2) Teaching practice data, including the number of views of each lesson, learning progress, video stay time and chapter completion;

(3) Students' learning behavior data such as login frequency, online discussion participation, homework submission and test scores;

(4) Platform technical support data, including system login logs, technical problem feedback records and resolution time.

Table 2 Selected variables and indicators.

Variable	Indicator Abbreviation	Indicator
Teaching Content Quality (LCQ)	CDR	Course Design Rationality
	TRR	Richness of Teaching Resources
	CUF	Content Update Frequency
Teacher Quality (TQ)	TPS	Teacher Professionalism
	TMI	Innovation in Teaching Methods
	TIF	Frequency of Teacher-Student Interaction
Student Engagement (SE)	LMI	Intensity of Learning Motivation
	EAP	Participation in Extracurricular Activities
	ODA	Activity Level in Online Discussions
Platform Technical Support (PTS)	PS	Platform Stability
	TSR	Responsiveness of Technical Support
	LDV	Degree of Learning Data Visualization
Educational Quality (Y)	SS	Student Satisfaction
	LA	Learning Achievement
	CR	Course Completion Rate

Through the collection and analysis of these data, we can comprehensively evaluate the educational quality of online open courses and provide data support for improving the teaching and learning outcomes.

3.2.2 Data Collection Methods

The big data capture technology is used as the data collection method. It automatically collects the required data from the MOOC platform of A University in Chengdu, China by writing crawlers and using data interfaces. The steps are as follows:

The first step is to define the target data set, including basic course information, and data on the teaching process, student learning behavior, and platform technical support. Then, the web crawler program is written to simulate the user to visit the platform page and extract the required data by parsing the HTML structure.

Step 2: Data capture:

(1) Send an HTTP request to obtain the page content;

(2) Parse page content and extract data;

(3) Store the extracted data.

In order to ensure the accuracy and completeness of the data, data cleaning and pre-processing are also required. This involves removing duplicate data, filling in missing values, and standardizing data formats.

3.3 Data Analysis Methods

(1) Data preprocessing

The preprocessing of data involves data cleaning, data integration, data transformation and data reduction. Data cleaning is used to deal with missing data, noisy data, and duplicate data to ensure data quality and consistency. Data integration consolidates data from different sources into a unified data set. Data transformation converts the data into a format suitable for analysis, and applies normalization processing.

(2) Data mining technology

This is used to extract valuable information and knowledge from a large amount of data. The data mining methods adopted in this study are classification and clustering. Classification is used to assign data to predefined categories, and common algorithms include decision trees, support vector machines, and neural networks. Clustering is used to group data into different clusters to discover patterns and structures in the data. Common algorithms include K-means, hierarchical clustering, and DBSCAN.

(3) Statistical analysis

Descriptive statistical analysis is used to summarize and describe the main features of the data. Common methods include mean, median, standard deviation, and frequency distribution. Inferential statistical analysis is used to infer population characteristics from sample data, and common methods include hypothesis testing, confidence intervals, and regression analysis.

4. DATA ANALYSIS AND RESULTS

4.1 Descriptive Statistical Analysis

The results of descriptive statistical analysis are shown in Table 3 below.

4.2 Analysis of Influencing Factors

4.2.1 Ccourse Content Factors

The factors associated with course content have an important impact on the educational quality of MOOCs, including the rationality of course design, the richness of teaching resources and the frequency of content updating. These factors directly affect students' learning experience and learning outcomes, and then affect the overall quality of their education. The

Table 3 Descriptive statistics analysis.

Indicator	Mean	SD	Min	Max
CDR	3.85	0.74	2.00	5.00
TRR	4.10	0.65	2.50	5.00
CUF	3.70	0.82	1.50	5.00
TPS	4.20	0.60	3.00	5.00
TMI	3.95	0.78	2.00	5.00
TIF	3.60	0.85	1.50	5.00
LMI	4.00	0.70	2.50	5.00
EAP	3.55	0.88	1.00	5.00
ODA	3.75	0.80	1.50	5.00
PS	4.30	0.55	3.00	5.00
TSR	3.90	0.75	2.00	5.00
LDV	3.65	0.83	1.50	5.00
SS	4.05	0.68	2.50	5.00
LA	3.80	0.77	2.00	5.00
CR	3.50	0.90	1.00	5.00

Table 4 Teaching content factors.

Indicator	Mean	SD	Min	Max
CDR	3.85	0.74	2.00	5.00
TRR	4.10	0.65	2.50	5.00
CUF	3.70	0.82	1.50	5.00
TRU	3.90	0.77	2.00	5.00
CDA	3.60	0.85	1.50	5.00
TRA	4.20	0.60	3.00	5.00
CCI	3.75	0.80	1.50	5.00
TUP	4.00	0.70	2.50	5.00

Table 5 Teacher factors.

Indicator	Mean	SD	Min	Max
TPS	4.20	0.60	3.00	5.00
TMI	3.95	0.78	2.00	5.00
TIF	3.60	0.85	1.50	5.00
TFT	4.00	0.75	2.50	5.00
TPC	4.10	0.70	2.50	5.00
TQF	3.70	0.82	1.50	5.00
TEP	4.05	0.68	2.50	5.00

results for the factors influencing course content are shown in Table 4 below.

Among the teaching content factors, the mean value of teaching resource richness (TRR) is 4.10, the standard deviation is 0.65, the minimum value is 2.50, and the maximum value is 5.00, indicating that the teaching resources of most courses are abundant and concentrated. The mean value of learning resource accessibility (TRA) is 4.20, the standard deviation is 0.60, the minimum value is 3.00, and the maximum value is 5.00, indicating that learning resources generally have high accessibility. The mean value of course design rationality (CDR) is 3.85, the standard deviation is 0.74, the minimum value is 2.00, and the maximum value is 5.00. The rationality of course design is high, but there is still some room for improvement. The mean value of content update frequency (CUF) is 3.70, the standard deviation is 0.82, the minimum value is 1.50, and the maximum value is 5.00, indicating that the frequency of course content updates is high, but the frequency of updating various courses differs greatly.

4.2.2 Teacher Factors

Factors associated with teachers themselves have an important bearing on the quality of the teaching in MOOCs. They include the professional level of teachers, the innovation of teaching methods and the frequency of teacher-student interaction. These factors directly affect students' learning experience and learning outcomes, and thus have an important impact on the overall quality of their education. The results of teacher-associated factors are shown in Table 5 below.

The mean value of teachers' professional level (TPS) is 4.20, the standard deviation is 0.60, the minimum value is 3.00, and the maximum value is 5.00, indicating that most teachers have a high professional level, and the distribution of this index is relatively concentrated. The mean value of teachers' curriculum preparation adequacy (TPC) is 4.10, the

Table 6 Student factors.

Indicator	Mean	SD	Min	Max
LMI	4.00	0.70	2.50	5.00
EAP	3.55	0.88	1.00	5.00
ODA	3.75	0.80	1.50	5.00
SLA	3.90	0.77	2.00	5.00
STM	3.65	0.83	1.50	5.00
PLC	3.80	0.75	2.00	5.00
LSA	3.70	0.78	2.00	5.00

standard deviation is 0.70, the minimum value is 2.50, and the maximum value is 5.00. Teachers are generally fully prepared for the curriculum. The mean value of teaching feedback timeliness (TFT) is 4.00, the standard deviation is 0.75, the minimum value is 2.50, and the maximum value is 5.00, indicating that most teachers can give timely feedback on students' learning problems.

However, the mean value of teacher-student interaction frequency (TIF) is 3.60, the standard deviation is 0.85, the minimum value is 1.50, and the maximum value is 5.00. The mean value of this indicator is low, and there are large differences between the values for different courses. The mean value of teaching method innovation (TMI) is 3.95, the standard deviation is 0.78, the minimum value is 2.00, and the maximum value is 5.00, indicating that teachers are relatively innovative in teaching methods, but there is still room for improvement. The mean value of teachers' online question answering frequency (TQF) is 3.70, the standard deviation is 0.82, the minimum value is 1.50, and the maximum value is 5.00. There are great differences in teachers' performance in terms of answering online questions. The mean value of teacher teaching enthusiasm (TEP) is 4.05, the standard deviation is 0.68, the minimum value is 2.50, and the maximum value is 5.00, indicating that most teachers are enthusiastic about teaching and actively participate in teaching activities.

4.2.3 Student Factors

Student factors are important factors that affect the educational quality of MOOCs. These factors are: learning motivation, participation in extracurricular activities and engagement in online discussion. These factors directly affect students' learning attitude and learning outcomes, and thus have an important impact on the overall educational quality. The results of student influence factors are shown in Table 6 below.

The mean value of learning motivation intensity (LMI) is 4.00, the standard deviation is 0.70, the minimum value is 2.50, and the maximum value is 5.00, indicating that most students have high learning motivation, and the distribution of this index is relatively concentrated. The mean value of autonomous learning ability (SLA) is 3.90, the standard deviation is 0.77, the minimum value is 2.00, and the maximum value is 5.00. Students generally have strong autonomous learning ability. The mean value of online discussion activity (ODA) is 3.75, the standard deviation is 0.80, the minimum value is 1.50, and the maximum value is 5.00, indicating that students' participation in online discussion is high, although there are some differences among different students.

However, with a mean of 3.55, a standard deviation of 0.88, a minimum of 1.00, and a maximum of 5.00, the EAP has a low mean and a large variation in the participation of different students. The mean value of learning time management (STM) is 3.65, the standard deviation is 0.83, the minimum value is 1.50, and the maximum value is 5.00, indicating that students have some problems in learning time management. The mean value of learning strategy application (LSA) is 3.70, the standard deviation is 0.78, the minimum value is 2.00, and the maximum value is 5.00. The students' performance in applying learning strategy is average. Finally, the mean value of peer learning collaboration (PLC) is 3.80, the standard deviation is 0.75, the minimum value is 2.00, and the maximum value is 5.00, indicating that students perform well in peer collaborative learning, but there is still room for improvement.

4.2.4 Platform Technical Factors

Platform technology factors are important factors that affect the educational quality of MOOCs, including platform stability, technical support response speed and learning data visualization degree. These factors directly affect students' learning experience and learning outcomes, and thus have an important impact on the overall quality of the online education. The platform technology factors are shown in Table 7 below.

The mean value of platform stability (PS) is 4.30, the standard deviation is 0.55, the minimum value is 3.00, and the maximum value is 5.00, indicating that most platforms have high stability and the distribution of this index is concentrated. The mean value of data security (DS) is 4.10, the standard deviation is 0.60, the minimum value is 2.50, the maximum value is 5.00, and the platform generally has high data security. The mean content transfer speed (CTS) is 4.20, the standard deviation is 0.65, the minimum value is 2.50, and the maximum value is 5.00, indicating that the platform performs well in terms of content transfer speed.

However, the Visualization Degree (LDV) of the learning data has a mean of 3.65, a standard deviation of 0.83, a minimum of 1.50, and a maximum of 5.00, and the mean of this indicator is low and there are substantial differences between different platforms. The mean value of technical support response speed (TSR) is 3.90, the standard deviation is 0.75, the minimum value is 2.00, and the maximum value is 5.00, indicating that the platform

Table 7 Platform technical factors.

Indicator	Mean	SD	Min	Max
PS	4.30	0.55	3.00	5.00
TSR	3.90	0.75	2.00	5.00
LDV	3.65	0.83	1.50	5.00
PUF	4.00	0.70	2.50	5.00
DS	4.10	0.60	2.50	5.00
MCC	3.85	0.78	2.00	5.00
CTS	4.20	0.65	2.50	5.00

Table 8 Regression analysis.

Variable	Coefficient (β)	Standard Error (SE)	t-value (t)	Significance Level (p)
Teaching Content Factor (X1)	0.35	0.08	4.375	0.000
Teacher Factor (X2)	0.30	0.07	4.286	0.000
Student Factor (X3)	0.25	0.09	2.778	0.006
Platform Technical Factor (X4)	0.20	0.06	3.333	0.001
Constant Term ($\beta 0$)	1.50	0.30	5.000	0.000

has certain problems in terms of technical support response speed. Mobile compatibility (MCC) has a mean value of 3.85, a standard deviation of 0.78, a minimum value of 2.00, and a maximum value of 5.00. Platform user-friendliness (PUF) has a mean of 4.00, a standard deviation of 0.70, a minimum of 2.50, and a maximum of 5.00, indicating that most platforms are relatively user-friendly, but there is still room for improvement.

4.3 Regression Analysis

This study analyzed the influence of teaching content, teacher quality, student factors and platform technology factors on MOOCs educational quality through multiple linear regression model. The results of regression analysis are shown in Table 8 below.

The effects of all four variables on educational quality were statistically significant ($p < 0.05$). The regression coefficient of teaching content factor (X1) is 0.35, indicating that teaching content has the greatest impact on educational quality, the standard error is 0.08, the T-value is 4.375, and the significance level is 0.000. The regression coefficient of teacher factor (X2) is 0.30, the standard error is 0.07, the T-value is 4.286, and the significance level is 0.000, indicating that teacher factor also has a significant impact on educational quality. The regression coefficient of student factor (X3) is 0.25, the standard error is 0.09, the T-value is 2.778, and the significance level is 0.006. The influence of student factor on educational quality is relatively small, but still significant. The regression coefficient of the platform technology factor (X4) was 0.20, the standard error was 0.06, the T-value was 3.333, and the significance level was 0.001, indicating that the platform technology factor had a significant impact on the educational quality although it was slightly smaller than other factors.

5. DISCUSSION

5.1 Results Discussion

Through regression analysis and descriptive statistics, this study explored the main factors affecting the educational quality of MOOCs. The results showed that teaching content, teacher quality, student participation and platform technology all had significant effects on the educational quality. First of all, teaching content has the greatest impact on educational quality, with a regression coefficient of 0.35 ($p = 0.000$), which indicates that the optimization of the course design, the richness of teaching resources and the frequency of content update can significantly improve students' learning outcomes. In descriptive statistical analysis, the higher mean values of teaching content, such as the rationality of course design (mean 3.85) and the richness of teaching resources (mean 4.10), further support this conclusion.

The influence of teacher quality on student learning outcomes was significant, with a regression coefficient of 0.30 ($p = 0.000$). High levels of teacher professionalism (mean 4.20), innovative teaching methods (mean 3.95) and frequency of teacher-student interaction (mean 3.60) are critical to improving the quality of teaching and learning. This shows that improving teachers' professional ability and teaching methods, and increasing the frequency of teacher-student interaction can significantly improve students' learning experience and learning outcomes.

The regression coefficient of student factors was 0.25 ($p = 0.006$), indicating that the degree of motivation to learn, participation in extracurricular activities, and active online discussion contributed significantly to the quality of teaching. According to the data analysis, students' level of motivation (mean 4.00) and independent learning ability (mean 3.90) are high, but their participation in extra-curricular activities (mean 3.55) is low, indicating that it is necessary to encourage students to participate and enrich the types of extra-curricular activities being offered in order to improve the overall learning outcomes.

The regression coefficient of the influence of platform technology factors on educational quality is 0.20 ($p = 0.001$), which is relatively small, but still statistically significant. The stability of the platform (mean 4.30) and content transfer speed (mean 4.20) are high, but the degree of learning data visualization (mean 3.65) and the speed of technical support response (mean 3.90) are relatively low, suggesting that technical support and data visualization tools need to be improved to better serve the teaching process.

5.2 Research Limitations

Although this study systematically analyzes the main factors affecting the educational quality of MOOCs and provides valuable theoretical and practical guidance, it also has certain limitations. First of all, the data source is limited to the MOOC platform of A University in Chengdu, China, and the sample range is relatively narrow. Hence, it cannot fully represent all MOOC platforms, limiting the generalizability of the research results. Secondly, although the regression analysis model adopted in this study reveals the influence of various factors on the quality of teaching, the explanatory and predictive power of the model is limited in that certain variables were not included, such as the background information of students, the teaching experience of teachers and the specific technical characteristics of the platform. Therefore, the model fails to fully take into account the complexity of the education ecosystem.

Finally, the research relies mainly on the analysis of quantitative data, and lacks an in-depth examination of the qualitative factors that may influence teaching, such as students' subjective experience, teachers' teaching style, the nature of teacher-student interactions etc. These factors have an important impact on the quality of teaching but are not fully explored in this study. Finally, the inevitable errors and biases associated with data collection and processing may affect the accuracy and reliability of the results.

Future research can further improve the comprehensiveness and explanatory power of the research by expanding the sample scope to include more MOOCs platforms of different types and backgrounds, and combining qualitative research methods to further explore the subjective and situational factors affecting teaching practice, so as to better serve the development and delivery of online education.

Funding

This work was supported by 2024 Annual Higher Education Teaching Reform Research and Practice Project in Henan Province, "Double-element Orientation and Five-chain Integration": Research and Practice on the Construction Path of Tourism Professional Cluster in Higher Vocational Schools (No. 2024SJGLX0719).

REFERENCES

1. Yang, H., & Loghej, H. (2019). Modelling and analysis of the influence of affective factors on students' learning efficiency improvement based on big data. *International Journal of Continuing Engineering Education and Life-Long Learning, 29*(4), 362-373. https://doi.org/10.1108/IJCEELL-06-2019-0100
2. Yao, D., & Lin, J. (2023). Identifying key factors influencing teaching quality: a computational pedagogy approach. *Systems, 11*(9), 455. https://doi.org/10.3390/systems11090455
3. Chen, Q., Wang, J., Kateb, F., & Kharabsheh, R. (2021). Informationisation construction of pharmacology continuing education: a case study on big data analysis of the aetiology of rheumatoid arthritis. Applied Mathematics and Nonlinear Sciences, 6(2), 209-214. https://doi.org/10.2478/amns.2021.2.00032
4. Liu, J. K., Yi, Y. Q., & Wang, X. T. (2022). Influencing factors for effective teaching evaluation of massively open online courses in the COVID-19 epidemics: An exploratory study based on grounded theory. *Frontiers in Psychology, 13*, 964836. https://doi.org/10.3389/fpsyg.2022.964836
5. Li, H. X., Zhao, C. L., Long, T. T., Huang, Y., & Shu, F. F. (2021). Exploring the reliability and its influencing factors of peer assessment in massive open online courses. British *Journal of Educational Technology, 52*(6), 2263-2277. https://doi.org/10.1111/bjet.13143
6. Le, D. L., Giang, T. V., Ho, D. K., & Pham-Huynh, H. N. (2023). Factors affecting the undergraduate student's satisfaction in short-term online courses: a case study of Vietnamese pedagogical students. *European Journal of Contemporary Education*, *12*(1), 105-117. https://doi.org/10.13187/ejced.2023.1.105
7. Zhang, J. R., Qiu, F. Y., Wu, W., Wang, J. Y., Li, R. Q., Guan, M. J., & Huang, J. (2023). E-Learning behavior categories and influencing factors of stem courses: a case study of the open university learning analysis dataset (OULAD). Sustainability, 15(10), 8235. https://doi.org/10.3390/su15108235
8. Kong, H., Fan, H., Zhao, Y., Zhai, C., Zhang, C., & Han, Y. (2018). Design of teaching quality evaluation model based on fuzzy mathematics and SVM algorithm. *Journal of Intelligent & Fuzzy Systems, 35*(3), 3091-3099. https://doi.org/10.3233/JIFS-169663
9. Fang, A., Xie, S., Cui, L., & Harn, L. (2019). Research on the structure and practice of internet environment of things based on big data analysis. *Ekoloji, 28*(107), 4239-4247. https://doi.org/10.5053/ekoloji.2019.107
10. El-Sakran, A., Salman, R., & Alzaatreh, A. (2022). Impacts of emergency remote teaching on college students amid COVID-19 in the UAE. International Journal of *Environmental Research and Public Health, 19*(5), 2979. https://doi.org/10.3390/ijerph19052979
11. Hsueh, N. L., Daramsenge, B., & Lai, L. C. (2022). Exploring the influence of students' modes of behavioral engagement in an online programming course using the partial least squares structural equation modeling approach. *Journal of Information Technology Education-Research, 21*, 403-423. https://doi.org/10.28945/5010
12. Zhou, L. Q., Xue, S. J., & Li, R. Q. (2022). Extending the Technology acceptance model to explore students' intention to use an online education platform at a university in China. SAGE Open, 12(1), 21582440221085259. https://doi.org/10.1177/21582440221085259
13. Lin, Y., & Yu, B. (2022). The evaluation of university course quality under the background of wireless communication and big data. Wireless Communications & Mobile Computing, 9639641. https://doi.org/10.1155/2022/9639641
14. Chen, C., (2024). Design of intelligent education decision support system based on big data analysis. Engineering Intelligent Systems, 30(4), 289-298.

15. Ji, S., & Tsai, S. (2021). A study on the quality evaluation of English teaching based on the fuzzy comprehensive evaluation of bat algorithm and big data analysis. Mathematical Problems in Engineering, 4418399. https://doi.org/10.1155/2021/4418399

16. Du, Z., & Su, J. (2021). Analysis of the practice path of the flipped classroom model assisted by big data in English teaching. Scientific Programming, 1831892. https://doi.org/10.1155/2021/1831892