Modern Management Based on Big Data IV
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA230183

Research on the Influencing Mechanism of Higher Education Quality in the Context of Big Data*

Tengjun GUO [a, c], Danli LI [b,1], Qingzhong SUN[a,2] and Yulan LIANG [c,3]
[a] *Jinan University, Guangzhou, P. R. China. 510632*
[b] *GuangDong Teachers College of Foreign Languages and Arts, Guangzhou, P. R. China.510640*
[c] *South China Normal University, Guangzhou, P. R. China. 510631*

Abstract. The development of big data has profoundly changed not only the way of human production and life but also the way we learn and educate ourselves. This research conducted an empirical analysis on the influencing mechanism of the quality of higher education using the structural equation model, finding that five factors "Active and Collaborative Learning (ACL)","Student and Faculty Interactive (SFI)","Level of Academic Challenge (LAC)","Supportive Campus Environment (SCE)" and "Enriched Educational Experiences (EEE)" influence the indicators "Knowledge Obtaining (KO)" ,"Ability Obtaining (AO)" and "Value Obtaining (VO)". SCE has the largest impact on KO, while the influence of ACL is relatively weak. On this basis, this paper, by combining the features of big data, presented four methods for improving the quality of higher education: first, enhancing the construction of the education data platform; second, energizing energize education reform via big data; third, accurately identifying the level of academic challenge and selecting the appropriate education model by the aid of the big data technology; and fourth, based on the information technology, building a harmonic teacher-student relationship and peer relationship.

Keywords. Information Technology, Big Data, Higher Education, Education Quality, Learning Engagement

1. Introduction

Big data, featured by convenience and efficiency, has imposed a significant influence on our lives. However, the popularization of big data has also impacted traditional higher education, especially in the aspects of teaching role cognition and education

[1] Author Tengjun GUO and Danli LI contributed equally to this work.

[2] Corresponding Authors, Qingzhong SUN, Jinan University, Guangzhou, P. R. China. 510632; E-mail: 1158454922@qq.com and Yulan LIANG, South China Normal University, Guangzhou, P. R. China. 510631; E-mail: 275557821@qq.com.

[3] Author Qingzhong SUN and Yulan LIANG contributed equally to this work.

* This research is supported by the Major Philosophical and Social Science Research Project of the Ministry of Education(21JZD061) and the Guangdong Province Undergraduate Teaching Quality and Teaching Reform Project"An Empirical Study on the Learning Involvement of Hong Kong and Macao Students in Mainland Universities -- A Case study of Jinan University".

methods. In the context of big data, higher education needs corresponding transformation.

Today, research on the quality of higher education worldwide is shifting from the macroscopic level to the microcosmic level, and researchers start to evaluate the effectiveness of education from the aspect of students' individual development. Learning engagement refers to the time and energy students put into effective learning activities and how they perceive the school's support of their study[1][2]. This theory emphasizes student-centered educational philosophy and pays attention to students' real learning experience and their growth and development. It is an important aspect for evaluating students' comprehensive quality and is regarded as a critical theoretical thesis for promoting quality evaluation of higher education to shift from scale orientation to connotation orientation.

This research explores the quality of higher education by studying the relationship between learning engagement and education quality in the context of big data. Besides, by analyzing the factors that influence college students' gains from learning and the acting mechanism of these factors in the context of big data, this paper presents the methods of improving the quality of higher education.

2. Theoretical basis

George officially proposed the theory of learning engagement in 2001, pointing out that learning engagement refers to the time and energy invested by students in academic and effective educational practices, as well as the degree of support provided by the campus environment for students' academic performance. It is the interaction between students' behavioral feelings and the academic environment[3]. Based on the theory of learning Engagement, the United States, Australia, and other countries have carried out surveys of college students' learning situation, in order to provide reference for improving students' learning effectiveness, among which the National Survey of Student Engagement in the United States, NSSE) had the most widespread impact.

On the basis of NSSE, a research team represented by Tsinghua University developed the Survey of Chinese College Students (CCSS). Driven by this project, some experts and scholars have carried out theoretical and practical research on college students' learning engagement, including examining Chinese higher education evaluation from the perspective of learning engagement, promoting CCSS questionnaire indicators and content localization process, and conducting research in combination with specific teaching practices. For sample, Tian Tian et al. built a five-dimensional second-order factor structure model of college students' learning engagement based on CCSS questionnaire data[4]. Sun Dongmei et al. found that interpersonal environment, cultural environment and learning motivation all have significant positive influences on educational harvest[5]. However, previous studies on engagement in learning did not take big data into account as an important factor, which is explored in this study.

From the existing research results on the influence mechanism of learning input on education quality, the research conclusion supports that all dimensions of learning input can predict education quality. For example, Li Xiongying et al. proved that learning input of students in "top-of-the line program" has a significant positive impact on learning harvest[6]. Xing Quanchao pointed out that activities such as "campus environment support degree" and "educational experience richness degree" have a

significant impact on undergraduate students' academic performance[7]. Wang Yashun used a multilayer linear model to explore the mechanism of the influence of college students' learning engagement on learning gain. The results show that college students' learning engagement has a higher explanatory rate on learning gain[8]. Liu Yunyun made an empirical analysis and found that hardware, interpersonal and cultural environment were significantly positively correlated with satisfaction at school[9]. Therefore, this paper proposes a research hypothesis: under the background of big data, college students' learning involvement also has a significant positive impact on the quality of higher education.

3. Research design

3.1. Research tool

This research employed CCSS questionnaire as the survey tool. The CCSS questionnaire constructs five-dimensional indicators of learning engagement: Level of Academic Challenge (LAC), Supportive Campus Environment (SCE), Active and Collaborative Learning (ACL), Student and Faculty Interactive (SFI) and Enriched Educational Experiences (EEE). In this paper, these five comprehensive and comparable indicators were selected as the factors for evaluating college students' learning engagement. The quality of education is first reflected in the quality of the education receivers. Therefore, this research focused on students' gains from learning and took the educational outcome-oriented view on value-added education quality evaluation, choosing three indicators, namely Knowledge Obtaining (KO), Value Obtaining (VO) and Ability Obtaining (AO) from indicator "Learning gains from self-report" of the CCSS questionnaire as the indicators to evaluate the education quality.

3.2. Research sample

In this study, samples were collected by grade stratification random sampling method, and questionnaire links were released on wechat and QQ groups in Guangdong universities that widely use big data management, by "Questionnaire Star" tool from March to April 2022. A total of 1764 questionnaires were collected, and 1394 valid questionnaires were obtained after deleting invalid data, with an effective rate of 79%.

4. Empirical analysis

4.1. Testing using the measurement model

4.1.1. Factor analysis

As for the exploratory factors, this research used Spss26.0 software and conducted suitability testing on pretest data with the aid of the KMO value and Bartlett's test of sphericity. The result shows KMO=0.925 and Bartlett's test P=0.00＜0.01, indicating that data in this research can be used for exploratory factor analysis. The method of principal component analysis-direct Oblimin was used to extract factors. After

retaining factors with a characteristic value larger than 1 and deleting the items with a communality smaller than 0.04, a loading value smaller than 0.40 and cross-loading of factors, five factors were finally extracted from the learning engagement scale. The final learning engagement scale extracted five factors with a communality between 0.598 and 0.942 and a cumulative variance explained of 62.13%, while the education quality scale extracted three factors with a communality between 0.560 and 0.913 and a cumulative variance explained of 65.91%, both scales reaching a satisfactory level.

As for the confirmatory factors, this research used Mplus7.0 software and conducted confirmatory factor analysis using the maximum likelihood (ML) method, so as to verify the construct validity of the scales. According to the results, the goodness-of-fit indexes of the learning engagement model are $\chi 2/df=2.168$, CFI=0.985, TLI=0.959, RMSEA=0.053, and SRMR=0.072; the goodness-of-fit indexes of the education quality model is $\chi 2/df=2.586$, CFI=0.912, TLI=0.907, RMSEA=0.049, and SRMR=0.069. This reveals that these two analysis models of the confirmatory factors have a good fitting performance, as shown in Table 1.

Table 1. Goodness-of-Fit Analysis

Goodness-of-Fitting Index	χ2/df	RMSEA	TLI	CFI	SRMR
Learning engagement model	2.168	0.053	0.959	0.985	0.072
Education quality model	2.586	0.049	0.907	0.912	0.069
Fitting criteria	＜3	＜0.08	＞0.90	＞0.90	＜0.10

4.1.2. Reliability and validity testing

As for reliability, SPSS26.0 was used for reliability testing, and the Cronbach's α coefficient is shown in Table 2. According to the result, the overall Cronbach's α value of the two scales is larger than 0.80, and the Cronbach's α value of each factor is larger than 0.70, indicating good reliability of the questionnaires. At the same time, the combined reliability (CR) of all variables is between 0.726-0.872, larger than 0.60 and reaching a statistically significant level. This reveals a good internal consistency of these two questionnaires and the influence factors, as shown in Table 2.

As for validity, the average variance extracted (AVE) was used to measure the convergent validity. The validity value of each factor is larger than 0.50, which is acceptable. In terms of the discriminant validity, the AVE square root of each factor is overall larger than the correlation coefficient of their respective correlation factors. This indicates a good discriminant validity of the factors and a good correspondence between the measured items, as shown in Table 3.

Table 2. Scale of Reliability and Convergent Validity of the Measurement Model

Scale	Factors	Items	Combined reliability (CR)	α coefficient	Scale α coefficient	Convergent validity
Learning engagement	LAC	9	0.801*	0.796	0.831	0.653
	SCE	10	0.793***	0.783		0.569
	ACL	11	0.760*	0.756		0.461
	SFI	9	0.778**	0.781		0.635
	EEE	12	0.745***	0.822		0.632

Education quality	KO	2	0.872**	0.849		0.547
	AO	10	0.741**	0.786	0.843	0.516
	VO	3	0.726***	0.854		0.620

Note: *** represents p＜0.001, ** represents p＜0.01, and * represents p＜0.05.

Table 3. Discriminant Validity Testing of the Measurement Model

Factors	1. LAC	2. SCE	3. ACL	4. SFI	5. EEE	6. KO	7. AO	8. SCO
1. LAC	0.772							
2. SCE	0.425	0.668						
3. ACL	0.324	0.311	0.751					
4. SFI	0.251	0.365	0.233	0.635				
5. EEE	0.310	0.326	0.224	0.259	0.826			
6. KO	0.266	0.351	0.313	0.367	0.282	0.752		
7. AO	0.339	0.239	0.390	0.392	0.262	0.311	0.846	
8. VO	0.412	0.347	0.408	0.345	0.277	0.372	0.269	0.854

Note: The values on the diagonal are the AVE square root values, and data in the lower triangular matrix are correlated coefficients of the factors.

4.2. Structural equation modeling

The first step of structural equation modeling is to set up the model to be estimated. Based on the research hypothesis, The statistical tool Mplus7.0 was used in this study to build an influence path model for college students' learning gains, as shown in Figure 1. The result of fitting index testing ($\chi 2/df=2.621$, RMSEA=0.045, CFI=0.914, TLI=0.909, SRMR=0.072) reveals that the fitting indexes of the model are in the acceptable scope and the model can well fit data, as shown in Table 4.

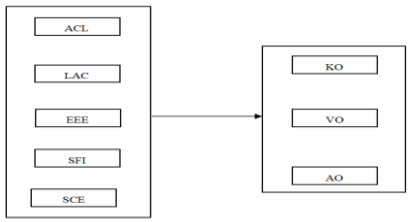

Figure 1. Hypothesis of China's higher education quality influence path model

Table 4. Fitting Effect Evaluation Table of the Structural Model

Fitting index	χ2/df	RMSEA	TLI	CFI	SRMR
Learning engagement model	2.621	0.045	0.909	0.914	0.072

4.3. Hypothesis path testing

The model in Figure 2 clearly shows the factors that influence the quality of higher education, the influence paths, and the structural relationships of these factors. For

college students, the five factors of LAC, SCE, ACL, SFI and EEE have a positive influence on KO, VO and AO. According to the results of the data analysis, in this model, except that the paths of LAC and VO, SFI and AO are not significant, there is a statistically significant correlation among 13 paths. There is a significant positive correlation in the paths between LAC and KO (β=0.561, p<0.05) and AO (β=0.455, p<0.001); there is no significant correlation in the paths between LAC and VO (β=0.496, p>0.05). There is a significant positive correlation in the paths between SCE and KO (β=0.652, p<0.001), AO (β=0.606, p<0.05) and VO (β=0.674, p<0.01). The paths between ACL and KO (β=0.295, p<0.05), AO (β=0.327, p<0.001) and VO (β=0.351, p<0.05) are significantly positive. There is a significant positive correlation in the paths between SFI and KO (β=0.373, p<0.05) and VO (β=0.432, p<0.001), but there is no significant relationship in the paths between SFI and AO (β=0.479, p>0.05). The paths between EEE and KO (β=0.599, p<0.01), AO (β=0.635, p<0.001) and VO (β=0.624, p<0.001) have a significant positive correlation. Specific data are shown in Figure 2.

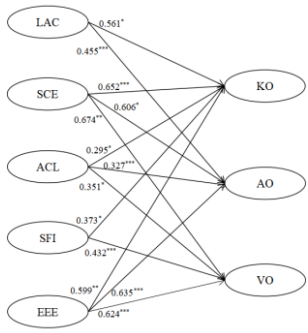

Figure 2. Structural equation model of higher education quality influence factors and influence paths

5. Discussion and suggestions

5.1. Discussion

5.1.1. All the factors of college students' learning engagement have a positive influence on the learning gains

According to the results of data analysis, college students' learning engagement has a significantly positive influence on their learning gains. In other words, students taking more challenging courses, spending more time and energy in their learning, increasing high-quality interactions with their peers and teachers, and actively taking part in various educational practices, all these activities can effectively improve their abilities such as synthetic analysis and critical thinking, and help the college students to have a better understanding of themselves and the world.

5.1.2. The influence of the factors of college students' learning engagement on their learning gains is different

The empirical research suggests that SCE is the most important factor that influences the quality of higher education, having an effect value above 0.6 with KO, VO and AO.

The second most influential factor is EEE, which has an effect value above 0.6 with AO and VO and an effect value of 0.59 with KO. In comparison, ACL and SFI have a relatively weak influence on college students' learning gains, with an influence value below 0.5.

5.1.3. SCE has the largest influence on college students' learning gains, while ACL has the smallest influence

The empirical research indicates that the influence path of SCE is the most significant factor for the quality of higher education. The main reason is that the campus environment has a powerful educational function and can unconsciously influence students' choice of values, way of thinking and behavioral tendencies[10]. As to ACL, analysis reveals a small influence coefficient between ACL and KO, VO and AO. This could be attributed to the lecture-based approach typically used by college instructors, which offers limited chances for students to engage in oral presentations, ask questions, or collaborate with peers. Consequently, ACL may not effectively enhance students' learning gains[11].

5.2. Suggestions

Today, the information technology represented by big data is developing quickly, which not only profoundly influences the reform and development of all sectors of society but also provides unprecedented opportunities for innovating the methods of higher education. Based on the results of the empirical study discussed above, we have the following suggestions of improving the quality of higher education by combining the features and advantages of information technology in the era of big data.

5.2.1. To enhance the construction of the education data platform

First, colleges need to set up an education data platform that integrates massive educational resources and information. In building this data platform, colleges should collect a magnitude of educational data and resources by screening and integrating data information and broadening the channels of data sources. At the same time, it is suggested colleges strengthen cooperation with enterprises, with the help of modern information technology, to set up a data platform integrating educational administration, student management, teaching and academic research. Furthermore, by using cloud computing technology, colleges can classify and analyze data on the platform, thereby gaining insight into the features of different student groups and seeking a basis for the innovation of education methods.

5.2.2. To energize education reform via big data

In educational practices, colleges can get an understanding of the number of participants and students' feedback on the practices with the aid of the big data platform, learn the advanced experience that helps enrich their educational pattern and content and keep broadening their thinking in the learning process. In social practices, colleges can also use big data technology to record the whole process of the out-of-school education practice in real time through various we-media platforms. They can also adopt a unified educational information-sharing system to accurately reflect

students' participation in the practices and their feedback, in an attempt to continuously improve the educational effect of the practices.

5.2.3. To accurately identify the level of academic challenge and select the appropriate education model depending on the big data technology

College administrators and teachers can collect and manage information such as the click rate of teaching content, the keyword search rate and the applicable level corresponding to students' abilities by using the big data platform. This helps the colleges to select appropriate educational content and reduce misguide from disqualified and wrong educational resources quickly and accurately. Colleges can also use the big data platform to search educational patterns of other colleges with the same major features and the same core disciplines, to learn from their beneficial experience and develop their respective educational patterns by combining their situations.

5.2.4. To build a harmonic teacher-student relationship and peer relationship based on information technology

In terms of the teacher-student relationship, teachers can use the big data platform to observe and understand each student's real situation of study and life and their mental state. Through data comparison and analysis, teachers can study the type of changes in students' thoughts, thereby communicating with students in a more targeted manner. As to the peer relationship, colleges can take full use of information technology to break boundaries in time and space, encourage more communication between students and create a harmonic and friendly peer atmosphere. This will further improve the effectiveness of collaborative learning and feasibly raise education quality.

6. Conclusions and limitations

Based on a sample of 1394 college students in China, this study proposed and tested the influence of students' learning involvement on the quality of college education in the era of big data. The results show that each dimension of college students' learning engagement has a positive influence on learning harvest, and the influence is different. Campus environment support has the biggest influence on college students' learning harvest, while active cooperative learning has the least influence. In conclusion, this study confirms that in the era of big data, students' learning involvement has a significant positive impact on the quality of higher education.

This study has some limitations that could be addressed in future studies. First, the sample size of this study is limited. More participants with diversified backgrounds could be included in future studies. Secondly, the data in the current study were collected with self-report and cross section. Although the measurements were widely used in previous studies and were proved to be reliable and valid in the current study, might still have led to a subjective bias and difficulties to track the study. Consequently, it is recommended for future studies to adopt multiple evaluation methods and data (e.g., field visits, tracking studies) with reliable measurements to evaluate the learning engagement of college students and the quality of higher education.

References

[1] KUH G D, “The National Survey of Student Engagement: Conceptual and Empirical Foundations”, New Directions for Institutional Research, vol.141,pp. 5-20,2009.

[2] ASTIN A W, ANTONIO A L. "Assessment for excellence : the philosophy and practice of assessment and evaluation in higher education”. Washington: Rowman & Littlefield Publishers pp. 21-22,2012.

[3][6] Li Xiongying, Qin Xiaoqing, “A Study on the Relationship between Learning Engagement and Learning Gain of Students in ‘Top-notch Program’ -- Also on the Promotion of Deep Learning for College Students”, Jiangsu Higher Education, vol.12, pp. 102-108, 2019.

[4] Tian Tian, Zhang Linghua, “Reresearch on the Learning Engagement Structure of College Students based on CCSS2016 Questionnaire”,Higher Education Research in Heilongjiang Province, vol.06, pp 119-125, 2018.

[5] Sun Dongmei, Liu Yunyun, “Research on the Influence Mechanism of Education Quality in research Universities: Based on the analysis of CCSS survey data of L University”, Fudan Education Forum, vol.06, pp. 56-64, 2018..

[7] Xing Quanchao, “An empirical study on the relationship between academic engagement and academic performance of undergraduate students based on CCSS survey ”, Chongqing University, pp.25, 2015.

[8] Wang Ya-shuang, “An empirical study on College students' Learning Engagement -- Based on the data of the National College Students' Learning Survey in 2012”, Higher education research in China,vol.01, pp. 32-36, 2013.

[9] Liu Yunyun, “Research on Influencing factors of college satisfaction based on CCSS survey data of L University”, Lanzhou University, pp.42, 2018.

[10] Wang Peipei, Yu Anlong, “Construction of Learning Evaluation System of Ideological and Political Courses in Colleges and Universities -- Exploration of Value-oriented innovation Path Based on moral cultivation”, Journal of Tianjin Normal University (Social Science Edition),vol.01, pp. 8-13, 2023.

[11] Li Yangyang, Liu Sannu-ya, Chen Zeng-zhao, FANG Jing, HE Xiu-ling, “Research on Collaborative Learning Model Construction from the perspective of Learning Space Integration”, Distance Education in China, vol.03, pp.42-50+60, 2022.