Does Greater Engagement in Online General Education Courses Lead to Better Academic Performance? Evidence from Chinese University Students

Yan Yang

Department of College English, Zhejiang Yuexiu University, Shaoxing, China
Email: 20152067@zyufl.edu.cn

Abstract

Although there is a plethora of online learning engagement studies, relatively little attention has been paid to the relationship between learning engagement and academic performance in the context of online general education courses. Accordingly, this study takes an online general education course offered by a university in eastern China as an example, proposes a model for evaluating online learning engagement and specific metrics for the model, and conducts a cluster analysis of the online learning engagement of 422 undergraduate students who took the course as sample data through a clustering *K*-means algorithm. Based on the relationship between learners' online learning behavioral engagement and academic performance, learners were classified into three categories: "active learners", "go-with-the-flow learners" and "passive learners". The study concludes that the classification of online learning engagement and academic performance is beneficial for teachers and administrators to grasp the whole learning process of learners in the context of online general education courses, clarify the types of online learning engagement and their characteristics, and provide data reference for students' personalized learning support service system, thus promoting the establishment of a high-quality school-based general education course system.

Keywords

Online General Education Courses, Online Learning Engagement, Academic Performance

1. Introduction

General education is highly valued in universities in China and abroad for its

ability to provide educated people with the knowledge and values that pass between different groups of people in a diverse society. In particular, after the publication of *General Education in a Free Society* (James, 1945), general education in the modern sense has flourished worldwide. Chinese universities have also gradually started to reform cultural quality education since the mid-1990s (Li, Yang, & Sun, 2001), aiming to provide more diverse choices for university students. However, the implementation of general education programs in many Chinese undergraduate institutions has suffered for a long time from a lack of faculty, a lack of diversity and systematization in the curriculum, class time that easily conflicts with other courses or activities, too little interaction between teachers and students in large classes, and a lack of student engagement in learning (Wang, 2008; Hu, 2017; Zhao, 2013; Yang & Ou, 2015).

As an important way to solve the problem of general education in China, online general education courses have become one of the first applications of educational integration inside and outside of universities to break down the walls. Relevant studies have found that over 70% of students wish to take public elective courses through the online learning system (Yang & Qu, 2015), as online learning is flexible and diverse, with dormitories and libraries becoming important places to learn general education courses, and students can earn credits by studying the videos of their chosen courses and completing the corresponding assignments, questions, discussions and examinations (Lei, 2016). However, while bringing various conveniences to teaching reform and innovative teaching methods, online general education courses also face a series of challenges. Firstly, it is difficult to ensure the management of the teaching process and teaching effectiveness across time and space. Secondly, the virtual and networked nature of teaching and learning has led to limitations in the assessment of learners. In online general education courses, teachers are unable to directly observe learners' learning behaviors, and the final assessment is based on the number of course visits and the final grade, which appears to be in line with the process evaluation approach, but in fact does not provide an in-depth understanding of learners' commitment to learning, making it difficult to fully reflect learners' learning status and effectiveness.

The above problems with online general education courses are in fact problems that exist in online learning in general. Studies have shown that due to the spatial and temporal nature of online learning, learners are prone to a lack of commitment and a high dropout rate in the process of online learning (Wei, 2012). In an era where online learning is widely used in higher education and is an indispensable way for learners to learn, the dilemma faced by online general education courses needs to be addressed urgently. Among the studies on online general education courses, there are more studies exploring the problems of learners in the learning process by means of theoretical thinking or questionnaire surveys, but relatively few studies have explored learners' learning situation through online learning behavior data, and there is a lack of clarity on learners' actual learning input in online general education and its relationship with aca-

demic performance. Based on this, this study analyzes learner engagement in online general education courses and the impact of different levels of learning engagement on learning performance from the perspective of learning engagement, and classifies the types of relationships between online learning engagement and learning performance, and proposes personalized learning support services for different types of students, with a view to providing ideas for solving the current dilemma of online general education courses.

2. Literature Review

2.1. Practice of Online General Education Courses

With the advent of online open courses, cross-school course selection, course sharing and the establishment of course consortia are technically guaranteed, providing an unprecedented opportunity to address the issue of general education courses in China. Platforms such as MOOC of Chinese Universities, Zhihuishu, Xuetang Online and Chaoxing Erya have joined forces with universities and social institutions to provide a wealth of online general education courses. According to official data from Zhihuishu Platform, over 10 million university students from more than 1900 institutions of higher education in China have used its online credit course services, and Chaoxing Erya has created a comprehensive literacy course system with six modules.

Along with the advancement of MOOCs, some problems have emerged in practice (Hao, 2013), and online general education courses are no exception. Based on questionnaires and interviews, Liu Cuiyin and others explored and summarized the problems in the implementation of university online general education courses and put forward corresponding countermeasures and suggestions (Liu, Zhang, & Shan, 2018). Chen Wenjuan et al. analyzed four aspects of the current situation of teaching online general education courses, students' learning, supervision and assessment, and students' feedback in two universities in Gansu Province by means of questionnaires, and the results showed that the main problems in the process of implementing online general education courses include: some students are forced to study online general education courses in order to complete the credit recognition requirements, which leads to their marginalization in course learning; the supervision of students in the process of conducting online general education courses is not perfect and the assessment method is too single; students do not attach much importance to online general education courses and their active learning is poor (Chen & Chen, 2018). Yao Yuanfeng and others used the method of data statistics and analysis to analyze students' selection, learning progress, learning characteristics and learning effects of online general education courses, and put forward suggestions for improvement (Yao & Zhao, 2016). Hao Xiaona wrote her own scale to measure the motivation and influencing factors of 220 undergraduate students in Erya's online general education course at Shanxi Normal University (Hou, 2017). Chen Shuwei investigated and analyzed the current situation of teaching and man-

agement of online general education courses in Zhejiang Agricultural Trade and Vocational College in the form of questionnaires and interviews, and found that the online general education courses in this institution had several prominent problems such as students' lack of attention to the courses, course management in formality and simple and lenient course assessment (Chen, 2016).

2.2. Relationship between Online Learning Engagement and Academic Performance

The concept of learning engagement was first introduced by Schaufeli (Schaufeli, Salanova, González-Romá, & Bakker, 2002), who argues that learning engagement is similar to work engagement in that it is a state of sustained, positive affect towards learning exhibited by the learner. It is characterized by energy, commitment and concentration. Numerous national and international studies have shown that learning engagement is an important factor in the ultimate learning performance of learners.

Foreign researchers have selected and classified the relationship between online learning engagement and learning performance from different perspectives. Kwon identified time management as a significant factor in the success of online learning and found a relationship between learners' level of action, time management and learning outcomes (Kwon, 2009). Jo et al. identified students' self-regulation, particularly their time management strategies, as an implicit psychological characteristic that drives normal engagement in online learning activities and leads to high performance (Jo, Yoon, & Ha, 2013). In their study, Bolliger et al. found a definite link between students' engagement awareness and their learning outcomes in online learning, with the stronger the engagement awareness and the longer the time invested in engaging in learning activities, the more likely they were to achieve higher learning outcomes (Bolliger & Halupa, 2018). Carver et al. analyzed learning data from 167 postgraduate students in nine online courses to see if there was a relationship between students' learning achievement and the amount of time they spent in each module of the course, including four areas: total course learning time, length of learning time for instructional videos, length of learning time for learning resources, and length of synchronous online sessions. The analysis revealed that the more time students spent in synchronous online sessions, the more likely they were to receive an A in the exam (Carver, Mukherjee, & Lucio, 2017).

Most of the domestic studies on online learning input assessment are based on data mining and analysis by combining the basic division dimensions of online learning input from existing studies.

Wang Yonghong used the yaahp hierarchical analysis software to determine the weights of the assessment index system, and thus constructed a complete index system for measuring the online learning engagement of university students. The system was based on three dimensions: online behavior, cognition and emotion, and is based on the application research of the mandatory course "Mobile Application Design and Development (Front-end)" of the online platform for

undergraduate education technology students in Northwest Normal University (Wang, 2020). Lang Bo et al. used a learning vectorization method to classify online learning behavior data and predict online learning performance by a genetic algorithm (Lang & Fan, 2019). Using data and models, Sun Faqin et al. investigated the primary, secondary, and irrelevant factors that influence online academic achievement through machine learning related knowledge (Sun & Feng, 2019). Zhao Xifen et al. studied the main influencing factors of online learning effectiveness using Northern Nationalities University as an example (Zhao, Li, & Liang, 2019). Guo Fangman et al. studied the correlation between learners' online learning behaviors and learning outcomes, as well as the differences between male and female students' online learning behaviors and learning outcomes (Guo & Liu, 2018). Chen Yuanyuan et al. further analyzed the relationship between adult learners' online learning behavior and learning outcomes based on data mining. Online learning behavior has been shown to have a strong link with learning performance, and learning behavior can be used as an important representation for predicting students' learning performance, and monitoring of learning behavior can guide students to engage in beneficial learning behaviors, thereby improving their effective learning ratio (Chen, Liu, & Dong, 2019).

From the above research status, it is clear that there are relatively few studies on university online general education courses, and most of the existing studies use questionnaire and interview methods. Although such research methods have their advantages, the final results may be less scientific and objective than those obtained through data analysis of learners' online learning records due to learners' subjective emotions in their responses. Of course, there are a few empirical studies that have used data statistics and analysis methods, but these data are basically at a large granularity of the total number of students who have completed the course videos, the total number of students who have participated in discussions, and the overall distribution of student performance. In addition, the majority of studies show a positive relationship between student engagement in online learning and learning performance: the higher the engagement, the higher the probability of achieving good grades. This positive effect relationship occurs in the context of mandatory university courses, adult education, or distance education, so would the same conclusions be drawn in the context of university online general education courses?

2.3. Research Questions

This study seeks to explore the relationship between online learning engagement and academic performance by collecting and analyzing various indicators and corresponding data on learners' learning engagement in the context of online general education courses. The research questions are as follows.

1) Is the relationship between learners' online learning engagement and academic performance in online general education a positive one?

2) Are there different types of online learning engagement in online general

education, and is there a non-positive relationship between online learning engagement and academic performance?

3) Is there a possibility of assessing the learning engagement characteristics of learners and giving timely and personalized learning support services according to their characteristics?

3. Research Design

3.1. Samples and Data Sources

The practical case of this paper selects an online general education course (Overview of English Countries) offered by a university in eastern China on the Wisdom Tree platform as the research object, which has been continuously built from a university-level high-quality online open course in 2017 to being recognized as the first batch of first-class model courses in Zhejiang Province in 2019, and the course is taught to students of all majors in the third year of undergraduate studies, and this study selects the learning behavior data of 432 students who participated in the online learning course in the fall of 2020-2021 as the research data, and the invalid sample was excluded, and the valid sample was 422 people. The invalid sample presented was students who had never participated in any learning activities in the course and had a final grade of 0. Of the 422 valid samples, the gender distribution was 10 male students and 412 female students, while the subject background distribution was higher for arts students with 395 students and for science students with 27 students. The online learning data involved in the study was extracted mainly through the backend. Among them, basic learner information (gender and major), days of attendance, number of resource views, number of homework assignments completed, frequency of using homework modules, number of postings and replies, number of words posted, number of videos watched can be directly exported from the backend, and the video viewing regurgitation ratio, homework scores and unit test scores are obtained by counting the corresponding data of each unit and then finding the average.

3.2. Study Procedures

A number of studies of learning engagement assessment scales have provided important references for this paper. For example, the National Survey of Student Engagement (NSSE) (Coates, 2007) was developed by Indiana University in the United States based on Coates' five-dimensional framework of learning engagement. The NSSE has become the primary reference for student engagement surveys in colleges and universities in North America and in many countries. Fredricks proposed that engagement in learning is a meaningful combination of behavioral, affective and cognitive dimensions (Fredricks & Paris, 2004), and Sun and Rueda developed the "Distance Learning Engagement Scale" based on the Fredricks Engagement Scale and the characteristics of distance learning (Sun & Rueda, 2012). Dixson developed the Online Student Engagement Survey

(OSES), which includes four dimensions: skill, emotion, engagement, and performance (Dixson, 2010). Li Shuang et al. classified distance learners' behavioral variables in the learning management system into four dimensions: online participation, active interaction, self-monitoring, and performance effort, and used the Distance Learner Engagement Scale to predict the role of behavioral variables generated by students in four courses at the National Open University on distance learning engagement (Li, Li, & Yu, 2018). Therefore, this paper intends to develop an online learning engagement model for online general education courses by collecting online learning behavior data and by analyzing the learning behavior data in four dimensions: participation, attention, interaction, and performance efforts, as shown in Figure 1.

In the model, "participation" refers to the time and effort learners put into online learning, the number of videos and texts they access on the online learning platform, etc. In this paper, the number of weeks of attendance, the number of days of attendance, and the number of resources browsing behaviors are used as the measure of "participation". The combination of weeks of attendance and days of attendance provides a more complete picture of learners' time commitment. "Attention" refers to the depth of learners' commitment in the learning process, mainly including the retention time of each study, the number of repetitions of wrong questions, the number of homework assignments completed, etc. The number of homework assignments completed is chosen here to reflect the learners' depth of learning; the frequency of using homework modules reflects to a certain extent the learners' persistence and concentration in completing homework assignments. The frequency of use of the homework module reflects to a certain extent the persistence and concentration of learners in completing homework assignments; the regurgitation ratio of video viewing refers to the ratio of the length of time spent by learners in watching the video to the original length of the video; a high regurgitation ratio means that learners have repeatedly watched the content of the video, which may be because learners are
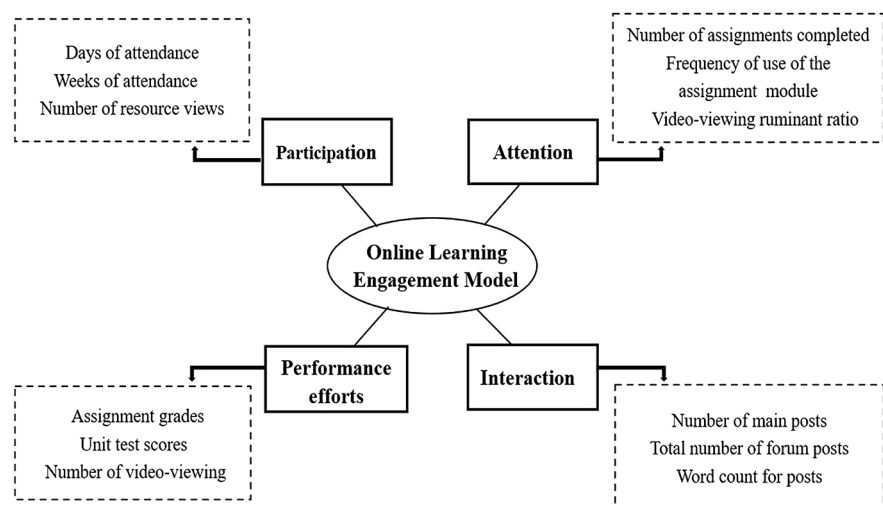
Figure 1. Online learning engagement model for online general education courses.

interested in the content and want to deepen their understanding through repeated viewing, or may be because learners are not concentrating when watching and need to replay to learn again. Of course, unexpected situations such as video lagging due to internet speed cannot be ruled out. "Interaction" refers to learners' communication with teachers and other learners in the learning process. The frequency and quality of communication and interaction can determine learners' enthusiasm for learning and reflect learners' commitment to online learning. Therefore, in this paper, the number of posts made by learners，the number of main posts and the number of words posted made by learners are chosen to reflect learners' communication and interaction with teachers and other learners through the forum. "Performance efforts" refer to the efforts made by online learners to earn course credits or achieve better grades on the indicators related to the final assessment, and the quality of learners' assignments，performance on unit tests and the number of videos watched are chosen as indicators for assessment here.

In this paper, data were collected using the online learning engagement model metric for online general education courses as a data base for measuring online learning engagement. Academic performance is the ability of learners to apply their newly acquired knowledge and skills, not only in terms of the acquisition of basic knowledge and skills, but also in terms of their ability to apply them flexibly (Zhong & Liang, 2006). Academic performance is a measure of learners' learning outcomes and is one of the most important items in the assessment of teaching quality (Zheng, Cao, Chen, & Wu, 2013). Although academic performance is not the whole content of learning performance, a quantitative result of academic performance as other assessment part of students' learning process has the characteristics of objectivity and comprehensiveness. Therefore, learners' summative exam scores are used as an academic performance metric in the paper.

Based on the above online learning input metrics and academic performance metrics, the paper uses data mining methods to analyze the data, explore the relationship between online learning inputs and academic performance, and try to clarify the types of online learning inputs and their characteristics of online learners, in order to provide ideas for personalized learning support services for teaching general education courses.

3.3. Methods

Data mining is the process of carefully analyzing large amounts of data to reveal meaningful new relationships, trends and patterns (Wang & Jiang, 2004). In this paper, cluster analysis is used, which divides a large collection of data points into classes such that the data in each class are maximally similar to each other and the data in different classes are maximally different from each other (He, Wu, & Cai, 2007). Moreover, the data object of the cluster analysis is unknown (Zou & Zhu, 2005). These algorithmic features are consistent with the purpose of this paper, which is to explore whether online learning inputs are positively corre-

lated with academic performance. Besides, SPSS 24.0 was used to do the clustering analysis and the *K*-means algorithm was chosen as the main clustering algorithm and the R language was used as the data processing method, which is a statistical, arithmetic and graphical based data processing method that is ideal for large data processing work (Song & Zhu, 2012).

4. Results and Discussion

4.1. Data Standardization

Based on the online learning input model proposed in the previous section, the data collected included data on 422 learners on a total of 12 behavioral variables across four dimensions of participation, attention, interaction and performance efforts, as well as summative exam results. As the range of values between the data varies widely, for example, the number of completed assignments ranges from 0 to 16, while the total number of resource views ranges from 0 to 2743. Without data transformation, the number of completed assignments will be ignored in the calculation, so the data needs to be uniformly transformed to the range [0, 1]. The specific conversion calculation is as follows.

$$p' = \frac{P - \min(N)}{\max(N) - \min(N)} \quad \text{wherein } P \in N$$

Compared to online learning engagement data for each dimension, learners' summative grades can be used as a more independent and important category classification feature, so they were converted to the range [0, 2] and calculated as follows:

$$P' = 2 * \frac{P - \min(N)}{\max(N) - \min(N)} \quad \text{wherein } P \in N$$

In the equation, *P* denotes the original data, *P'* denotes the transformed data and *N* denotes the original data set.

4.2. Determination of the Number of *K*-Means Cluster Classes

The choice of *K*-value in the *K*-means clustering algorithm directly affects the effect of clustering. To obtain a better clustering effect, the elbow method is applied to select the *K*-value. The core index of this method is SSE (sum of the squared errors)

$$\text{SSE} = \sum_{i=1}^{k} \sum_{p \in C_i} |p - m_i|^2$$

In the equation, C_i is the *i*th cluster, *p* is the sample points in C_i, m_i is the center of mass of C_i (the mean of all samples in C_i) and SSE is the clustering error of all samples, representing how well the clustering works.

The core idea of the elbow method is that as the number of clusters *K* increases, the sample will be more finely divided and the degree of aggregation of each cluster will gradually increase, so the sum of squared errors and SSE will natu-

rally gradually become smaller. Moreover, when K is smaller than the true number of clusters, the SSE decreases significantly because increasing K increases the degree of aggregation of each cluster. When K reaches the true clustering number, the return on the degree of aggregation obtained by increasing K again will rapidly become smaller, so the decline in SSE will plummet and then level off as the value of K continues to increase, meaning that the graph of SSE versus K is the shape of an elbow, and the value of K corresponding to this elbow is the true clustering number of the data.

In this paper, the 422 samples of standardized data were firstly selected using the elbow method to find the best clustering number K. According to previous research results of online cluster analysis, most of the catechism learner groups are divided into 3 - 6 (Qiao & Jiang, 2020), so in this paper, the value range of K was set to 2 - 8, and each K value was clustered and the corresponding SSE was noted, and the relationship between K and SSE is shown in Figure 2.

It can be seen that the corresponding K value in the graph is 3. Therefore, the optimal number of clusters for this dataset should be chosen as 3.

4.3. Clustering Results and Analysis

The standardized 422 sample data were subjected to K-means cluster Analysis in SPSS 24.0, and the number of cases in each cluster was obtained by setting the cluster number K value to 3 as shown in Table 1.

It can be seen that the number of cases included in the three clusters is 232, 85 and 105, which represent 55%, 20% and 25% of the total sample size, respectively.

From Table 2, it can be seen that learners from category 1 have a high overall performance, learners from category 2 have a low overall performance, and learners from category 3 have a medium overall performance. In order to more clearly determine the specific score levels of the three categories of learners'
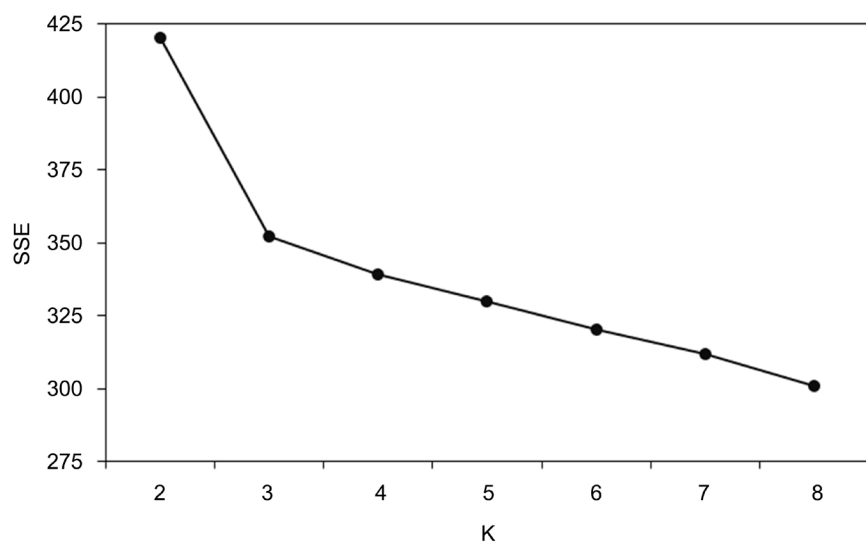

Figure 2. Relationship between K and SSE.

Table 1. Number of cases in clusters.

Cluster 1	232.000
Cluster 2	85.000
Cluster 3	105.000
Effective	422.000
Ineffective	0.000

Table 2. Clustering center results.

	Clusters		
	1	2	3
Weeks of attendance	0.983	0.942	0.912
Days of attendance	0.748736	0.717436	0.674876
Number of resource views	0.924284	0.884736	0.820494
Number of assignments completed	0.995746	0.802834	0.797363
Frequency of use of the assignment module	0.998473	0.928477	0.847857
Video-viewing ruminant ratio	0.084636	0.098463	0.093726
Number of main posts	0.018863	0.003483	0
Total number of forum posts	0.084736	0.038445	0.00
Word count for posts	0.074632	0.038473	0.026351
Assignment grades	0.948574	0.894736	0.784363
Unit test results	0.897433	0.847463	0.749483
The number of video-viewing	1.000000	0.864642	0.765893
Summative exam results	0.867564	0.007564	0.728746
Overall results	1.943773	1.384735	1.645332

learning inputs in each dimension and the overall performance, the mean values of the learning behavior scalars and the mean values of the overall performance of each cluster member's inputs in each dimension were compared, as shown in Table 3.

As can be seen from the table above, the first category of learners has an overall average score of 92. In the participation dimension, the average number of weeks and days of attendance and the average number of resource browsing are significantly higher than the last two categories of learners, especially the average number of days of attendance, which is nearly twice as high as the third category of learners, and is basically close to the total number of activities; in the concentration dimension, the average number of homework assignments completed and the average frequency of homework modules used are also significantly higher than the other two categories of learners, and the video-viewing ruminant ratio mean is smaller than the last two categories of learners. In the interaction dimension, the mean value of the number of main posts, the mean value of the total number of forum posts, and the mean value of the number of words posted were also significantly higher than those of the last two categories of learners. In the performance effort dimension, the mean values of all learning behaviors were significantly greater for the first category of learners than for the second and third categories of learners.

The second category of learners has a mean score of 74, which is the lowest score among the third category of learners. In the participation dimension, the

Table 3. Mean value of each dimensional variable.

Variables	Category 1	Category 2	Category 3
Average number of weeks of attendance	8.54	8.02	7.45
Average number of days of attendance	63.67	46.76	23.97
Average number of resource views	478.89	479.43	278.45
Average number of assignments completed	9.45	7.34	5.32
Average frequency of use of assignment modules	10.62	8.91	7.27
Video-viewing ruminant ratio mean	1.068473	1.274634	1.184737
Average number of main posts	1.34	1.12	0.00
Average value of total number of forum posts	3.67	3.42	0.45
Average word count of posts	7.85	3.45	0.89
Assignment grade mean	94.5643	89.7632	87.5632
Mean value of unit test results	92.6533	83.3464	81.562
Average number of video views	56.00	48.42	46.89
Mean value of summative exam results	89.57	0.89	85.72
Mean value of overall results	92.08	74.26	84.35

average number of weeks and days of attendance for this category of learners is smaller than that of the first category of learners, but larger than that of the third category of learners, but the average number of resource views is the highest, slightly higher than that of the first category of learners; in the concentration dimension, the average number of homework assignments completed and the average frequency of homework modules used rank second among the three categories of learners, but the average video-viewing ruminant ratio is the highest; in the interaction dimension, the average number of forum posts and the total number of forum posts for this category of learners rank second, but only slightly lower than that of the first category of learners. The average value of the total number of forum posts and the average value of the total number of forum posts ranked second, but only slightly lower than the first category of learners; in the performance effort dimension, the average value of homework scores, the average value of unit test scores and the average value of the number of videos watched were also in the middle level, significantly lower than the first category of learners and slightly higher than the third category of learners, and the average value of summative exams was the lowest among the three categories of learners and less than 1 point, so it can be inferred that most of the learners in this category did not take the test. This inference was confirmed after checking that most of the students in this category had zero test time.

The third category of learners had a mean summative score of 84, which is in the middle of the three categories of learners. As can be seen, this category of learners is smaller than the first two categories of learners in all four dimensions (except for the second ranking of video-viewing ruminant ratio mean) and belongs to the category of the lowest participation, but the mean value of summative exam scores is lower than that of the first category of learners and significantly higher than that of the second category of learners.

In summary, the first category of learners has a high level of participation and motivation in the course activities throughout the learning process, and their

Table 4. Correlation analysis of clustering results.

	Category 1	Category 2	Category 3
Correlation coefficient r value	0.52	−0.23	−0.37
Significant level p value	0.01	0.02	0.01

final grades are high, so they can be regarded as "active learners", while the second category of learners has a medium level of participation and motivation in the course activities among the three categories of learners, and their overall grades are the lowest among the three categories of learners because they are not involved in the final course exams. The second group of learners has the lowest overall grade among the three groups of learners because they do not participate in the final exams, so this group of learners is considered as "go-with-the-flow learners". The third category of learners has the lowest level of participation and motivation in learning activities in the course, although the average overall grade is at the middle level, so they can be considered as "passive learners".

4.4. Correlation Analysis of Clustering Results

The results of the correlation analysis of the data in categories 1 - 3 are shown in Table 4. The significant level of correlation between the sum of the converted values of the factors of academic performance and learning engagement in each cluster is within 0.02. The correlation coefficient of the first category is positive, and there is a positive correlation between learning engagement and academic performance, accounting for 55% of the total sample size, while the correlation coefficients of the second and third categories are negative, with a negative correlation between learning engagement and academic performance, but the second category of learners has a lower overall score because they did not take the summative examination, which can be considered an anomaly. Therefore, if only the third category of learners is classified as having a negative relationship between learning engagement and academic performance, it accounts for 25% of the total sample, which shows that there is a non-positive relationship between learning engagement and academic performance for a quarter of the learners in the online general education course.

5. Conclusions and Implications

This study takes a representative online general education course offered by a university in eastern China as an example, refers to the existing research on online learning engagement and combines the characteristics of online general education courses to construct an online learning input model for online general education courses. The K-means clustering algorithm was used to classify the learners into three categories: "active learners", "go-with-the-flow learners" and "passive learners". Only about 55% of the learners had a positive relationship between online learning engagement and academic performance, about 25% of the learners did not have a positive impact on their learning performance, and

about 20% of the learners were anomalous because they did not take the summative exam, and it was not possible to determine whether they had a positive relationship. In addition, the correlation between gender, subject background and achievement was not analyzed as only 10 of the 224 students in this study were male and 90% of the students were in Arts.

One of the most important reasons for the high online engagement but low summative test scores for "go-with-the-flow learners" is that most students do not take the final exam, making it impossible to accurately determine their learning effectiveness. However, it may also be due to the fact that the learners are only committed to online learning in order to complete the number of online learning clicks required by the school, but do not really invest time in learning, or they may have a weak learning foundation or inappropriate learning methods. And for the "passive learners", there are two possible reasons for the low online commitment and high final exam scores: first, the learners only pay attention to the final exam scores and do not seriously commit to online learning; second, learners do not like or are not accustomed to online learning and learn mainly through paper books, face-to-face classes, offline communication with teachers and other non-online learning methods, and the data indicators of online learning input do not fully reflect learners' learning input. To address the situation of these two types of learners, teachers and administrators should obtain relevant information about these two types of learners in a timely manner, regularly check their learning progress, and intervene in a timely manner to prevent learners from not participating in course learning activities for a long time and completing tasks unexpectedly during concentrated periods or late in the course. Interviews and questionnaires should also be used at different stages of learning to understand learners' specific situations in more depth and to adopt targeted and personalized learning support services. In addition, the final course assessment should not only be based on superficial data, but should also explore more indicators reflecting students' deep learning, such as learners' concentration on learning, interaction after forum discussion and video viewing, and the quality of homework completion.

In addition, during the implementation of the study, the writer noticed that students did not attach enough importance to online general education courses because they are elective courses, for example, 20% of the learners in the "go-with-the-flow learners" did not take the summative examination which is an important part of the final performance. The root causes for this are many local undergraduate institutions in China lack a deep understanding of general education, lack scientific planning of the online curriculum system, and lack standardization of course management, and tend to treat general education as an education to expand knowledge, which is usually developed in unsystematic individual courses that are easily detached from the realities of life and thus fail to arouse students' interest, resulting in general education being promoted mainly at the knowledge level (Zhao & Wei, 2017). Therefore, in the process of promoting the

reform of general education, local undergraduate institutions in China can take advantage of the spring breeze of "Internet+", that is, the rapid development of MOOC in China, to build a perfect school-based special general education system based on high-quality and diverse online courses, so that teachers can teach and students can learn efficiently. At the same time, the responsibilities of teachers of online general education courses are clearly defined and an assessment system is established, and a human-computer integrated learning support service system is set up. Learning support services for online courses are a guarantee of learners' commitment to learning. Research has shown that teachers have an important role in guiding students' participation, and that teachers can promote learners' learning engagement by issuing encouraging messages, asking thought-provoking questions, and giving feedback when they do (Liu, Zhang, & Liu, 2017). On the one hand, learning support services are provided through personalized recommendations and adaptive learning paths using the big data and artificial intelligence of online course platforms, while on the other hand, teachers help learners answer questions, correct assignments, and implement manual learning support services through the Internet. At the school level, administrators should develop corresponding regulations, assessment standards and mechanisms according to school-based conditions to motivate teachers, clarify their responsibilities and improve their learning support services, so as to increase the learning input of online general education learners and guarantee learning quality.

Although the author has made a lot of efforts in conducting the study and achieved some results, there are still some shortcomings in this paper, mainly including the following two points: firstly, the sample is not representative. Online general education courses contain many types, and different courses differ in terms of the number of students taking them and the way they are assessed. This study only uses one course as an example for analysis, and the sample size is small, so the conclusions drawn may have certain limitations. In future studies, the research sample should be expanded and different types of courses should be selected for comparative analysis, with a view to obtaining more comprehensive research results. Secondly, the data collection of learners' online learning engagement assessment is not yet comprehensive, for example, the completeness of resource browsing, the number of weeks of attendance and the time interval of assignment submission, and the relevance of the content of forum postings to the course are not yet collected due to the author's limited authority and capacity as they are closely related to the software development of the online learning platform. In future research, more data mining and analysis techniques will be attempted to further expand the collection of learning input data, to dig deeper into the types of learners' online learning input in the context of online general education courses, and to provide reference and basis for the development of personalized human-computer integrated learning services support, as well as to lay the foundation for the improvement of the school-based special general education curriculum system.

Acknowledgements

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Bolliger, D. U., & Halupa, C. (2018). Online Student Perceptions of Engagement, Transactional Distance, and Outcomes. *Distance Education, 39,* 299-316. https://doi.org/10.1080/01587919.2018.1476845

Carver, L. B., Mukherjee, K., & Lucio, R. (2017). Relationship between Grades Earned and Time in Online Courses. *Online Learning, 21,* 303-313. https://doi.org/10.24059/olj.v21i4.1013

Chen, S. W. (2016). Optimizing the Teaching Management of Online Liberal Studies Courses in Higher Vocational Colleges: A Case Study of Zhejiang Agricultural Trade Vocational College. *Read and Write Periodical, 13,* 50-51+4.

Chen, W. J., & Chen, Q. Q. (2018). Study on the Optimization of Teaching Management of Online General Education Courses. *Research of Agricultural Modernization, No. 4,* 88-91.

Chen, Y. Y., Liu, S. F., & Dong, K. (2019). Analysis of Online Learning Behavior and Learning Effect of Adult Learners Based on Data Mining. *Journal of Anhui Radio & TV University, No. 1,* 38-42.

Coates, H. (2007). A Model of Online and General Campus-Based Student Engagement. *Assessment & Evaluation in Higher Education, 32,* 121-141. https://doi.org/10.1080/02602930600801878

Dixson, M. D. (2010). Creating Effective Student Engagement in Online Courses: What Do Students Find Engaging? *Journal of the Scholarship of Teaching & Learning, 10,* 1-13.

Fredricks, J. A., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74,* 59-109. https://doi.org/10.3102/00346543074001059

Guo, F. X., & Liu, Q. (2018). A Study on the Correlation between Online Learning Behavior and Learning Effect—Based on the Teaching Practice of the Flipped Classroom of Blackboard. *Higher Education Sciences, No. 1,* 8-13.

Hao, D. (2013). Literature Analysis of the Current Status of MOOC Research in China. *Journal of Higher Education, No. 11,* 42-50.

He, L., Wu, L. D., & Cai, Y. C. (2007). Survey of Clustering Algorithms in Data Mining. *Application Research of Computers, No. 1,* 10-13.

Hou, X. N. (2017). *The Research of Learning Motivation for Network General Curriculum of Normal University Students (D).* Linfen: Shanxi Normal University.

Hu, B. Y. (2017). Practical Exploration on General Courses in Zhihuishu MOOC Platform Network in Colleges and Universities—Taking Workplace Communication as an Example. *Journal of Hubei Correspondence University, No. 9,* 144-145.

James, B. C. (1945). *General Education in a Free Society*. Cambridge, MA: Harvard University Press.

Jo, I., Yoon, M., & Ha, K. (2013). Analysis of Relations between Learner's Time Management Strategy, Regularity of Learning Interval, and Learning Performance: A Learning Analytics Approach. *The E-Learning Korea 2013,* Seoul, 3-5 September 2013, 201-213.

Kwon, S. (2009). The Analysis of Differences of Learners' Participation, Procrastination, Learning Time and Achievement by Adult Learners' Adherence of Learning Time Schedule in e-Learning Environments. *Journal of Learner-Centered Curriculum and Instruction, 9,* 61-86.

Lang, B., & Fan, Y. N. (2019). Method of Using Learning Vector Classification Samples to Predict Online Achievements. *Computer Systems & Applications, No. 3,* 215-222.

Lei, X. Y. (2016). A Study on the Current Situation of the Application of Erya's General Education Curriculum in Teacher Training Colleges: The Case of Zhoukou Normal College. *Journal of Shangqiu Vocational and Teaching College, 15,* 125-127.

Li, M. L., Yang, L., & Sun, H. T. (2001). Survey and Analysis of the Current Situation of Liberal Studies Education in China's Colleges and Universities—Take the Four Universities of Peking University, Tsinghua University, National People's Congress and Beijing Normal University as Examples. *Research on Education Tsinghua University, No. 2,* 125-133.

Li, S., Li, R. Q., & Yu, C. (2018). Evaluation Model of Distant Student Engagement: Based on LMS Data. *Open Education Research, 24,* 91-102.

Liu, B., Zhang, Y. L., & Liu, J. L. (2017). Study on the Effect of Teacher Support on Online Learners' Learning Engagement. *China Educational Technology, No. 11,* 63-68, 80.

Liu, C. Y., Zhang, Y., & Shan, X. T. (2018). Current Situation, Problems and Improvement Measures for the Implementation of University Online General Education Courses. *New Curriculum Research, No. 3,* 129-131.

Qiao, L., & Jiang, F. G. (2020). Clustering Analysis of MOOC Learner Groups-Taking the MOOC of "STEM Course Design and Case Study" for Example. *Modern Education Technology, 30,* 100-106.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The Measurement of Engagement and Burnout: A Two Sample Confirmatory Factor Analytic Approach. *Journal of Happiness Studies, 3,* 71-92.
https://doi.org/10.1023/A:1015630930326

Song, J., & Zhu, L. (2012). Mass Data Processing Platform Design and Implementation Based on Cloud Computing. *Telecommunication Engineering, No. 4,* 566-570.

Sun, Q. F., & Feng, R. (2019). A Research on Online Learning Achievement Factors Based on Learning Analysis. *China Educational Technology, No. 3,* 48-54.

Sun, J. C. Y., & Rueda, R. (2012). Situational Interest, Computer Self-Efficacy and Self-Regulation: Their Impact on Student Engagement in Distance Education. *British Journal of Educational Technology, 43,* 191-204.
https://doi.org/10.1111/j.1467-8535.2010.01157.x

Wang, G. H., & Jiang, P. (2004). Survey of Data Mining. *Journal of Tongji University, No. 2,* 246-252.

Wang, Y. J. (2008). Problems and Countermeasures of Comprehensive Quality Elective Courses in Colleges. *Education Exploration, No. 4,* 49-50.

Wang, Y. H. (2020). *Research on the Construction Application of Online Learning Investment Evaluation Index System for College Students (D)*. Lanzhou: Northwest Normal University.

Wei, S. P. (2012). An Analysis of Online Learning Behaviors and Its Influencing Factors: A Case Study of Students' Learning Process in Online Course "Open Education Learning Guide" in the Open University of China. *Open Education Research, 18,* 81-90, 17.

Yang, H. Q., & Ou, Z. H. (2015). Exploring the Reform of University-Level Public Elective Courses Based on "MOOC". *Journal of Agricultural University of Hebei (Agriculture and Forestry Education), 17,* 78-81.

Yao, Y., & Zhao, H. (2016). Exploration and Practice of Introducing Online General Education Courses in Colleges—An Example from Anyang Normal College. *Journal of Anyang Normal University, No. 4,* 110-114.

Zhao, X. (2013). Exploring the Reform of Public Elective Courses in Higher Education Based on the Platform of "Chaoxing Erya General Education Course". *Mechanical Vocational Education, No. 2,* 57-59.

Zhao, X. F., Li, Q. L., & Liang, C. M. (2019). Analysis of the Factors Influencing the Learning Effect of Online Learning Platform—Take Northern University for Nationalities as an Example of the Analysis of the Causes of Factors Influencing the Learning Effect of Online Learning Platform. *Thin Tank Era, No. 3,* 165+172.

Zhao, Y., & Wei, K. Y. (2017). Study on the Current Situation and Countermeasures of the Construction of General Education Curriculum System in Local Universities. *China Adult Education, No. 9,* 105-107.

Zheng, Q. H., Cao, L., Chen, L., & Wu, Y. F. (2013). An Investigation on Contributing Factors of Students' Learning Performance in E-Learning. *Open Research Education, No. 6,* 69-73.

Zhong, K. D., & Liang, Y. J. (2006). The Research on the Relationship of Structure Character of Web-Based Learning Community and Learning Performance. *Open Research Education, No. 6,* 69-73.

Zou, Z. W., & Zhu, J. W. (2005). Research and Summary of Data Mining Algorithms. *Computer Engineering and Design, No. 9,* 2304-2307.