Unfolding the learning behaviour patterns of MOOC learners with different levels of achievement

Shuang Li[1,2]* , Junlei Du[1] and Jingqi Sun[1]

*Correspondence:
lilybnu@bnu.edu.cn
[2] School of Educational
Technology, Beijing Normal
University, 19, Xinjiekouwai
Street, Haidian District,
Beijing, China
Full list of author information
is available at the end of the
article

Abstract

In an open and flexible context of Massive Open Online Courses (MOOCs), learners who take final assessments exhibit the motivation for performance goals. The learning trajectories of this group usually provide more clues for course design and teaching improvement in that this group tend to interact more fully with course learning activities and resources for better learning outcomes. This study focused on such learners to investigate their learning engagement, time organization, content visit sequences, and activity participation patterns by applying statistical analysis, lag sequence analysis, and other data mining methods. This study examined the data of 535 learners taking the assessment in a MOOC to detect the differences in learning engagement and the above learning patterns amongst three groups of learners with different achievement levels, labeled *failed*, *satisfactory* and *excellent*. We found differences in both learning engagement and learning patterns among the three groups. The results indicated that for the learners to be successful, they require a certain degree of task completion as a basic guarantee for passing the course, effective session workload organization, reasonable learning content arrangement, and more cognitive engagement (rather than investing more time and energy). Based on the outcomes, implications for personalized instructional design and intervention to promote academic achievement in MOOCs are discussed.

Keywords: Online learning, MOOC, Learning engagement, Learning behaviour patterns, Differences, Learning achievement, Learning analytics

Introduction

Since 2012, Massive Open Online Courses (MOOCs) have developed rapidly with the wide use of the Internet and active involvement of numerous universities. In the past decade, on the one hand, more open and flexible MOOCs have attracted a large number of groups with different educational backgrounds, learning motivations, and learning styles to participate in MOOCs. Class Central showed that the modern MOOC movement has impacted 180 million learners and that MOOC providers launched over 16,300 courses, 1180 micro-credentials, and 67 MOOC-based degrees by the end of 2020 (Shah, 2020). On the other hand, the low completion rate and learner engagement of MOOCs have been the main concerns for their quality

and effectiveness (Xing & Du, 2019). In recent years, an increasing number of studies have begun to focus on course design and effective teaching strategies in MOOCs to enhance learners' engagement and effectiveness (Jung & Lee, 2018). However, the diversity of MOOC learners' learning backgrounds, characteristics, and motivations is undoubtedly the greatest challenge in MOOC design and teaching. Analysing various learners' needs and learning behaviour patterns in MOOCs is the key to optimizing MOOC design and teaching.

MOOC participants sign up for different reasons such as the desire to learn, curiosity about MOOCs, personal challenges, and the desire to collect certificates of completion (Hew & Cheung, 2014). Among them, considerable MOOC participants do not come to study, but register for MOOCs for casual interest in a topic, teaching methods, or temporary interests of teachers and course learners (Chen et al., 2020). There are also some participants with specific purposes in MOOCs. They usually do not follow the requirements and suggestions of the course, but instead focus on their objectives in their own way and pace (Davis et al., 2016). The above participants in MOOCs are less likely to learn according to the design of MOOCs, so the design and teaching of MOOCs have limited influence on them. Therefore, research on the learning behaviour patterns of these learners has limited implications for MOOC design and teaching. In order to support the improvement on MOOC design and teaching, it is suggested to dig and analyse the learning behaviour patterns of those learners who attach more importance to course design and teaching, such as those with a higher course completion degree or a stronger desire to get good grades. We believe that learners who take final assessments of MOOCs are examples of these types of learners. In an open and flexible context of MOOCs, learners who take final assessments of MOOCs exhibit the motivation for performance goals, such as course certificates and better course grades. Driven by these performance goals, they value course requirements and suggestions, and tend to interact more fully with the course resources and activities for good learning outcomes (de Barba et al., 2020; Lan & Hew, 2020; Maldonado-Mahauad et al., 2018). So, the learning trajectories of this group usually provide more clues for course design and teaching improvement. Therefore, in-depth research on these learners' learning patterns and characteristics can provide valuable insight for the design and teaching of MOOC in an open and flexible context. However, the literature review reveals that few studies have investigated the learning processes and experiences of this group. Most past studies have either explored the learning processes of dropouts or lurkers (Bozkurt et al., 2020; Chaker & Bachelet, 2020) or contrasted the learning characteristics of those who have completed a course with those who have not (Charo et al., 2020; Lan & Hew, 2020; Peng & Xu, 2020). Although these studies are important in motivating learners' learning persistence and engagement, they have failed to demonstrate the effective learning behaviour patterns of different learners in MOOCs. This topic should be explored further with empirical research on learners' learning behaviour patterns in MOOCs, with a focus on the differences, effectiveness, and influencing factors of their learning patterns.

Above all, this paper scrutinises learners who take final assessments in MOOCs, and conducts an in-depth and systematic analysis of their learning behaviour patterns based on different performance levels. More targeted insight into learning processes and rules is provided to improve the design and instruction of MOOCs for educators.

Related work

Previous studies have made valuable observations about MOOC learning patterns by analyzing learning trace data from different perspectives, such as engagement patterns (Ferguson & Clow, 2015; Kizilcec et al., 2013; Poquet et al., 2020), activity distribution (de Barba et al., 2020; Lan & Hew, 2020; Wen & Rosé, 2014), behaviour sequences (Maldonado-Mahauad et al., 2018; Rizvi et al., 2020; van den Beemt et al., 2018), content visit sequences (Guo & Reinecke, 2014; van den Beemt et al., 2018), and self-regulated learning (de Barba et al., 2020; Kizilcec et al., 2017; Maldonado-Mahauad et al., 2018).

Since the growth of MOOCs around the world, engagement patterns of MOOC participants have been a hot topic in the fast-developing research field of learning analytics and MOOCs. Many researchers explored engagement patterns in MOOCs based on the identified engagement indicators, and then defined subpopulations of MOOC participants according to their behavioural profiles accordingly (Tseng et al., 2016; Ye et al., 2015). Kizilcec et al. advanced this work further by mining engagement patterns across time. According to MOOC learners' weekly engagement trajectories in assessments, videos, and quizzes in three computer science courses, these authors clustered MOOC learners into four groups, labeled as Auditing, Completing, Disengaging, and Sampling learners (Kizilcec et al., 2013). Ferguson and Clow (2015) followed and extended the methodology of Kizilcec and his colleagues to extract engagement patterns in FutureLearn MOOCs. In their study, Ferguson and Clow added data about commenting for extracting engagement patterns to reflect the importance of discussion in FutureLearn MOOCs, and they finally found seven clusters. Based on a bottom-up approach (Kizilcec et al., 2017), the above studies have provided helpful insights into diversified learning patterns of MOOC learners from the perspective of engagement.

With the recognition of the diversity of MOOC learners' motivations and characteristics, and their impact on quality of their commitment and achievement (Li & Baker, 2018; Rizvi et al., 2020), more and more studies have begun to adopt a top-down approach to probe into MOOC learning patterns combined with learners' performance, as shown in Table 1. These studies firstly grouped MOOC learners according to their behaviour profiles or achievement, and then explored the learning patterns and strategies of different groups by applying learning analytics, self-reported questionnaires and interviews (de Barba et al., 2020; Liu et al., 2021a, 2021b; Rizvi et al., 2018). In comparison with the studies based on a bottom-up approach, these studies could reveal the learning characteristics of MOOC learners with different performances, and support individualized design and teaching of MOOC.

As shown in Table 1, the prior studies categorized MOOC participants according to different criteria. Most of the researchers grouped learners based on whether they completed the course or not. However, they defined completion of the course according to different criteria. For example, Rizvi et al. (2018) labeled learners as completers based on whether learners marked each activity as completed. Similarly, Lan and Hew (2020) also labeled completers by learners' self-report. Charo et al. (2020) defined completers by whether learners continued to work on the MOOC until the last days of the course. Besides, some researchers grouped learners based on whether they attained certificates (Guo & Reinecke, 2014; Liu et al., 2021a), course grades (Wen & Rosé, 2014), engagement (Li & Baker, 2018; Maldonado-Mahauad et al., 2018), achievement and engagement

Li *et al. Int J Educ Technol High Educ*   (2022) 19:22

Page 4 of 20

Table 1 Related literature on learning patterns of MOOC learners with different performances

Authors (year)	Categories of learners	Analytical dimensions
Anderson et al. (2014)	Bystander; Viewer or Collector; All-rounder; Solver	Engagement in assignment, questions, lectures, and forum
	High-achievers; Low-achievers	
Guo and Reinecke (2014)	Certificate learners; Non-certificate learners	Course navigation strategies
Wen and Rosé (2014)	None; Fail; Pass; Distinction	Learning session topic distribution
Li and Wan (2016)	Completers; Non-completers	Percentage completed; Interactions in forum
Maldonado-Mahauad et al. (2018)	Completers; Non-completers	Interaction sequences in terms of video-lecture and assessment behaviours from a SRL perspective
	Comprehensive learners; Sampling learners; Targeting learners	
Li and Baker (2018)	Disengagers; Auditors; Quiz-takers; All-rounders	Behavioral engagement; Cognitive engagement
Rizvi et al. (2018)	Completers; Non-completers	Participatory behaviours; Temporal learning pathways; Learning activities sequence
van den Beemt et al. (2018)	Certificate students; Non-certificate students	Video Watching behaviour trace Quiz submission process
	Samplers; Disengagers; Venturers; Accomplishers	
Handoko et al. (2019)	Completers; Non-completers	Self-regulated learning strategies
Charo et al. (2020)	Completers; Non-completers	Self-regulated learning strategies; Interaction with other students; Engagement with the subject
de Barba et al. (2020)	Auditors; Failed; Passed	Session distribution patterns; Session activity patterns
Lan and Hew (2020)	Completers; Non-completers	Learning activities participation: Behavioural engagement; Emotional engagement; Cognitive engagement
Poquet et al. (2020)	Residents; Visitors	Discussion topics; Discussion topic sequences
Rizvi et al. (2020)	Markers; Partial-markers; Non-markers	Engagement; Temporal learning paths; Learning activities sequence
Liu et al. (2021a)	Certificate achievers; Explorers	Behaviours related to certificate achievement; Behaviour frequency differences; Behaviour sequence differences
Liu et al. (2021b)	Completers; Non-completers	Cognitive engagement

(de Barba et al., 2020). The diverse classifications of learners undoubtedly provide a variety of perspectives for observing and exploring the MOOC learning patterns.

The results of these studies shed a light on the differences in learning engagement, activity distribution, learning process, learning sequence and self-regulated learning strategies among learners with different academic performances. In general, the prior research demonstrates that high-performance learners, such as completers and certificate achievers, usually presented a higher level of course commitment and were involved in multiple activities (Anderson et al., 2014; De Barba et al., 2020; Lan & Hew, 2020; Maldonado-Mahauad et al., 2018) compared with low-performance learners. Barba et al. (2020) found that compared with active learners whose final grade was below 65, which they defined as failed learners, active learners who had the final grade above 65, which they defined as passed learners, tended to mix up activities

within weekly sessions. Lan and Hew (2020) uncovered that MOOC completers watched more videos, posted more forum messages, and did more peer interactive activities than non-completers. Liu et al. (2021b) further found that completers tended to post comments related to the course' content while non-completers usually posted off-task content in forum, which indicated that completers learned the course with more cognitive engagement.

Going deeper, several studies further uncover the difference in the learning process of MOOC learners with different performances by extracting behaviour sequences from learning trace (Maldonado-Mahauad et al., 2018; van den Beemt et al., 2018; Wen & Rosé, 2014). Wen and Rosé (2014) characterized MOOC learning session of interaction as a probabilistic combination of interaction and interaction sequence patterns by using topical N-gram models and investigated how students' final grade is related to their learning session topic distribution along the course based on the dataset of two Coursera MOOCs. They found that high achievement groups' learning session distribution reflected the course schedule better than that of the low achievement group. Based on MOOC learners' log files, Maldonado-Mahauad et al. (2018) identified six interaction sequence patterns with highest frequency by Process Mining techniques, and examined how these interaction sequence patterns vary according to whether or not the group of learners completed the course. They found that compared with non-completers, completers exhibited more interaction patterns and got involved in more assessment activities which either appear combined with a video-lecture activity or appear in a successional way. In contrast, the learning process of non-completers is usually dominated by video-lecture activity. In line with this, Rizvi et al. (2020) found that learners who marked all activities as completed tended to have longer learning processes and more diversified activities than learners who marked no activities as completed. Van den Beemt et al. (2018) advanced this work by combining behavioral processes with learning sequence, and found that non-certificate learners showed less structured behaviour with some quiz submissions or videos being skipped compared with certificate learners. Similar results were also found in their follow-up study on learning sequences of four groups with different engagement extracted according to video-watching behaviours, that is, the high-engagement cluster (labeled as accomplishers) showed more steady video watching and quiz submission behaviors in course order than low-engagement clusters (labeled as disengagers, venturers, and samplers). It is worth noting that van den Beemt and his team also revealed that even certificate students and accomplishers appeared to work in an increasingly disordered way as time progressed (van den Beemt et al., 2018).

Poquet et al. (2020) extended this comparative study of learning process to forum activity, and explored the difference in communication topic patterns of two sup-populations with different forum commitment (residents and visitors) based on a large dataset of four edX MOOCs. Their results showed that resident posters tend to engage in more diverse discourse, and generated more socio-cognitive and informational queries as the course processed compared with visitors. Liu et al. (2021a) further revealed the difference in behavior patterns of certificate achievers and explorers who did not achieve certificate in the learning process of integrating diversified activities by lag sequential analysis. It was found that certificate achievers had more bidirectional and repeated behaviour sequences compared with explorers.

In all, the findings of previous researches on the learning process of MOOC learners indicate that high-performance learners usually exhibit longer learning sequences and more diversified learning activities compared with low-performance learners. However, there is no consensus on whether high-performance learners tend to follow the structure of course materials to visit course content. Some studies suggest that learners preferred to non-linear and unstructured learning pathways through the course content (Guo & Reinecke, 2014; Li & Wan, 2016; Rizvi et al., 2020). But other studies indicated that high-performance learners tended to follow the structure of course material in course order (Maldonado-Mahauad et al., 2018; Van den Beemt et al., 2018).

Self-regulated learning (SRL) plays an important role in MOOC learning (Hew & Cheung, 2014; Jansen et al., 2020; Littlejohn et al., 2016). This leads to the increasing attention on the self-regulated learning process and strategies of learners with different performances. Several studies investigated the self-regulated learning process of MOOC learners by survey instruments and found that high-performance learners usually show more self-regulated learning activities (Charo et al., 2020; Handoko et al., 2019). Charo et al. (2020) found that completers reported better sores than non-completers in all three SRL phases: forethought, performance, and self-reflection. Some studies combined self-regulated learning theory or survey data with learning analytics methods to further reveal how SRL strategies manifest in online behaviours (De Barba et al., 2020; Maldonado-Mahauad et al., 2018). Maldonado-Mahauad et al. (2018) clustered MOOC learners based on six interaction patterns and observed how learners with different SRL profiles are distributed across the different clusters. They found that the learners with a high SRL profile, which they labeled as comprehensive learners and targeting learners, either followed the sequential structure provided by the course design or strategically sought out specific information to pass the course assessments. By contrast, the learners with a low SRL profile tended to behave in irregular ways. De Barba et al. (2020) examined session distribution and session activity pattern of three MOOC groups with different performances from a time management perspective. The results showed that learners with higher levels of time management had longer and more sessions across the course, and prioritized more than one activity within a session early in the course.

Although existing studies have enriched the comprehension of learning patterns with different performances, few studies investigated the learning process of MOOC learners with different levels of achievement in-depth. Most of existing studies mainly focus on the difference in learning patterns of MOOC completers and non-completers in the purpose of increasing completion rate of MOOCs rather than improving achievement of MOOC learners. So, the implications provided by the prior research for achievement improvement of MOOC learners are limited. There is still little knowledge of the difference in activity distribution, content visit sequences, behavior patterns, and learning strategies among learners with different levels of achievement. Thus, learning patterns of MOOC learners with different levels of achievement need further empirical investigation from a process-based perspective.

The purpose of the present study

This research aims to investigate different learning engagement and patterns of final-assessment-taking learners of MOOCs with different achievement levels by mining their learning behaviours recorded in MOOC log data. To this end, five research questions (RQ) were investigated:

RQ1: What is the overall learning engagement level of final-assessment-taking learners?

RQ2: Are there differences in course learning engagement among final-assessment-taking learners with different levels of achievement?

RQ3: Are there differences in time organization in terms of learning time and session distribution across the course among final-assessment-taking learners with different levels of achievement?

RQ4: Are there differences in content visit sequences among final-assessment-taking learners with different levels of achievement?

RQ5: Are there differences in activity participation patterns among final-assessment-taking learners with different levels of achievement?

Methods

Participants and data

The data were based on a finance course on a MOOC platform called XuetangX in 2018. This course was accessible to the public, and not part of college or university courses. The course consisted of ten chapters, with one chapter updated each week. Before the course, there was an orientation week labelled as Week0. Learning resources of each chapter, such as teaching videos and self-test quizzes, were gradually released to learners as the course processed. Learners were also encouraged to participate in the course forum to interact with others. The final assessement was scheduled for week 11. After the course, all the learning resources were still available to learners for a while. More than 12,000 learners took the course, but only 535 learners took the final assessment. The course grade was determined by the assignment scores and the final assessment score, each accounting for 50% of the total grade.

Notably, the course data used in this study was provided by XuetangX just for academic research. The course dataset does not contain identity information of teachers and learners because it has undergone data masking before sharing. Furthermore, we deliberately omitted the course name to avoid any possible negative impact of the analysis results on the course teachers or learners. In all, there is no conflict of interest.

In this study, the 535 learners who persisted in the course and took the final assessment were chosen as samples for investigating the learning patterns of final-assessment-taking learners. To determine the differences in the learning patterns of these 535 subjects with different levels of achievement, they were classified into three groups according to their final grades: (1) *failed*, referring to learners who had a course grade below 59.9; (2)

Li *et al. Int J Educ Technol High Educ*   (2022) 19:22

Page 8 of 20

satisfactory, referring to learners who had a grade between 60 and 89.9; and (3) *excellent*, referring to learners whose grade was above 90. Table 2 shows the number and proportion of the three types of learners.

Data processing methods and procedures

This study explored five research questions through learning analytics, which has become very popular in recent years. The learning analytics method entails the measurement, collection, analysis, and reporting of data about learners and their contexts in order to understand and optimise learning and the environments in which it occurs (Siemens & Gasevic, 2012). Learning analytics is commonly used in research on learning behaviour patterns (de Barba et al., 2020; van den Beemt et al., 2018; Jovanović et al., 2017). Compared with traditional investigations, it can reveal learning processes and learning behaviour patterns in a more comprehensive and objective way. To answer the fiver research questions, this study adopted statistical analysis, lag sequential analysis (LSA), and other techniques to explore the learning engagement, time organization, content visit sequences, and activity participation patterns of the above three types of final-assessment-taking learners at different levels of performance (*failed*, *satisfactory* and *excellent*).

First, descriptive statistical analysis was carried out on the overall performance of the 535 samples based on the six indicators of learning engagement, including session frequency, total time, video completion rate, quiz accuracy rate, quiz completion rate, and forum post count.

Second, ANOVA was used to establish whether there were significant differences among the three groups in the above learning engagement indicators.

Third, the three groups' differences were investigated in terms of time organization, based on the distribution of their logging time and weekly sessions. Heat maps of learning time for each week of the course and each day of the week were drawn to support the visual analysis of time organization among the three groups. Sessions, which are regarded as a period of time devoted to a cluster or sequence of tasks without interruption (Tough, 1971), are the approach learners take to organize their online study time (de Barba et al., 2020; Whipp & Chiarelli, 2004). A series of activities can be considered a session each time after learners log into a MOOC. The time learners spend on each session reflects how they organize their time and their learning workload. Referring to the definition of de Barba et al. (2020), in the present study, the session boundary was defined as 60 min by analysing the distribution of inter-activity periods of all course learners' interactions within the system. The first and third quartiles were taken from the session samples, which were above 3 min, as standards. Sessions longer than the first

Table 2 Number and percentage of the three groups

	Count	Percentage	Grade ranges
Failed	100	18.7	0–59.9
Satisfactory	239	44.7	60–89.9
Excellent	196	36.6	90–100
Total	535	100	0–100

quartile were defined as large, while those shorter than the third quartile were defined as small, with others defined as medium. Next, the frequency of small sessions, the frequency of medium sessions, and the frequency of large sessions per week was calculated among the three groups, and a weekly session distribution pattern was constructed to support the significance test of differences in terms of session distribution.

Fourth, the learners' content visit sequences were mined according to their transitions between different chapters and among knowledge points within each chapter. Based on the direction of learners' visit transition and whether it spanned several chapters, we defined four categories of visit transition: *next, proximity backtracking, forward leap*, and *leap backtracking*. Then, we calculated the average frequency of the above four kinds of learning transitions between chapters for the three groups. The transitions whose average frequency were less than once were eliminated. Meanwhile, we calculated each group's average number of transitions among knowledge points in each chapter. In this way, the content visit sequence pattern of the three groups was constructed.

Finally, the differences in the activity participation patterns among the three groups were examined by exploring their significant sequences of transitions between activities and behaviours. Lag sequential analysis was employed to explore participation patterns at the activity and behavioural levels. Lag sequential analysis is a statistical method primarily used to test the probability of transitions between different behaviours. Since activities related to videos, quizzes, and forums are the central activities of the course in this research, significant sequences between the activities and between the specific behaviours within activities of the three groups were mined to construct their activity participation patterns. Table 3 portrays the codes of relevant activities and behaviours.

Results

RQ1: what is the overall learning engagement level of final-assessment-taking learners?

Descriptive statistical analysis was first carried out on the overall performance of the 535 samples based on the six indicators of learning engagement. Table 4 depicts the outcomes. The overall learning engagement of final-assessment-taking learners did not meet the expectation of the course. The total time that final-assessment-taking learners spent on the course was actually less than half of the recommended time of the course

Table 3 Coding table of activities and behaviours

Activity code	Behavioural code	Description
V	vse	Pause the video
V	va	Adjust the settings of the video
V	vp	Play the video
V	vd	Drag the video
Q	qv	View the quiz
Q	qc	Complete the quiz
Q	qs	Submit the quiz
Q	qa	Refer to quiz answers
F	fv	View the forum
F	fr	Reply in the forum
F	fp	Post in the forum

Table 4 Descriptive statistical analysis of learning engagement indicators (n = 535)

Variables	Min	Max	Median	Mean	Std
Session time (minutes)	0.00	81.15	22.56	23.20	12.13
Total time (hours)	0.00	90.46	10.24	12.89	11.75
Video completion rate	0.00	100.00	100.00	83.50	31.57
Quiz accuracy rate	0.00	100.00	89.55	81.47	23.89
Quiz completion rate	0.00	1.00	0.62	0.51	0.45
Forum post count	0.00	41.00	0.00	1.59	2.95

Table 5 Group differences in the learning engagement indicators

	Session frequency	Total time (hour)	Video completion rate	Quiz accuracy rate	Quiz completion rate	Forum post count
Failed	30.54 ± 32.11	10.96 ± 9.79	63.01 ± 43.66	52.41 ± 33.89	0.43 ± 0.43	0.26 ± 1.17
Satisfactory	45.90 ± 34.41	17.54 ± 13.02	86.13 ± 28.07	81.90 ± 15.69	0.66 ± 0.41	1.03 ± 3.19
Excellent	23.44 ± 23.20	8.19 ± 8.46	90.75 ± 22.89	95.76 ± 6.34	0.36 ± 0.46	2.94 ± 2.76
F	30.763***	41.214***	29.993***	183.882***	29.234***	39.935***
LSD	2 > 1, 3	2 > 1 > 3	3, 2 > 1	3 > 2 > 1	2 > 1, 3	3 > 2 > 1

***Refers to $p < 0.001$; 1 = *failed*; 2 = *satisfactory*; 3 = *excellent*

(30 h). Their learning behaviours mostly focused on watching videos, with half of the quizzes not finished and low levels of participation in the forums.

RQ2: are there differences in course learning engagement among final-assessment-taking learners with different levels of achievement?

ANOVA was performed to examine the differences among the three groups based on the six engagement indicators. The results revealed significant differences ($p < 0.001$) in all learning engagement indicators of the three groups. The least significant difference was further used for the post-test for specific differences among different categories. The findings indicated that *excellent* presented a higher quiz accuracy rate, a more forum post count, and a less total time compared with the other two groups. *Satisfactory* presented a higher session frequency, a longer total time, and a higher quiz completion rate compared with the other two groups. *Failed* presented a lower video completion rate, a lower quiz accuracy rate, and a less forum post count compared with the other two groups. Table 5 portrays the results of the significance test of differences.

RQ3: are there differences in time organization in terms of learning time and session distribution across the course among final-assessment-taking learners with different levels of achievement?

Learning time distribution across the course

Heat maps of learning time distribution across the course were drawn according to the weekly distribution of time spent on the course among the three groups. As shown in Fig. 1, although the three groups had a common feature in that they distributed more learning time in the middle and later periods of the course, the three groups' learning time distributions were obviously different from each other. *Satisfactory* chose to study
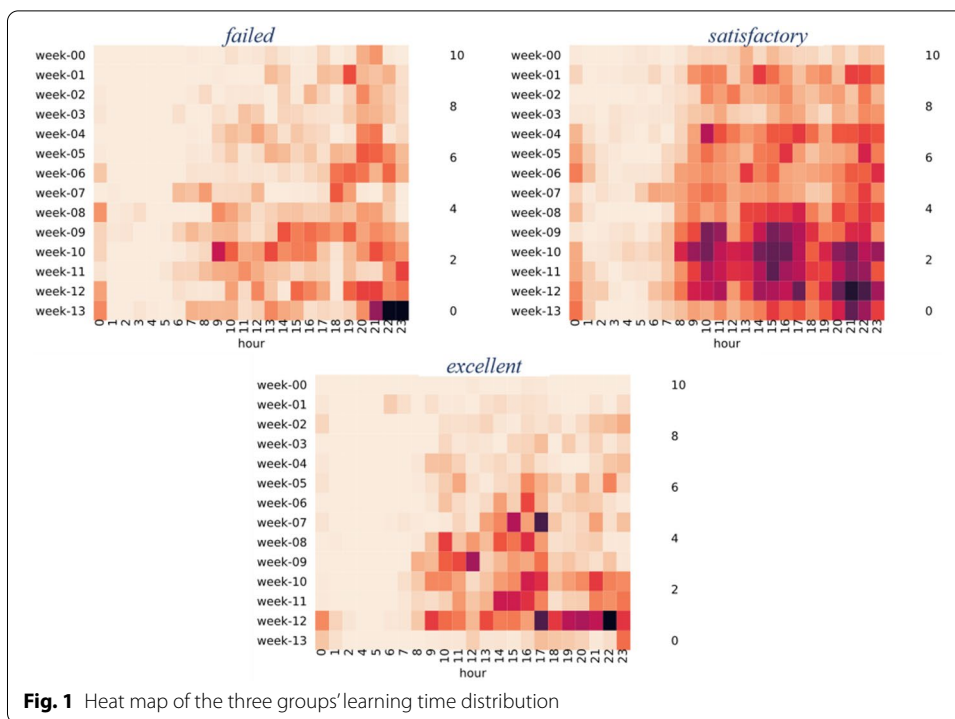
**Fig. 1** Heat map of the three groups' learning time distribution

every morning, afternoon, and evening. There was a dramatic rise in the frequency of their learning behaviours in the middle and later periods of the course. The other two groups did not show the same characteristics. *Excellent* participants were more inclined to study during the day, while *failed* participants were more inclined to study at night. It was not until the weeks near the final assessment that these two groups exhibited the characteristic of learning during all three periods throughout the day.

Weekly session distribution

We calculated the average frequency of the small, medium, and large sessions for each group per week, and described the weekly session distribution pattern of the three groups accordingly. As seen in Fig. 2, the session distribution patterns of *failed* and *satisfactory* were similar. The frequency of small-sized sessions was significantly higher than that of medium-sized and large-sized sessions. The frequency of large sessions is less than or equal to that of medium sessions. The pattern of *excellent* shows some differences. There were fewer small sessions, and there were more large sessions than medium-sized ones.

RQ4: are there differences in content visit sequences among final-assessment-taking learners with different levels of achievement?

According to the analysis of content visit sequences across chapters and within each chapter, we constructed the content visit sequence patterns of the three groups, as seen in Fig. 3. The arrow lines refer to the transitions between chapters, and the circle size refers to the average number of transitions among knowledge points in the corresponding chapter. In Fig. 3, the orientation chapter at the beginning of the course was named

Fig. 2 Weekly session distribution for the three groups

Fig. 3 Content visit sequence patterns
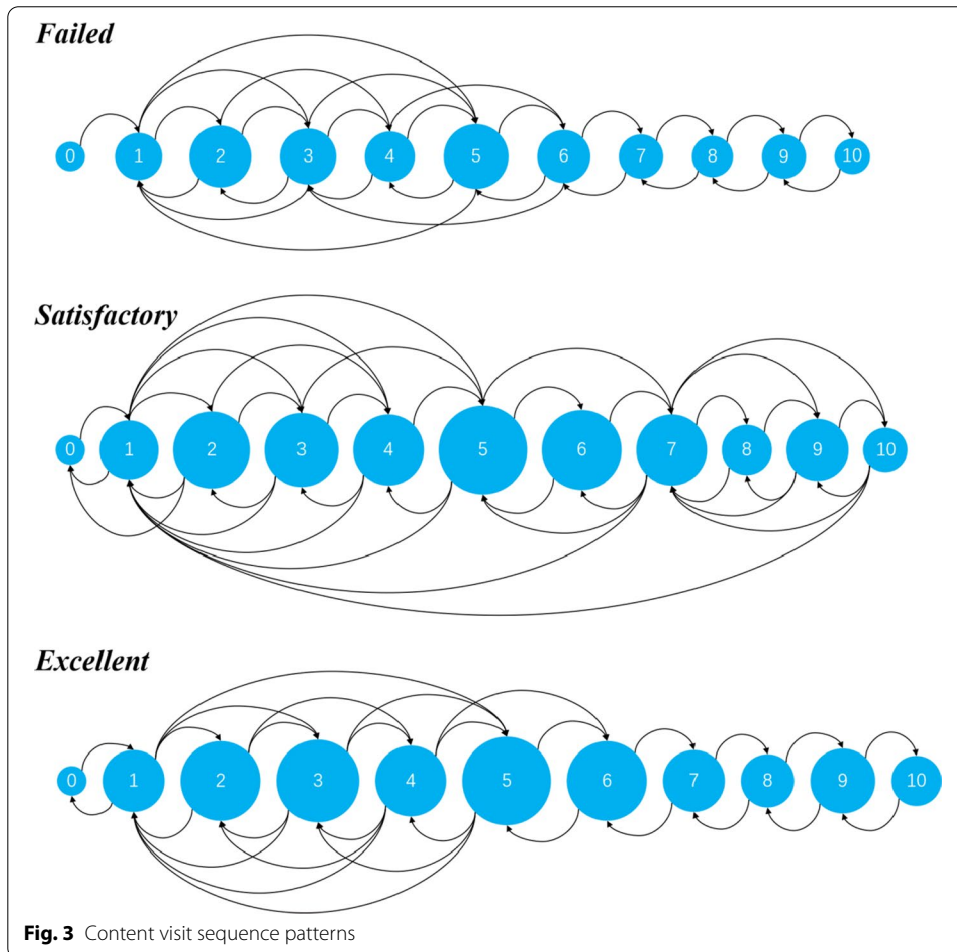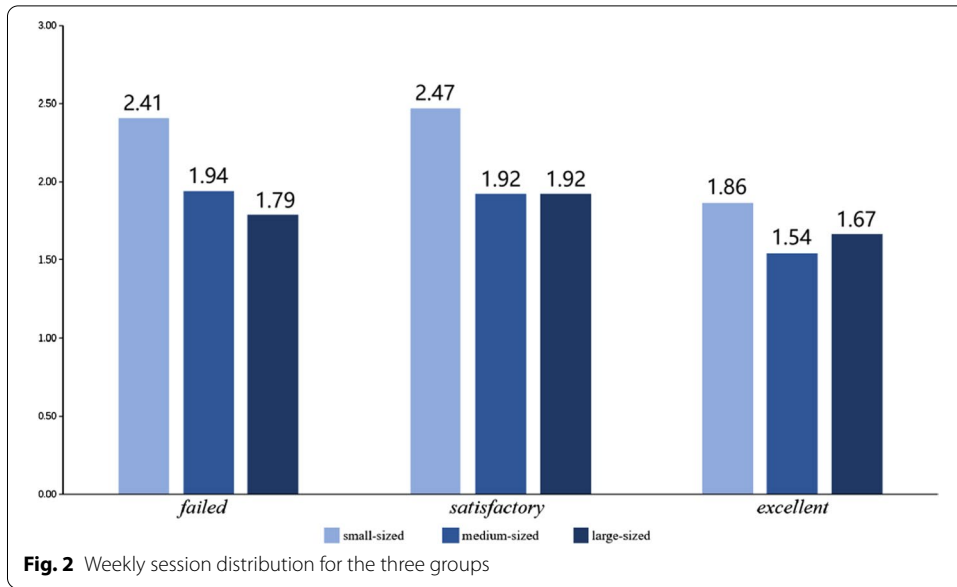
0, and the ten chapters of the course were labeled from 1 to 10. According to the direction of transitions (forward or reverse) and whether the transitions were between adjacent chapters, the transitions were divided into four types: *next*, *proximity backtracking*, *forward leap*, and *leap backtracking*.

Clearly, the three groups had almost all obvious *next* and *proximity backtracking* transitions, which means that they basically followed the learning sequences as designed and reviewed in time. In addition, all the three groups had more cross-chapter leaps among the first few chapters of the course, which means that they tended to backtrack what they had learned in previous chapters or to explore the released subsequent chapters when they studied the first few chapters of the course. To some extent, this finding suggests that all the three groups made great efforts to connect the knowledge of different chapters during their study of the first few chapters.

According to Fig. 3, there are some differences in the content visit pattern among the three groups. Firstly, the circle size of the pattern of *failed* is smaller than that of the other two groups. It is indicated that this group's visit transitions among knowledge points within each chapter were less than those of the other two groups, which means that this group's substantial engagement in the chapter content was limited. By contrast, the bigger circle size of the other two group's patterns suggested that *satisfactory* and *excellent* invested more energy into the chapter content.

In terms of cross-chapter visit sequences, compared with the other two groups, *satisfactory* irregularly visited chapters with more *forward leap* transitions and *leap backtracking* transitions. To some extent, this irregular visit pattern may indicate that the learning of this group is not concentrated and solid enough. Compared with *satisfactory*, *failed* and *excellent*'s cross-chapter visit sequences are more steady and align to the course order. However, considering the small size of the circles in the pattern of *failed*, it can be inferred that the substantial engagement of this group in chapter content is very limited despite their seemingly structured content visit pattern. In addition, there was another difference in the content visit pattern between *failed* and *excellent*. In the pattern of *excellent*, there is a *proximity backtracking* transition from the first chapter to the orientation chapter while *failed*'s pattern doesn't have. The orientation chapter provided the content of the course guide. So, the backtracking from the first chapter to the orientation chapter indicates that *excellent* preferred to review the course guide before starting their learning, which shows that *excellent* attached great importance to the course guide and tended to goal-oriented learning.

RQ5: are there differences in activity participation patterns among final-assessment-taking learners with different levels of achievement?

Lag sequential analysis was used to identify the three groups' significant transition sequences between the three kinds of MOOC activities (i.e., video activities, quiz activities, and forum activities) and significant behavioural sequences in each kind of activity. Inspired by Taylor et al. (2020), we further distinguished different significance levels of behaviour sequences by [1.96, 20), [20, 80), and [80, $+\infty$), as there were a number of learning behaviour sequences whose significance levels were higher than 1.96. Different lines were employed to represent sequences of different significance levels. According to the LSA outcomes, the activity participation patterns of

final-assessment-taking learners with different levels of achievement were constructed, as shown in Fig. 4. In addition, the significance levels of self-transition of each activity were all much greater than those between two activities among the three groups. Therefore, the self-transition of activity was excluded when conducting lag sequence analysis at the activity level in order to reveal significant sequences between two different activities of each group, and to detect differences in sequences at the activity level across the three groups.

As seen in Fig. 4, there were some differences in the three groups' learning behaviour transition sequences, both at the activity level and at the behavioural level.

At the activity level, the three groups all had mutual transitions between V and Q. The three groups all followed the course-designed learning path to some extent; that is, doing quizzes after watching teaching videos and reviewing teaching videos during or after doing quizzes. However, it is clear that the significance levels of *satisfactory* and *excellent* in related sequences were higher than that of *failed*. In addition, compared with the other two groups, *satisfactory* presented a significant sequence of

Fig. 4 The activity participation patterns of the three groups
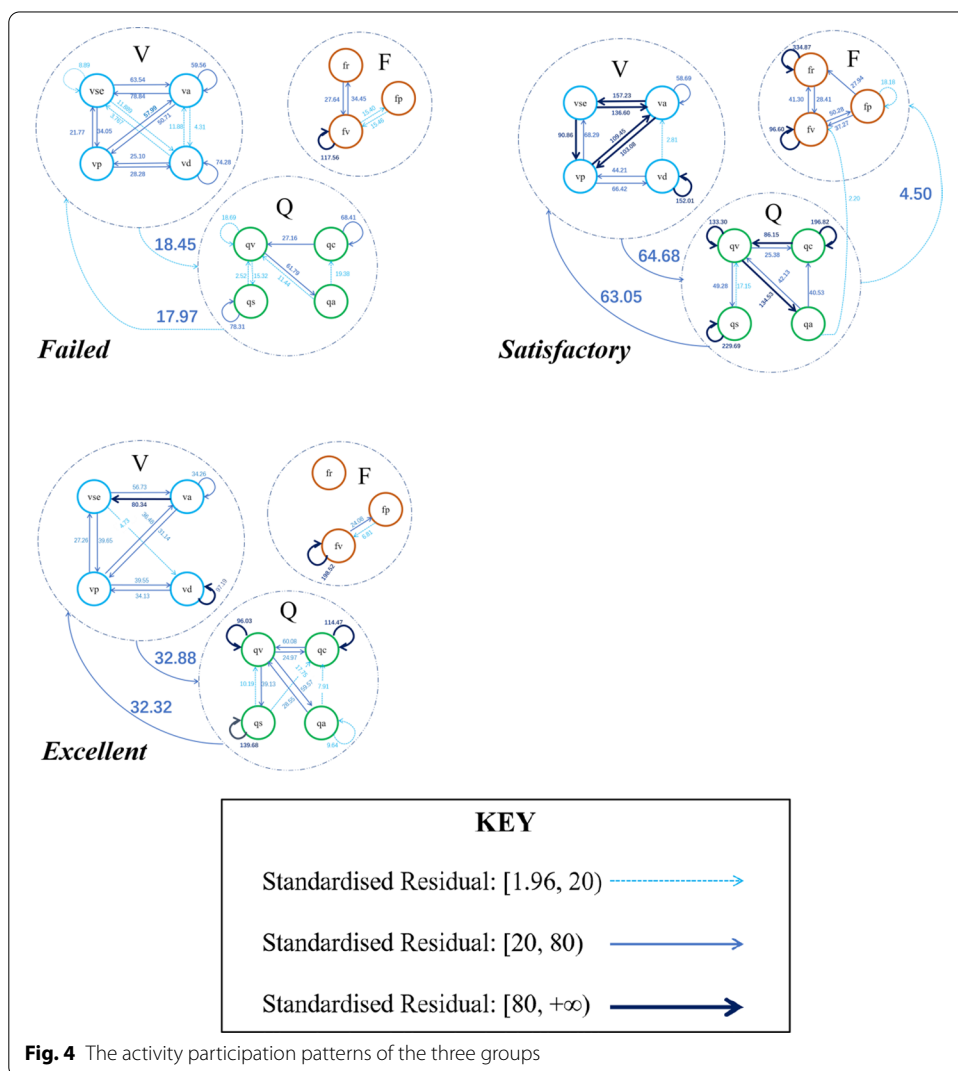
Q → F, indicating that they were more likely to use the forum to seek to help with problems encountered in quizzes.

At the behavioural level, the three groups all exhibited more significant sequences in video and quiz activities. However, there were still differences in the three groups' significant learning behaviour sequences and their significance levels. The groups that generated the most kinds of transition sequences related to Video, Quiz and Forum activities were *failed*, *excellent* and *satisfactory*, respectively. Related to the Video activity, *satisfactory* had vd → vd, with a high significance level, while *excellent* had both significant vd → vd and vse → vd, indicating that *excellent* was able to grasp their learning speed while watching videos because of their stronger purpose. Related to the Quiz activity, *excellent*, with plentiful learning sequences, had a unique learning path of qs → qc, which means that they sometimes chose to continue doing the quizzes after submitting one. This characteristic of continuing to practice showed their high degree of concentration and engagement in quiz activities. In addition, *excellent* had the lowest significance level of qv → qa, while *satisfactory* had a relatively high significance level of qv → qa. The fact that *satisfactory* often looked at the answers after viewing quizzes indicated that they could not wholly grasp the relevant knowledge. *Satisfactory* lacked adequate practice and tended to regard quizzes as one of the resources while *excellent* were able to better grasp the relevant knowledge. In terms of forum activity, *satisfactory* was the most active group. They had multiple sequences of both posting and replying in the forum, while *excellent* primarily focused on interactions through posting. *Satisfactory* engaged more in communication with each other, while *excellent* tended to treat the forum as one way to seek help.

Discussion

The first research question revealed that the overall learning engagement of final-assessment-taking learners was lower than the recommendation or expectation of the course in terms of their time engagement and involvement in resources and interactions. The final-assessment-taking learners of the three achievement levels were all more inclined to learn by watching videos, which is the most attractive learning method in a MOOC, with less dependence on deep cognitive engagement. Further, Davis et al. (2016) found that learners mostly focus their attention on lectures and assessments, with little concern for discussion. Gitinabard et al. (2019) showed that the chance of learners engaging with course materials is much higher than the chance of them participating in discussions after they submit assignments. The issue of how to promote learners' learning engagement in MOOCs is still a challenge that must be addressed in MOOC design and teaching, even for learners who have strong learning motivation for better performance. Some studies have explored teaching strategies to motivate MOOC learners' engagement (Chen et al., 2017, 2020). Lan and Hew (2020) found that active learning, course resources, and instructor accessibility are essential factors to boost both completers' and non-completers' learning engagement. De Barba et al. (2016) discovered that situational interest has a stronger impact on learners' participation than general learning motivation. Courses can be designed to create more situations to stimulate learners' motivation for further learning, constantly activate learners' situational interest, and enhance their future learning. In addition,

we must reflect on the learning task load in MOOC design. For MOOC learners with limited time resources, it is not better to provide them with more resources. In contrast, it is necessary to streamline the design of learning paths, tasks and resources, so that learners can arrange their learning more reasonably.

The study also demonstrates differences in learning engagement, time organization, content visit sequences, and activity participation patterns among the three groups at different levels of achievement. The investment of the *satisfactory* group entailed a lot of energy, time, and feelings throughout the entire course. Their learning duration and frequency were much higher than those of the other two groups. They learned at different stages of the course and had a certain proportion of concentrated learning. They presented more chains of backtracking transitions and more sequences of transitions between activities and behaviours. They interacted with others in the forum actively. All of the above indicated that the *satisfactory* group had stronger learning motivation and active engagement in learning behaviours. However, their learning efficiency was limited, and their learning behaviours were less dedicated and targeted. They had more leaps across chapters, as well as disordered learning sequences both at the activity level and at the behavioural level, which may be why their performance and achievement were not good enough. By contrast, *excellent's* learning behaviours were more effective and concentrated. They were able to achieve the highest degree of task completion and engagement in limited learning time. They organized a greater proportion of long-term learning to ensure deep learning. They adopted more structured and step-by-step content visit patterns and had more regular learning behaviour patterns. *Failed* individuals performed worst on task participation and engagement. Their large sessions were least common in the weekly session distribution. Their content visit pattern show lower frequencies of transition between knowledge points within each chapter, which means that their learning of each chapter may not be solid. Their activity participation pattern lacks meaningful learning sequences. Their learning behaviours were either disordered or simple.

Comparing the learning engagement and learning behaviour patterns among the three groups, enough substantial engagement in the course is an important guarantee for learners to pass it. Substantial engagement refers to the completion degree of tasks, rather than superficial indices such as learning time and session frequency. Some researchers have found a positive relationship among learning duration, session frequency, and learning achievement (Jiang et al., 2015). For example, de Barba et al. (2020) found that successful learners had more frequent and longer sessions across a course. However, other studies have negated the opinion that learners who spend more time learning are certain to obtain higher achievement. Van den Beemt et al. (2018) affirmed that successful learners' steady learning behaviours are much more related to the order of watching videos than to the actual timing. Enough learning time is only one of the conditions required to complete tasks. The time learners with different learning abilities need to complete tasks will be correspondingly different. Excellent learning ability can allow learners to complete tasks quickly, so they can achieve a high degree of task completion in a short period of time. In this research, the significant difference between the *failed* group and the other two groups was their low participation in tasks, which is the main reason why they failed the course.

Further, van den Beemt et al. (2018) showed that non-certificate learners exhibited irregular patterns, such as more skipping quizzes, while learners with high pass rates had the best-regulating learning behaviour.

In comparing the learning engagement and learning behaviour patterns between *excellent* and *satisfactory*, *excellent* had more effective learning organization and more cognitive engagement. Their effective learning organization was embodied in their organization and arrangement of learning time and learning content. Consistent with the findings of de Barba et al. (2020), this study found that successful learners (i.e., *satisfactory* and *excellent*) arranged more large sessions each week than *failed* learners. *Excellent* arranged more large sessions each week than *satisfactory*, which created a better environment for focused and in-depth learning. In addition, compared with *satisfactory*, who had more transitions in learning, *excellent* adopted a more solid and step-by-step learning method, and showed more concentration in learning. Finally, *excellent* made more cognitive efforts and had more specific purposes in task participation. For example, they took the initiative to adjust their pace of watching videos, insisted on completing quizzes, and mainly posted in the forum. The above kinds of performance all reflect an excellent level of cognitive engagement and self-regulated learning skills in MOOC learning. Hence, although a better learning foundation is one reason why *excellent* obtained higher scores in the course, sufficient substantial learning engagement, more cognitive engagement, and higher self-regulated learning are the fundamental reasons.

Conclusions

This study focused on MOOC learners who take final assessments, labelled final-assessment-taking learners, to investigate their learning engagement, time organization, content visit sequences, and activity participation patterns by applying learning analytics methods to examine the learning data of 535 learners who took final assessments in a typical MOOC in China. In particular, this study examined the differences in learning engagement and the above learning patterns among the three groups of learners with different levels achievement defined as *failed*, *satisfactory* and *excellent*. There were differences in both learning engagement and learning patterns among the three groups. The results indicated that success is determined by a certain degree of task completion as a basic guarantee for passing the course, effective session workload organization, reasonable learning content arrangement, and more cognitive engagement, rather than investing more time and energy. A higher degree of task completion, a greater proportion of large sessions, gradual learning, and activity participation patterns that reflect more cognitive engagement will encourage MOOC learners to obtain excellent academic achievement.

The results of this study, on the one hand, support the prediction of learning achievement based on learning engagement and behavioural patterns and, on the other, provide useful cues for personalised MOOC design and teaching. Teachers can predict learners' learning achievement based on the learning engagement and learning behaviour patterns of the three groups and learners' learning performance. In this way, they can provide targeted intervention and support in time. In addition, teachers can consider how to optimise MOOC design according to learners' different needs, referring to the characteristics of the three groups' learning behaviour patterns in this study. For example, they can design tests to evaluate learners' prior ability before the course, and provide

Li *et al. Int J Educ Technol High Educ*   (2022) 19:22

Page 18 of 20

learners with personalised learning advice, resources, and even personalised learning paths according to their results. They can support learning behaviours across chapters if learners pass the chapter test, and identify learners who need support and intervention through learning methods.

Recommendation for future research

This study has several limitations. First, this study chose learners in only one MOOC as the sample, which might not be typical enough. The findings need to be tested in more learners and courses, especially in courses with different teaching styles. This could help us to discover more factors that influence learners' learning behaviour patterns. Second, this study explored learners' learning behaviour data in the course through learning analytic methods. More research techniques, such as questionnaires and interviews, should be used to investigate learners' subjective learning motivation and plans; doing so would be meaningful to confirm the findings of this study and to extract comprehensive learning behaviour patterns among MOOC learners. In addition, future studies should explore how learners' prior learning foundation, learning habits, learning styles, and other factors affect their learning behaviour patterns and performance, which have a strong impact on the explanations of the results in this study. Future research is suggested to scrutinise how other factors influence learners' learning behaviour patterns.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Research Center of Distance Education, School of Educational Technology, Beijing Normal University, Beijing, China.
[2]School of Educational Technology, Beijing Normal University, 19, Xinjiekouwai Street, Haidian District, Beijing, China.

References

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. In *Proceedings of the 23rd international conference on World Wide Web* (pp. 687–698).
Bozkurt, A., Koutropoulos, A., Singh, L., & Honeychurch, S. (2020). On lurking: Multiple perspectives on lurking within an educational community. *The Internet and Higher Education, 44*, 100709. https://doi.org/10.1016/j.iheduc.2019.100709
Chaker, R., & Bachelet, R. (2020). internationalizing professional development: Using educational data mining to analyze learners' performance and dropouts in a French MOOC. *International Review of Research in Open and Distributed Learning, 21*(4), 199–221.

Charo, R., Maite, A.-S., & Guillermo, M. (2020). Self-regulation of learning and MOOC retention. *Computers in Human Behavior, 111*, 106423. https://doi.org/10.1016/j.chb.2020.106423

Chen, B., Fan, Y., Zhang, G., & Wang, Q. (2017). Examining motivations and self-regulated learning strategies of returning MOOCs learners. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 542–543). https://doi.org/10.1145/3027385.3029448

Chen, Y., Gao, Q., Yuan, Q., & Tang, Y. (2020). Discovering MOOC learner motivation and its moderating role. *Behaviour & Information Technology, 39*(12), 1257–1275. https://doi.org/10.1080/0144929X.2019.1661520

Davis, D., Chen, G., Hauff, C., & Houben, G. J. (2016). Gauging MOOC learners' adherence to the designed learning path. *International Educational Data Mining Society, 16*, 54–61.

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning, 32*(3), 218–231. https://doi.org/10.1111/jcal.12130

de Barba, P. G., Malekian, D., Oliveira, E. A., Bailey, J., Ryan, T., & Kennedy, G. (2020). The importance and meaning of session behaviour in a MOOC. *Computers & Education, 146*, 103772. https://doi.org/10.1016/j.compedu.2019.103772

Ferguson, R., & Clow, D. (2015). Consistent commitment: Patterns of engagement across time in Massive Open Online Courses (MOOCs). *Journal of Learning Analytics, 2*, 55–80.

Gitinabard, N., Heckman, S., Barnes, T., & Lynch, C. F. (2019). What will you do next? A sequence analysis on the student transitions between online platforms in blended courses. In *International conference on educational data mining 2019.* https://arxiv.org/abs/1905.00928

Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. In *Proceedings of the first ACM conference on Learning@ Scale conference* (pp. 21–30). https://doi.org/10.1145/2556325.2566247

Handoko, E., Gronseth, S. L., McNeil, S. G., Bonk, C. J., & Robin, B. R. (2019). Goal setting and MOOC completion: A study on the role of self-regulated learning in student performance in massive open online courses. *International Review of Research in Open and Distributed Learning.* https://doi.org/10.19173/irrodl.v20i4.4270

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of Massive Open Online Courses (MOOCs): Motivations and challenges. *Educational Research Review, 12*, 45–58. https://doi.org/10.1016/j.edurev.2014.05.001

Jansen, R. S., van Leeuwen, A., Janssen, J., Conijn, R., & Kester, L. (2020). Supporting learners' self-regulated learning in massive open online courses. *Computers & Education, 146*, 103771.

Jiang, Z., Zhang, Y., & Li, X. (2015). Learning behavior analysis and prediction based on MOOC data. *Journal of computer research and development, 52*(3), 614.

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*(1), 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001

Jung, Y., & Lee, J. (2018). Learning engagement and persistence in Massive Open Online Courses (MOOCS). *Computers and Education, 122*(1), 9–22. https://doi.org/10.1016/j.compedu.2018.02.013

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18–33.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). https://doi.org/10.1145/2460296.2460330

Lan, M., & Hew, K. F. (2020). Examining learning engagement in MOOCs: A self-determination theoretical perspective using mixed method. *International Journal of Educational Technology in Higher Education, 17*(1), 1–24. https://doi.org/10.1186/s41239-020-0179-5

Li, Q., & Baker, R. (2018). The different relationships between engagement and outcomes across participant subgroups in massive open online courses. *Computers & Education, 127*, 41–65. https://doi.org/10.1016/j.compedu.2018.08.005

Li, Q., & Wan, F. (2016). A case study of the characteristics of MOOCs completers: Taking an online professional training MOOC for example. In *2016 IEEE 16th international conference on advanced learning technologies* (pp. 503–505). https://doi.org/10.1109/ICALT.2016.2

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education, 29*, 40–48.

Liu, B., Wu, Y., Xing, W., Cheng, G., & Guo, S. (2021a). Exploring behavioural differences between certificate achievers and explorers in MOOCs. *Asia Pacific Journal of Education.* https://doi.org/10.1080/02188791.2020.1868974

Liu, Z., Mu, R., Liu, S., Peng, X., & Liu, S. (2021b). Modeling temporal association of cognition-topic in MOOC discussion to track learners' cognitive engagement dynamics. In *Proceedings of the eighth ACM conference on Learning @ Scale* (pp. 319–322). https://doi.org/10.1145/3430895.3460170

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior, 80*, 179–196.

Peng, X., & Xu, Q. (2020). Investigating learners' behaviors and discourse content in MOOC course reviews. *Computers and Education, 143*, 103673. https://doi.org/10.1016/j.compedu.2019.103673

Poquet, O., Jovanovic, J., & Dawson, S. (2020). Differences in forum communication of residents and visitors in MOOCS. *Computers and Education, 156*, 103937. https://doi.org/10.1016/j.compedu.2020.103937

Rizvi, S., Rienties, B., & Rogaten, J. (2018). Temporal dynamics of MOOC learning trajectories. In *Proceedings of the first international conference on data science, e-learning and information systems* (vol. 39, pp. 1–6). https://doi.org/10.1145/3279996.3280035

Rizvi, S., Rienties, B., Rogaten, J., & Kizilcec, R. F. (2020). Investigating variation in learning processes in a Future Learn MOOC. *Journal of Computing in Higher Education, 32*, 162–181. https://doi.org/10.1007/s12528-019-09231-0

Shah, D. (2020). The second year of the MOOC: A review of MOOC stats and trends in 2020. Class Central. Retrieved March 3, 2021, from https://www.classcentral.com/report/the-second-year-of-the-mooc/

Siemens, G., & Gasevic, D. (2012). Guest editorial-learning and knowledge analytics. *Journal of Educational Technology & Society, 15*(3), 1–2.

Taylor, O., Keatley, D. A., & Clarke, D. D. (2020). A behavior sequence analysis of perceptions of alcohol-related violence surrounding drinking establishments. *Journal of Interpersonal Violence, 35*(9–10), 1982–1997.

Tough, A. M. (1971). *The adult's learning projects. A fresh approach to theory and practice in adult learning* (2nd ed.). Ontario Institute for Studies in Education.

Tseng, S. F., Tsao, Y. W., Yu, L. C., Chan, C. L., & Lai, K. R. (2016). Who will pass? Analyzing learner behaviors in MOOCs. *Research and Practice in Technology Enhanced Learning, 11*(1), 1–11.

van den Beemt, A., Buijs, J., & van der Aalst, W. (2018). Analysing structured learning behaviour in massive open online courses (MOOCs): An approach based on process mining and clustering. *International Review of Research in Open and Distributed Learning, 19*(5), 37–60. https://doi.org/10.19173/irrodl.v19i5.3748

Wen, M., & Rosé, C. P. (2014). Identifying latent study habits by mining learner behavior patterns in massive open online courses. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management* (pp. 1983–1986). https://doi.org/10.1145/2661829.2662033

Whipp, J. L., & Chiarelli, S. (2004). Self-regulation in a web-based course: A case study. *Educational Technology Research and Development, 52*(4), 5. https://doi.org/10.1007/BF02504714

Xing, W., & Du, D. (2019). Dropout prediction in MOOCs: Using deep learning for personalized intervention. *Journal of Educational Computing Research, 57*(3), 547–570. https://doi.org/10.1177/0735633118757015

Ye, C., Kinnebrew, J. S., Biswas, G., Evans, B. J., Fisher, D. H., Narasimham, G., & Brady, K. A. (2015). Behavior prediction in MOOCs using higher granularity temporal information. In *Proceedings of the second (2015) ACM conference on Learning@ Scale* (pp. 335–338). https://doi.org/10.1145/2724660.2728687

Publisher's Note