Computers & Education 140 (2019) 103595

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course

Jacqueline Wong[a,*], Mohammad Khalil[b], Martine Baars[a], Björn B. de Koning[a], Fred Paas[a,c]

[a] Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, the Netherlands
[b] Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway
[c] Early Start Research Institute, University of Wollongong, Australia

ARTICLE INFO

Keywords:
Self-regulated learning (SRL)
Massive open online course (MOOC)
Clickstream data
Sequential pattern mining
Learning analytics

ABSTRACT

Self-regulated learning (SRL) refers to how learners steer their own learning. Supporting SRL has been shown to enhance the use of SRL strategies and learning performance in computer-based learning environments. However, little is known about supporting SRL in Massive Open Online Courses (MOOCs). In this study, weekly SRL prompts were embedded as videos in a MOOC. We employed a sequential pattern mining algorithm, Sequential Pattern Discovery using Equivalence classes (cSPADE), on gathered log data to explore whether differences exist between learners who viewed the SRL-prompt videos and those who did not. Results showed that SRL-prompt viewers interacted with more course activities and completed these activities in a more similar sequential pattern than non SRL-prompt viewers. Also, SRL-prompt viewers tended to follow the course structure, which has been identified as a behavioral characteristic of students who scored higher on SRL (i.e., comprehensive learners) in previous research. Based on the results, implications for supporting SRL in MOOCs are discussed.

1. Introduction

Massive Open Online Courses (MOOCs), like most online courses, offer learners the flexibility of self-paced learning without the constraints of time and place (Jung, Kim, Yoon, Park, & Oakley, 2019). To enable self-paced learning, many activities in MOOCs are asynchronous in nature, whereby learners watch a series of videos, take quizzes, or participate in discussion forums. Yet, unlike online courses that offer credits, MOOCs have no enrolment restrictions and can be taken by any interested individual at little or no cost (Liyanagunawardena, Adams, & Williams, 2013). Therefore, MOOCs have a much larger and more diverse learner population than other online learning environments. In that respect, designing instructions to support the highly diverse learners in MOOCs is important but challenging (Kovanović et al., 2019). Particularly, a key challenge relates to the issues of providing learners of different characteristics with adequate guidance to steer their own learning. The processes involved in steering one's own learning, such as planning, self-monitoring, and self-reflecting, is known as self-regulated learning (SRL; Zimmerman, 2002).

While demonstrations of positive effects of SRL strategies on learning in MOOCs continue to be important (Lee, Watson, & Watson, 2019), it is increasingly argued (e.g., Kovanović et al., 2019) that understanding how learners make use of the provided support is essential for enhancing the overall MOOC learning experience. Therefore, the main objective of the current study is to gain an

* Corresponding author. Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, the Netherlands.

E-mail address: wong@essb.eur.nl (J. Wong).

https://doi.org/10.1016/j.compedu.2019.103595
Received 7 February 2019; Received in revised form 28 May 2019; Accepted 29 May 2019
Available online 01 June 2019

understanding of how MOOC learners make use of the SRL supports by exploring sequences of learners' activities in a MOOC embedded with videos prompting SRL.

In the following sections of this paper, we provide a theoretical background on MOOCs and SRL. Then, we propose the use of learning analytics to identify sequences of learner activities as a means to examine how learners interact with course activities in an SRL-supported MOOC. The current study is presented in the third section followed by the method in the fourth section and results in the fifth section. Finally, in the last section, we discuss our findings in relation to the model of SRL and previous studies.

2. Theoretical background

2.1. MOOCs and self-regulated learning

MOOCs appeal to many learners because not only are the courses offered for free or at a relatively low cost for certification but they are also considerably autonomous. Most MOOCs hosted on Coursera, a major MOOC platform, are typically designed for learners to engage with the course activities in a linear manner. The course activities are organized by weeks. In a typical course week, course activities may comprise of a series of video lectures to deliver the course content followed by discussion prompts to stimulate thinking of key concepts, and an assessment to assess the knowledge gained from the video lectures. However, learners do not have to follow the designed sequence of the course activities as they have the autonomy to pursue their own sequence of learning activities, such as when to watch a video lecture, and which course activity to engage in after watching a video lecture. Moreover, learners are free to skip any of the course activities. Therefore, learning in MOOCs requires learners to self-monitor their learning and engage with the course activities in a way that would optimize their learning. This implies that learning in MOOCs requires learners to self-regulate their learning.

Self-regulated learning (SRL) is a broad framework that encompasses motivational, metacognitive, cognitive, affective, and behavioral aspects of learning (Panadero, 2017). Several SRL models have been developed to conceptualize the phases and processes of SRL (for review, see Panadero, 2017). Based on Puustinen and Pulkkinen's (2001) review, SRL models are typically illustrated by processes operating in cyclical phases. For example, in Zimmerman's SRL model, the learning process functions in three cyclical phases: forethought, performance, and self-reflection (Zimmerman & Moylan, 2009). Learners start with the forethought phase where they are involved with task analysis and self-motivation beliefs. They set goals and make plans before embarking on a learning task. Self-motivation beliefs influence these goals and plans. After the forethought phase, learners proceed to the performance phase where they carry out their plans by exercising self-control and self-observation. To effectively learn, learners manage their time, structure their environment, and apply useful learning strategies. In addition, they monitor their learning progress. The self-reflection phase comes after the performance phase. Learners evaluate their learning progress based on the information derived from metacognitive monitoring in the performance phase and the feedback they are given. Learners reflect on their goals, plans and strategies, and make use of these information to form new goals and plans.

In view of the number of processes involved in SRL, it is not surprising that SRL is linked to academic success in both traditional (Dent & Koenka, 2016) and online higher education (Broadbent & Poon, 2015). Similarly, several meta-analyses have shown a positive relationship between SRL-supports, SRL strategies, and academic achievement (Dignath & Büttner, 2008; Zheng, 2016). In MOOCs research, SRL is gaining increasing attention. According to Lee, Watson, and Watson's (2019) review, the studies on SRL in MOOCs are mainly exploratory and correlational. Furthermore, a small but increasing number of studies have examined interventions to support SRL in MOOCs. Therefore, other than examining whether supporting SRL enhance learning in MOOCs, it is of interest to understand how MOOC learners behave in an SRL-supported MOOC.

Investigating SRL as a process is made possible by the emergence of online environments where data on learners' activities in a course (e.g., watched a video lecture and completed an assignment) can be collected (Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, Morales, & Munoz-Gama, 2017). When viewed as a process, SRL can be conceptualized as a series events or actions generated by the learners during learning (Bannert, Reimann, & Sonnenberg, 2014).

SRL events can be operationalized at three levels (Winne, 2010). The first level is the frequency of the observed actions. For example, the number of times a learner watched a video. The second level is the transition state. For example, the activity that a learner begins after ending the previous activity. The third is the sequence of transitions that regularly occurs. For example, learners proceed to a discussion after viewing a video followed by completing a quiz. With the extensive use of computer-based learning environments, Winne (2010) postulated that traces being behavioral manifestations of motivational, cognitive, and metacognitive events measure SRL more adequately than self-reports and think-aloud protocols.

2.2. Applying learning analytics in MOOCs

Learning analytics refers to the process of measuring, collecting, analyzing, and reporting of data about the learners and the learning environment to help understand and optimize learning and the learning environment (Siemens, 2012). When learners click on the online course activities, MOOCs, as with all IT systems, are able to capture and store data on learners' interactions with the course activities. Therefore, the data set generated by MOOC learners is considerably large, providing a strong basis for the use of learning analytics to examine learner behavior in MOOCs.

2.2.1. Identifying MOOC learners' engagement patterns

Studies that applied learning analytics to identify patterns of learner engagement in MOOCs suggest that learners engage with the

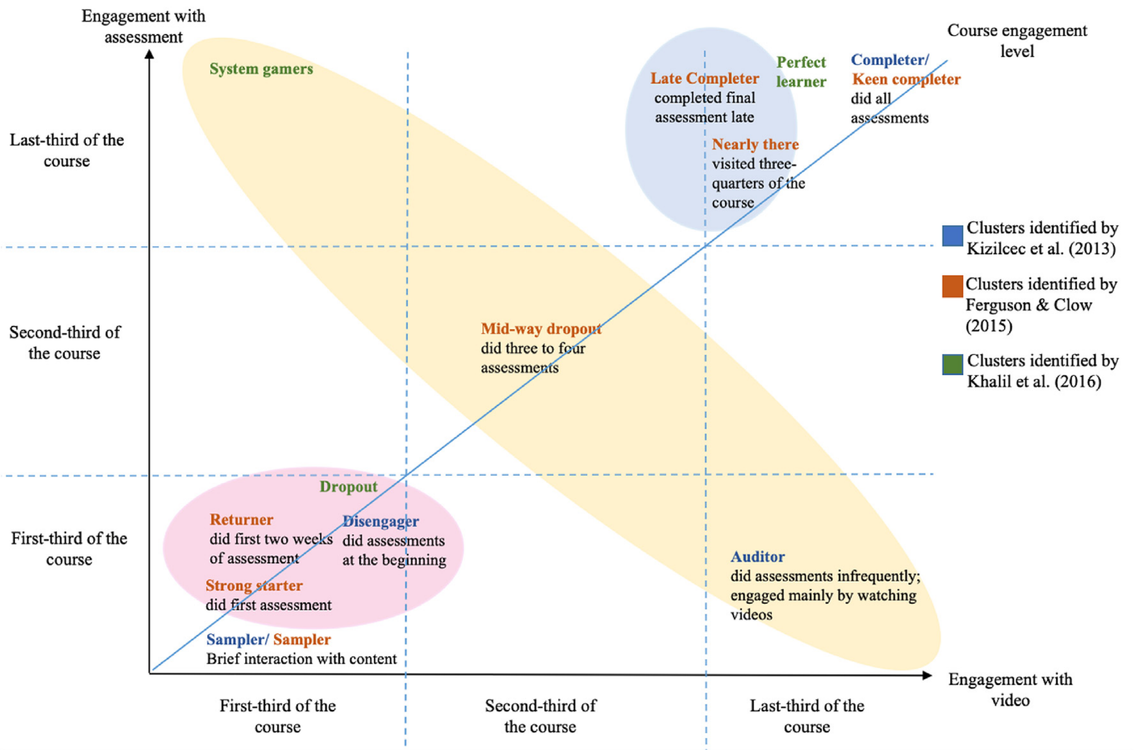

Fig. 1. Course engagement of learner clusters based on level of engagement with videos and assessments in MOOCs.

course materials in a myriad of ways. While MOOCs may differ in the type and number of activities, video lectures and assessments are the two main activities that are featured in most MOOCs. Therefore, based on learners' engagement with the videos and assessments, the cluster of learners can be mapped onto a course engagement continuum. Fig. 1 maps the different clusters of MOOC learners identified in studies by Kizilcec, Piech, and Schneider (2013), Ferguson et al. (2015), and Khalil, Kastl, and Ebner (2016).

The learners clustered at the top-end of the course engagement continuum can be generalized as ‘completer’. Kizilcec et al. (2013) defined ‘completer’ as learners who completed the majority of the assessments provided in the course. In Khalil and Ebner's (2016) study, the cluster of learners who were highly engaged in videos and assessments were labeled as ‘perfect learner’. However, according to Ferguson et al. (2015), there were two clusters of ‘completer’: ‘keen’ and ‘late’. Both ‘keen and late completer’ referred to learners who were highly engaged throughout the course with the only difference being that ‘keen completers’ submitted almost all assessments on time while ‘late completers’ submitted assessments late. Although late submission is not of consequence in MOOCs, it can be argued that submitting assessments on time is a behavioral indicator of SRL (Zimmerman, 1990). Another cluster that is close to the top-end of the course engagement continuum is the cluster of ‘nearly there’ identified by Ferguson et al. (2015). Learners in the ‘nearly there’ cluster consistently completed assessments during the course but dropped out just before the course ended. In Fig. 1, ‘nearly there’ and ‘late completer’ are grouped by a blue oval to indicate the group of learners that is considerably engaged in their learning but may need some support to not only complete the course but also complete it in a timely manner.

The yellow oval in Fig. 1 indicates clusters mapped onto the midpoint of the course engagement continuum. Ferguson et al. (2015) described learners who dropped out after visiting half of the course and completing about three to four assessments as ‘mid-way dropout’. While there are many reasons for dropping out halfway through the MOOCs, one of the reasons often mentioned by learners was poor time-management (Eriksson, Adawi, & Stöhr, 2017). Two other distinctive clusters were ‘auditor’ (Kizilcec et al., 2013) and ‘system gamer’ (Khalil & Ebner, 2016). ‘Auditor’ engaged with the course mainly by watching videos whereas ‘system gamer’ engaged with assessments more than the videos and did what is required to get a grade. By engaging with only one type of course activity, learners in these two clusters do not fully benefit from what the whole course has to offer.

The rest of the clusters map onto the lower end of the course engagement continuum. Kizilcec et al. (2013) and Ferguson et al.'s (2015) referred to learners who briefly interacted with the course activities as ‘sampler’. Many ‘samplers’ try out the course by watching one or two videos at the beginning of the course. Therefore, their course engagement level is the lowest among all clusters. Apart from ‘sampler’, there are four other clusters (i.e., ‘dropout’, ‘strong starter’, ‘returner’, and ‘disengager’) at the lower end of the course engagement continuum as indicated by the pink oval in Fig. 1. Compared to the learner clusters in the yellow oval, learner clusters in the pink oval are even less engaged in the course and these learners typically dropped out at an early stage of the course. In the courses examined by Kizilcec et al. (2013), three major reasons reported by learners for disengaging were work conflict, course workload, and personal commitments. To some extent, SRL strategies such as planning could be a way to manage course workload and personal commitments.

Together, the findings of the above-described studies suggest that learners engage in MOOCs in many different ways. Apart from learners who were trying out the course (i.e., 'sampler') and learners who were highly engaged in the course (i.e., 'keen completer' and 'perfect learner'), there are learners who failed to complete the course and either drop out at a very early stage, half-way through the course, or even just before the course ended. SRL is linked to learners' engagement with the course (Littlejohn, Hood, Milligan, & Mustain, 2016). According to SRL theories, self-regulated learners engage in the course activities in a way to achieve their objectives, suggesting that the manner in which learners engage in the course activities would inform us about how learners are self-regulating their learning or not (Bannert et al., 2014).

2.2.2. *Examining self-regulated learning in MOOCs*

Studies examining SRL in MOOCs suggest that SRL has an impact on the way learners engage with the course activities in MOOCs and their learning goals. Kizilcec, Pérez-Sanagustín, and Maldonado (2017) examined the role of SRL in MOOCs by relating six self-reported SRL strategies (i.e., goal setting, strategic planning, self-evaluation, task strategies, elaboration, and help-seeking) with learner behavior and goal achievement. The results showed that learners with higher scores on goal setting and strategic planning were more likely to achieve personal course goals. The self-reported SRL strategies (except help-seeking) were related to frequent revisiting of course materials. However, the self-reported SRL strategies measured were not directly observable by the data. Therefore, it is not clear whether learners set goals or self-explained while watching videos. Moreover, the SRL strategies were found to differ by individual characteristics such as age and educational level, suggesting that the way learners engage with the course materials could be influenced by SRL-related individual differences.

Extending Kizilcec and Cohen's (2017) study, Maldonado-Mahauad et al. (2017) used process mining and clustering methods to examine SRL in MOOCs. Based on the most frequent interaction sequences in the MOOC studied, they identified six common interaction sequences. Two of the common interaction sequences consisted of interactions with either videos (i.e, videos only) or assessments (i.e., assessment only). The other four sequences consisted of combinations of interactions with videos and assessments: i) 'assessment-try to video' sequences in which learners tried assessments before searching for information in content videos, ii) 'video to assessment-pass' sequences in which learners started with content videos and ended with passing assessments, iii) 'video-complete to assessment-try' sequences in which learners viewed content videos before trying assessments, and vi) 'explore' sequences in which learners engaged with assessments and videos without completing any of them.

Using the six common interaction sequences, three learner clusters were formed: sampling, comprehensive, and targeting. The self-reported SRL scores of comprehensive and targeting clusters were similar and higher than the sampling cluster, suggesting that the comprehensive and targeting learners were more self-regulated. While both comprehensive and targeting learners were equally likely to complete the course, they interacted with the course in different ways. Comprehensive learners tend to follow the sequential structure of the course materials and they performed more 'videos only' and 'video to assessment-pass' sequences than targeting learners. Targeting learners tend to look for information to pass the course assessment and performed more 'video-complete to assessment-try' sequences than comprehensive learners. Learners in the sampling clusters had low self-reported SRL and low course activities. They also behave in irregular ways and were least likely to complete the course. Their profile is similar to the 'sampler' cluster in Kizilcec et al. (2013) and Ferguson et al.'s (2015) studies (see Fig. 1).

These findings are in line with research that showed that for MOOC learners to achieve their objectives, they need to be capable of regulating their own learning (Littlejohn et al., 2016). For learners who lack the capacity to self-regulate their learning in MOOCs, support would be required (Kizilcec & Cohen, 2017).

2.2.3. *Supporting SRL in MOOCs*

Research on supporting SRL in MOOC is emerging. In traditional online environments, SRL supports, in general, have a positive effect on academic performance (Zheng, 2016). In a recent review, Wong et al. (2019) identified a number of SRL-supports implemented in online learning environments (e.g., prompts, feedback, and integrated support system). Aligned with Zheng's (2016) study, results revealed that SRL supports positively enhance SRL strategies and academic outcome. Prompts were the most widely investigated form of SRL support because of the ease and simplicity of adding prompts to systems. Out of the fourteen studies that examined prompts, only one study (i.e., Kizilcec, Pérez-Sanagustín, & Maldonado, 2016) examined prompts in MOOCs in the form of recommending SRL strategies at the start of the course. Research in MOOCs has also investigated prompting retrieval practice at the end of videos and providing a study planning module (Davis, Chen, Van der Zee, Hauff, & Houben, 2016).

Both Kizilcec et al. (2016) and Davis et al. (2016) did not find any significant differences in course performance or engagement with course activities between learners who were provided the SRL support and those not provided with the support. However, when Davis et al. (2016) compared learners who were provided with the study planning module and interacted with it and those who were not provided with the study planning module, significant differences in course grade and engagement level were found. Learners who interacted with the study planning module logged into the course more often, spent more time in the course, accessed more course materials, and achieved higher final grades than the highly engaged learners who did not interact with the study planning module and the highly engaged learners who were not provided with the study planning module. This suggests that the study planning module increases course activities and course performance of learners in MOOCs only if the learners used the study planning module.

Kizilcec et al. (2016) and Davis et al. (2016) offer two important implications for the current study. First, it is important to examine differences in course engagement between learners who were provided with the SRL support and interacted with it and those who were provided with the SRL support and did not interact with it. Such a comparison will allow us to examine the influence of compliance to the SRL support on course engagement. Secondly, in addition to the frequency of course activities (e.g., percentage of assessments passed, number of logins, and number of videos viewed), studies can examine the sequences in which learners interacted with the course activities to obtain a more detailed insight into how learners navigated through the MOOCs when provided with an SRL support, i.e., how learners interact with the SRL support (in our study).

2.2.4. Sequential pattern mining

To examine the sequences in which learners interacted with the course activities and the SRL support, we employed Sequential Pattern Mining (SPM), a technique 'borrowed' from learning analytics and data mining disciplines where it is used to determine frequently occurring sequences in large databases. SPM has been widely used in market research, such as identifying shopping behaviors to predict what customers are most likely to buy in sequence (e.g., Desai & Ganatra, 2015). Applications in education have so far been limited to detecting patterns relating to the use of course materials in Learning Management Systems (LMS) (Poon, Kong, Wong, & Yau, 2017).

One of the algorithms for discovering sequential patterns is *cSPADE* (Zaki, 2001). Using *cSPADE* to conduct SPM can help us to identify the frequency of the sequences based on the total number of performed events and the number of transitions that were captured by the log file (Chen, Resendes, Chai, & Hong, 2017). Several insights can be gained through the identification of sequences of learner activities and the extent to which the sequences occurred. First, we can identify the order of activities that were followed by most of the learners in the MOOC and come to understand whether learners are adhering to certain patterns when learning in a MOOC. Second, we can examine which course activities are present or absent in the sequences of learner activities. The high presence of certain course activities in the sequences of learner activities may suggest that learners perceive the course activity to have high utility for learning, such as participating in discussion forum after a video may help to deepen understanding of the subject matter.

Therefore, examining sequences of learner activities using *cSPADE* adds a new dimension to analyzing learner behavior in MOOCs beyond the frequency of activities in previous studies (e.g., Davis et al., 2016). When applied to an SRL-supported MOOC, *cSPADE* can help us to examine learners' use of the SRL support and allow comparison of sequences of learner activities between MOOC learners who used the SRL support and those who did not.

3. Current study

The current study aimed to describe and compare learner behavior by analyzing clickstream data obtained from a MOOC with SRL-prompt videos using *cSPADE*. The MOOC is an introduction to the topic of Serious Gaming. Given that the MOOC learners have autonomy over which course activities they want to interact with (e.g., assessment, videos, and discussions), an understanding of learners' use of SRL-prompt videos and sequences of learner activities in relation to SRL in MOOCs can be obtained by identifying and comparing the sequences of learner activities between learners who decided to view the SRL-prompt videos (i.e., SRL-prompt viewers) and learners who decided not to view the SRL-prompt videos (i.e., SRL-prompt non-viewers). The main research question of the present study was:

What are the differences in sequences of learner activities between SRL-prompt viewers and non-viewers in a MOOC embedded with SRL-prompt videos?

Findings from previous studies revealed differences in the interaction sequences with course activities between learners with different levels of SRL (Kizilcec et al., 2017; Maldonado-Mahauad et al., 2017). Learners with high levels of SRL tend to interact with more course activities but the patterns in which they interact with the course activities varied. For example, comprehensive learners had more 'videos only' sequences than targeting learners. Therefore, it is of interest to explore whether similar differences exist between SRL-prompt viewers and non-viewers when applying the *cSPADE* analysis. Moreover, there are currently only a few studies that examined SRL supports in MOOCs (Lee et al., 2019). While the study of Davis et al. (2016) showed that learners who interacted with the study planning module accessed more course materials and spent more time in the course than those who did not interact with the study planning module, it is not clear whether the SRL-prompt viewers would also be more engaged in the course activities than SRL non-prompt viewers.

4. Method

4.1. Participants

During the period from April to June 2017, 655 learners enrolled in the Serious Gaming MOOC. Of which, 222 learners enrolled in the version of the Serious Gaming MOOC with weekly SRL video prompts embedded. The Serious Gaming MOOC was hosted on Coursera whereby the terms of use and privacy policy included agreeing to exposure to variations in course content and data collection for education research purposes. When learners enrolled in the course version with the SRL video prompts embedded, they were informed that modifications were made to the current version they were enrolled in. Through a partnership between Coursera and Erasmus University Rotterdam, de-identified data gathered in this MOOC were made available to us for research purposes. Learners enrolled in the course were identified using a hashed user ID linked to a course session ID unique to the course version with SRL video prompts embedded.

Completing the demographic information was not mandatory on the Coursera platform and from the available demographics data collected on the platform during registration, it appears that overall 44% of the enrolled learners were female. The mean age of the learners was 38.15 years ($SD = 10.58$). The majority of learners had a higher education degree (36.13% held a bachelor's degree and 36.97% held a master's degree). Almost half of the learners reported having full-time employment (47.19%). Despite the few demographics data collected pertaining to the sample of learners in this analysis, there is no reason to assume that the overall

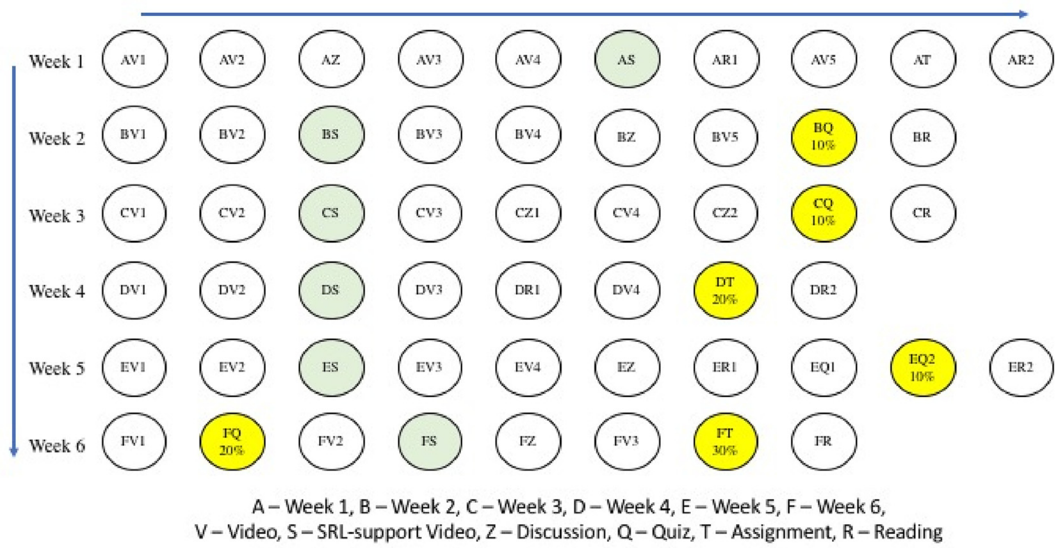

Fig. 2. Overview of the number of course items per week in the Serious Gaming MOOC.

demographics differ for the sample of learners in this study because enrolled learners were randomly assigned to this version of the MOOC (i.e., MOOC embedded with SRL-prompt videos). From the data, 103 learners were identified as active learners (i.e., learners who completed at least one activity) and were included in the analyses. The 103 learners were further grouped as i) learners who watched at least one of the weekly SRL-prompt videos (SRL-prompt viewers, $n = 39$) and ii) learners who did not watch any of the SRL-prompt videos (SRL-prompt non-viewers, $n = 64$).

4.2. Study context

The Serious Gaming MOOC was a six-week course that was launched on Coursera in April 2016. It was one of the first three MOOCs developed by the university and the average completion rate of the MOOC is 3.4%. The Serious Gaming MOOC was selected for three reasons: i) the MOOC utilized all possible types of course activities (i.e., video lectures, reading, quizzes, peer assignments, and discussion forums), ii) the MOOC is offered by the university in which the research is conducted and iii) the instructors of the MOOC gave us consent to embed SRL video prompts in the course.

The Serious Gaming MOOC was designed to introduce serious games (i.e., games that train people to deal with challenges or to create social awareness), such as, a web-based video game 'Darfur is Dying' (http://www.gamesforchange.org/game/darfur-is-dying/) that raises awareness of the situation of refugees in the Darfur region of Sudan. The modules were: introduction and definition (Week 1), theoretical framework on the enjoyment of playing (Week 2), different types of serious games (Week 3), persuasive gaming (Week 4), impact of serious games (Week 5), and future of serious games (Week 6). The expected workload was three to five hours per week.

Fig. 2 gives an overview of the course structure and course activities (e.g., video, discussion, and quizzes). Each week, learners had access to eight or nine course activities. A graded quiz or an assignment was due every week except for the first week. In total, there were 54 activities including the six SRL-prompt videos. To facilitate the analyses, we coded each of the activities using letters and numbers. The first letter represented the week, the second letter represented the type of activity, and the number represented the order of that particular type of activity. For example, the third (3) content video (V) in week 1 (A) was AV3. The course was designed to be navigated in a linear manner. To pass the course, four quizzes and two peer-graded assignments had to be completed. The percentages for the final grade allocated to each graded item are shown in Fig. 2.

4.2.1. Supporting SRL with video prompts

Fig. 3 shows a screenshot of one of the SRL-prompt videos. Three questions were presented in each SRL-prompt video according to the three phases of Zimmerman's SRL model: planning (e.g., Am I setting goals to ensure that I have a good understanding of the course materials?), monitoring (Am I concentrating on learning the materials in this course?), and reflection (e.g., Do I understand all the key points of this week's course material?). To help learners apply the questions to themselves, an overlay appeared after each question to elicit learners' responses on a 5-point Likert scale from 1 (*not at all*) to 5 (*all the time*). The overlay served as a pause after each prompt to give learners the time to comprehend and think as prompted by the questions.

4.3. Clickstream data pre-processing

A MOOC's clickstream data contains information about the MOOC activities that learners did. We followed Wickham's (2014)

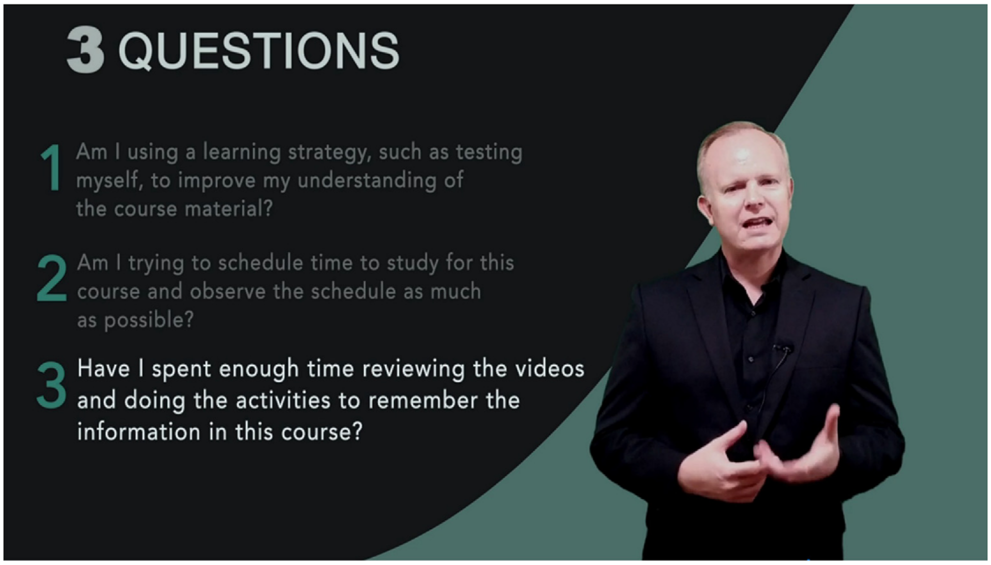

Fig. 3. Screenshot of one the SRL-prompt videos.

method for "tidying up" the data: i) structuring the tables of the data (i.e. rows and columns) so that variables become columns, observations become rows, and type of observational units become distinct tables; ii) labeling nulls and empty records to *NA*; iii) adequately format date and time; and iv) sorting and aggregating values of the tables to achieve the objectives of the analysis phase. After tidying the data, we extracted the required data for the analysis. To do so, we wrote Structured Query Language (SQL) queries and selected the following data for each interaction: (a) timestamp, (b) code of the course activity, and (c) anonymous user id. We then assigned the short abbreviation (e.g., AV2) for each course activity as described in the section "Study context".

4.4. Sequential pattern mining using cSPADE

Sequential Pattern Mining (SPM) is used to determine iterations of occurring frequencies using statistical algorithms that generate sequences with shared patterns. Sequences are defined by various measurements such as occurrence frequency and sequence length. SPM depends on the chosen algorithm to generate the sequences. The underlying main concept is that the generated sequence is identified by a list of items in a sequence S such that a sequence consists of $S = < I_1, I_2, I_3, \ldots, I_\mathrm{n} >$. In our study, I_n, represents learners' interaction with the course activities in the MOOC (e.g., watching the second video in Week 1, followed by participating in Week 1's discussion forum, then watching the fifth video in Week 1).

We employed an SPM algorithm called Sequential PAttern Discovery using Equivalence classes (*cSPADE;* Zaki, 2001) since it offered a direct and simple implementation. The algorithm in *cSPADE* uses a vertical id-list format in the databases, where every sequence is connected with a list of objects in which it appears and sequential patterns are explored based on a predefined minimal support threshold (Slimani & Lazzez, 2013). The support level (i.e. support threshold) is a measurement variable in every SPM algorithm that can be determined by the researcher depending on the number of sequences to be identified and research objectives (for further explanations, see Slimani & Lazzez, 2013). The higher the level of support for the sequences, the more frequent the sequences occurred in the dataset (i.e., more learners interacted with the course activity in that sequence). A value between 1.0 and 0.01 for the support level was used in the current study as a threshold to identify common sequential pattern of learner activities from intersections on the id-lists. The c*SPADE* analysis was performed separately for the SRL-prompt viewers and non-viewers in order to retrieve sequential patterns of learner activities from the clickstream data of the two groups.

4.5. Visualization of interaction patterns

This study used two types of visualizations to describe and represent the analyzed data: alluvial diagram and interaction map. An alluvial diagram shows learner transitions from the weekly SRL-prompt videos to the activities they chose to engage in next. The purpose of depicting transitions using an alluvial diagram was to build an understanding of what the learners did after watching the SRL-prompt videos. Visualizing each transition enabled us to identify (ir)regularities across all transitions.

Interaction maps were used to examine differences in the type and number of activities interacted with by the SRL-prompt viewers and non-viewers. Through the visualization of the activities interacted by the learners, we can understand how learners connected the various course activities.

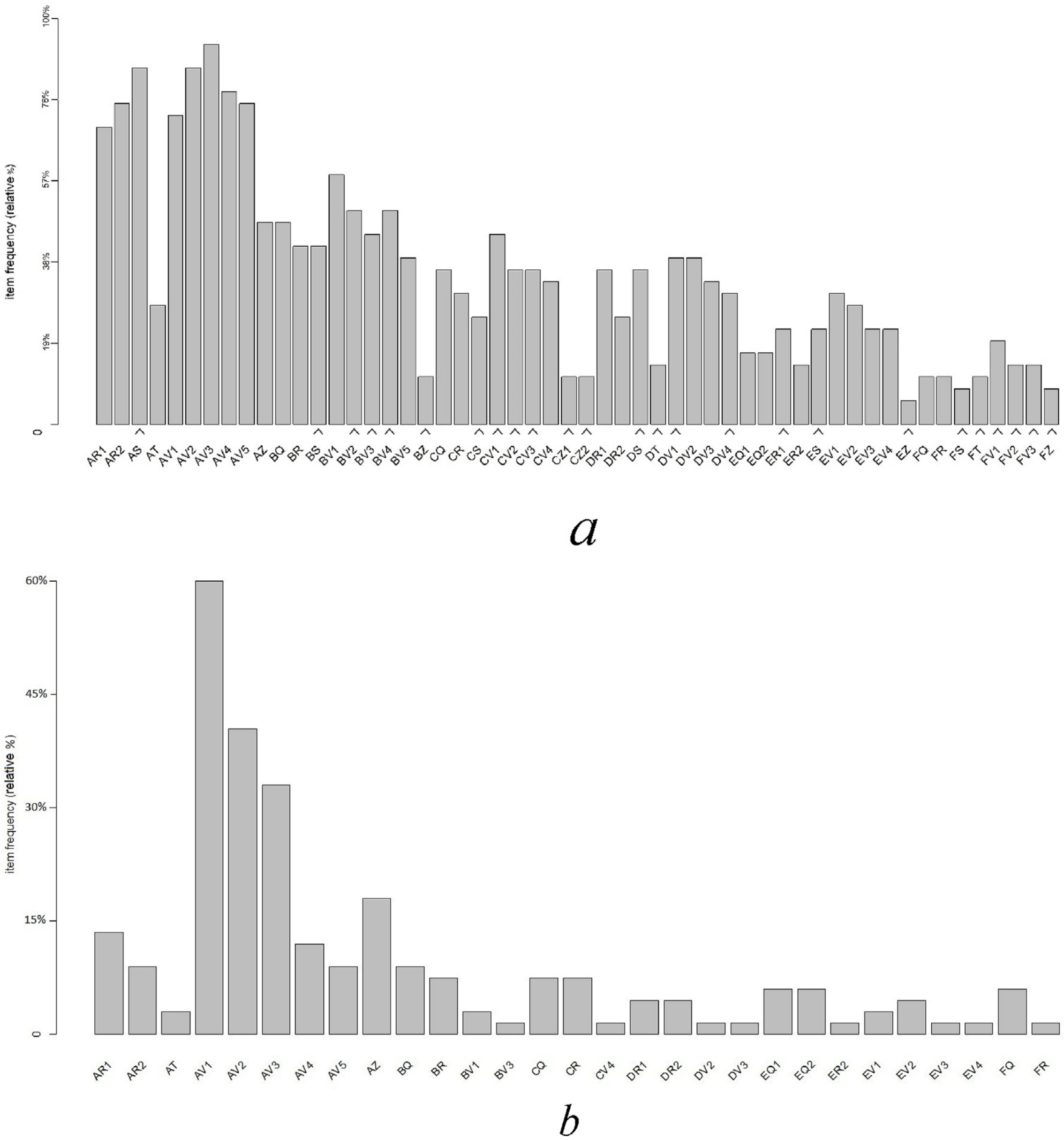

Fig. 4. ***a***. Frequency (%) of SRL-prompt viewers' (n = 39) interaction with course activities. ^ indicates the distinctive course activities that only SRL-prompt viewers interacted with. ***b***. Frequency (%) of SRL-prompt non-viewers' (n = 64) interaction with course activities.

5. Results

5.1. Learners' interaction with course activities

Fig. 4a and Fig. 4b show the frequency of course activities interacted by the SRL-prompt viewers and non-viewers respectively. When comparing the two figures, three observations are made regarding the differences between SRL-prompt viewers and non-viewers' activities. First, there were not only more interactions with course activities by SRL prompt-viewers than non-viewers, but the frequency of interactions per course activity was also higher for SRL-prompt viewers than non-viewers. Based on Fig. 4, SRL-prompt viewers interacted with all 54 possible activities but SRL-prompt non-viewers interacted with only 29 activities. The arrows below the course activities in Fig. 4a indicate the course activities that only the SRL-prompts viewers interact with. When looking at the activities by course week, SRL-prompt non-viewers interacted with fewer than half of the activities in most weeks except in Week 1 and 5 (i.e. activities abbreviated with letter A and E), suggesting that they were more active at the beginning of the course and towards the end of the course.

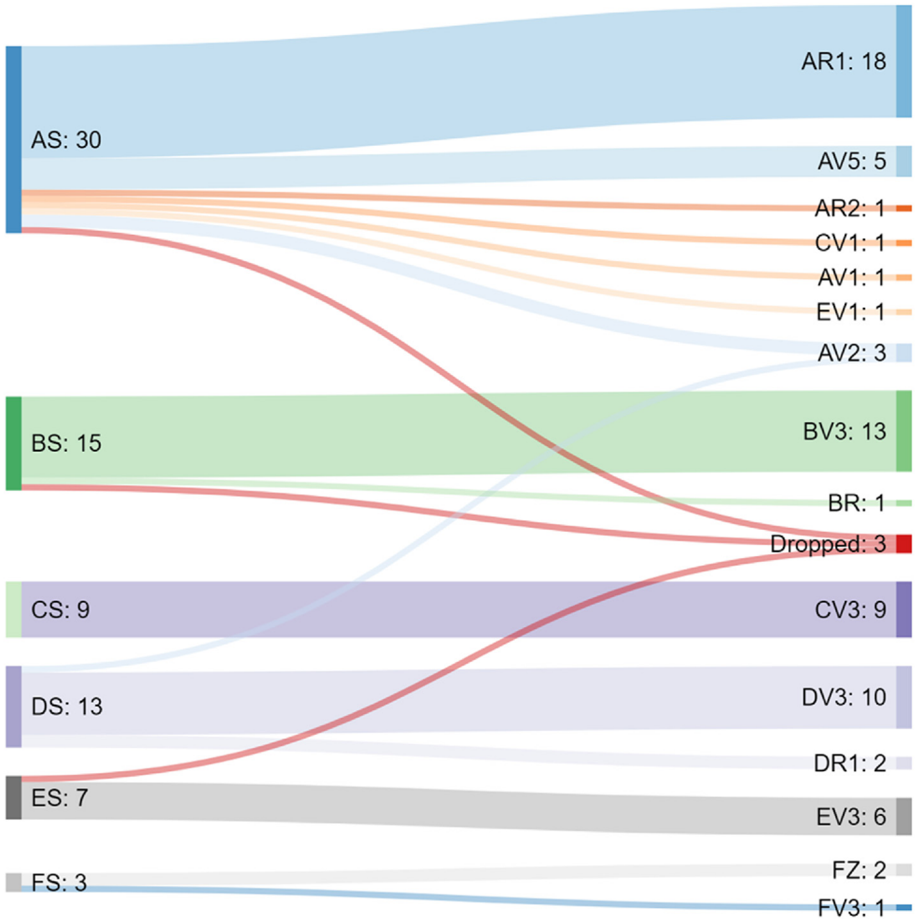

Fig. 5. Students' transition from SRL-prompt videos to the next course activity.

Second, SRL-prompt viewers participated in all discussion prompts (i.e., AZ, BZ, CZ1, CZ2, EZ, and FZ), whereas SRL-prompt non-viewers participated only in the first week's discussion prompt (AZ). Third, SRL-prompt viewers worked on all graded assessments (i.e., four quizzes, BQ, CQ, EQ2, and FQ; two peer-reviewed assignments, DT and FT), whereas SRL-prompt non-viewers worked on four quizzes, but not the peer-reviewed assignments.

5.2. Visualization of learners' interaction with course activities

Fig. 5 is an alluvial diagram to visualize the transitions from the weekly SRL-prompt videos to the next activity. The majority of the learners who viewed the SRL-prompt videos also went on to interact with the activity that followed the SRL-prompt videos in that week as structured in the course. For example, 18 of the 30 learners who viewed the first SRL-prompt video (AS) continued with the course by reading the course material for week 1 (AR1) and 13 of the 15 learners who viewed the second SRL-prompt video (BS) went on to view the third content video of that week (BV3). There are also several learners who did not interact with the activity placed after the SRL-prompt videos. The learners either moved forward to other activities for that week (e.g., BS → BR1, DS → DR1, and FS → FV3) or moved from Week 1 SRL-prompt videos to activities of other weeks (i.e., AS → CV1 and AS → EV1). Three backward transitions were observed (i.e., AS → AV1, AS → AV2, and DS → AV2).

5.3. Sequential pattern of learner activities

The three most frequent sequential patterns of learner activities (i.e., sequences with the highest support level) for each sequence length and the corresponding support levels for the SRL-prompt viewers and non-viewers are illustrated in Table 1. When the support level for *cSPADE* was set at 0.40, frequent sequential patterns of learner activities emerged for the SRL-prompt viewers but not for the SRL-prompt non-viewers. As shown in Table 1, the support level for the sequential patterns of learner activities for SRL-prompt viewers ranged between 0.41 and 0.64. To retrieve sequential patterns of learner activities for the SRL-prompt non-viewers, the support level was gradually reduced till 0.01 at which point sequential patterns of learner activities emerged. As shown in Table 1, the support level for the sequential patterns of learner activities for SRL-prompt non-viewers ranged from 0.02 to 0.27.

Table 1
Frequent sequential patterns of learner activities for SRL-prompt viewers and SRL-prompt non-viewers.

SRL-prompt viewers	Support level	SRL-prompt non-viewers	Support level
3-item sequence			
< AV3, AV4, AS >	0.6410	< AV1, AV2, AV3 >	0.2656
< AV3, AV4, AV5 >	0.6153	< AV1, AV2, AZ >	0.1250
< AV2, AV3, AV4 >	0.6153	< AV1, AZ, AV3 >	0.1250
4-item sequence			
< AV1, AV3, AV4, AS >	0.5897	< AV1, AV2, AZ, AV3 >	0.1250
< AV1, AV2, AV3, AV4 >	0.5641	< AV1, AV2, AV3, AV4 >	0.1093
< AV1, AV3, AV4, AV5 >	0.5384	< BQ, EQ2, EQ1, FQ >	0.0625
5-item sequence			
< AV1, AV2, AV3, AV4, AS >	0.5384	< AV1, AV2, AZ, AV3, AV4 >	0.0625
< AV1, AV2, AV3, AV4, AV5 >	0.4871	< BQ, CQ, EQ2, EQ1, FQ >	0.0468
< AV2, AV3, AV4, BV2, BV4 >	0.4615	< CR, DR1, DR2, EV4, FR >	0.0156
6-item sequence			
< AV2, AV3, AV4, BV1, BV2, BV4 >	0.4615		
< AV1, AV2, AV3, AV4, AS, AR2 >	0.4358		
7-item sequence			
< AV2, AV3, AV4, AS, BV1, BV2, BV4 >	0.4102		

MOOC videos were more commonly featured in the sequential patterns of learner activities identified by the algorithm compared to other activities such as quizzes and discussion. Several sequences (with varying sequence length) include the first week's SRL-prompt video (i.e., AS). Even the longest sequential pattern of learner activities (i.e., 7-item sequence) includes the first SRL video in the middle between Week 1 and Week 2's activities. Additionally, each time the first week's SRL-prompt video (AS) was clicked, it was preceded by the course introduction video (AV4). In this sense, SRL-prompt viewers appear to follow the order of the videos presented in the MOOC.

The results suggest that the sequential patterns of learner activities for SRL-prompt non-viewers are more varied than those of SRL-prompt viewers. For instance, sequential patterns of learner activities of SRL-prompt non-viewers included pure quizzes in the 4-item (i.e., < BQ, EQ2, EQ1, FQ >) and 5-item (i.e., < BQ, CQ, EQ2, EQ1, FQ >) sequences but these sequences are not identified in the data of SRL-prompt viewers.

5.4. Visualization of sequences of learner activities

Interaction maps were used to examine the activities that the learners in the two groups interacted with. Figs. 6a and 6b show the connection between the activities for SRL-prompt viewers and non-viewers respectively. A larger node size indicates a higher frequency of events. Although there were more SRL-prompt non-viewers, the node sizes and ties connecting the nodes as shown in Fig. 6b indicate weaker connections among the activities. This suggests that SRL-prompt viewers made more connections among the activities than SRL-prompt non-viewers.

Similarly, the links connecting the nodes show that SRL-prompt viewers' activities adhere to the structure of the course more than SRL-prompt non-viewers. This further strengthens the observation from the *cSPADE* analysis that SRL-prompt viewers' follow the course structure more closely whereas SRL-prompt non-viewers' approaches to learning are more fragmented and selective.

6. Discussion

To investigate whether there were differences in learner behavior in a MOOC with SRL-prompt videos, we first identified SRL-prompt viewers and non-viewers in the MOOC to be able to describe and compare their learning activities. Only slightly more than one-third of the active learners viewed at least one SRL-prompt video, suggesting that learners in the examined MOOC are less inclined to interact with activities that are related to learning in general but remotely related to the course content (i.e., SRL-prompt videos titled ‘Supporting Your Learning’). This finding aligns with Davis et al.'s (2016) study which showed that learner engagement with ‘added-on’ interventions is lower than with the actual course content. However, SRL-prompt viewers on average interacted with more activities than SRL-prompt non-viewers.

The differences in the support level, length of sequences, and learner activities from the *cSPADE* analysis suggest that SRL-prompt viewers and non-viewers engage in different learning behaviors. Despite the fact that there were fewer SRL-prompt viewers, the sequential patterns of learner activity had higher support levels than SRL-prompt non-viewers. Similarly, a 7-item sequence emerged only from the SRL-prompt viewers' data. These findings suggest that SRL-prompt viewers' learning behaviors as manifested in their sequential patterns of activities are more similar to each other. In contrast, SRL-prompt non-viewers are more diverse and random in their learning behaviors in the MOOC as suggested by the low support level for their sequences. The findings align with Maldonado-Mahauad et al.'s (2017) study which showed that learners who reported higher levels of SRL either followed the sequential structure of the course materials designed by the instructor or strategically sought specific information to complete the course. In contrast, learners who reported lower SRL tend to interact with the course materials in an irregular fashion and are unlikely to complete the course. Therefore, the more orderly sequential patterns of SRL-prompt viewers' activities seem to suggest that SRL-prompt viewers

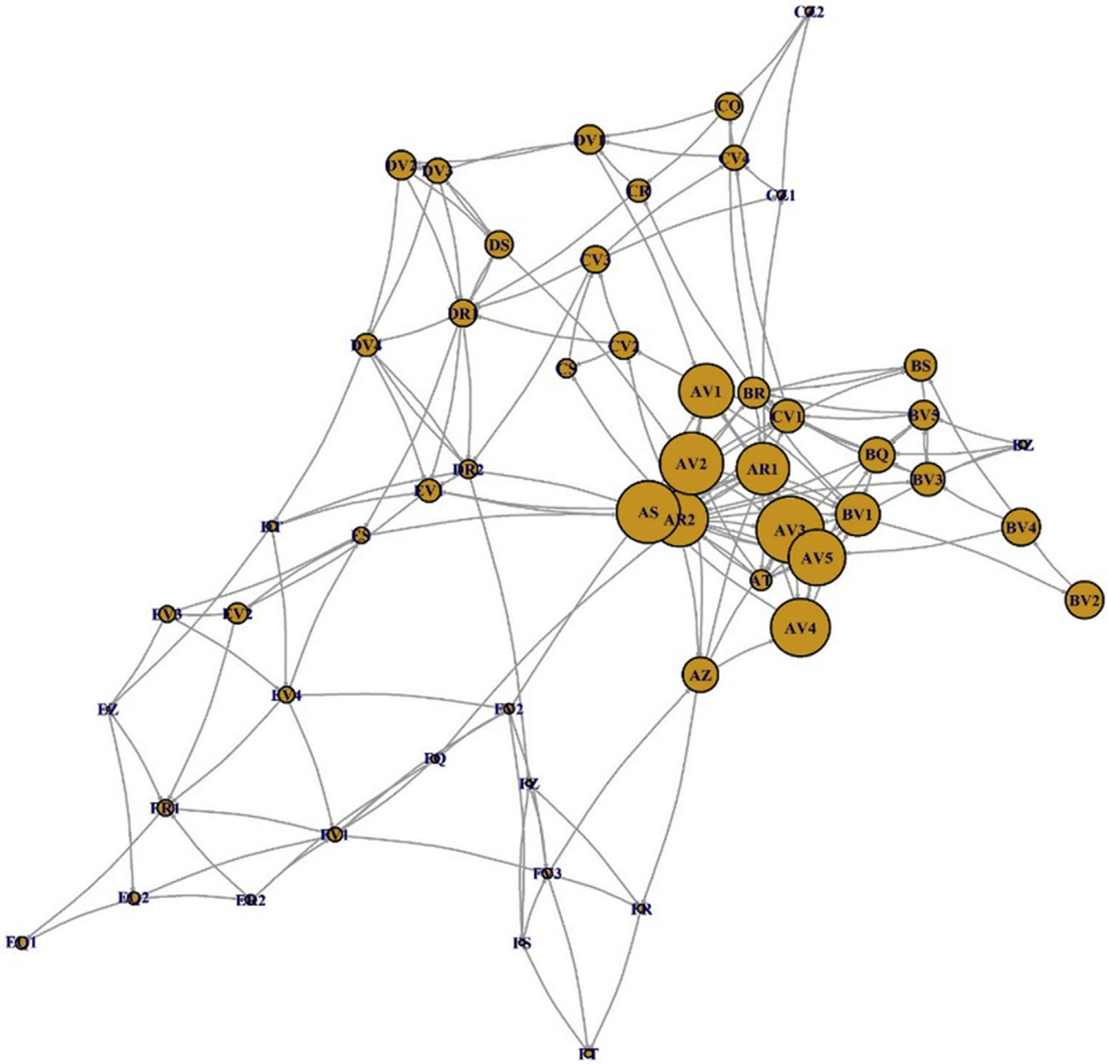

Fig. 6a. Connections among course activities generated by SRL-prompt viewers.

are better at regulating their learning compared to SRL-prompt non-viewers whose sequential pattern of activities were less regular in the MOOC.

In the *cSPADE analysis,* both SRL-prompt viewers and non-viewers exhibited sequential patterns of learner activities that consisted of ‘only videos’. Maldonado-Mahauad et al. (2017) observed that sequential patterns of ‘only videos’ were associated with three SRL strategies: studying, rehearsing, and repeating. Given that the support level for ‘only videos’ sequences was higher for the SRL-prompt viewers than for non-viewers, it could be that SRL-prompt viewers are more diligently using the three SRL strategies than non-viewers. Another interesting finding was that ‘only assessment’ sequences were found only in SRL-prompt non-viewers. Maldonado-Mahauad et al.'s (2017) observed that the ‘only assessment’ sequence was associated with SRL strategies of elaboration and evaluation. However, the support level for this sequence was very low (i.e., 0.06 and 0.05) in the group of SRL-prompt non-viewers. This suggests that only a few SRL-prompt non-viewers are using some SRL strategies.

The differences between the SRL-prompt viewers and non-viewers’ engagement with discussion prompts is also noteworthy. Discussion prompts provide learners the opportunity to use cognitive SRL strategies, such as elaboration or organization (Boekaerts, 1997). By making use of discussion prompts as a learning activity, learners can deepen their understanding of the subject matter. The results showed SRL-prompt viewers interacted with more discussion prompts than SRL-prompt non-viewers. Therefore, SRL-prompt viewers seem to be more actively involved in their learning which is an important characteristic of a self-regulated learner (Zimmerman & Moylan, 2009). However, the quantitative analysis used in this study (here: how often discussion forums were used), do not provide information on what strategies the SRL-prompt viewers used and why SRL-prompt non-viewers skipped the discussion prompts (Lackner, Khalil, & Ebner, 2016). Future studies could use text mining or content analysis to extract information from learners' posts to examine how the discussion posts indicate the use of SRL strategies, which would be an interesting avenue for future

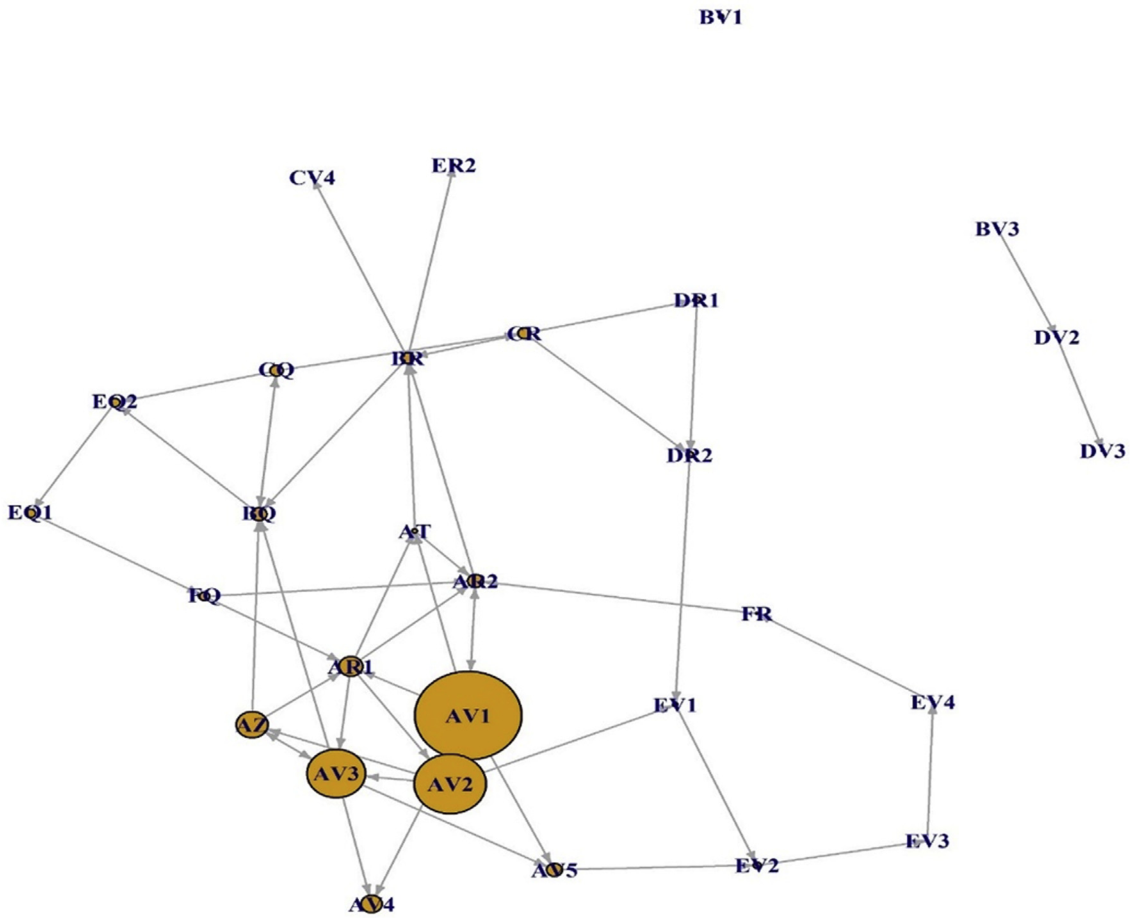

Fig. 6b. Connections among course activities generated by SRL-prompt non-viewers.

research.

Besides differences in participation of discussion prompts, SRL-prompt viewers and non-viewers also differed in their course assessment choices (i.e., four quizzes and two peer-reviewed assignments). Peer- and self-assessments relate to SRL subprocesses of self-monitoring and self-evaluation in Zimmerman's SRL model (Zimmerman & Moylan, 2009). In addition, Maldonado-Mahauad et al. (2017) found that completers were more engaged in assessments and reported higher levels of SRL than non-completers. SRL-prompt viewers worked on all assessment (regardless of quizzes or peer-reviewed assignments) while SRL-prompt non-viewers only worked on quizzes. Such a difference questions whether SRL-prompt viewers are more persistent in completing the course since they attempt more assessments. However, correlations with learners' SRL profile and course grades will be required to examine the effects of watching SRL-prompt videos on learning outcomes. An analysis of the transitions following course assessment items will be needed to examine how learners make use of the course assessment items for SRL, for instance whether learners revisited content videos after obtaining a low score on a course assessment.

6.1. Limitations of the study

There are at least three limitations in this study. First, repetition of the course activities and duration of each activity are not accounted for in the *cSPADE* algorithm. Therefore, the algorithm should be further developed to model the sequential patterns of learner activities to better understand learning behaviors: 1) whether there are behavioral differences when repetitions are considered, 2) whether SRL-prompt viewers and non-viewers differ in their time spent on the course activities, and 3) whether learners adapt their learning over time when self-regulating their learning in MOOCs. Second, the study compared learners who decided to view the SRL-prompt videos and those who decided not to view. Therefore, it is possible that the SRL-prompt viewers are a self-selected group and they could be the more self-regulated or conscientious learners to begin with. To better understand the differences between learners' sequential patterns of learner activities and SRL-support, future work should include groups with different conditions (e.g. without the SRL support or with another type of SRL support) in the analysis. Finally, the current analysis is based on only clickstream data. Without learners' self-reported SRL, we are unable to relate the observed sequences of learners with the use of

SRL strategies that are unobservable from the clickstream data (e.g., goal setting). Future work could also include triangulation of clickstream data, learners' self-report of motivation and SRL, and course grades, to better understand how various sequential patterns of learner activities relate to learner characteristics and learning outcomes. Nonetheless, the analyses detailed in this paper are valuable in helping us to develop the algorithm used. This will not only increase the efficiency of the analytical procedures in the future but also help us to work with MOOC designers and instructors to enhance the content as well as the activities of the MOOCs that were interacted with most by the learners and to design better interventions to support learning in MOOCs.

7. Conclusion

SRL is a broad and complex construct and supporting SRL in MOOCs is inherently challenging as the group of MOOC users consists of highly diverse learners. The current study explored sequences of learner activities in a MOOC that provided learners with videos prompting learners to think about how well they are planning, monitoring, and reflecting. The findings suggest differences in sequential patterns of learner activities between those who viewed the SRL-prompts and those who did not. SRL-prompt viewers tend to follow the sequential structure of the course provided by the instructor and whereas this was less so in the group of SRL-prompt non-viewers. This result is similar to the interaction sequences found in Maldonado-Mahauad et al.'s (2017) study in which ‘Comprehensive learners’ tended to follow the sequential structure of the course provided by the instructor. In addition, SRL-prompt viewers interacted with more course activities (e.g., all discussion prompts and assessments), indicating a higher level of involvement in their learning. As mentioned in the introduction, MOOCs make use of the affordances of the Internet and technology to structure and deliver instructions. Under such conditions, learners are required to make certain learning-related decisions (e.g., what to learn). If learners are not equipped with the necessary (self-regulated) learning skills and if the learning environment does not provide sufficient support, learners may not be able to achieve learning success. Therefore, it is important for future studies to explore the use of learning analytics and build on current methodology to examine learners' behavior in SRL-supported MOOCs. By doing so, we will be able to gain a more thorough understanding of how learners can be better supported to self-regulate their learning in MOOCs.

To conclude, this study adds to the field of SRL research by exploring the use of a sequential pattern mining algorithm, *cSPADE*, to identify sequential patterns of learner activities in an SRL-supported MOOC. The current analysis provides useful insights into learners' learning behavior and regulation of learning in relation to SRL support. Future work should continue to build on these findings to advance our understanding of supporting SRL in MOOCs.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgments

The author's (Jacqueline Wong) research is supported by the Leiden-Delft-Erasmus Centre for Education and Learning. The authors would like to thank Prof. Dr. Jeroen Jansz, Dr. Mijke Slot, Drs. Kris Stabel, and the Coursera Partner Support Team for supporting the implementation of the study. The authors express gratitude to the reviewers whose insightful comments and suggestions helped improve this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2019.103595.

References

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. https://doi.org/10.1007/s11409-013-9107-6.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7*(2), 161–186. https://doi.org/10.1016/S0959-4752(96)00015-1.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007.

Chen, B., Resendes, M., Chai, C. S., & Hong, H. Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments, 25*(2), 162–175. https://doi.org/10.1080/10494820.2016.1276081.

Davis, D., Chen, G., Van der Zee, T., Hauff, C., & Houben, G. J. (2016). Retrieval practice and study planning in moocs: Exploring classroom-based self-regulated learning strategies at scale. *European conference on technology enhanced learning* (pp. 57–71). . https://doi.org/10.1007/978-3-319-45153-4_5.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8.

Desai, N. A., & Ganatra, A. (2015). Buying scenario and recommendation of purchase by constraint based sequential pattern mining from time stamp based sequential dataset. *Procedia Computer Science, 45*, 166–175. https://doi.org/10.1016/j.procs.2015.03.113.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x.

Eriksson, T., Adawi, T., & Stöhr, C. (2017). “Time is the bottleneck”: A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education, 29*(1), 133–146. https://doi.org/10.1007/s12528-016-9127-8.

Ferguson, R., Clow, D., Beale, R., Cooper, A. J., Morris, N., Bayne, S., et al. (2015). Moving through MOOCS: Pedagogy, learning design and patterns of engagement. *Proceedings of the 10th European conference on technology enhanced learning (EC-TEL '15)* (pp. 70–84). . https://doi.org/10.1007/978-3-319-24258-3_6.

Jung, E., Kim, D., Yoon, M., Park, S., & Oakley, B. (2019). The influence of instructional design on learner control, sense of achievement, and perceived effectiveness in a supersize MOOC course. *Computers & Education, 128*, 377–388. https://doi.org/10.1016/j.compedu.2018.10.001.

Khalil, M., & Ebner, M. (2016). What massive open online course (MOOC) stakeholders can learn from learning analytics? *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy, 1–30*. Cham: Springer International Publishing https://doi.org/10.1007/978-3-319-17727-4_3-1.

Khalil, M., Kastl, C., & Ebner, M. (2016). Portraying MOOCs learners: A clustering experience using learning analytics. *Proceedings of the European Stakeholder Summit on experiences and best practices in and around MOOCs (EMOOCS 2016)* (pp. 265–278). .

Kizilcec, R. F., & Cohen, G. L. (2017). Eight-minute self-regulation intervention raises educational attainment at scale in individualist but not collectivist cultures. *Proceedings of the National Academy of Sciences, 114*(17), 4348–4353. https://doi.org/10.1073/pnas.1611898114.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Recommending self-regulated learning strategies does not improve performance in a MOOC. *Proceedings of the third ACM conference on learning@ scale* (pp. 101–104). . https://doi.org/10.1145/2876034.2893378.

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. https://doi.org/10.1016/j.compedu.2016.10.001.

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

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., de Vries, P., Hatala, M., ... Gašević, D. (2019). Examining communities of inquiry in Massive Open Online Courses: The role of study strategies. *The Internet and Higher Education, 40*, 20–43. https://doi.org/10.1016/j.iheduc.2018.09.001.

Lackner, E., Khalil, M., & Ebner, M. (2016). How to foster forum discussions within MOOCs: A case study. *International Journal of Academic Research in Education, 2*(2), 1–13. https://doi.org/10.17985/ijare.31432.

Lee, D., Watson, S. L., & Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australasian Journal of Educational Technology, 35*(1), 28–41. https://doi.org/10.14742/ajet.3749.

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. https://doi.org/10.1016/j.iheduc.2015.12.003.

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning, 14*(3), 202–227. https://doi.org/10.19173/irrodl.v14i3.1455.

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2017). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior, 80*, 179–196. https://doi.org/10.1016/j.chb.2017.11.011.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*(422), 1–28. https://doi.org/10.3389/fpsyg.2017.00422.

Poon, L. K., Kong, S. C., Wong, M. Y., & Yau, T. S. (2017). Mining sequential patterns of students' access on learning management system. *International conference on data mining and big data* (pp. 191–199). . https://doi.org/10.1007/978-3-319-61845-6_20.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269–286. https://doi.org/10.1080/00313830120074206.

Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 4–8). ACM. https://doi.org/10.1145/2330601.2330605.

Slimani, T., & Lazzez, A. (2013). Sequential mining: Patterns and algorithms. *International Journal of Computer and Electronics Research, 2*(5), 639–647.

Wickham, H. (2014). Tidy data. *Journal of Statistical Software, 59*(10), 1–23. https://doi.org/10.18637/jss.v059.i10.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*(4), 267–276. https://doi.org/10.1080/00461520.2010.517150.

Wong, J., Baars, M., Davis, D., Van der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human-Computer Interaction, 35*, 356–373. https://doi.org/10.1080/10447318.2018.1543084.

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Machine Learning, 42*(1), 31–60. https://doi.org/10.1023/A:1007652502315.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202. https://doi.org/10.1007/s12564-016-9426-9.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.). *The educational psychology series. Handbook of metacognition in education* (pp. 299–315). New York, NY: Routledge.