Computers in Human Behavior 86 (2018) 174–180

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

# Effect of uninterrupted time-on-task on students' success in Massive Open Online Courses (MOOCs)

Youngjin Lee

*University of Kansas, 1122 W. Campus Rd. (Room 413) Lawrence, KS 66045, USA*

ARTICLE INFO

*Article history:*
Available online 23 April 2018

*Keywords:*
Massive open online course (MOOC)
Log file analysis
Academic success
Self-regulated learning (SRL)
Time-on-task

ABSTRACT

This study investigated the relationship between uninterrupted time-on-task and academic success of students enrolled in a Massive Open Online Course (MOOC). The variables representing uninterrupted time-on-task, such as number and duration of uninterrupted consecutive learning activities, were mined from the log files capturing how 4286 students tried to learn Newtonian mechanics concepts in a MOOC. These variables were used as predictors in the logistic regression model estimating the likelihood of students getting a course certificate at the end of the semester. The analysis results indicate that the predictive power of the logistic regression model, which was assessed by Area Under the Precision-Recall Curve (AUPRC), depends on the value of off-task activity threshold time, and the likelihood of students getting a course certificate increases as students were doing more uninterrupted learning activities over a longer period of time. The findings from this study suggest that a simple count of learning activities, which has been used as a proxy for time-on-task in previous studies, may not accurately describe student learning in the computer-based learning environment because it does not reflect the quality, such as uninterrupted durations, of those learning activities.



Since Carroll (1963) proposed a model of school learning, time-on-task, the amount of time students are actively engaged in learning, has been considered an important variable that can explain academic success of students. Earlier studies conducted with undergraduate students in traditional face-to-face courses found that self-reported days of study per week (Allen, Lerner, & Hinrichsen, 1972) and self-reported hours of study per week (Wagstaff & Mahmoudi, 1976) were positively correlated with the GPA of students. Similarly, Wagner, Schober, and Spiel (2008) reported that secondary school students who spent more time working on homework obtained a better grade.

Although self-reported survey is the most common form of data collection method employed in the research investigating time-on-task, it inherently contains errors because students report their time-on-task after they completed the learning tasks, instead of while performing them. Creating a journal of time-on-task while performing the learning task can improve the accuracy of self-reported time-on-task to some extent. This approach, however, may generate a different type of error because creating a journal entry disrupts the natural flow of learning processes of students. To address these issues, researchers started analyzing log files of computer-based learning environments when they need to estimate time-on-task of students. Since computer-based learning environments, such as Massive Open Online Courses (MOOCs), can capture learning behaviors of students in detail without interrupting their learning processes, it is relatively easy to estimate the time-on-task of students. Wellman and Marcinkiewicz (2004) used a frequency of accessing Web pages of an online course as a proxy for the time-on-task of college students, and found that it was positively correlated with the achievement of students measured by pre- and post-tests. Similarly, Cho and Shen (2013) reported that the amount of time spent in the Learning Management System (LMS) is positively correlated with the total points students earned in the course although they did not explain how the LMS computed the time-on-task values they analyzed in their study.

Even though log file analysis allows researchers to better estimate the time-on-task of students in computer-based learning environments, it also presents a new challenge. If students stop using the learning environment, engage in an alternative task for a while, and return to what they were doing, the log file would not be able to recognize this off-task activity because the learning session in the system is preserved as long as the Web browser window remains open. In order to address this issue, time-on-task values

*E-mail address:* yjlee@ku.edu.

https://doi.org/10.1016/j.chb.2018.04.043
0747-5632/© 2018 Elsevier Ltd. All rights reserved.

longer than a pre-determined threshold, which typically ranges from 10 to 60 min, are often discarded from the analysis (Ba-Omar & Petrounias, 2007; del Valle & Duffy, 2007; Munk & Drlík, 2011; Wise, Speer, Marbouti, & Hsiao, 2012). Although Kovanović et al. (2015)'s experiments show that the threshold value for determining off-task activities has an impact on the subsequent data analysis, many studies analyzing log files of computer-based learning environments did not take into account off-task activities when they examined the relationship between time-on-task and students' academic success. In addition, most studies focusing on time-on-task were descriptive in nature, and did not investigate whether time-on-task can predict the academic success of students.

The study reported in this paper analyzed the log files of a MOOC to infer uninterrupted time-on-task of students by excluding off-task activities from the analysis, and examined the effect of uninterrupted time-on-task on the academic success of students in the MOOC. More specifically, this study has two research questions: (1) Does the threshold value determining off-task activities have an impact on the predictive power of the model estimating the likelihood of students getting a course certificate in the MOOC?; and (2) What is the relationship between uninterrupted time-on-task inferred from the log files of a MOOC and students' success in getting a course certificate? The rest of the paper is organized as follows. After reviewing related work, the method section describes the MOOC and its log files analyzed in this study, and explains in detail how the predictive models estimating the likelihood of students getting a course certificate were developed. The results section compares the predictive power of the developed models, and examines the relationship between uninterrupted time-on-task and the likelihood of students getting a course certificate, followed by discussion.

## 1. Related work

### 1.1. Completion rate and attrition in MOOCs

Completion rate and attrition in MOOCs have been extensively studied in recent years. Breslow et al. (2013) examined how 154,763 students who signed up for a physics MOOC in spring 2012. They found that 15% of registrants tried to complete the first homework assignment, 6% of them passed the midterm exam, and only 5% of them earned the course certificate at the end of the semester. Similarly, Ho et al. (2014) reported that about 5% of registrants were able to earn the course certificate when they examined the completion rate of students enrolled in seventeen HarvardX and MITx MOOCs between fall 2012 and summer 2013. Jordan (2015) investigated the completion rates of more than hundred MOOCs having a different grading scheme and varying length of study. She found that the completion rate of MOOCs varied widely (from 0.7% to 52.1%), depending on the length of study (higher completion rates for shorter MOOCs), start date (higher completion rates for recent MOOCs), and assessment type (higher completion rates when MOOCs adopt an automatic grading scheme). When Goldberg et al. (2015) investigated the relationship between participants' level of education and engagement in their completion of a MOOC on dementia, they found that 38% of registrants were able to complete the course, and the discussion board activity was a significant predictor for their course completion. Crossley, Paquette, Dascalu, McNamara, and Baker (2016) examined whether students' online activity and the language they produce in the discussion forum can predict the course completion. Of 320 students who completed at least one graded assignment and produced at least 50 words in discussion forum, 187 students were able to complete the course successfully (completion rate = 58%). In Hone and SaidEl (2016)'s survey study of 379 participants, 32% of survey respondents were able to complete an entire course, and there was no significant difference in the course completion rate by gender, level of study (undergraduate or postgraduate) or MOOC platform used by students. In Pursel, Zhang, Jablokow, Choi, and Velegol (2017)'s study, 5.6% of students who registered for a MOOC on creativity and innovation were able to complete the course with a statement of accomplishment. They also found that registering after the course launch date is negatively correlated with the course completion while the frequency of watching a video, posting a message and providing a comment are positively correlated with the course completion.

Several studies indicate that study time or time management is the most common reason for disengaging from MOOCs. Nawrot and Doucet (2014) reported that about 70% of the survey respondents ($N = 508$) indicated that bad time management (e.g., bad time organization, conflicting real life responsibilities, too much time consuming course, left behind due to illness or work) was the main reason for their MOOC withdrawal decision. Kizilcec and Halawa (2015) analyzed the survey responses of 1968 students who were sampled from twenty MOOCs to investigate reasons for attrition. Their analysis revealed four reasons for disengaging: time issues, course difficulty, format and content, and goals and expectations (in the order of significance). 84% of survey respondents mentioned that they did not have enough time for the course, and half of those respondents also indicated that they are easily distracted from the course. Zheng, Rosson, Shih, and Carroll (2015)'s interview with 18 students taking a MOOC revealed that high workload, challenging course content, lack of time, lack of pressure, no sense of community, social influence and lengthy course start-up were the factors relevant to course drop-out. Similarly, Eriksson, Adawi, and Stöhr (2016) found that the learner's ability to find and manage time was the most frequently mentioned reason for course drop-out when they interviewed 34 students selected from two MOOCs with a different completion rate. Other reasons for disengagement include the learner's perception of course content, the learner's perception of course design, and the learner's social situation and characteristics.

### 1.2. Effect of time-on-task on academic performance in computer-based learning environments

In the last decade, researchers studied the effect of time-on-task on academic performance of students learning in the computer-based learning environment. Lustria (2007) found that college students who spent more time in using interactive Web sites containing health related information performed better in the comprehension test. Louw, Muller, and Tredoux (2008) examined the importance of various predictors such as pre-existing level of mathematics ability, degree of access to computers outside of school, time spent on computers both inside and outside of school and on the computer-based tutoring system used in the study, computer literacy, confidence in using information technology, motivation for mathematics, degree of enjoyment of learning mathematics, intention to study after school, language used at home and parental encouragement. Of these predictors, they found that time spent on the computer-based tutoring environment was the strongest variable predicting the academic success of high school students participating in their study. When Krause, Stark, and Mandl (2009) investigated how 137 college students learned statistics in the computer-based learning environment, they found that the time-on-task is significantly positively correlated with the learning outcome of students. Macfadyen and Dawson (2010)'s log file analysis of an LMS showed that number of log-ins and time spent in the LMS can explain more than 30% of variance in the final grade of college students. Cho and Shen (2013) reported that time-on-task logged in the LMS, along with effort regulation, can predict students' academic achievement in the course. Goldhammer et al. (2014) investigated how the effect of time-on-task on the learning performance in the computer-based learning environment is moderated by task difficulty and student skill using linear mixed models. When students were solving problems in the computer-based learning environment, time-on-task increased with task difficulty, and was positively related to the performance of students. However, the positive effect of time-on-task decreased as skill levels of students is increasing. Landers and Landers (2015) studied the effect of time-on-task in the gamified learning environment. They reported that college students who learned in the gamified learning environment spent much more time using the learning environment than students who used non-gamified learning environment, which in turn improved their academic performance in the course. Although these studies reported time-on-task as an important predictor for academic success of students learning in the computer-based learning environment, none of them took into account off-task activities of students, which may have an impact on the predictive power of time-on-task.

## 2. Method

### 2.1. Context

This study analyzed the log files that captured how 12,981 students who signed up for a MOOC called "edX 8.MReVx Mechanics Review (hereafter, 8.MReVx)" interacted with various learning resources during the Summer 2014 semester. 83% of 12,981 registrants were male and 17% of them were female. Age of registrants varied from 15 to 75: 45% of them were under 26, 39% of them were in the range of 26 and 40, and 16% of them were age 41 and above. 25% of registrants had an advanced degree, 35% of them had a college degree, and 38% of them had a secondary diploma or less. Geographic distribution included the US (27% of registrants), India (18%), UK (3.6%), Brazil (2.8%) and others.

8.MReVx, which was offered by MIT, is designed to provide a comprehensive overview of Newtonian mechanics and greater expertise with problem-solving. 8.MReVx provides various learning resources, such as e-texts, videos, discussion boards, wiki, checkpoints, weekly homework problems, quizzes, midterm exam and final exam, to help students learn Newtonian mechanics concepts. Students did not need to use an external resource, such as textbook, because the course was designed to be self-contained; it provided all required information through its e-texts and wiki. Checkpoints are easier problems embedded within e-texts as formative assessments whereas homework problems, quizzes, midterm exam and final exam are more difficult problems provided throughout the 12-week long semester as summative assessments. Students were given a week to work on formative and summative assessment problems which were due on Sunday at midnight. The achievement of students was determined by checkpoints (8%), homework problems (34%), quizzes (36%), midterm exam (7%) and final exam (16%), and students who scored more than 60% of the maximum possible points received a course certificate. For further explorations of course structure and available learning resources, see the archived course at https://courses.edx.org/courses/MITx/8.MReVx/2T2014/info.

### 2.2. Procedures

Of 12,981 students who registered in 8.MReVx, this study focused on students who solved at least one assessment problem ($N = 4286$) because problem solving is the most important learning activity for earning a certificate in this course; in 8.MReVx, all available points were allocated to solving formative and summative assessment problems. Of these 4286 students who solved at least one assessment problem, only 434 students earned a course certificate at the end of the semester.

Fig. 1 shows a snippet of log files analyzed in this study. Each row is a timestamped database transaction representing a specific learning activity experienced by one particular student enrolled in the MOOC. One can easily see that the student solved a problem ([check_problem_correct] in row 14 in Fig. 1), watched the same video twice ([play_video] in row 15 and 17 in Fig. 1), and then solved another problem ([check_problem correct] in row 19 in Fig. 1). In total, there were 12,981 log files like the one shown in Fig. 1 because each log file captured what one student did while taking the course. These log files were imported into an SQLite (https://www.sqlite.org) database, and Python (https://www.python.org) and R (https://www.r-project.org) scripts were developed to pre-process the merged log files in the imported database.

Pre-processing of the merged log files consists of the following steps. First, examine "Event" and "Resource Name" columns in the database to create a list of unique problems available in the entire course. Second, for each student in the database, examine how many problems he or she tried to solve. Third, select a subset of students who attempted to solve at least one problem. Fourth, compute the time-on-task for all learning activities of selected students (e.g. [play_video] in row 17 in Fig. 1), by subtracting its timestamp value (e.g., 6/24/14 16:05:21.327 in row 17 in Fig. 1) from the timestamp value of the subsequent learning activity (e.g., 6/24/14 16:13:28.982 in row 18 in Fig. 1). Finally, using the computed time-on-task values, create variables representing the frequency and the duration of uninterrupted time-on-task such as (1) number of learning chunks per week ($N_{LC}$); (2) number of learning activities per week ($N_{LA}$); (3) duration of all learning chunks per week ($T_{LC}$); and (4) median of duration of learning chunks per week ($MedT_{LC}$).

Figure 2 explains how these variables are defined and computed in this study. Each row in Fig. 2 represents a specific learning activity, such as solving a problem or watching a video, one particular MOOC student experienced during one week period. A learning chunk is defined as a group of consecutive learning activities not interrupted by an off-task activity, an activity whose time-on-task value is larger than a predetermined threshold value (e.g., 10, 30 or 60 min). Fig. 2 indicates that the student had three learning chunks this week ($N_{LC} = 3$) that are separated by two off-task activities. Number of learning activities per week is the count of individual learning activity whose time-on-task value is smaller than the predetermined off-task activity threshold time ($N_{LA} = 10$ in Fig. 2). Duration of all learning chunks can be obtained by summing up the duration of individual learning chunk observed in the week (i.e., $T_{LC} = T_{LC,1} + T_{LC,2} + T_{LC,3}$; $T_{LC,1} = TOT_1 + TOT_2 + TOT_3$;

| | A | B | C | D |
|---|---|---|---|---|
| 1 | Timestamp | Event | Resource Name | Display Name |
| 2 | 6/24/14 15:47:52.603 | [goto_page] | [info] | [info] |
| 3 | 6/24/14 15:48:17.947 | [goto_page] | [courseware] | [courseware] |
| 4 | 6/24/14 15:48:19.519 | [goto_page] | [HWK_3] | [HWK_3] |
| 5 | 6/24/14 15:48:25.259 | [goto_page] | [general_forum] | [general_forum] |
| 6 | 6/24/14 15:48:57.918 | [read_discussion] | [general_forum] | [TODO] |
| 7 | 6/24/14 15:57:26.043 | [goto_page] | [courseware] | [courseware] |
| 8 | 6/24/14 15:57:27.977 | [goto_page] | [HWK_3] | [HWK_3] |
| 9 | 6/24/14 15:57:47.130 | [goto_page] | [Pulley_Acceleration] | [Pulley_Acceleration] |
| 10 | 6/24/14 15:57:51.771 | [problem_get] | [Pulley_Acceleration_Part_2] | [Pulley_Acceleration] |
| 11 | 6/24/14 15:57:51.772 | [problem_get] | [Pulley_Acceleration] | [Pulley_Acceleration] |
| 12 | 6/24/14 15:57:51.931 | [problem_get] | [Pulley_Acceleration_Part_3] | [Pulley_Acceleration] |
| 13 | 6/24/14 15:57:51.977 | [problem_get] | [Pulley_Acceleration_Part_4] | [Pulley_Acceleration] |
| 14 | 6/24/14 15:59:13.091 | [check_problem_correct] | [Pulley_Acceleration_2_1] | [Pulley_Acceleration] |
| 15 | 6/24/14 16:00:25.198 | [play_video] | [3ec6b4bf400647e392e3a0721b0d0bec] | [Pulley_Acceleration] |
| 16 | 6/24/14 16:02:34.980 | [pause_video] | [3ec6b4bf400647e392e3a0721b0d0bec] | [Pulley_Acceleration] |
| 17 | 6/24/14 16:05:21.327 | [play_video] | [3ec6b4bf400647e392e3a0721b0d0bec] | [Pulley_Acceleration] |
| 18 | 6/24/14 16:13:28.982 | [pause_video] | [3ec6b4bf400647e392e3a0721b0d0bec] | [Pulley_Acceleration] |
| 19 | 6/24/14 16:14:19.478 | [check_problem_correct] | [Pulley_Acceleration_Part_2_2_1] | [Pulley_Acceleration] |

**Fig. 1.** An example of 8.MReVx log files analyzed in the study.

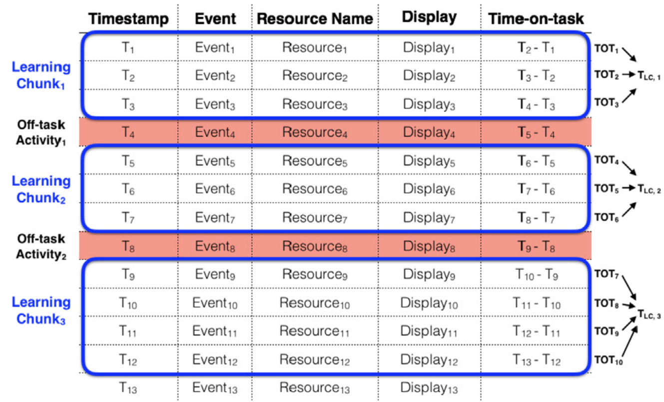


**Fig. 2.** Learning chunk ($LC_i$), off-task activity, time-on-task ($TOT_i$), and duration of learning chunk ($T_{LC,i}$) identified in the log file.

$T_{LC,2} = TOT_4 + TOT_5 + TOT_6$; $T_{LC,3} = TOT_7 + TOT_8 + TOT_9 + TOT_{10}$). Finally, median of duration of learning chunks ($MedT_{LC} = Median(\{T_{LC,1}, T_{LC,2}, T_{LC,3}\})$ captures, on average, how long students were engaged in meaningful learning activities before getting distracted by an off-task action.

Log transform was applied to these variables in order to mitigate heavy skewness and incompatible ranges. Table 1 summarizes descriptive statistics of unscaled and log-transformed variables pre-processed with different off-task activity threshold values.

### 2.3. Predictive models estimating likelihood of students getting course certificate

In order to answer the research questions, three logistic regression models (base, main effect and interaction) were developed. The response variable of logistic regression model was whether or not students earned a course certificate at the end of the semester. The base model predicts a probability for getting a course certificate based on two explanatory variables, number of learning chunks per week ($N_{LC}$) and number of learning activities per week ($N_{LA}$). The base model uses frequencies of uninterrupted learning activities as a proxy for the time-on-task of MOOC students as was done in the previous study. The main effect model includes two additional duration-based explanatory variables, duration of all learning chunks per week ($T_{LC}$) and median of duration of learning chunks per week ($MedT_{LC}$), in order to investigate whether the duration-based time-on-task variables proposed in this study can improve the predictive power of logistic regression model. Finally, the interaction model adds interaction terms to the main effect model in order to examine the importance of interactions between main effect explanatory variables. To investigate the effect of off-task threshold value on the predictive power of logistic regression model, each model was fit to data sets pre-processed with three different off-task activity threshold values frequently used in educational data mining research (10, 30 and 60 min). Table 2 summarizes the resulting nine predictive models compared in this study in terms of the off-task threshold value and the explanatory variables.

**Table 1**
Mean, median, standard deviation and interquartile range of unscaled and log-transformed variables representing the frequency and the duration of uninterrupted time-on-task.

| Variable | Mean | Median | SD | IQR |
|---|---|---|---|---|
| Off-task threshold = 10 min | | | | |
| $N_{LC}$ | 8.75 | 5.00 | 10.13 | 9.00 |
| $N_{LA}$ | 112.48 | 66.00 | 121.73 | 135.25 |
| $T_{LC}$ | 6826.12 | 3626.35 | 7947.68 | 8041.12 |
| $MedT_{LC}$ | 658.48 | 491.57 | 618.97 | 543.15 |
| $\log(N_{LC})$ | 0.82 | 0.78 | 0.37 | 0.60 |
| $\log(N_{LA})$ | 1.83 | 1.83 | 0.46 | 0.75 |
| $\log(T_{LC})$ | 3.55 | 3.56 | 0.54 | 0.79 |
| $\log(MedT_{LC})$ | 6.12 | 6.20 | 0.93 | 1.08 |
| Off-task threshold = 30 min | | | | |
| $N_{LC}$ | 5.40 | 3.50 | 5.51 | 5.00 |
| $N_{LA}$ | 116.35 | 67.94 | 126.34 | 137.62 |
| $T_{LC}$ | 10638.27 | 5265.48 | 13310.41 | 12161.47 |
| $MedT_{LC}$ | 1501.39 | 1159.74 | 1371.65 | 1461.46 |
| $\log(N_{LC})$ | 0.69 | 0.65 | 0.31 | 0.43 |
| $\log(N_{LA})$ | 1.84 | 1.84 | 0.46 | 0.75 |
| $\log(T_{LC})$ | 3.70 | 3.72 | 0.59 | 0.83 |
| $\log(MedT_{LC})$ | 6.88 | 7.06 | 1.08 | 1.28 |
| Off-task threshold = 60 min | | | | |
| $N_{LC}$ | 4.53 | 3.00 | 4.24 | 4.00 |
| $N_{LA}$ | 117.30 | 68.00 | 127.53 | 138.75 |
| $T_{LC}$ | 13134.35 | 6316.93 | 17047.45 | 14899.79 |
| $MedT_{LC}$ | 2029.12 | 1583.05 | 1914.13 | 2152.61 |
| $\log(N_{LC})$ | 0.65 | 0.60 | 0.28 | 0.37 |
| $\log(N_{LA})$ | 1.85 | 1.84 | 0.46 | 0.75 |
| $\log(T_{LC})$ | 3.77 | 3.80 | 0.61 | 0.86 |
| $\log(MedT_{LC})$ | 7.18 | 7.37 | 1.14 | 1.37 |

## 3. Results

### 3.1. Effect of off-task threshold value on predictive power of logistic regression model

In order to investigate whether an off-task activity threshold value has an impact on the predictive power of logistic regression model, data sets pre-processed with three different off-task activity threshold values (10, 30 and 60 min) were divided into training and test sets, each logistic regression model (base, main effect and interaction) was fit to the three training sets, and the predictive power of each model was evaluated on the three test sets. When creating training and test sets, which consist of 80% and 20% of the original data sets, stratified random sampling was used to ensure that the ratio of positive (students who earned a course certificate) to negative instances (students who did not earn a course certificate) in both sets is similar. Since the test sets were not used in building logistic regression models, which were fit only to the training sets, they can play a role of future unseen data providing an unbiased measure of predictive power of the model.

Accuracy and Area Under the receiver operating characteristic Curve (AUC) are two most frequently used performance metrics when comparing predictive power of binary classification models. However, accuracy and AUC are not appropriate performance measurements for the data set analyzed in this study because it is heavily imbalanced to negative instances. Because only 434 students, out of 4286 students who solved at least one problem, got a course certificate, the proportion of negative instance, students who did not earn a course certificate, is approximately 0.90 in both training and test sets. Therefore, a simple model that always predicts everyone will not get a course certificate will achieve a 90% accuracy. Despite the high accuracy, however, such a model is useless because it does not provide any meaningful information about who will get a course certificate. Likewise, AUC is not an appropriate performance measure because it would give an overly optimistic result based on the high percentage of correct classifications of the majority class. In order to address these issues, this study used Area Under the Precision-Recall Curve (AUPRC) (Davis & Goadrich, 2006) when comparing the predictive power of logistic regression model.

Table 3 compares AUC and AUPRC of nine logistic regression models tested in this study. As explained above, AUC values do not change much because of the large number of negative instances in

**Table 2**
Nine logistic regression models compared in the study.

| Model | Type | Off-task Threshold | Explanatory Variables |
|---|---|---|---|
| $M_1$ | Base | 10 min | $\log(N_{LC})$, $\log(N_{LA})$ |
| $M_2$ | Base | 30 min | $\log(N_{LC})$, $\log(N_{LA})$ |
| $M_3$ | Base | 60 min | $\log(N_{LC})$, $\log(N_{LA})$ |
| $M_4$ | Main effect | 10 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$ |
| $M_5$ | Main effect | 30 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$ |
| $M_6$ | Main effect | 60 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$ |
| $M_7$ | Interaction | 10 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$, $\log(N_{LC})$:$\log(N_{LA})$, $\log(N_{LC})$:$\log(T_{LC})$, $\log(N_{LC})$:$\log(MedT_{LC})$ |
| $M_8$ | Interaction | 30 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$, $\log(N_{LC})$:$\log(N_{LA})$, $\log(N_{LC})$:$\log(T_{LC})$, $\log(N_{LC})$:$\log(MedT_{LC})$ |
| $M_9$ | Interaction | 60 min | $\log(N_{LC})$, $\log(N_{LA})$, $\log(T_{LC})$, $\log(MedT_{LC})$, $\log(N_{LC})$:$\log(N_{LA})$, $\log(N_{LC})$:$\log(T_{LC})$, $\log(N_{LC})$:$\log(MedT_{LC})$ |

*Note*. ":" in the explanatory variable name indicates an interaction between the corresponding main effect variables.

the data. All nine logistic regression models show a very similar AUC value ranging from 0.922 to 0.938 regardless of the off-task activity threshold value or the explanatory variables included in the model. On the other hand, AUPRC values are comparable only in the same model type (base, main effect and interaction). In addition, prediction models including more explanatory variables achieve a larger AUPRC value. On average, AUPRC of the main effect models is about 32% larger than that of the base models, indicating that the two duration-related time-on-task variables proposed in this study ($T_{LC}$ and $MedT_{LC}$) are important in discriminating students who earned a course certificate from students who did not. On the other hand, the difference between main effect and interaction models is smaller. On average, AUPRC of the interaction models is approximately 4% larger than that of the main effect models.

### 3.2. *Relationship between uninterrupted time-on-task and students' success in MOOC*

This study elects to examine the importance of explanatory variables in the interaction model with 60-min threshold ($M_9$) in relation to the student's success in getting a course certificate because this model achieved the largest AUPRC value. The results from logistic regression analysis suggest that the more learning activities ($N_{LA}$) and learning chunks ($N_{LC}$) students have and the longer their average learning chunk ($MedT_{LC}$) is, the more likely they will earn a course certificate; when $\log(N_{LA})$, $\log(N_{LC})$ and $\log(MedT_{LC})$ increase by one standard deviation, the log odds of students getting a course certificate increase by 2.27, 1.37 and 0.62, respectively (see Table 4).

Of three interactions, only the interaction between number of learning chunks per week and number of learning activities per week, $\log(N_{LC})$:$\log(N_{LA})$, is statistically significant. The large, negative value of the interaction coefficient indicates that students are less likely to get a course certificate if they are engaged in many learning activities in a large number of learning chunks. Rather, their chance of getting a course certificate will increase when they are engaged in many learning activities in fewer learning chunks. This is a sensible result because number of learning chunks is also proportional to the frequency of off-task activities and students are less likely to be engaged in meaningful learning when they are frequently distracted by other tasks.

Of three interactions, only the interaction between number of learning chunks per week and number of learning activities per week, $\log(N_{LC})$:$\log(N_{LA})$, is statistically significant. The large, negative value of the interaction coefficient indicates that students are less likely to get a course certificate if they are engaged in many learning activities in a large number of learning chunks. Rather, their chance of getting a course certificate will increase when they are engaged in many learning activities in fewer learning chunks. This is a sensible result because number of learning chunks is also proportional to the frequency of off-task activities and students are less likely to be engaged in meaningful learning when they are frequently distracted by other tasks.

**Table 3**
AUC and AUPRC of nine logistic regression models compared in this study.

| Model | AUC | AUPRC |
|---|---|---|
| $M_1$ | 0.925 | 0.547 |
| $M_2$ | 0.923 | 0.522 |
| $M_3$ | 0.922 | 0.517 |
| $M_4$ | 0.935 | 0.701 |
| $M_5$ | 0.931 | 0.696 |
| $M_6$ | 0.932 | 0.692 |
| $M_7$ | 0.936 | 0.704 |
| $M_8$ | 0.934 | 0.731 |
| $M_9$ | 0.938 | 0.739 |

**Table 4**
Summary of logistic regression analysis for variables predicting students' success in getting a course certificate.

| Explanatory Variable | Estimate | Standard Error | *p*-value |
|---|---|---|---|
| $\log(N_{LC})$ | 1.37 | 0.33 | $<0.0001^*$ |
| $\log(N_{LA})$ | 2.27 | 0.32 | $<0.0001^*$ |
| $\log(T_{LC})$ | 0.78 | 0.44 | 0.078 |
| $\log(MedT_{LC})$ | 0.62 | 0.24 | $0.011^*$ |
| $\log(N_{LC})$:$\log(N_{LA})$ | −1.45 | 0.25 | $<0.0001^*$ |
| $\log(N_{LC})$:$\log(T_{LC})$ | 0.30 | 0.24 | 0.21 |
| $\log(N_{LC})$:$\log(MedT_{LC})$ | 0.24 | 0.16 | 0.12 |

*Note*. $^*p < .05$; ":" in the explanatory variable name indicates an interaction between the corresponding main effect variables.

## 4. Discussion and future works

Uninterrupted time-on-task is closely related to Self-Regulated Learning (SRL) because it involves allocating time and effort to improve learning performance. Students with higher SRL ability would have longer time-on-task that is uninterrupted by off-task activities than students who do not possess enough SRL ability. Previous studies found that SRL is important for students to be successful especially in the computer-based learning environment where there is no instructor or peers who can guide and support their learning processes (Puzziferro, 2008; Sun & Rueda, 2011). These studies used survey instruments, such as Motivated Strategies for Learning Questionnaire (MSLQ) (Duncan & McKeachie, 2005), which include questions measuring SRL ability of students (e.g., I make good use of study time for this course). However, self-reported survey instruments may not be able to estimate uninterrupted time-on-task accurately because they are relying on students' perception of their self-regulatory processes aggregated over more than one learning activity (Zimmerman, 2008). One contribution of this study is that it operationalized SRL ability of students in terms of *observable* learning behaviors and examined how these learning behaviors are correlated with the academic performance of students by analyzing log files of a computer-based learning environment.

This study found that students who did more learning activities and had more and longer learning chunks per week were more likely to get a course certificate, which is not a surprising result. However, what is interesting is the interaction between number of learning activities and number of learning chunks per week; the likelihood of getting a course certificate increases when the same number of learning activities occurred in *fewer* learning chunks. Since most LMS provide statistics on how many times students accessed a specific learning resource in the course, number of learning activities students are engaged in (e.g., number of clicks on the Web page containing a specific instructional material) is often used as a proxy for their time-on-task. The findings from this study suggest that frequency of learning activities alone does not provide enough information about how students self-regulate their learning and their time-on-task. Rather, it is important to examine how learning activities are grouped to form more meaningful learning experiences as students are interacting with learning resources in the computer-based learning environment.

In most MOOCs (and other learning environments), important learning activities have a due date and time. For instance, students taking 8.MReVx were required to complete all assigned learning activities by Sunday at midnight each week. A recent study that examined a problem completion rate of students enrolled in a MOOC found that students who were able to successfully complete the course started working on their weekly homework problems very early (Lee, Y, 2018), suggesting that uninterrupted time-on-task of students may change over the one-week assignment cycle. Thus, as a future work, it would be meaningful to investigate whether incorporating daily uninterrupted time-on-task into the prediction model can improve its predictive power.

This study focused on the relationship between the amount of uninterrupted time-on-task and students' success in acquiring a course certificate in the MOOC. Although getting a course certificate is important, especially from the perspective of MOOC instructors and providers, not all people take MOOCs to earn a certificate (Kizilcec, Piech, & Schneider, 2013). Therefore, it would be important to investigate how uninterrupted time-on-task is related to other variables of success, such as participation in discussion or problem solving in subsequent weeks. Similarly, it would be interesting to compare uninterrupted time-on-task of students having a different intention of enrollment.

In this study, the variables representing uninterrupted time-on-task were averaged over weeks. As a result, these variables do not capture how uninterrupted time-on-task of MOOC students are changing over the course of the semester. As a future work, we plan to extend the statistical model developed in this study into a multilevel model in which regression coefficients vary by week. By comparing logistic regression coefficients from each week, we will be able to better understand how the effect of uninterrupted time-on-task changes over time, and how it is related to the success of students in the MOOC.

## 5. Limitations of study

Learning is a complex cognitive activity that can take many different forms depending on the subject matter. Solving homework and quiz problems on the physics concepts covered each week was the most important learning activity for students enrolled in the MOOC examined in this study. Although problem solving is not limited to mathematics and science, because any higher-order thinking can be considered a problem solving activity (Veresov, 2004, 2010; Vygotsky, 2012), solving mathematics or science problems in the computer-based learning environment is different from applying higher-order thinking skills in other subject domains, such as literature or social studies, emphasizing different pedagogies (e.g., socio-constructivism). First of all, mathematics or science problems students are trying to solve in MOOCs almost always have one correct answer whereas problem solving in other subject domains may have more than one correct answer. Second, while solving mathematics or science problems in MOOCs, students are allowed to submit their answer only a few times, and their incorrect answers are often penalized, which is usually not the case in solving problems in other subject domains. In MOOCs on mathematics or science, moreover, students are expected to solve problems without getting external helps because of the honor code of MOOCs. On the other hand, students enrolled in MOOCs on literature or social studies are expected to develop their understanding by engaging in meaningful discussions with their peers and the instructor. Problem solving in this case must be studied in connection with higher mental functions, such as abstract thinking, logical memory and voluntary attention (Veresov, 2004, 2010; Vygotsky, 2012), by carefully observing the entire learning processes of students. Consequently, the findings from this study may not be generalized to other subject domains in which collaboration with other people are encouraged because different learning activities and pedagogies would affect how students interact with learning resources in MOOCs.

In this study, only 434 out of 12,981 registrants were able to get a course certificate, which falls on the lower end of the range of completion rates Jordan (2015) reported in her study. Since students taking MOOCs with a higher completion rate may show different learning behaviors, the findings from this study may not be applicable to such MOOCs, either. It would be interesting to conduct a similar study examining uninterrupted time-on-task of students who are learning in the MOOC with a higher completion rate.

In determining uninterrupted time-on-task, this study used off-task activity threshold values frequently used in educational data mining research (10, 30 and 60 min). Since the choice of these values is not based on a strong theoretical foundation, care should be taken when interpreting the findings from this study. Depending on the nature of content knowledge being acquired and the pedagogies employed in the course, different off-task activity threshold values could result in an optimal predictive power of the model. More in-depth replication studies in different subject domains are warranted to better understand the effect of off-task activity threshold values on the predictive power of the model incorporating uninterrupted time-on-task as a predictor variable.

## Acknowledgement

The author thanks Prof. David Pritchard at MIT for providing the log files of 8.MReVx analyzed in this study.

## References

Allen, G. J., Lerner, W. M., & Hinrichsen, J. J. (1972). Study behaviors and their relationships to test anxiety and academic performance. *Psychological Reports, 30*(2), 407–410. http://doi.org/http://doi.org/10.2466/pr0.1972.30.2.407.

Ba-Omar, H., & Petrounias, I. (2007). A framework for using Web usage mining to personalise e-learning. In *Proceedings of the seventh IEEE international conference on advanced learning technologies* (pp. 937–938). http://doi.org/http://doi.org/10.1109/ICALT.2007.13.

Breslow, L., Pritchard, D. E., DeBoer, J. D., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research Practice in Assessment, 8*, 13–25.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64*(8), 723–733.

Cho, M.-H., & Shen, D. (2013). Self-regulation in online learning. *Distance Education, 34*(3), 290–301. http://doi.org/10.1080/01587919.2013.835770.

Crossley, S., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 6–14). http://doi.org/10.1145/2883851.2883931.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on world wide Web* (pp. 233–240). http://doi.org/10.1145/1143844.1143874.

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated Strategies for learning Questionnaire. *Educational Psychologist, 40*(2), 117–128. http://doi.org/10.1207/s15326985ep4002_6.

Eriksson, T., Adawi, T., & Stöhr, C. (2016). "Time is the bottleneck": A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education, 29*(1), 133–146. http://doi.org/10.1007/s12528-016-9127-8.

Goldberg, L. R., Bell, E., King, C., OMara, C., McInerney, F., Robinson, A., et al. (2015). Relationship between participants' level of education and engagement in their completion of the understanding dementia massive open online course. *BMC Medical Education, 15*(1), 60–67. http://doi.org/10.1186/s12909-015-0344-z.

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology, 106*(3), 608–626. http://doi.org/10.1037/a0034716.

Hone, K. S., & Said El, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education, 98*(C), 157–168. http://doi.org/10.1016/j.compedu.2016.03.016.

Ho, A. D., Reich, J., Nesterko, S. O., Seaton, D. T., Mullaney, T., Waldo, J., et al. (2014). HarvardX and MITx: The first year of open online courses, fall 2012-summer 2013. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.2381263.

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *The International Review of Research in Open and Distributed Learning, 16*(3). http://doi.org/10.19173/irrodl.v16i3.2112.

Kizilcec, R. F., & Halawa, S. (2015). Attrition and achievement gaps in online learning. In *Proceedings of the second ACM conference on learning @ scale* (pp. 57–66). http://doi.org/10.1145/2724660.2724680.

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). http://doi.org/10.1145/2460296.2460330.

Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R., & Hatala, M. (2015). Does time-on-task estimation matter? Implications for the validity of learning analytics findings. *Journal of Learning Analytics, 2*(3), 81–110. http://doi.org/10.18608/jla.2015.23.6.

Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*(2), 158–170. http://doi.org/10.1016/j.learninstruc.2008.03.003.

Landers, R. N., & Landers, A. K. (2015). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming, 45*(6), 769–785. http://doi.org/10.1177/1046878114563662.

Lee, Y. (2018). Using self-organizing map and clustering to investigate problem-solving patterns in the massive open online course: An exploratory study. *Journal of Educational Computing Research* (in press) https://doi.org/10.1177/0735633117753364.

Louw, J., Muller, J., & Tredoux, C. (2008). Time-on-task, technology and mathematics achievement. *Evaluation and Program Planning, 31*(1), 41–50. http://doi.org/10.1016/j.evalprogplan.2007.11.001.

Lustria, M. L. A. (2007). Can interactivity make a difference? Effects of interactivity on the comprehension of and attitudes toward online health content. *Journal of the American Society for Information Science and Technology, 58*(6), 766–776. http://doi.org/10.1002/asi.20557.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599. http://doi.org/10.1016/j.compedu.2009.09.008.

Munk, M., & Drlík, M. (2011). Impact of different pre-processing tasks on effective identification of users' behavioral patterns in Web-based educational system. *Procedia Computer Science, 4*, 1640–1649. http://doi.org/10.1016/j.procs.2011.04.177.

Nawrot, I., & Doucet, A. (2014). Building engagement for MOOC students. In *Proceedings of the 23rd international conference on world wide Web* (pp. 1077–1082). http://doi.org/10.1145/2567948.2580054.

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2017). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning, 32*(3), 202–217. http://doi.org/10.1111/jcal.12131.

Puzziferro, M. (2008). Online technologies self-efficacy and self-regulated learning as predictors of final grade and satisfaction in college-level online courses. *American Journal of Distance Education, 22*(2), 72–89. http://doi.org/10.1080/08923640802039024.

Sun, J. C.-Y., & Rueda, R. (2011). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology, 43*(2), 191–204. http://doi.org/10.1111/j.1467-8535.2010.01157.x.

del Valle, R., & Duffy, T. M. (2007). Online learning: Learner characteristics and their approaches to managing learning. *Instructional Science, 37*(2), 129–149. http://doi.org/10.1007/s11251-007-9039-0.

Veresov, N. (2004). Zone of proximal development (ZPD): The hidden dimension? In A. Ostern, & R. Helia-Ylikallio (Eds.), *Languages as culture Tensions in time and space* (pp. 13–30).

Veresov, N. (2010). Introducing cultural historical theory: Main concepts and principles of genetic research methodology. *Cultural-Historical Psychology, 2010*(4), 83–90.

Vygotsky, L. (2012). *Thoughts and language*. Cambridge, MA: MIT Press.

Wagner, P., Schober, B., & Spiel, C. (2008). Time students spend working at home for school. *Learning and Instruction, 18*(4), 309–320. http://doi.org/10.1016/j.learninstruc.2007.03.002.

Wagstaff, R., & Mahmoudi, H. (1976). Relation of study behaviors and employment to academic performance. *Psychological Reports, 38*(2), 380–382. http://doi.org/10.2466/pr0.1976.38.2.380.

Wellman, G. S., & Marcinkiewicz, H. (2004). Online learning and time-on-task: Impact of proctored vs. un-proctored testing. *Journal of Asynchronous Learning Networks, 8*, 93–104.

Wise, A. F., Speer, J., Marbouti, F., & Hsiao, Y.-T. (2012). Broadening the notion of participation in online discussions: Examining patterns in learners' online listening behaviors. *Instructional Science, 41*(2), 323–343. http://doi.org/10.1007/s11251-012-9230-9.

Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). Understanding student motivation, behaviors and perceptions in MOOCs. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1882–1895). http://doi.org/10.1145/2675133.2675217.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183. http://doi.org/10.3102/0002831207312909.