An Analysis of Course Characteristics, Learner Characteristics, and Certification Rates in MITx MOOCs

Nergiz Ercil Cagiltay[1], Kursat Cagiltay[2], and Berkan Celik[3]
[1]Atilim University, Ankara, Turkey, [2]Middle East Technical University, Ankara, Turkey, [3]Van Yuzuncu Yil University / Middle East Technical University, Ankara, Turkey

Abstract

Massive Open Online Courses (MOOCs), capable of providing free (or low cost) courses for millions of learners anytime and anywhere, have gained the attention of researchers, educational institutions, and learners worldwide. Even though they provide several benefits, there are still some criticisms of MOOCs. For instance, MOOCs' high dropout rates or predominantly elite participation are considered to be important problems. In order to develop solutions for these problems, a deeper understanding of MOOCs is required. Today, despite the availability of several research studies about MOOCs, there is a shortage of in-depth research on course characteristics, learner characteristics, and predictors of certification rates. This study examined MOOC and learner characteristics in detail and explored the predictors of course certification rates based on data from 122 Massachusetts Institute of Technology MOOCs (MITx) on edX platform as well as data about the 2.8 million participants registered in these MOOCs. The results indicated that as the number of courses offered and the number of learners enrolled increased in years, there was a decrease in the certification rates among enrolled learners. According to our results, the number of average chapters completed, total forum messages, and mean age predicted course certification rates positively. On the other hand, the total number of chapters in a course predicted the course certification rates negatively. Based on these results, shorter and more interactive MOOCs are recommended by considering the needs of the learners, course content design, and strategies encouraging the enrolled students to enter the courses.

Keywords: online learning, massive open online courses, MOOCs, MITx

Introduction

MOOCs have the potential to support traditional education activities both in- and out-of-class, such as homework and exercises, as well as individuals' lifelong learning. There are several benefits of MOOCs, as they are open-access and offered at little or no cost, with thousands of participants able to enroll and earn credits or receive certificates without constraints of space or time (De Barba, Kennedy, & Ainley, 2016; Porter, Graham, Spring, & Welch, 2014). MOOCs serve learners from all over the world, and there is no limit to learners' age, educational level, individual characteristics, or culture. Currently, more than 800 universities worldwide offer MOOCs and the number of these courses exceeds 10,000 (Shah, 2018b). Thus, the number of learners registered in MOOCs is huge compared to traditional courses. According to Shah (2018a) the top five MOOC providers (and number of registrations) are Coursera (37 million), edX (18 million), XuetangX (14 million), Udacity (10 million), and FutureLearn (8.7 million). This massive number of learners comprises people from diverse backgrounds with different motivations (DeBoer, Stump, Seaton, & Breslow, 2013; Kizilcec & Schneider, 2015). Deng, Benckendorff, and Gannaway (2019) have reported that MOOC learners' age distribution is mainly between 25 and 65 years. Even though MOOCs are open to everyone, the majority of learners hold higher education degrees, and most are male (Christensen et al., 2013).

MOOCs offer the possibility of providing free education for everybody; however, they have some limitations in terms of their, effectiveness, and benefits for both learners and educational organizations. Learners' behavior in MOOCs, instructional design of MOOCs, assessment processes, and interactions among learners and instructors are significantly different from traditional educational platforms. For instance, MOOC learners are rarely able to obtain direct and timely feedback from instructors (Kop, Fournier, & Mak, 2011). Furthermore, a study performed on four edX MOOCs reported that on average, certificate earners skipped 22% of the course content and made use of non-linear navigation (Guo & Reinecke, 2014). In the same study, it was also noted that older learners and those from lower learner-teacher ratio countries showed more comprehensive and non-linear navigation. Hence, MOOC learners present different behaviors than do learners in traditional courses. Another significant finding showed that less than 10% of the enrolled learners tended to complete their MOOC (Ho et al., 2015; Jordan, 2014). Providing course content in such a way so as to address these different individual requirements is a great challenge for MOOC instructional designers (Adair et al., 2014; Beaven, Codreanu, & Creuze, 2014). In addition, understanding the culture of learning in MOOCs is a complex process, and learners also face problems in adapting to these platforms (Loizzo & Ertmer, 2016). Compared to traditional education, learning outcomes from MOOCs, learners' purposes for enrolling in them (Watson et al., 2016), and their educational preferences are changing significantly (Watson, Watson, Yu, Alamri, & Mueller, 2017). A recent study reported that the same engagement measures may result in different achievement levels for different learner groups (Li & Baker, 2018). In other words, the instructional design of MOOCs is critical and essential (Yang, Shao, Liu, & Liu, 2017). Hone and El Said (2016) found that the MOOC content affected learners' retention and perceived effectiveness. Transactional interaction between learner and content, as well as the structure and assessment of course design factors are reported as significant predictors of learner control and sense of progress in MOOCs (Jung, Kim, Yoon, Park, & Oakley, 2019). Aparicio, Oliveira, Bacao, and Painho (2019) reported that gamification is contributing significantly to the overall success of MOOCs by reducing dropout rates and improving learner satisfaction and user experience. Therefore, because of its very nature, the

instructional design of MOOCs also needs to incorporate new approaches, rather than traditional ones (Adair et al., 2014; Margaryan, Bianco, & Littlejohn, 2015; Rodriguez, 2012). Currently, MOOC quality is reported as suboptimal (Margaryan et al., 2015), while quality instructional design can improve learning outcomes in distance learning (Hsu & Shiue, 2005).

In recent years, several studies have been conducted to better understand MOOC learners (e.g., Cagiltay, Esfer, & Celik, 2020; Hew & Cheung, 2014; Khalil & Ebner, 2014). However, there have been a limited number of studies offering a bigger picture on MOOCs. Jordan (2015) analyzed 221 MOOCs from different providers and the results indicated decreased average total enrolments in these courses over time, but an increase in completion rates. Jordan proposed some significant predictors for course completion rates, which were reported as positively correlated with the start date and assessment type, and negatively correlated with the course length. Ho et al. (2014) analyzed 17 Harvard University and MIT MOOCs offered in 2013, and reported a decrease from 3.2% to 2.5% in course certification for registered learners and an increase in the registration rates However, there is not sufficient evidence to confirm the results of these studies. In order to see the big picture, there is a need to analyze a larger number of courses, whereas previous studies only cover a limited number of MOOCs.

Accordingly, this study investigated 122 MITx courses with approximately 2.9 million learners, in order to provide feedback for MOOC developers on how to improve their courses. Specifically, this study analyzed the data provided by MITx, categorized these MOOCs into 15 course subjects classified into three course levels, and revealed the predictors of course certification rates. Although these MOOCs were from just one specific MOOC provider, they included heterogeneous learners from different backgrounds and countries. Therefore, this study is not limited to a particular region or country and can be generalized globally to some degree. In this sense, this study provides a unique contribution to open and distributed learning.

Methodology

Research Questions

The current study focused on the following six research questions:

1. How are the courses and number of enrolled learners distributed, in terms of subject areas?

2. How are the courses distributed in terms of course levels?

3. How is learner activity distributed in terms of subject areas and course levels?

4. How is learner activity and course certification distributed, in terms of course levels?

5. How is learner activity distributed in terms of specific course subjects?

6. What are the predictors of course certification rates?

Research Method

This quantitative study utilized descriptive and correlational research methods. Since it is difficult to confine educational events within controlled laboratory conditions, some types of educational research questions call for descriptions in order to explain the data (Knupfer & McLellan, 1996). The main focus of descriptive studies is to depict patterns rather than answer questions that ask why (Neuman, 2014), as they aim to describe and interpret what is happening (Cohen, Manion, & Morrison, 2007). In correlational research, the associations among variables are explored without any manipulation, and the variables can be used for prediction (Fraenkel & Wallen, 2009).

Data Collection and Analysis Process

The data were obtained from MITx MOOCs on the edX platform. In total, the available course data from 122 MITx MOOCs offered between 2012 and 2016 were obtained. The data provided by MITx was organized to represent details of each course. The level of each course was taken from the MITx website and combined with the course data. Then, the data were analyzed using descriptive statistics (mean, standard deviation, frequency, and percentage) and inferential statistics (multiple linear regression). Multiple linear regression (MLR) requires absence of outliers among the independent variables (predictor variables) and on the dependent variable (outcome variable), normality, linearity, and homoscedasticity of residuals, absence of multicollinearity, and independence of errors assumptions (Tabachnick & Fidell, 2007). Before carrying out an MLR analysis, its assumptions were checked. The data from four MOOCs, identified as outliers, were removed from the analysis. Due to the residuals not being completely normally distributed, a Box-Cox transformation was applied to the outcome variable, namely course certification rates. Thus, it was ensured that residuals were completely normally distributed. The homoscedasticity of residuals was evaluated by checking the scatterplot of the residuals, which showed no obvious pattern. Variance inflation factor (VIF) values were checked to determine whether there was multicollinearity between predictor variables. All the predictor variables had VIF values less than 3, ranging from 1.00 to 2.09, and no multicollinearity was detected. The Durbin-Watson value was checked for the independence of errors and found to be 1.81. To summarize, all assumptions were met for the multiple regression analysis. As a result, the MLR analysis was carried out with 118 MOOCs.

Results

Research Question One

Most of the courses (22.13%, n = 27) were offered in the computer science subject area; accordingly, the highest number of learners were also registered in these courses (38.24%; n = 1,107,780), followed by engineering subjects (15.57%; n = 19) and business and management (14.75%; n = 18). The mean age of the learners enrolled in each course ranged from 28.27 to 32.49 years. Table 1 shows the numbers and percentages of the analyzed courses according to subject area.

Table 1

Number of Courses According to Subject Area

	Courses		Enrolled learners		
Subject area	*n*	%	*n*	%	Mean age
Biology and life sciences	13	10.66	173,944	6.01	30.53
Education and teacher training	6	4.92	131,063	4.52	32.49
Business and management	18	14.75	358,048	12.36	31.81
Physics	9	7.38	80,940	2.79	29.67
Social sciences	9	7.38	101,665	3.51	31.44
Data analysis and statistics	5	4.10	200,318	6.92	30.61
Engineering	19	15.57	307,348	10.61	27.67
Math	3	2.46	91,002	3.14	29.00
Computer science	27	22.13	1,107,780	38.24	29.34
Communication	2	1.64	106,694	3.68	30.34
Humanities	3	2.46	119,658	4.13	31.07
Philosophy and ethics	2	1.64	43,786	1.51	30.07
Chemistry	3	2.46	23,424	0.81	28.27
History	2	1.64	42,151	1.46	32.23
Art and culture	1	0.82	8,718	0.30	29.76
Total	122	100	2,896,539	100	30.29

Research Question Two

Most of the courses (45%) were at introductory level and 60% of the learners registered for these courses, followed by intermediate (31%) and advanced level (24%) courses. In all these course levels, the mean age of the learners was 30.03 years. Table 2 shows the number of courses and their percentages according to course level.

Table 2

Number of Courses According to Course Level

Courses			Enrolled learners		
Course level	*n*	%	*n*	%	Mean age
Introductory	55	45.08	1,748,008	60.35	30.08
Intermediate	38	31.15	680,112	23.48	30.18
Advanced	29	23.77	468,419	16.17	29.84
Total	122	100	2,896,539	100	30.03

Research Question Three

As shown in Table 3, about 50% to 60% of the learners viewed the course content only once and the remainder left the course after enrolling. Around 5% to 15% of the learners completed half of their course.

Finally, approximately 2% to 4% completed their course. Rates of course completion and certification were very similar.

Table 3 also reveals that the percentage of learners enrolled in education and teacher training courses was 4.52% and was lowest for chemistry courses (0.81%). In data analysis and statistics courses, 66.81% of the enrolled learners viewed the course one or more times. This rate was lowest for the math and physics courses (around 50%). It should be noted from Table 3 that the certificate rates were similar (on average 2%–3%) for the courses with higher or lower course effort (e.g., data analysis and statistics courses at 210 hours, and art and culture courses at 30 hours, respectively). The certificate rate was highest for history courses (4.70%) and lowest for math courses (1.25%).

Table 3

Percentage of Learner Activity According to Subject Area

Subject area		Enrolled learners	Viewed course at least once	Viewed half the course	Completed course	Received certificate
Biology and life sciences	n	173,944	104,906	17,003	5,916	5,676
	%	6.01	60.31	9.77	3.40	3.26
Education and teacher training	n	131,063	74,442	19,306	4,688	4,382
	%	4.52	56.80	14.73	3.58	3.34
Business and management	n	358,048	194,664	46,168	15,180	13,806
	%	12.36	54.37	12.89	4.24	3.86
Physics	n	80,940	40,141	8,400	1,366	1,150
	%	2.79	49.59	10.38	1.69	1.42
Social sciences	n	101,665	55,130	10,191	4,482	4,410
	%	3.51	54.23	10.02	4.41	4.34
Data analysis and statistics	n	200,318	133,836	19,252	5,185	4,878
	%	6.92	66.81	9.61	2.59	2.44
Engineering	n	307,348	183,207	30,846	5,928	5,563
	%	10.61	59.61	10.04	1.93	1.81
Math	n	91,002	45,784	10,436	1,188	1,136
	%	3.14	50.31	11.47	1.31	1.25
Computer science	n	1,107,780	681,693	93,299	44,879	44,556
	%	38.24	61.54	8.42	4.05	4.02
Communication	n	106,694	56,645	12,563	3,414	3,413
	%	3.68	53.09	11.77	3.20	32.0
Humanities	n	119,658	65,489	10,344	2,492	2,418
	%	4.13	54.73	8.64	2.08	2.02
Philosophy and ethics	n	43,786	26,555	2,809	1,082	1,082
	%	1.51	60.65	6.42	2.47	2.47
Chemistry	n	23,424	14,318	1,128	433	433
	%	0.81	61.13	4.82	1.85	1.85
History	n	42,151	22,887	3,559	1,983	1,983
	%	1.46	54.30	8.44	4.70	4.70
Art and culture	n	8,718	4,879	798	254	254
	%	0.30	55.96	9.15	2.91	2.91

Note. Percentage figures represent the percent of the total number of learners.

The completion rates were higher for the courses in subjects such as education and teacher training, business and management, social sciences, computer science, and history compared to those related to biology and life sciences, physics, data analysis and statistics, engineering, math, communication, humanities, philosophy and ethics, chemistry, and art and culture.

Research Question Four

For those learners enrolled in an introductory level course, 59.92% viewed the course at least once (Table 4). This ratio was 58.05% for the intermediate and 56.01% for the advanced level courses. Of the learners enrolled in the introductory level courses, 9.09% viewed at least half of their course. This ratio was 11.18% and 10.94% for the intermediate and advanced course levels, respectively.

Table 4

Percentage of Learner Activity According to Course Level

Course level	Enrolled learners	Viewed course at least once	Viewed half the course	Completed course	Received certificate
Introductory					
n	1,748,008	1,047,383	158,854	61,702	60,613
%	60.35	59.92	9.09	3.53	3.47
Intermediate					
n	680,112	394,821	76,007	24,484	23,802
%	23.48	58.05	11.18	3.60	3.50
Advanced					
n	468,419	262,372	51,241	12,284	10,725
%	16.17	56.01	10.94	2.62	2.29

Note. Percentage figures represent the percent of the total number of learners.

The completion rate for the advanced level courses was 2.62%, 3.60% for the intermediate, and 3.53% for the introductory level courses. The majority of the learners who completed their course also received a certificate. These findings show that there is a huge gap between enrollment and certification rates. Thus, it is quite possible to refer to those who register but never look at the course content as MOOC window shoppers.

Research Question Five

As shown in Table 5, the highest average number of forum messages was in the humanities courses (4,004.67) and lowest was in the education and teacher training courses (0.50). The highest number of total forum messages was in computer science courses (51,696) and lowest was in education and teacher training courses (3). However, when the percentages of forum messages were compared to the number of enrolled students, the result was highest in history (10.29%) and humanities (10.04%) courses, and lowest for education ad teacher training (0%) courses.

Table 5

Learner Activity According to Subject Area

Subject area	Courses	Forum messages		Enrolled students and percentage of forum messages		Total chapters		Chapters visited per course
	n	n	M	n	%	n	M	M
Biology and life sciences	13	9,971	767.00	173,944	5.73	203	15.62	3.74
Education and teacher training	6	3	0.50	131,063	0.00	45	7.50	2.69
Business and management	18	25,185	1,399.17	358,048	7.03	204	11.33	3.93
Physics	9	2,322	258.00	80,940	2.87	97	10.78	2.54
Social sciences	9	4,980	553.33	101,665	4.90	120	13.33	3.62
Data analysis and statistics	5	3,814	762.80	200,318	1.90	84	16.80	3.41
Engineering	19	7,692	404.84	307,348	2.50	325	17.11	3.38
Math	3	2,241	747.00	91,002	2.46	34	11.33	2.35
Computer science	27	51,696	1,914.67	1,107,780	4.67	368	14.15	3.78
Communication	2	6,629	3,314.50	106,694	6.21	28	14.00	4.23
Humanities	3	12,014	4,004.67	119,658	10.04	25	8.33	2.31
Philosophy and ethics	2	1,798	899.00	43,786	4.11	27	13.50	3.26
Chemistry	3	763	254.33	23,424	3.26	78	26.00	3.91
History	2	2,411	1,205.50	23,424	10.29	40	20.00	4.89
Art and culture	1	360	360.00	8,718	4.13	10	10.00	3.18

Depending on the subject area, the average length of the courses ranged from 6 to 15 weeks, and the average weekly workload ranged between 5 and 14 hours. Course effort, which was calculated by multiplying the course length and average hours required per week, were between 30 hours and 210 hours. The highest total number of chapters was in the computer science courses (368) and the lowest was in the art and culture courses (10). The mean of the average completed chapters was highest in history courses (4.89) and lowest in humanities courses (2.31). Table 6 reveals that the highest number of forum messages pertained to the introductory courses (90,158) and the lowest to the advanced courses (15,096). There was a similar pattern for the total number of chapters and average chapters completed, which were higher in the introductory courses and lower in the advanced courses.

Table 6

Learner Activity According to Course Level

	Courses	Forum messages		Chapters		
Course level	*n*	n	*M*	*n*	*M*	Completed
Introductory	55	90,158	1,639	847	15	204
Intermediate	38	26,625	701	487	13	132
Advanced	29	15,096	521	354	12	88

Research Question Six

The variables of (a) viewing the course once, (b) viewing half the course, (c) total forum messages, (d) total number of chapters, (e) average chapters completed, (f) length of the course in weeks, and (g) mean age of users were used to predict earned certification rates in MITx courses. The descriptive statistics regarding these variables are given in Table 7.

Table 7

Descriptive Statistics of Outcome and Predictor Variables

Variable	*M*	*SD*	Min	Max
Certification rate	2.74	2.27	.05	10.03
Certification rate (transformed)	1.25	.41	.38	2.07
Viewed the course once	14,413	14,479	231	65,380
Viewed half of the course	2,394	2,246	72	9,748
Total forum messages	1,092	1,688	0	9,018
Total number of chapters	14.05	7.56	4.00	50
Average number of chapters completed	3.42	.87	1.69	6.41
Length of the course	9.52	3.86	3.00	18
Learners' mean age	29.85	1.90	25.33	35.48

A stepwise multiple linear regression was carried out to explore the predictors, and to determine which contributed most to the course certification rates in the 118 MITx courses. The results of the multiple regression showed four significant models (see Table 8), in which four predictors explained 54.3% of the variance in course certification rates ($F(4,111) = 33.03$, $p < .05$). It was found that average chapters completed ($\beta = .80$, $p < .05$), total number of chapters ($\beta = -.41$, $p < .05$), total forum messages ($\beta = .18$, $p < .05$), and mean age ($\beta = .15$, $p < .05$) significantly contributed to the model. Viewing the course once ($p > .05$), viewing half the course ($p > .05$), and length of the course ($p > .05$) did not make significant contributions.

Table 8

Stepwise Multiple Linear Regression Results

Model	Predictor	B	SE B	β	t	sr²	R²	ΔF
1	Average chapters completed	.28	.04	.59	7.72[*]	.34	.343	59.56
2	Average chapters completed	.43	.04	.92	10.62[*]	.50	.502	35.99
	Total number of chapters	-.03	.01	-.52	-6.00[*]	.16		
3	Average chapters completed	.40	.04	.84	9.34[*]	.37	.525	5.42
	Total number of chapters	-.03	.01	-.49	-5.75[*]	.14		
	Total forum messages	.00	.00	.16	2.33[*]	.02		
4	Average chapters completed	.37	.04	.80	8.66[*]	.31	.543	4.52
	Total number of chapters	-.02	.01	-.41	-4.42[*]	.08		
	Total forum messages	.00	.00	.18	2.59[*]	.03		
	Mean age	.03	.02	.15	2.13[*]	.02		

The average number of chapters completed, total forum messages, and mean age positively predicted the course certification rates. On the other hand, the total number of chapters predicted the course certification rates negatively. When the unique contributions of the predictors were examined to determine how each one explained the variance in course certification rates, the average number of chapters completed explained 31%, the total number of chapters explained 8%, the total number of forum messages explained 3%, and mean age explained 2% of the variance.

Discussion and Conclusion

Course and Learner Characteristics

The purpose of this study was to provide a deeper understanding of the MITx MOOCs presented on the edX platform and the predictors of the course certification rates in these courses. Between 2012 and 2016, both the number of MITx courses and the number of enrolled learners increased. This is consistent with MOOCs being offered by more than 800 universities worldwide, and the number of MOOCs having exceeded 10,000 (Shah, 2018b). However, the current study found a decrease in the certification rates among the enrolled learners. These findings conflict with Jordan's (2015) comprehensive study that reported a decrease in enrollments over time and an increase in course completion percentages.

Based on the number of courses and the total registered learners, computer science as well as business and management courses were the most popular. According to Shah (2018b) technology courses ($n = 17$) dominated Class Central's list of the all-time, top 50 MOOCs; business courses ($n = 6$) were very popular as well. Moreover, the top 10 most popular courses on Coursera included a significant number of computer science and similar courses (Young, 2018), and the trend was similar in 2017, with cutting-edge tech skills being those most demanded in online education (Sinha, 2017). Thus, it can be inferred that MITx MOOCs are effective in satisfying individual's learning needs in both computer science and business and management courses.

While Cohen, Shimony, Nachmias, and Soffer (2019) reported that only 50% of registered learners actually started their course, among the courses analyzed in the current study, the rate was slightly higher, with 50% to 62% of learners starting their courses. However, the completion rates in this study were slightly lower compared to the earlier results reported as 4.6% (Pardos, Bergner, Seaton, & Pritchard, 2013) and 5.6% (Despujol, Turró, Castañeda, & Busquets, 2017) for edX courses, and 8% to 10% for MOOCs in general (Cohen et al., 2019; Jordan, 2014). In addition, completion rates were lower for both gender groups in the advanced level courses as compared to the introductory level courses. The results of the current study support the findings in the relevant literature. In general, course completion rates were lower in the MOOCs; however, in this study, it was shown that once learners were able to complete their course, it was most likely they would also receive certificates. Accordingly, when course completion rates improved, certification rates also improved.

In general, the majority of the enrolled learners had bachelor's or master's degrees, or both. Enrolled learners with a middle or high school diploma (equivalent to secondary, high school, or junior secondary/high/middle school) was 24%. Learners having completed primary or elementary school or who had no formal education accounted for 10% of those enrolled in the courses, as small minority of enrolled learners. These results parallel those in earlier studies, indicating that the majority of MOOC learners had some educational background including a higher degree (Bayeck, 2016; Christensen et al., 2013; Macleod, Haywood, Woodgate, & Alkhatnai, 2015). Furthermore, in the intermediate and advanced level courses, the percentages of learners having bachelor's degree, master's degree, or a doctorate, were higher than in the introductory courses. On the other hand, in the introductory level courses, the percentage of learners having an associate degree and secondary or high school diploma were higher than the participants in the intermediate and advanced level courses. The results also indicated that the rates of completing at least half of the course increased with higher education levels. Based on these results, MOOCs benefit the educated population, but they have not yet satisfied initial expectations because they do not serve the potential needs of those from lower educational levels.

Regarding course levels, the (a) total number of courses, (b) number of enrolled learners, (c) course length, (d) percentage of learners who viewed their course at least once, and (e) percentage who viewed half the course were higher for the introductory level courses than the advanced courses. The situation was similar for the total number of chapters, number of forum messages, and average number of completed chapters. Similarly, in the advanced level courses, course completion and certification rates were lower than the introductory and intermediate level courses. The introductory level courses were more popular considering the number of courses and registered learners with higher completion rates. Additionally, compared to the intermediate and advanced level courses, in the introductory level courses there was a higher number of learners who had completed secondary or high school, or who had no formal education. Accordingly, a course that provides basic, introductory information could be added to existing introductory level courses. Thus, learners with lower education levels could be supported by MOOCs, for a variety of purposes, such as returning to school and preparing for college.

Predictors of Certification Rates

This study revealed that the average number of chapters completed, total number of forum messages, and mean age positively predicted course certification rates. On the other hand, the total number of chapters in

a course negatively predicted course certification rates. Overall, the results of the current study support earlier results in the literature. When the average number of chapters completed increased, a learner was closer to completing the course. Consistent with this finding, Hone and Said (2016) commented that the MOOC learners who passed the midpoint of a course were likely to complete it.

Similar to our results, previous research reported that viewing online forums and participating in online discussions were significant influencing factors for predicting MOOC completion rates (Bonafini, 2017; Goldwasser, Mankoff, Manturuk, Schmid, & Whitfield, 2016). Social interactions in online groups are crucial for successful learning (Barak, Watted, & Haick, 2016). Furthermore, engaging with forum comments is related to greater commitment to the course materials (Ferguson & Clow, 2015). MOOC forums can have different functions; for example, in MOOC forums, learners can socialize with their peers, while also asking questions about the course material and exams. In this way, forums take on the traditional classroom role of offering assistance during office hours or talking to and helping a classmate understand a challenging subject (Diver & Martinez, 2015). However, it is important to mention that discussions might be problematic in MOOCs due to the massive number of participants. Regarding age, in a study conducted with 33,938 learners in a MOOC offered in 2013, Greene, Oswald, and Pomerantz (2015) found that older participants were less likely to leave a MOOC. A similar finding was reported by Morris, Hotchkiss, and Swinnerton (2015), who indicated that older learners were more likely to complete a MOOC. Concerning the length of MOOCs, Jordan (2014) revealed that completion rates were negatively associated with course length. However, in this study, course length was not significant in predicting course certification rates, but the total number of chapters was. As the chapters in MITx courses are also aligned with the course sections, they are related to the volume of the course, and thus indirectly related to the course length. The associations mentioned above were observed in the courses investigated in this study. For example, computer science, business and management, and history courses can be seen as more successful in terms of course completion and certification rates. According to the number of forum messages, learners in these courses were also more active. However, the chemistry courses could be evaluated as less successful considering their course completion rates and less interaction in terms of the number of forum messages. Accordingly, it can be concluded that in some courses, the level of interaction and success was higher than others. There could be different reasons for this (e.g., instructional design features of the courses, how the content is represented, usability features of the courses). Jung et al. (2019) provided evidence that instructional design factors were more often linked to successful MOOC experiences, compared to content and demographic factors, and that instructional components were significant predictors of learning in MOOCs. According to Deshpande and Chukhlomin (2017), content, accessibility, and interactivity influenced participants' motivation to learn. Furthermore, Junjie (2017) found that knowledge outcome was a strong predictor of continuance intention of MOOC learners. Furthermore, learners' attitudes, motivations, and backgrounds could also be influencing factors. Hence, these factors need to be analyzed in depth to improve the interaction level and success of MOOCs.

Implications

Several recommendations for MOOCs regarding course length, course activities, learner needs, course content design, and learner motivation can be represented in the acronym SINCE: Shorter and more Interactive courses that consider learners' Needs, as well as Course design and strategies that encourage

enrolled students to Enter their course. These recommendations, explained in detail below, might be useful for people who are interested in designing and developing MOOCs.

Shorter courses. Certification rates increase when the number of chapters learners complete and the number of forum messages learners post also increase. On the other hand, certification rates decrease when the number of chapters in a course increases. Therefore, courses should be divided into smaller sections while also keeping the total number of chapters in the course low, so that courses can be designed to be an easily digestible size. Similarly, as the number of learners who completed half of their course is high, perhaps the optimal length for a course is half as long.

Interactive courses. As learners enter the course for the first time, they should encounter interactive components, such as forums or activities, to boost learner motivation and retain learners in the course; MOOC learners should also be rewarded for the chapters they complete. In addition, since the number of forum messages predicted certification rates positively, forum activities should be promoted and course discussions should be carefully designed.

Needs. Learner profiles (e.g., education level, gender) as well as their needs and expectations should be analyzed carefully in relation to the offered MOOCs. Course content and design should be updated regularly, taking these needs into consideration.

Content of courses. Completion rates are lower in some course subjects, such as chemistry, while higher in others, such as business and management. Course content should be critically analyzed in terms of its appropriateness for the MOOC format. This issue should also be further researched in order to find more effective ways of developing and designing MOOCs.

Entering courses. Since most of the enrolled learners do not even enter the course, such window shoppers should be encouraged to enter the course through use of motivational strategies. Regular reminders to enter the course could be sent to learners. Moreover, as in the design of a shopping mall, participant attraction strategies could be employed and new course marketing approaches could be tested and evaluated.

Limitations and Further Research

This study has a number of limitations. First, it was limited to 122 MITx courses offered between 2012 and 2016. As well, the current study was based on the log data kept by MITx only, so the scope of this study was limited to the content of these log data.

Future research studies should consider using multiple data sources, as well as merging the log data with learners' self-reported data in order to get a more comprehensive view of course and learner characteristics, and how these influence course certification rates. Future work could also include and compare other MOOC portals with regards to course characteristics, learner characteristics, and certification rates. Third, the recommendations this study yielded for MOOCs regarding course length, course activities, learner needs, course content design, and learner motivation could be empirically tested in future studies. Finally, research could explore ways to motivate learners to enter their course, as many learners who register for a MOOC never attempt to start the course.

Acknowledgements

We are grateful to MITx for allowing us to obtain the course data.

References

Adair, D., Alman, S. W., Budzick, D., Grisham, L. M., Mancini, M. E., & Thackaberry, A. S. (2014). Many shades of MOOCs. *Internet Learning, 3*(1), 53–72. doi: 10.1.1.883.3241

Aparicio, M., Oliveira, T., Bacao, F., & Painho, M. (2019). Gamification: A key determinant of massive open online course (MOOC) success. *Information & Management*, *56*(1), 39–54. doi: 10.1016/j.im.2018.06.003

Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education*, *94*, 49–60. doi: 10.1016/j.compedu.2015.11.010

Bayeck, R. (2016). Exploratory study of MOOC learners' demographics and motivation: The case of students involved in groups. *Open Praxis*, *8*(3), 223–233. doi: 10.5944/openpraxis.8.3.282

Beaven, T., Codreanu, T., & Creuzé, A. (2014). Motivation in a language MOOC: Issues for course designers. In M. Elena & B. Elena (Eds.), *Language MOOCs: Providing learning, transcending boundaries* (pp. 48–66). Berlin, Germany: De Gruyter Open.

Bonafini, F. C. (2017). The effects of participants' engagement with videos and forums in a MOOC for teachers' professional development. *Open Praxis, 9*(4), 433–447. doi: 10.5944/openpraxis.9.4.637

Cagiltay, K., Esfer, S., & Celik, B. (2020). Insights into a nationwide pdMOOC portal: bilgeis.net of Turkey. In K. Zhang, C. J. Bonk, T. C. Reeves, & T. H. Reynolds (Eds.), *MOOCs and open education in the global south: Challenges, successes, and Opportunities* (pp.130–139). New York, NY: Routledge. doi: 10.4324/9780429398919

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. (2013). *The MOOC phenomenon: Who takes massive open online courses and why?* (Working paper). Retrieved from Elsiver website: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2350964

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. New York, NY: Routledge.

Cohen, A., Shimony, U., Nachmias, R., & Soffer, T. (2019). Active learners' characterization in MOOC forums and their generated knowledge. *British Journal of Educational Technology*, *50*(1), 177–198. doi: 10.1111/bjet.12670

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

DeBoer, J., Stump, G. S., Seaton, D., & Breslow, L. (2013). Diversity in MOOC students' backgrounds and behaviors in relationship to performance in 6.002 x. In *Proceedings of the Sixth Learning*

International Networks Consortium Conference (Vol. 4, pp. 16–19). Retrieved from
https://linc.mit.edu/linc2013/proceedings/Session2/Session2DeBoer-Stump.pdf

Deng, R., Benckendorff, P., & Gannaway, D. (2019). Progress and new directions for teaching and learning in MOOCs. *Computers & Education*, *129*, 48–60. doi: 10.1016/j.compedu.2018.10.019

Deshpande, A., & Chukhlomin, V. (2017). What makes a good MOOC: A field study of factors impacting student motivation to learn. *American Journal of Distance Education*, *31*(4), 275–293. doi: 10.1080/08923647.2017.1377513

Despujol, I. M., Turró, C., Castañeda, L., & Busquets, J. (2017). Effect of free certificate discontinuation in completion rates of MOOC. In *European Conference on Massive Open Online Courses* (pp. 182–187). Cham, Switzerland: Springer. doi: 10.1007/978-3-319-59044-8_21

Diver, P., & Martinez, I. (2015). MOOCs as a massive research laboratory: Opportunities and challenges. *Distance Education*, *36*(1), 5–25. doi: 10.1080/01587919.2015.1019968

Ferguson, R., & Clow, D. (2015). Examining engagement: Analysing learner subpopulations in

massive open online courses (MOOCs). In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 51–58). Poughkeepsie, NY, USA: ACM. doi: 10.1145/2723576.2723606

Fraenkel, J. R., & Wallen, N. E. (2009). *How to design and evaluate research in education* (7th ed.). New York, NY: McGraw-Hill.

Goldwasser, M., Mankoff, C., Manturuk, K., Schmid, L., & Whitfield, K. E. (2016). Who is a student: Completion in Coursera courses at Duke University. *Current Issues in Emerging eLearning*, *3*(1), 125–138. Retrieved from https://scholarworks.umb.edu/ciee/vol3/iss1/8/

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal, 52*(5), 925–955. doi: 10.3102/0002831215584621

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

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review, 12*, 45–58. doi: 10.1016/j.edurev.2014.05.001

Ho, A., Chuang, I., Reich, J., Coleman, C., Whitehill, J., Northcutt, C., . . . & Petersen, R. (2015). HarvardX and MITx: Two years of open online courses fall 2012-summer 2014 (Working Paper). Retrieved from HarvardX website: doi: 10.2139/ssrn.2586847

Ho, A., Reich, J., Nesterko, S., Seaton, D., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The first year of open online courses, fall 2012-summer 2013 (Working Paper). Retrieved from HarvardX and MITx website: doi: 10.2139/ssrn.2381263

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

Hsu, Y., & Shiue, Y. (2005). The effect of self-directed learning readiness on achievement comparing face-to face and two-way distance learning instruction. *International Journal of Instructional Media, 32*(2), 143–156. Retrieved from https://eric.ed.gov/?id=EJ698905

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, *15*(1), 133–160. doi: 10.19173/irrodl.v15i1.1651

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), 341–358. doi: 10.19173/irrodl.v16i3.2112

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. doi: 10.1016/j.compedu.2018.10.001

Junjie, Z. (2017). Exploring the factors affecting learners' continuance intention of MOOCs for online collaborative learning: An extended ECM perspective. *Australasian Journal of Educational Technology*, *33*(5). doi: 10.14742/ajet.2914

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction*, *22*(2), 6. doi: 10.1145/2699735

Khalil, H., & Ebner, M. (2014). MOOCs completion rates and possible methods to improve retention: A literature review. In *EdMedia+ Innovate Learning* (pp. 1305–1313). Waynesville, NC: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/147656

Knupfer, N. N. & McLellan, H. (1996). Descriptive research methodologies. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1196–1212). New York, NY: Macmillan.

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *International Review of Research in Open and Distributed Learning, 12*(7), 74–93. doi: 10.19173/irrodl.v12i7.1041

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

Loizzo, J., & Ertmer, P. A. (2016). MOOCocracy: The learning culture of massive open online courses. *Educational Technology Research and Development, 64*(6), 1013–1032. doi: 10.1007/s11423-016-9444-7

Macleod, H., Haywood, J., Woodgate, A., & Alkhatnai, M. (2015). Emerging patterns in MOOCs: Learners, course designs and directions. *TechTrends*, *59*(1), 56–63. doi: 10.1007/s11528-014-0821-y

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses. *Computers & Education*, *80*, 77–83. doi: 10.1016/j.compedu.2014.08.005.

Morris, N. P., Hotchkiss, S., & Swinnerton, B. (2015). Can demographic information predict MOOC learner outcomes? In *Proceedings of the European MOOC Stakeholder Summit*, (pp. 199–207) Mons, Belgium. Retrieved from http://eprints.whiterose.ac.uk/86184

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Essex, England: Pearson Education.

Pardos, Z. A., Bergner, Y., Seaton, D. T., & Pritchard, D. E. (2013). Adapting bayesian knowledge tracing to a massive open online course in edX. *EDM*, *13*, 137–144. doi: 10.1.1.668.6548

Porter, W. W., Graham, C. R., Spring, K. A., & Welch, K. R. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers & Education, 75*, 185–195. doi: 10.1016/j.compedu.2014.02.011

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance and E-learning, 2012*(1). Retrieved from https://files.eric.ed.gov/fulltext/EJ982976.pdf

Shah, D. (2018a, December 11). By the numbers: MOOCs in 2018 (Blog post). *Class Central.* Retrieved from https://www.class-central.com/report/mooc-stats-2018

Shah, D. (2018b, August 20). Class central's top 50 MOOCs of all time (2018 edition; Blog post). *Class Central.* Retrieved from https://www.class-central.com/report/top-50-moocs-2018-edition

Sinha, N. (2017, December). Year in review: 10 most popular courses in 2017 (Blog post). *CourseraBlog*. Retrieved from https://blog.coursera.org/year-review-10-popular-courses-2017

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson Education.

Watson, S. L., Loizzo, J., Watson, W. R., Mueller, C., Lim, J., & Ertmer, P. A. (2016). Instructional design, facilitation, and perceived learning outcomes: An exploratory case study of a human trafficking

MOOC for attitudinal change. *Educational Technology Research and Development*, *64*(6), 1273-1300. doi: 10.1007/s11423-016-9457-2

Watson, S. L., Watson, W. R., Yu, J. H., Alamri, H., & Mueller, C. (2017). Learner profiles of attitudinal learning in a MOOC: An explanatory sequential mixed methods study. *Computers & Education*, *114*, 274–285. doi: 10.1016/j.compedu.2017.07.005

Yang, M., Shao, Z., Liu, Q., & Liu, C. (2017). Understanding the quality factors that influence the continuance intention of students toward participation in MOOCs. *Educational Technology Research and Development*, *65*(5), 1195–1214. doi: 10.1007/s11423-017-9513-6

Young, J. R. (2018, June). How blockbuster MOOCs could shape the future of teaching. *EdSurge*. Retrieved from https://www.edsurge.com/news/2018-06-19-how-blockbuster-moocs-could-shape-the-future-of-teaching