Exploring Demographics and Students' Motivation as Predictors of Completion of a Massive Open Online Course

Qing Zhang[1], Fernanda Cesar Bonafini[2], Barbara B. Lockee[3], Kathryn W. Jablokow[4], and Xiaoyong Hu[5]
[1,3]Virginia Tech, [2,4]The Pennsylvania State University, [5]South China Normal University

Abstract

This paper investigates the degree to which different variables affect the completion of a Massive Open Online Course (MOOC). Data on those variables, such as age, gender, English proficiency, education level, and motivation for course enrollment were first collected through a pre-course survey. Next, course completion records were collected via the Coursera database. Finally, multiple binomial logistic regression models were used to identify factors related to MOOC completion. Although students were grouped according to their preferences, working in groups did not affect students' likelihood for MOOC completion. Also, other variables such as age, the institution hosting the MOOC, academic program alignment with students' needs, and students' intention to complete the course all affected their probability of MOOC completion. This study contributes to the literature by indicating the factors that influence the probability of MOOC completion. Results show that older participants (age > 50 years old) have higher probability of completing the MOOC. Students' MOOC completion also increases when the MOOC provides experiences that add to students' current academic backgrounds and when they are hosted by institutions with a strong academic reputation. Based on these factors, this study contributes to research methods in MOOCs by proposing a model that is aligned with the most important factors predicting completion as recommended by the current MOOC literature. For the next phase of assigning learners to work in groups, findings from this study also suggest that MOOC instructors should provide assistance for group work and monitor students' collaborative processes.

Keywords: MOOC completion, demographics, motivation, intention of completion, groups in MOOCs

Introduction

The Massive Open Online Course (MOOC) is a popular online learning platform in which millions of students enroll. MOOCs offer educational opportunities for people who otherwise could not afford a formal education (Dillahunt, Wang, & Teasley, 2014). Unfortunately, MOOCs are also known for their high attrition rates (Ho et al., 2014; Lim, Coetzee, Hartmann, Fox, & Hearst, 2014; Malan, 2013). According to Liyanagunawardena, Adams, and Williams (2013), most MOOCs have a completion rate of less than 10%. Many factors cause dropouts in MOOCs, including overall lack of intention to complete the course (Cross, 2013; Onah, Sinclair, & Boyatt, 2014), lack of time and support (Hew & Cheung, 2014; Kellogg, Booth, & Oliver, 2014), change of job or location (Gütl, Rizzardini, Chang, & Morales, 2014; Onah et al., 2014), and language barriers (Schulze, 2014).

Many studies have been conducted to identify factors that contribute to MOOC completion; however, the findings vary across those studies. Cisel (2014) indicated that learner performance in MOOCs was highly correlated with the learner's geographic location, employment status, and time constraints, and that unemployed learners from high Human Development Index (HDI) countries were more likely to complete the course. Other variables that have been examined for their effects on MOOC completion include years of education (Guo & Reinecke, 2014; Schulze, 2014), friends' performance in a MOOC (Brown et al., 2015), prior online learning experience (Morris, Hotchkiss, & Swinnerton, 2015), English proficiency (Engle, Mankoff, & Carbrey, 2015; Konstan, Walker, Brooks, Brown, & Ekstrand, 2015; Schulze, 2014), number of posts and number of videos watched (Bonafini, 2017; Bonafini, Chae, Park, & Jablokow, 2017), gender (Bayeck, Hristova, Jablokow, & Bonafini, 2018; Breslow et al., 2013; Konstan et al., 2015; Schulze, 2014), and age (Breslow et al., 2013; Guo & Reinecke, 2014; Konstan et al., 2015; Morris et al., 2015; Schulze, 2014; Zhang et al., 2016). Most of the studies agree that there is a positive relationship between age and MOOC completion rates. Zhang et al. (2016) concluded that learners with age over 40 years who intended to complete the course achieved higher MOOC completion rates. In addition, Morris, Hotchkiss, and Swinnerton (2015) found that unemployed and older learners who had higher levels of education and previous online learning experiences tended to achieve higher course completion rates.

In spite of the numerous studies conducted to study learner motivations for enrolling in MOOCs (Belanger & Thornton, 2013; Gil-Jaurena, Callejo-Gallego, & Agudo, 2017; Konstan et al., 2015; Macleod, Haywood, Woodgate, & Alkhatnai, 2014; Radford, Coningham, & Horn, 2015; Zhong, Zhang, Li, & Liu, 2016), only a few of them have investigated the influence of motivation on MOOC completion. Konstan, Walker, Brooks, Brown, and Ekstrand (2015) concluded that most of the reasons that learners enrolled in a MOOC, such as university/instructor-related reasons or access to educational institutions-related reasons, did not affect course completion, but that learners' self-reported intention of completing the MOOC was a significant predictor of course completion.

In consideration of the inconsistent findings for predictors of MOOC completion from the existing literature, this paper presents a MOOC completion model that includes relevant variables to identify the most useful predictors pertaining to MOOC completion. This model identifies relevant characteristics of MOOC completers and non-completers, which can further inform the design and development of future MOOCs.

Literature Review

Students' motivation for taking a MOOC has been identified as a crucial factor for course engagement, which keeps learners persisting in the course (Xiong et al., 2015). Motivation factors, which contribute to sustained student engagement, include interest in the topics (Dillahunt et al., 2014; Hew & Cheung, 2014; Kizilcec & Schneider, 2015), curiosity about MOOCs (Hew & Cheung, 2014; Zheng, Rosson, Shih, & Carroll, 2015), current job needs (Christensen et al., 2013), the opportunity to connect with others (Belanger & Thornton, 2013), preparation for future jobs (Kizilcec & Schneider, 2015; Zheng, Rosson, Shih, & Carroll, 2015), relevance to current academic programs, interest in earning a certificate, and interest in the professor or the institution that offers the MOOC (Kizilcec & Schneider, 2015). Xiong et al. (2015) categorize these motivations as intrinsic motivations (interest related), extrinsic motivations (external rewards related, e.g. earning course completion certificate), and social motivations (taking this course with friends and connecting with others). Upon finding out that intrinsic and extrinsic motivations are significant predictors of learner engagement, and learner engagement correlates positively with retention, Xiong et al. (2015) propose forming a student learning community and providing incentives (e.g., certificates) as motivation factors to enhance learner engagement and retention.

Plenty of studies suggest that the use of group work could improve learning interaction and engagement, and that it has potential to enhance learning in MOOCs (Arendale & Hane, 2014; Berger & Wild, 2016; Hiltz, 1998; Jones, 1997; Williams, Duray, & Reddy, 2006, Wen, 2016). By working with others in a MOOC, students could learn from and assist one another in the learning process (Yuan & Powell, 2013). In their study, Guàrdia, Maina, and Sangrà (2013) found that collaborative work and peer assistance and assessment were effective MOOC design principles. Kulkarni, Cambre, Kotturi, Bernstein, and Klemmer (2015) found that "the more geographically diverse the discussion group, the better the students performed" (p. 1126).

A number of grouping approaches in MOOCs have been implemented in recent years. These approaches can be summarized into two categories: random grouping and criteria-based grouping. Random grouping is done by assigning learners into groups randomly (Zheng, Vogelsang, & Pinkwart, 2015). Whereas criteria-based grouping is performed based on different grouping mechanisms. For instance, Wen (2016) formed teams based on the transactive discussion within a large community and further deployed an automated agent to support team discussion. Zheng, Vogelsang, and Pinkwart (2015) created MOOC groups based on learner's preferred collaboration media and demographic information, including gender, time zone, and language. Sinha (2014) proposed to assign MOOC students to teams based on their connections with other learners in a social network.

In addition to motivation and grouping factors, learner's intention for completing a MOOC was identified as a significant estimator of their actual completion of a MOOC (Bonafini et al., 2017; Koller, Ng, Do, & Chen, 2013; Konstan et al., 2015). For instance, Koller, Ng, Do, and Chen (2013) concluded that learners with the intention of completing a MOOC achieved higher completion rates when compared to those who did not. Bonafini, Chae, Park, and Jablokow (2017) found that student's desire for certification had an amplifying effect on students' MOOC completion, as well as on the number of videos watched by the students. These studies inform us that a learner's commitment in completing the course plays an

importance role in terms of improving learner engagement and retention in MOOCs. In this sense, it seems that the higher level of goal commitment a learner sets for oneself when the tasks are achievable, the better performance the learner will achieve (Locke, 1982).

The existing literature lays the foundation for incorporating pertinent variables to build a MOOC completion model for a particular MOOC such as learner demographics, motivation for enrollment, intention of completion, and working in groups.

Methodology

The purpose of this paper is to develop a multiple binomial logistic regression model that distinguishes significant variables affecting MOOC completion. The completion level is treated as a binary dependent variable with the result of either completing the course or not. Independent variables include age, gender, education level, motivation for taking the MOOC, working in groups, and intentions of completing the course. Participants were recruited to work in small online groups by matching their grouping preferences, such as their preferred language, media to communicate, and intention of completing the course. Students who had group preferences that could not to be matched by their preferences were placed in the control group. This study investigates the following research questions:

1. What are the characteristics of MOOC learners who participated in this study?

2. What are the learners' preferences related to working in groups?

3. What are the learners' motivations for taking this MOOC?

4. Which demographics and motivational factors predict the probability of MOOC completion?

Participants

Participants in this study were recruited from a MOOC offered through the Coursera platform from July to August, 2014 (Jablokow, Matson, & Velegol, 2014). Prior to the beginning of the course, an invitation for participating in online groups was sent out to MOOC learners. Learners who responded with interest in working in online groups received a pre-course survey, which inquired about their demographic information, reasons for taking this course, and grouping preference, among other questions. Participants were assigned into groups following the order of their preferred language to communicate within a team, intention of completion, and mode of communication (synchronous text, asynchronous text, or synchronous video and audio) (Zhang et al., 2016). Some of the synchronous groups were formed based on converted time zones. Participants whose grouping preference could not be satisfied or matched with others such as preferred language to speak in an online team or preferred time to work with others, were assigned into a control group. Students who were assigned to the control group received no instructional guidance or monitoring for group work.

After the online groups were formed, a general group work instruction email was sent out to the participants. In consideration of the large number of Chinese participants who volunteered for this grouping study, the email instruction was also translated into Chinese. Various online tools were suggested for different types of group communication such as MOOC discussion forums and email for learners' asynchronous communication, and Skype and QQ (a Chinese instant messaging tool) for learners' synchronous communication. Additionally, ZOOM (a video conferencing tool for large group discussions) was offered by the research team to learners for free use.

Data Sources

Pre-course survey. At the beginning of this course, a pre-course survey was sent to participants to collect their demographic information, such as gender, age, level of education, level of English proficiency, previous online learning experience, and employment status.

Post-course survey. At the end of this course, a post-course survey was sent to participants to gather feedback of their experiences of working in online groups in this MOOC.

Completion data. Learners in this MOOC were required to submit at least six assignments in order to obtain a certificate of completion. For learners who opted to earn a certificate of completion with distinction, twelve additional peer reviews were required. Learners who failed to meet these requirements were not awarded a completion certificate. Original course completion data was retrieved from Coursera with three levels of completion: none, normal, and distinction. These three levels of completion were recoded as a binary variable showing two levels of course completion: Complete (the combination of normal completion and completion with distinction) and Non-Complete.

Data Analysis

The pre-course survey data was exported from Qualtrics, students' completion records were collected through Coursera, and various data sets were retrieved and combined together in an SQL database. The data analysis and its graphical representation were computed using ArcMAP, SPSS, and R-Studio. R-Studio was used to run multiple binomial logistic regression models in order to identify the predictors that affect learners' MOOC completion. Within the model, MOOC completion is defined as a binary dependent variable, and all the independent variables are defined as categorical variables.

Independent variables were drawn from existing literature as shown in Table 1. We included learner demographics in our model as suggested by the research of Bayeck, Hristova, Jablokow, and Bonafini (2018), Breslow et al. (2013), Cisel (2014), and Engle, Mankoff, and Carbrey (2015), which include age, gender, education level, English proficiency, and employment status. We also included as parameter estimates: learners' motivations for taking MOOCs, as suggested in the research of Belanger and Thornton (2013), Brown et al. (2015), Dillahunt, Wang, and Teasley (2014), and Kizilcec and Schneider (2015). Motivations for taking MOOCs included interest in the subject, interest in the institution and professor that provides the course, building social connection with others, employment opportunities, earning a certificate, and friends' taking the course. Other variables identified from the literature contained the intention of completing the course (Engle et al., 2015; Koller et al., 2013; Konstan et al., 2015) and

participation in online groups (Kulkarni, Cambre, Kotturi, Bernstein, & Klemmer, 2015; Sinha, 2014; Wen, 2016; Zheng, Vogelsang, & Pinkwart, 2015).

Table 1

Variable Literature and Data Sources

	Variable	Literature Source	Data Source
1	English level	Engle et al. (2015); Konstan et al. (2015); Schulze (2014); Zhang et al. (2016).	Pre-course survey
2	Education level	Engle et al. (2015); Guo and Reinecke (2014); Schulze (2014); Zhang et al. (2016).	Pre-course survey
3	Age	Bonafini, Chae, Park, and Jablokow (2017); Breslow et al. (2013); Guo and Reinecke (2014); Konstan et al. (2015); Morris et al. (2015); Schulze (2014); Zhang et al. (2016).	Pre-course survey
4	Gender	Bayeck et al. (2018); Bonafini (2017); Breslow et al. (2013); Konstan et al. (2015); Schulze (2014).	Pre-course survey
5	Intention to complete	Bonafini (2017); Koller et al. (2013); Konstan et al. (2015).	Pre-course survey
6	Groups	Kulkarni et al. (2015); Sinha (2014); Wen (2016); Zheng, Rosson, Shih, and Carroll (2015).	Grouping database
7	Previous online Learning experience	Morris et al. (2015); Zhang et al. (2016).	Pre-course survey
8	Employment status	Cisel (2014); Morris et al. (2015).	Pre-course survey
9	Personal interest	Belanger and Thornton (2013); Dillahunt et al. (2014); Hew and Cheung (2014); Kizilcec and Schneider (2015).	Pre-course survey
10	Connect with others	Belanger and Thornton (2013); Zheng, Rosson, Shih, and Carroll (2015).	Pre-course survey
11	Institution	Kizilcec and Schneider (2015).	Pre-course survey
12	Professor	Kizilcec and Schneider (2015).	Pre-course survey
13	Earn certificate	Bonafini (2017); Kizilcec and Schneider (2015);	Pre-course survey

14	Related to academic program	Kizilcec and Schneider (2015).	Pre-course survey
15	Current job	Christensen et al. (2013); Kizilcec and Schneider (2015); Zheng, Rosson, Shih, and Carroll (2015).	Pre-course survey
16	Future job	Kizilcec and Schneider (2015); Zheng, Rosson, Shih, and Carroll (2015).	Pre-course survey
17	Friend take	Brown et al. (2015).	Pre-course survey
18	MOOC completion	Dependent variable.	Coursera database

Findings

This section presents findings from the statistical analyses conducted to examine the characteristics of MOOC learners who participated in this study, their preferences of working in groups, their motivations of taking this MOOC, and which demographics and motivations factors predict MOOC completion. Statistical analyses were performed using SPSS and R-Studio.

Learner Characteristics

To address our first research question "What are the characteristics of MOOC learners who participated in this study?" we analyzed participants' demographics. Demographics show that students who participated in this study (n = 655) came from all over the world (see Figure 1 for participants' locations on a world map). Table 2 presents the top ten countries where the learners were located. Chinese learners accounted for the largest number of volunteers participating in this grouping study (25.2%), followed by learners from the United States (17.9%).
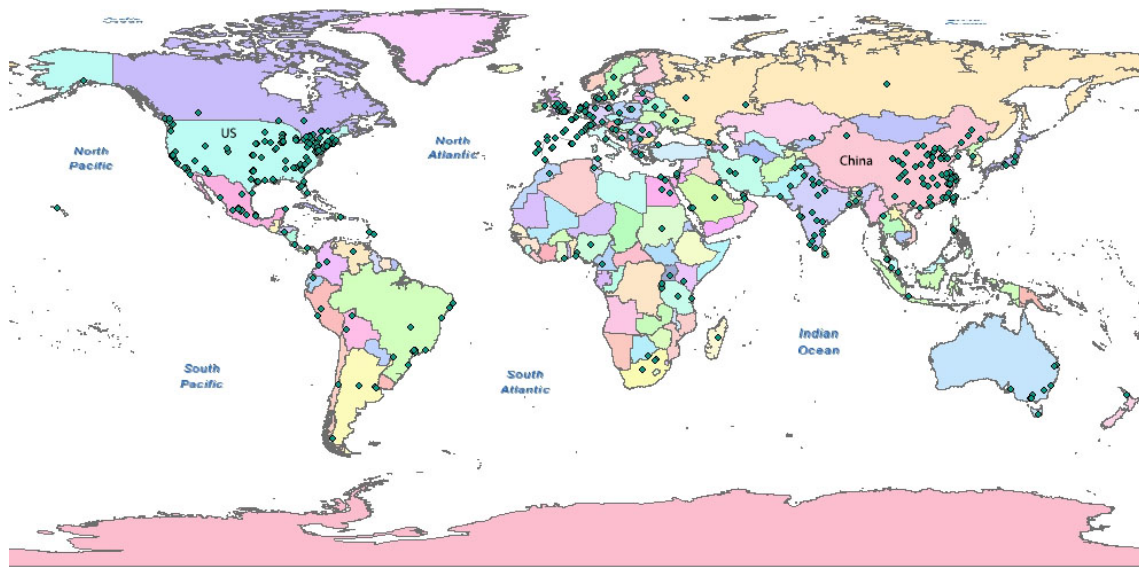
Figure 1. Location of the participants.

Table 2

Distribution of Participants by Country

Country	Frequency	%	Cumulative %
China	165	25.2	56.6
United States	117	17.9	74.5
India	59	9.0	83.5
Mexico	24	3.7	87.2
Canada	18	2.7	89.9
Brazil	16	2.4	92.4
Taiwan	15	2.3	94.7
Egypt	12	1.8	96.5
Spain	12	1.8	98.3
Nigeria	11	1.7	100.0
Others	206	31.5	31.5
Total	655	100.0	

When analyzing participants' demographics, we noticed that our participants are comprised of a larger percentage of female (61.3%) than male (38.7%) learners. Whereas, there is almost an equal number of female (48%) and male (52%) learners from the total population of 39069 who enrolled in this MOOC. Data on 31 out of 655 participants were excluded from our original dataset due to missing records for several variables. Valid records of 624 participants were used in this study.

Participants classified themselves as belonging to one of four English levels: Poor (6.1%), Basic (23.7%), Fluent (41.7%), or Native (28.5%). Participants were categorized into six age levels, which were ages 10-19 (10.8%), 20-29 (35.8%), 30-39 (20.5%), 40-49 (14.7%), 50-59 (11.7%), and 60 and above (6.5%). Sixty-five percent of the participants indicated an intention to complete the entire course when given the options of choosing either to complete all, most, some, or none of the course modules in the pre-course survey. For comparison purpose, thirty percent of the total population planned to complete this MOOC.

Learner Grouping Preferences

To address our second research question "What are the learners' preferences related to working in groups?" we used the grouping preference question in the pre-course survey that asked participants to rank their preferences regarding working in groups by marking the most important factor as 1 and the least important as 9. Results synthesized in Table 3 show that participants' first preference for participating in online groups was to work with people whose native language was the same as theirs. Participants' second grouping preference was to be grouped with others who had similar intentions of completing the course (e.g., complete the whole course, most of the course modules and assignments, or none of those). Their third preference was to be grouped with others who had a similar availability to join group meetings. Although the researchers grouped learners according to their identified preferences, students indicated in their post-course survey that many participants had difficulties in arranging online meetings due to the time zone differences and schedule conflicts.

Table 3

Ranking of Participants' Grouping Preferences

	Ranking	Mean	*SD*
Language spoken	1	2.79	2.073
Intention to complete	2	3.15	1.900
Similar schedule to work Together	3	4.15	2.505
Education level	4	4.29	1.901
Country living is different	5	4.65	2.389
Age	6	5.15	2.202
Similar occupation	7	6.25	2.085
Country living is the same	8	6.93	1.680
Gender	9	7.65	1.579

Note. SD = Standard deviation.

Learners' Motivations for Taking this MOOC

To address our third research question "What are the learners' motivations for taking this MOOC?" we used the motivation question from the pre-course survey. The motivation question was stated as follows: "Please rate the importance of the following reasons for you to enroll in this course on a scale of 1-5 (1 as not at all important, 5 as absolutely critical) in the statements below." Statements listed included: "I am interested in taking a course from this particular institution;" "I am interested in taking a course from this particular professor(s);" I am interested in earning a certificate;" "I am interested in connecting with other students;" "I have friends taking this course;" "The course relates to my current academic program;" "The course relates to my current job;" and "The course will be helpful for me to get a new job."

Results show that participants rated taking this course because of their friends as most the important reason, with a mean score of 4.2 as shown in Table 4. Other important factors that emerged from participants' responses were: because of the MOOC professors ($\bar{x} = 3.03$), institution offering the MOOC ($\bar{x} = 2.37$), and participants' personal interest ($\bar{x} = 2.27$). Table 4 seems to suggest that learners tend to be more socially and extrinsically motivated since they enrolled in the course because their friends were also taking it.

Table 4

The Importance of Motivation Factors for Enrolling in MOOC

	Mean	Std. Error	SD
Friends	4.20	.031	.782
Professors	3.03	.042	1.068
Institution	2.37	.050	1.264
Personal Interest	2.27	.045	1.147
New Job	2.05	.046	1.181
Current Job	2.05	.046	1.181
Academic Program	2.05	.046	1.181
Connect with Others	2.04	.043	1.088
Earn Certificate	1.35	.032	.807

Note. SD = Standard deviation.

Demographics and Motivation Factors Predicting the Probability of MOOC Completion

Stepwise binomial logistic regression was used to build answers to the fourth research question: "Which demographics and motivational factors predict the probability of MOOC completion?" In this procedure, an interactive process was used for variable selection. The investigators started by performing a saturated model to map out which factors may affect the probability of MOOC completion. Then, parameter estimates were removed when identified as nonsignificant (p-value greater than 0.05). After excluding these nonsignificant parameters, the model was refitted and the p-values of the remaining parameter estimates were rechecked to assure that all variables with significant p-value were included in the model. The lowest

Akaike Information Criterion (AIC) (Akaike, 1973) was used to decide for the model that contained the best predictor subset.

The saturated model contains demographics parameter estimates such as education level (Education), age (Age), gender (Gender), employment status (Employment: full time/part time/not working), and English proficiency (English_Level). The model also includes the parameter estimates: students assigned to work in groups according to their preferences (Groups), students' motivation for taking the MOOC such as personal interest (Personal_Int), interest in connecting with others (Connect_w_Others), course offered by a certain institution (Institution) or professor they like (Professor), relationship of MOOC content to their academic program (Academic_Pgm), relationship of MOOC content to their current job responsibilities (Current_Job), MOOC fostering a potential skill participants might need in their future job (Future_Job), intention of completion (Intent_Completion), participants' desire to earn a certificate (Earn_Certificate), and friends' participation in the same MOOC (Friends).

Results from the saturated model (Model 1) displayed in Table 5 present Gender, English_Level, Education, Employment, Personal_Int, Connect_w_Others, Earn_Certificate, Current_Job, Future_Job, Friends, and Groups as not significant factors in predicting the probability of learners' MOOC completion. On the other hand, the variables Age, Institution, Professor, Academic_Pgm, and Intent_Completion are significant (p-value < 0.05) when considering course completion (AIC = 752.46 and G^2 = 750.12).

Table 5

AIC Comparison Among Models

	Model	AIC
Model 1	Completion ~ Gender + English_Level + Education + Employment + Age + Personal_Int + Connect_w_Others + Institution + Professor + Earn_Certificate + Academic_Pgm + Current_Job + Future_Job + Friends + Intent_Completion + Groups	752.46
Model 2	Completion ~ Age + Institution + Professor + Academic_Pgm + Intent_Completion	718.12
Model 3	Completion ~ Age + Institution + Academic_Pgm + Intent_Completion	717.45

The researchers reran the model with only the significant predictors labeled as Model 2 in Table 5. Results show all variables as significant (AIC = 718.12 and G^2 = 750.12) with exception of the variable Professor (p > 0.05), indicating that learners' desire of taking this MOOC with a specific professor is not a significant factor affecting course completion when compared to other factors such as student age, the institution hosting the MOOC, MOOC content related to the student's current academic program, and the student's intention to complete the course.

The investigators removed the variable Professor from the model and reran the analysis (Model 3). Results from multiple binomial logistic regression on Model 3 (Table 6) presented Age5 (p = 0.00404), Age6 (p = 0.00306), Institution3 (p = 0.01757), Institution5 (p = 0.01513), Academic_Pgm3 (p = 0.04061),

Intent_Completion5 (p = 0.03448) as statistically significant when considering MOOC completion (AIC = 717.45 and G^2 = 750.12). Model 3 also presents an improvement of fit with a lower AIC when compared to previous models as shown in Table 5.

Table 6

Model 3: Multiple Binomial Logistic Regression Presenting Age, Institution, Academic Program, and Intention to Complete the MOOC as Significant Variables

Coefficients	Estimate	z value	p-value
(Intercept)	-3.01157	-4.166	3.10E-05 ***
Age2	0.14526	0.426	0.67022
Age3	0.39474	1.061	0.28885
Age4	0.65473	1.701	0.08901
Age5	1.17022	2.875	0.00404 **
Age6	1.40091	2.961	0.00306 **
Institution2	0.52938	1.831	0.06708
Institution3	0.68319	2.375	0.01757 *
Institution4	0.30471	0.873	0.3824
Institution5	1.00574	2.429	0.01513 *
Academic_Pgm2	-0.08525	-0.316	0.752
Academic_Pgm3	0.53921	2.047	0.04061 *
Academic_Pgm4	-0.0878	-0.247	0.80479
Academic_Pgm5	0.352	0.728	0.46671
Intent_Completion2	-1.42697	-1.185	0.23612
Intent_Completion3	0.48371	0.692	0.48911
Intent_Completion4	1.13618	1.704	0.08844
Intent_Completion5	1.40575	2.114	0.03448 *
Null deviance	750.12 on 629 degrees of freedom		
Residual deviance	681.45 on 612 degrees of freedom		
AIC	717.45		

Note: *** p< 0.001; ** p<0.01; * p<0.05

As shown in Table 7, the odds of completing a MOOC for participants who are at Age5 (50 to 59 years old) over the odds of completing a MOOC for participants who are at Age1 (up to 19 years old) is exp (1.17022)

= 3.22, meaning that the probability of MOOC completion increases by a multiplicative factor of 3.22 for participants between the ages of 50 to 59 in comparison to participants less than 19 years old.. Likewise, the odds of completing a MOOC for participants who are at Age6 (above 60 years old) over the odds of completing a MOOC for participants who are at Age1 (up to 19 years old) is exp (1.40091) = 4.06, meaning that for older participants the probability of MOOC completion is even bigger, increasing by a multiplicative factor of 4.06.

Table 7

Odds Ratio of Multiple Binomial Logistic Regression Coefficients in Model 3

Coefficients	Odds ratio	Confidence interval (2.5%, 97.5%)	
(Intercept)	0.04921413	0.0099133	0.1808451
Age2	1.15633887	0.60284086	2.312948
Age3	1.48400145	0.72460047	3.1392124
Age4	1.92461899	0.91496897	4.1674627
Age5	3.22269393	1.47095305	7.2976927
Age6	4.05889711	1.61618906	10.4110257
Institution2	1.69787955	0.96343306	2.9999831
Institution3	1.9801884	1.12999262	3.4980814
Institution4	1.35623809	0.68222608	2.6866481
Institution5	2.73392595	1.2105576	6.1669373
Academic_Pgm2	0.91828417	0.53801261	1.5525023
Academic_Pgm3	1.71464723	1.02195189	2.8742068
Academic_Pgm4	0.91594201	0.44841206	1.8157699
Academic_Pgm5	1.42191421	0.53264162	3.6086101
Intent_Completion2	0.24003458	0.01122428	2.0920954
Intent_Completion3	1.62208169	0.45749282	7.7080725
Intent_Completion4	3.1148375	0.95372359	14.15117
Intent_Completion5	4.07858993	1.25417009	18.4701751

The odds of completing a MOOC for students who perceive the institution hosting the MOOC as moderately important (*Institution3*) is 1.98 times greater over the odds of students who perceive the institution hosting the MOOC as not important at all (*Institution1*). Similarly, the odds of completing a MOOC for students who perceive the institution hosting the MOOC as very important (*Institution5*) over the odds of students

who perceive the institution hosting the MOOC as not important at all (*Institution1*) is 2.73. These results indicate that each one-point increase in the scale of importance for the institution hosting the MOOC is associated with MOOC completion increasing by a multiplicative factor of 1.98 and 2.73, respectively, for the ones who perceive the institution as moderately important and very important.

The odds of completing a MOOC for students who perceive that it is moderately important for the MOOC to be aligned with their academic program (*Academic_Pgm3*) over the odds of students who perceive that is not important at all for the MOOC be aligned with their academic program (*Academic_Pgm1*) is exp (0.53921) = 1.71. This means that each one-point increase in the scale of importance for an academic program is associated with the MOOC completion increasing by a multiplicative factor of 1.71.

The odds of completing a MOOC for students who strongly agree with the statement of intention to complete (*Intent_Completion5*) over the odds of students who indicated no intention to complete (*Intent_Completion1)* is exp (1.40575) = 4.08. This means that the probability of MOOC completion increases by factor of 4.08 for participants who are initially strongly committed with the intention to complete the course. The researchers also explored the interaction effect between the independent variables, however, none of these interactions were significant.

Discussion and Conclusion

This study shows that age, the institution hosting the MOOC, alignment with students' academic needs, and students' intention to complete the course can affect the probability of students' completion of a MOOC. The results are in line with the literature (e.g., Morris et al., 2015; Schulze, 2014; Zhang et al., 2016) in showing that older participants tend to achieve a higher course completion rate. This study extends the literature by indicating that the age of participant relates to MOOC completion, and older students (age > 50 years old) present a higher probability of completing a MOOC when compared with young ones.

It also sheds light on the importance of MOOCs providing experiences that add to students' current academic experiences as well as the importance of MOOCs being hosted by institutions with high academic reputation. As the majority of MOOC students are college degree holders (Christensen et al., 2013; Despujol, Turró, Busquéis, & Cañero, 2014), it makes perfect sense that when students expect that a MOOCs content will add knowledge to their current academic experiences, it increases their probability of completing the MOOC. This result adds to the literature that points out that students tend to register in MOOCs to learn new things, gain understanding of the subject matter, and to develop professional skills (Belanger & Thornton, 2013; Christensen et al., 2013).

In order to fulfill students' desire to register in a MOOC that is aligned with their academic needs, this study suggests a focus on making MOOC goals and content as clear as possible for its audience. By doing so, a MOOC can attract students who are looking for an experience that is aligned with their academic expectations, avoiding simply curious enrollments, which may diminish subsequent students' dropout. With this, MOOC providers should explicitly inform their potential students about the characteristics of MOOC content and how students may use the knowledge that will be acquired in that MOOC.

Based on the idea that "MOOCs enable learning with the best" (Davis et al., 2014, p. 6), it is intuitively known that an institution's reputation may motivate students' enrollment in MOOCs. However, this study advances the field by showing how much the reputation of an institution has the potential to affect the probability of students completing a MOOC. From an alternate perspective, it is also possible that the creation of a MOOC may enhance an institution's reputation as reported by Jansen and Schuwer (2015).

In addition to discussing the variables that relate to MOOC completion, it is also important to discuss the examined variables that did not influence the likelihood of completion. Results from the multiple binomial regression model show that variables such as gender, student personal interest, connection with others, friends and groups did not play a role on the probability of students' completion in this MOOC. Although "taking course because of friends who also took it" was rated by learners as the most important motivation factor for enrolling in this course, it did not appear as a significant predictor of MOOC completion. This factor may boost MOOC registration as suggested by Schulze (2014), but not MOOC completion as reported in this study.

Another surprising result from this study is the lack of effect on completion when students work in online groups. This result contradicts the literature (e.g., Kulkarni, et al., 2015; Williams et al., 2006), and to understand this result it is important to look at the design of the group work implemented in this MOOC. The lack of support and monitoring of students' group work process may explain why group work failed to increase MOOC completion rates. Moreover, students' group work activities were not facilitated or assessed by MOOC instructors or researchers.

Assigning learners to work in groups in MOOCs presents many challenges because of the heterogeneity in learner population, such as differences in education levels, cultural backgrounds, and study schedules. There seems not to be a perfect grouping mechanism that satisfies the needs of all learners. In the next stage of our research, we hope to record learners' interactions and learning behaviors as they engage in group activities in various social media applications (e.g. Skype and discussion forums) and use these data to understand how learners could benefit from MOOCs. These additional data could also help us to improve the grouping interventions, and eventually provide a better MOOC experience to the learners.

It is also worth noting that some learners didn't meet online with others regardless of being assigned into groups (as reported by participants in their post-course survey), a factor which may have contributed to the lack of effect that groupwork had on course completion. Feedback provided by students in the post-course survey informed the lack of monitoring students' group activity. We hypothesize that this could be one of the reasons why assigning learners to work in groups did not work in this study. Thus, further implications of this study suggest to MOOC instructors assigning teaching assistants (TAs) and/or group leaders to student groups as ways of providing assistance and monitoring their work process. These TAs could be recruited from learners who have completed the MOOC previously and are willing to assist others in taking the course. Another way to foster participants' group work would be assigning roles to each group member such as group leader and meeting coordinator. Meanwhile, data on the communication and interaction among team members in both synchronous and asynchronous media will be collected and analyzed to inform the design and facilitation of group work in the next phase of this grouping research. In the end, the authors expect that MOOC instructors and MOOC providers should be aware of students' motivations for

enrollment and the demographics that impact the likelihood of students' MOOC completion so that this information may be used to shape the course content and format to better support learners.

Limitations

Although this study identifies variables that impact MOOC completion, it is not possible to infer the reasons why those variables are significant. We can speculate that age plays an important role in affecting MOOC completion since older people may have more time to take the course and may have better time management and self-regulation skills. However, more investigation is needed to gain a deeper understanding of the reasons why older people have a higher probability of completing MOOCs.

Another limitation of this study is its small sample size compared to the large number of students who enrolled in this MOOC. Because of effect size and subject taught in this MOOC, the findings may not be generalizable to other MOOCs. This effect size could be overcome with studies comprising multiple MOOC cohorts. In the next phase of this study, the investigators aim to implement follow up interviews with students to collect feedback about their group work process and suggestions on how to improve their group work experiences. Future plans also include researching indicative variables of course completion as described in Pursel, Zhang, Jablokow, Choi, and Velegol (2016) such as course activities, number of videos watched, and number of posts made in the discussion forum as predictors of MOOC completion.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Caski (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.

Arendale, D. R., & Hane, A. R. (2014). Holistic growth of college peer study group participants: Prompting academic and personal development. *Research & Teaching in Developmental Education, 31*(1), 7-29. Retrieved from http://search.proquest.com/docview/1658878111?accountid=13158

Bayeck, R. Y., Hristova, A., Jablokow, K., & Bonafini, F. C. (2018). Exploring the relevance of single-gender group formation: What we learn from a massive open online course (MOOC). *British Journal of Educational Technology, 49*(1), 88-100. doi: 10.1111/bjet.12528

Belanger, Y., & Thornton, J. (2013). *Bioelectricity: A quantitative approach: Duke university's first MOOC.* Duke University. Retrieved from http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/6216/Duke_Bioelectricity_MOOC_Fall2012.pdf

Berger, D., & Wild, C. (2016). Turned on, tuned in, but not dropped out: Enhancing the student experience with popular social media platforms. *European Journal of Law and Technology, 7*(1). Retrieved from http://ejlt.org/article/view/503/639

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

Bonafini, F. C., Chae, C., Park, E., & Jablokow, K. W. (2017). How much does student engagement with videos and forums in a MOOC affect their achievement? *Online Learning, 21*(4), 223-240. doi: 10.24059/olj.v21i4.1270

Breslow, L., Pritchard, D. E., DeBoer, J., 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*(1), 13-25. Retrieved from https://files.eric.ed.gov/fulltext/EJ1062850.pdf

Brown, R., Lynch, C., Wang, Y., Eagle, M., Albert, J., Barnes, T., ... & McNamara, D. S. (2015, June). Communities of performance & communities of preference. In *Proceedings of EDM Workshops* (pp. 12-19). Madrid, Spain. Retrieved from http://ceur-ws.org/Vol-1446/edm2015ws_proc.pdf

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). The MOOC phenomenon: Who takes massive open online courses and why? *Social Science Research Network Journal.* http://dx.doi.org/10.2139/ssrn.2350964

Cisel, M. (2014). Analyzing completion rates in the first French xMOOC. In *Proceedings of the European MOOC Stakeholder Summit 2014, 26.* Retrieved from
https://www.openeducationeuropa.eu/sites/default/files/legacy_files/asset/From-field_37_6.pdf

Cross, S. (2013). *Evaluation of the OLDS MOOC curriculum design course: Participant perspectives and experiences.* OLDS MOOC Project, Milton Keynes. Retrieved from
http://oro.open.ac.uk/37836/1/EvaluationReport_OLDSMOOC_v1.0.pdf

Davis, H. C., Dickens, K., Leon Urrutia, M., Vera, S., del Mar, M., & White, S. (2014). MOOCs for universities and learners: An analysis of motivating factors. In *Proceedings of the 6th International Conference on Computer Supported Education CSEDU* (pp. 105-116). Spain: Springer. Retrieved from
https://eprints.soton.ac.uk/363714/1/DavisEtAl2014MOOCsCSEDUFinal.pdf

Despujol, I. M., Turró, C., Busquéis, J., & Cañero, A. (2014, October). Analysis of demographics and results of student's opinion survey of a large scale MOOC deployment for the Spanish speaking community. In *Frontiers in Education Conference (FIE), 2014 IEEE* (pp. 1-8). Madrid, Spain: IEEE.

Dillahunt, T., Wang, B., & Teasley, S. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open and Distributed Learning, 15*(5). http://dx.doi.org/10.19173/irrodl.v15i5.1841

Engle, D., Mankoff, C., & Carbrey, J. (2015). Coursera's introductory human physiology course: Factors that characterize successful completion of a MOOC. *The International Review of Research in Open and Distributed Learning, 16*(2). http://dx.doi.org/10.19173/irrodl.v16i2.2010

Gil-Jaurena, I., Callejo-Gallego, J., & Agudo, Y. (2017). Evaluation of the UNED MOOCs implementation: Demographics, learners' opinions and completion rates. *The International Review of Research in Open and Distributed Learning, 18*(7). http://dx.doi.org/10.19173/irrodl.v18i7.3155

Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles: A pedagogical approach from the learner's perspective. *eLearning Papers, 33*, 1-6. Retrieved from http://r-libre.teluq.ca/596/1/In-depth_33_4.pdf

Guo, P. J., & Reinecke, K. (2014, March 4-5). Demographic differences in how students navigate through MOOCs. In *Proceedings of the First ACM Conference on Learning@ Scale Conference* (pp. 21-30)*.* Atlanta, Georgia. Retrieved from https://cs.rochester.edu/hci/pubs/pdfs/edX-MOOC-demographics-and-navigation_LAS-2014.pdf

Gütl, C., Rizzardini, R. H., Chang, V., & Morales, M. (2014). Attrition in MOOC: Lessons Learned from Drop-Out Students. In L. Uden, J. Sinclair, YH. Tao, D. Liberona (Eds.), *Learning technology for education in cloud. MOOC and big data* (pp. 37-48). Springer, Cham.

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

Hiltz, S. R. (1998). *Collaborative learning in asynchronous learning networks: Building learning communities.* Invited address at the WEB98 Conference, Orlando, Florida. Retrieved from https://files.eric.ed.gov/fulltext/ED427705.pdf

Ho, A. D., Reich, J., Nesterko, S., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The first year of open online courses (HarvardX and MITx Working Paper No. 1). Retrieved from https://dash.harvard.edu/handle/1/11987422

Jablokow, K., Matson, J. V., & Velegol, D. (2014). *A multidisciplinary MOOC on creativity, innovation, and change: Encouraging experimentation and experiential learning on a grand scale.* In *Proceedings of the 2014 ASEE Annual Conference & Exposition,* Indianapolis, IN. Retrieved from http://www.asee.org/public/conferences/32/papers/9669/view

Jansen, D., & Schuwer, R. (2015). Institutional MOOC strategies in Europe*. Status report based on a mapping survey conducted in October - December 2014. The European Association of Distance Teaching Universities.* Retrieved from http://www.eadtu.eu/documents/Publications/OEenM/Institutional_MOOC_strategies_in_Europe.pdf

Jones, B. M. (1997, March). *Study groups: Collaboration and conflagration.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. Retrieved from http://files.eric.ed.gov/fulltext/ED408362.pdf

Kellogg, S., Booth, S., & Oliver, K. (2014). A social network perspective on peer supported learning in MOOCs for educators. *The International Review of Research in Open and Distributed Learning, 15*(5). http://dx.doi.org/10.19173/irrodl.v15i5.1852

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 (TOCHI), 22*(2), 6. http://dx.doi.org./10.1145/2699735

Koller, D., Ng, A., Do, C., & Chen, Z. (2013). Retention and intention in massive open online courses: In depth. *Educause Review, 48*(3), 62-63. Retrieved from https://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses-in-depth

Konstan, J. A., Walker, J. D., Brooks, D. C., Brown, K., & Ekstrand, M. D. (2015). Teaching recommender systems at large scale: Evaluation and lessons learned from a hybrid MOOC. *ACM Transactions on Computer-Human Interaction (TOCHI), 22*(2), 10. https://doi.org/10.1145/2728171

Kulkarni, C., Cambre, J., Kotturi, Y., Bernstein, M. S., & Klemmer, S. R. (2015, February). Talkabout: Making distance matter with small groups in massive classes. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada,* 1116-1128. Retrieved from http://www.cs.cmu.edu/~chinmayk/assets/pdfs/talkabout-cscw.pdf

Lim, S., Coetzee, D., Hartmann, B., Fox, A., & Hearst, M. A. (2014, March). Initial experiences with small group discussions in MOOCs. In *Proceedings of the First ACM Conference on Learning@ Scale Conference, Atlanta, Georgia, USA*, 151-152. Retrieved from https://dl.acm.org/citation.cfm?id=2567854

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

Locke, E. A. (1982). Relation of goal level to performance with a short work period and multiple goal levels. *Journal of Applied Psychology, 67*(4), 512-514. http://dx.doi.org/10.1037/0021-9010.67.4.512

Macleod, H., Haywood, J., Woodgate, A., & Alkhatnai, M. (2014). Emerging patterns in MOOCs: Learners, course designs, and directions. *TechTrends, 59*(1), 56–63. http://dx.doi.org/10.1007/s11528-014-0821-y

Malan, D. J. (2013). Implementing a massive open online course (MOOC). *Journal of Computing Sciences in Colleges, 28*(6), 136-137. Retrieved from https://dl.acm.org/citation.cfm?id=2460182

Morris, N. P., Hotchkiss, S., & Swinnerton, B. (2015). *Can demographic information predict MOOC learner outcomes?* In *Proceedings of the EMOOC Stakeholder Summit,* Belgium. Retrieved from https://www.researchgate.net/publication/278392770_Can_demographic_information_predict_MOOC_learner_outcomes

Onah, D. F., Sinclair, J., & Boyatt, R. (2014). Dropout rates of massive open online courses: Behavioural patterns. In *EDULEARN14 Proceedings* (pp. 5825-5834). Barcelona: IATED Academy. Retrieved from http://wrap.warwick.ac.uk/65543/

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning, 32*, 202–217. doi: 10.1111/jcal.12131

Radford, A. W., Coningham, B., & Horn, L. (2015). MOOCs: Not just for college students—How organizations can use MOOCs for professional development. *Employment Relations Today, 41*(4), 1–15. http://dx.doi.org/10.1002/ert.21469

Schulze, A. S. (2014). *Massive open online courses (MOOCs) and completion rates: Are self-directed adult learners the most successful at MOOCs?* (Unpublished doctoral dissertation). Pepperdine University, California, United States. Available from ProQuest Dissertations & Theses Global. (1549976283)

Sinha, T. (2014, February). Together we stand, together we fall, together we win: Dynamic team formation in massive open online courses. In *The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014), Bangalore, India,* 107-112. doi: 10.1109/ICADIWT.2014.6814694.

Wen, M. (2016). *Investigating virtual teams in massive open online courses: Deliberation-based virtual team formation, discussion mining and support* (Doctoral dissertation, Carnegie Mellon University). Retrieved from https://www.lti.cs.cmu.edu/sites/default/files/wen,%20miaomiao%20-%20CMU-LTI-16-015%20-%202016.pdf

Williams, E. A., Duray, R., & Reddy, V. (2006). Teamwork orientation, group cohesiveness, and student learning: A study of the use of teams in online distance education. *Journal of Management Education*, *30*(4), 592-616. https://doi-org.ezproxy.lib.vt.edu/10.1177/1052562905276740

Xiong, Y., Li, H., Kornhaber, M. L., Suen, H. K., Pursel, B., & Goins, D. D. (2015). Examining the relations among student motivation, engagement, and retention in a MOOC: A structural equation modeling approach. *Global Education Review*, *2*(3). Retrieved from https://ger.mercy.edu/index.php/ger/article/view/124

Yuan, L., & Powell, S. (2013). *MOOCs and open education: Implications for higher education* [White Paper]. Retrieved from http://www.smarthighered.com/wp-content/uploads/ 2013/03/MOOCs-and-Open-Education.pdf

Zhang, Q., Peck, K. L., Hristova, A., Jablokow, K. W., Hoffman, V., Park, E., & Bayeck, R. Y. (2016). Exploring the communication preferences of MOOC learners and the value of preference-based groups: Is grouping enough? *Educational Technology Research and Development*, *64*(4), 809-837. https://doi.org/10.1007/s11423-016-9439-4

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 - CSCW '15*, (pp. 1882-1895). Vancouver: ACM. http://dx.doi.org/10.1145/2675133.2675217

Zheng, Z., Vogelsang, T., & Pinkwart, N. (2015). The impact of small learning group composition on student engagement and success in a MOOC. In *Proceedings of the 8th International Conference of Educational Data Mining* (pp. 500-503).

Zhong, S.-H., Zhang, Q.-B., Li, Z.-P., & Liu, Y. (2016). Motivations and challenges in MOOCs with eastern insights. *International Journal of Information and Education Technology, 6*(12), 954–960. http://dx.doi.org/10.7763/ijiet.2016.v6.824