



TYPES OF PARTICIPANTS' BEHAVIOURS IN A MASSIVE OPEN ONLINE COURSE

Tali Kahan, Tal Soffer, Rafi Nachmias, Tel Aviv University, Israel

Introduction

In recent years, there has been a proliferation of Massive Open Online Courses (MOOCs). Initiatives like Coursera, edX and Udacity provide platforms for higher education institutions to develop and deliver online courses to the general public. These courses are usually offered free of charge, with no preconditions or commitment, and attract massive numbers of registrants from around the world. From a pedagogical perspective, the courses (sometimes referred to as xMOOCs) follow a cognitive-behaviourist approach (Rodriguez, 2012; Daniel, 2012) and consist of diverse learning resources such as reading materials, video lectures, discussion forums, and assessments. Some courses give participants a statement of accomplishment or certificate upon successful completion of course requirements.

A central criticism about MOOCs refers to the relatively low completion rates of participants (Kizilcec, Piech & Schneider, 2013) with 10% or less of the course registrants earning a statement of accomplishment (Daniel, 2012; Breslow et al., 2013). Ho et al. (2014) argued that certification rates are a misleading representation of the diverse ways in which registrants are engaging with MOOCs; certificate earning is only one possible way, while others may include watching videos, reading texts, focusing on assessments, and so on. Kizilcec et al. (2013) argued that the categorization of MOOCs learners into those who pass the class, and everyone else, makes no allowances for learners who choose to stay engaged with the course and participate in some aspects of it, without earning a certificate.

Subsequently, the purpose of this study was to gain more insights into different types of participants' behaviours in MOOCs, by analyzing non-certificate earners' (NCEs) and certificate earners' (CEs) behaviours. Using a data mining methodology, the participants of a MOOC were classified into clusters, based on their activity in the main learning resources of the course: video lectures, discussion forums, and assessments. Thereafter, the different types of engagement and behaviours that emerged in the clusters were analyzed.

Background

Participants' engagement in MOOCs

Large-scale participation in MOOCs, together with data collected from the participants' interaction with the course, open up opportunities for studying student engagement at scale (Ramesh, Goldwasser, Huang, Daume III & Getoor, 2014). Several studies have examined the types of engagement and persistence of participants in MOOCs, based on different criteria. Kizilcec et al. (2013) identified four prototypical trajectories of engagement, based on the participants' actions in regard to the videos and assessments: learners who completed the majority of assessments; learners who did assessments infrequently, if at all, and engaged by watching video lectures; learners who did assessments at the beginning of the course, but then had a marked decrease in engagement; and learners who watched video lectures for only one or two assessment periods. Ho et al. (2014) presented another classification with four types: registrants who never access the courseware; non-certified registrants who access less than half of the available chapters; non-certified registrants who access more than half of the available chapters; and registrants who earned a certificate. Halawa, Greene and Mitchell (2014) identified four common patterns of persistence in MOOCs: continuous persistence, continuous persistence with extended absences, bursty persistence, and drop out.

Videos in MOOCs

Videos are a dominant medium in MOOCs, and are central to the student learning experience (Guo, Kim & Rubin, 2014; Diwanji, Simon, Marki, Korkut & Dornberger, 2014). The courses usually consist of several short videos, interspersed with quizzes. Glance, Forsey and Riley (2013) described the pedagogical principles underlying this structure as short videos that allow students to control the pace, to pause, to rewind, to explore and to return to the content. In addition, the format of short videos that cover a concept, associated with quizzes that provide formative assessment, enable the learners to engage in Mastery Learning - the possibility to achieve mastery of a concept before moving on to the next one, and Retrieval Learning - the act of enhancing long-term, factual memory by recalling information from short-term memory. Several studies have examined the use of videos by MOOCs participants. Some studies focused on the categorization of factors that affect students' engagement in the videos (Guo et al., 2014; Diwanji et al., 2014). Breslow et al. (2013) related to characteristics of video use by CEs.

Discussion forums in MOOCs

Discussion forums in MOOCs support social learning (Brinton et al., 2013). Online forums help to create a learning community through which learners generate knowledge (Li, 2004) and are intended to promote learners' motivation, reflection, and engagement (Onah, Sinclair & Boyatt, 2014). In MOOCs, the discussion forums constitute a primary means of interaction among the course participants and the instructors (Onah et al., 2014) with peer communication, support, and assessment becoming central, due to the large-scale participation versus the low number of instructors in a course (Onah et al., 2014). Nevertheless, when examining participants' engagement in MOOCs forums, the usage of

forums has been described as quite low, in general, with the discussion often involving a minority of course participants (Onah et al., 2014; Breslow et al., 2013; Ho et al., 2014). In addition, several studies have indicated that CEs are significantly more active in the forums than NCEs (Breslow et al., 2013; Kizilcec et al., 2013; Ho et al., 2014).

Assessments in MOOCs

MOOCs in their essence are open courses, thus, the course assessments are optional and may serve different participant goals, such as self-testing or receiving course credit. Due to the large number of participants in MOOCs, conducting assessments by humans is impossible (Glance et al., 2013; Sandeen, 2013) and different models of assessments are evolving. Among them are the automated quizzes with instant feedback, peer assessments, self-assessments, and machine grading. The automated quizzes are appropriate for subjects that can be assessed by objective means, whereas other methods are more appropriate for writing-based assessments (Glance et al., 2013; Sandeen, 2013). Breslow et al. (2013) noted that with free and easy registration to MOOCs, the courses include a large number of participants who may not have any interest in completing the course assessments. Ho et al. (2014) stated that many registrants engage in MOOCs without completing the assessments for credit.

The study

This study examined the different types of participants' behaviours in a MOOC, based on the participants' activity in the main learning resources of the course. The research questions were:

1. What are the types of engagement in the course?
2. What are the types of behaviours in regard to the course video lectures, discussion forums, and assessments?

Methodology

The study was conducted using a data mining methodology. Educational data mining is an emerging methodology that deals with developing methods for exploring the types of data that come from educational settings, and using those methods to better understand students and their learning environments (www.educationaldatamining.org).

The study examined one MOOC on plant biology, which was offered by Tel Aviv University in Coursera. The data mining was applied on a data set that documented the participants' actions during the course. Using cluster analysis, the course participants were classified into clusters, based on their activity in the main learning resources of the course. The different types of engagement and behaviours that emerged in the clusters were analyzed.

The studied MOOC

The course began on October 2013 and lasted 7 weeks. Each week, a different topic was covered via the course learning resources, which consisted of: professor announcements, reading recommendations, 50 short video lectures, around 40 interactive in-video questions, 7 discussion forums, 6 quizzes, and a final exam.

The video lectures were uploaded to the course website on a weekly basis, ranging from 5 to 9 new lectures per week. The participants could watch the videos online, and/or download them and watch them offline. The quizzes interspersed among the videos were presented in online mode only, and contained both open and close-ended questions, that were followed by feedback. Some of the questions contained hyperlinks to resources outside the course website, and some referred the students to the course discussion forums.

The discussion forums were arranged into 7 sub-forums according to topics, e.g. lectures, assignments, study groups. The participants were able to observe the discussions and participate by initiating new threads, adding posts, or commenting on their peers' posts. In addition, a voting mechanism allowed the participants to vote for posts in order to bring more attention to thoughtful and helpful posts.

The course assessments were based on weekly quizzes and a final exam. The quizzes consisted of 5 to 6 mostly close-ended questions. The participants were given 100 attempts to complete each quiz, and were required to submit each quiz within one week of its release, in order to receive credit for it (a Late Days mechanism allowed for late submissions without losing points). The quizzes remained accessible throughout the course, enabling the participants to use them for practice. The final exam was published during the last week of the course and consisted of close-ended questions. The exam due date was 2 weeks after its release, with a time limit of 3 hours to complete it. Participants who completed the course with a grade of at least 78% (comprised of 50% quizzes and 50% final exam) received a statement of accomplishment.

Participants

32,007 people registered for the course, and 68.4% of the registrants (21,889) started it. 10.6% of the participants who started the course (2,319) completed it and received a certificate.

Analysis

The data examined in this study was received from Coursera in database tables that contained over 1 million records of participants' actions during the course. Using MySQL queries, a set of variables were computed for each participant, displaying their use of the course components. A table was then created from these variables with 21,889 records (one record per participant who started the course), in order to perform a cluster analysis. Following a correlation test, 8 variables were chosen for the analysis: video variables: unique video lectures viewed online, unique video lectures downloaded, and unique in-video questions answered; discussion forum variables: total thread views, total threads opened, total comments; and

assessment variables: unique quizzes submitted, and final exam submitted. A Two-Step cluster analysis was applied separately for the NCEs and for the CE groups. This separation was essential in order to gain better insights into the different behaviours between the groups and within each group.

Findings

The clusters

The cluster analyses were applied in order to identify types of participants' behaviours during the course. The analyses revealed 8 clusters in total; 4 clusters per group. The average silhouette measure was 0.7 for the NCEs group analysis, and 0.6 for the CE group analysis. Table 1 presents the clusters and the average value of each cluster variable (for the variable called Exam Submitted, the mode is presented).

Table 1: The clusters found in the cluster analyses

Cluster id	Non-Certificate Earners (N=19,570)				Certificate Earners (N=2,319)			
	NCE-1	NCE-2	NCE-3	NCE-4	CE-1	CE-2	CE-3	CE-4
Cluster size	13,317 (68%)	1,935 (10%)	3,187 (16%)	1,131 (6%)	523 (23%)	1,461 (63%)	257 (11%)	78 (3%)
Unique videos downloaded	0.92	45.69	1.33	10.79	35.55	1.43	47.19	12.99
Unique videos viewed online	3.79	5.13	29.92	31.75	12.18	48.61	45.95	43.09
Unique video questions answered	2.12	2.60	21.43	20.71	4.21	33.54	32.21	29.67
Total thread views	0.75	1.32	4.34	21.59	19.58	24.74	34.10	271.18
Total threads opened	0.00	0.00	0.00	0.33	0.14	0.14	0.18	2.68
Total comments	0.04	0.03	0.19	1.28	0.34	0.75	0.75	18.72
Unique quizzes submitted	0.20	0.65	2.30	4.51	6.00	6.00	6.00	5.82
Exam Submitted	No (100%)	No (100%)	No (100%)	Yes (72.9%)	Yes (100%)	Yes (100%)	Yes (100%)	Yes (100%)

Types of engagement

The clusters obtained in the analyses were examined for the general level of engagement during the course, which was measured by overall activity in regard to learning resources. Five major types of engagement were identified. The first type appeared in cluster NCE-1, which characterized 13,317 participants with very low average values for all variables, which indicated very low activity in regard to all the main learning resources of the course. These participants barely experienced the course, and dropped out almost instantly; clearly, all of them were NCEs. They were thus named the Tasters. The second type of engagement emerged in cluster NCE-2, which was characterized by 1,935 NCEs who demonstrated, similarly to the first cluster, very low levels of engagement in all variables, except for an extensive downloading of video lectures. It is unknown whether they watched the videos offline or only stored them (perhaps for future use). They were thus named the Keepers.

A third type of engagement emerged in cluster NCE-3, which was characterized by 3,187 NCEs who demonstrated moderate activity in viewing the videos online and answering the in-video questions, and low activity in the forums and the assessments. They were named

the Partially Persisting. The fourth type of engagement appeared in cluster NCE-4, in which the participants demonstrated moderate to high average values in most of the variables. This cluster contained 1,131 participants who, on average, watched and downloaded a large portion of the videos, submitted a significant number of quizzes, and submitted the exam (most of them). Despite their high level of engagement during the course, they did not earn a statement of accomplishment. Therefore, they were named the Committed participants.

The fifth and last type of engagement during the course was apparent in clusters CE-1 to CE-4, which contained participants who used the videos and the assessments thoroughly, and used the forums in varied degrees. These clusters contained 2,319 participants, and all of them were CEs. They were named the Completers.

Video usage

The clusters resulting from the analyses were examined (excluding the first one due to negligible participation during the course) and three types of behaviours were identified in regard to the videos: watching the videos online, downloading the videos, and a combination of the two. The first type appeared in clusters NCE-3 and CE-2, and characterized 4,648 participants, from both the NCEs and the CEs groups, who viewed the videos mostly online. They were thus named the Onliners. The second type appeared in clusters NCE-2 and CE-1, and characterized 2,458 participants, from both the NCEs and the CEs groups, who mostly downloaded the videos. They were named the Offliners. The remaining clusters (NCE-4, CE-3 and CE-4) characterized the third type of video usage, which was a combination of viewing the videos online and downloading them. The total number of participants in these clusters was 1,466. They were named the Combiners.

Participation in the discussion forums

When analyzing the clusters (excluding the first one due to negligible participation during the course), two prominent types of behaviours in regard to the discussion forums were identified: passive participation and active participation. The first type was common in most clusters (NCE-2, NCE-3, CE-1, CE-2 and CE-3) and characterized 7,363 participants who mainly observed the forums to varied degrees, without taking an active part in the on-going discussions. The NCEs among them observed the forums very slightly, whereas the CEs observed them to a moderate degree. These participants were named the Observers. The second type of participation in the forums was much less common, and appeared in clusters NCE-4 and CE-4. These clusters contained 1,209 participants who, in addition to observing the forums, participated actively in the discussions. The NCEs among them observed the forums to a moderate degree and were active very slightly, whereas the CEs participated intensively both passively and actively. They were named the Active Participants.

Participation in the assessments

Three levels of participation in the assessments were identified in all clusters (excluding the first one due to negligible participation during the course): negligible, moderate, and, extensive. Cluster NCE-2 characterized 1,935 participants with negligible usage of the

assessments. Cluster NCE-3 characterized 3,187 participants with low usage of the quizzes, and no submission of the exam. These two clusters consisted of NCEs only (by definition). The remaining clusters (NCE-4, CE-1 to CE-4) characterized 3,450 participants who used the assessments extensively. It should be noted that some of them did not earn a certificate.

Table 2 summarizes the types of participants' behaviours that emerged in the analyses.

Table 2: Types of participants' behaviours

Cluster id	Non-Certificate Earners (N=19,570)				Certificate Earners (N=2,319)			
	NCE-1	NCE-2	NCE-3	NCE-4	CE-1	CE-2	CE-3	CE-4
Cluster size	13,317 (68%)	1,935 (10%)	3,187 (16%)	1,131 (6%)	523 (23%)	1,461 (63%)	257 (11%)	78 (3%)
Engagement	Tasters	Keepers	Moderately Persisting	Committed	Completers			
Videos		Offliners	Onliners	Combiners	Offliners	Onliners	Combiners	
Discussion forums		Slight Observers		Moderate Observers	Moderate Observers			Extensive Observers
				Slightly Active				Extensively Active
Assessments		Negligible	Moderate	Extensive	Extensive			

Discussion

Five major types of engagement during the course were identified, based on the participants' activity in regard to the main learning resources. Among the NCEs, a significant number of participants were Tasters, who barely experienced the course. However, with reference to criticism regarding relatively low completion rates in MOOCs, significant numbers of participants who demonstrated some level of engagement during the course were identified, ranging from the Moderately Persisting participants to the highly Committed participants, who used a large portion of the course resources. In addition, a group of NCEs, who expressed an interest in downloading a significant portion of the video lectures, was identified (the Keepers). It is unknown whether they watched the videos offline, or only stored them, perhaps intending to use them in the future.

Regarding video usage, three types of behaviours were identified: online, offline, and combined. These types may indicate different participant preferences for watching the videos, and may have significant implications on the learning process, as they facilitate different pedagogies. The Onliners may benefit from the scaffolding that exists online, such as the in-video quizzes that support Mastery Learning and Retrieval Learning. The Offliners, on the other hand, may benefit from the flexibility of watching the videos anywhere, also in cases where there is a lack of proper internet connectivity. It was found that the majority of the participants were Onliners. Interestingly, Offliners were found among the CEs group. These participants did not make an exhaustive use of the pedagogical affordances inherent in the online videos but still earned a certificate.

Regarding the discussion forums, two prominent types of participation were identified: passive and active. Consistent with the literature, the vast majority of discussion participants were found to be Observers, whereas only a small group, consisting mostly of CEs, held the

on-going discussions. One of the benefits of online learning, especially with great promise in MOOCs, is the possibility to be part of a large learning community. These findings raise questions regarding the extent to which the social learning potential of MOOCs is realized.

Finally, three levels of participation in the assessments were identified: negligible, moderate, and extensive. The CEs obviously used the assessments thoroughly. Among the NCEs who were engaged with the course to some extent, the quizzes were used in varied degrees, being more popular than the exam.

This study was conducted on one MOOC. More research is required on other MOOCs, in varied disciplines, target audiences, and structures. This study is part of a large-scale study which examines diverse aspects of participants' behaviours in MOOCs.

References

1. Breslow, L.; Pritchard, D.E.; DeBoer, J.; Stump, G.S.; Ho, A.D. and Seaton, D.T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. In *Research & Practice in Assessment*, 8, (pp. 13-25).
2. Brinton, C.G.; Chiang, M.; Jain, S.; Lam, H.; Liu, Z. and Wong, F.M.F. (2013). Learning about social learning in MOOCs: From statistical analysis to generative model. In *IEEE Transactions on Learning Technologies*, 7(4), (pp. 346-359).
3. Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. In *Journal of Interactive Media in Education*, 3. Retrieved from <http://jime.open.ac.uk/article/view/259>
4. Diwanji, P.; Simon, B.P.; Marki, M.; Korkut, S. and Dornberger, R. (2014). Success factors of online learning videos. In *International Conference on Interactive Mobile Communication Technologies and Learning*, (pp. 125-132).
5. Gance, D.G.; Forsey, M. and Riley, M. (2013). The pedagogical foundations of massive open online courses. In *First Monday*, 18(5). Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/4350/3673>
6. Guo, P.J.; Kim, J. and Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference*, (pp. 41-50).
7. Halawa, S.; Greene, D. and Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. In U. Cress & C.D. Kloos (eds.), *Proceedings of the European MOOC Stakeholder Summit*, (pp. 58-65).
8. Ho, A.D.; Reich, J.; Nesterko, S.; Seaton, D.T.; Mullaney, T.; Waldo, J. and Chuang, I. (2014). HarvardX and MITx: The first year of open online courses. In *HarvardX and MITx Working Paper No. 1*.

9. Kizilcec, R.F.; Piech, C. and Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, (pp. 170–179).
10. Li, Q. (2004). Knowledge building community: Keys for using online forums. In *TechTrends*, 48(4), (pp. 24–29).
11. Onah, D.F.O.; Sinclair, J. and Boyatt, R. (2014). Exploring the use of MOOC discussion forums. In *Proceedings of London International Conference on Education*, (pp. 1-4).
12. Ramesh, A.; Goldwasser, D.; Huang, B.; Daume III, H. and Getoor, L. (2014). Learning latent engagement patterns of students in online courses. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*.
13. Rodriguez, C.O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. In *European Journal of Open, Distance and E-learning*, 2012(II). Retrieved from <http://www.eurodl.org/?p=archives&year=2012&halfyear=2&article=516>
14. Sandeen, C. (2013). Assessment's place in the new MOOC world. In *Research & Practice in Assessment*, 8(1), (pp. 5-12).