What do We Know about MOOC Students Caring More about Content than Platform? - One Step toward Defining MOOC Learner Success

Yuan Wang ELLE.WANG@COLUMBIA.EDU
Columbia University, New York, NY 10027 USA

Abstract

Massive open online courses (MOOCs) has been highly praise as a disruptive innovation in online learning. However, the difficulty in defining learning success and evaluating course effectiveness has also been recognized as an ongoing challenge. Many have noted that traditional course metrics such as course completion and grades alone do not apply as well in the MOOC context. The present study proposed to look at a new group of variables "content-interests scores" and "platform-interest scores" to study how these two scores relate to different aspects of MOOC learners, in efforts to better define MOOC course success in the long run. This work-in-progress paper presented a combination of methods applying a principal component analysis and multiple regression models. The initial findings suggested that individual variables such as professional background, native language as well as mastery-goal orientation all showed significant influence on either "content-interests scores" or "platform-interest scores". Discussions and implication on these findings were also included.

1. Introduction

1.1 Background

MOOCs, massive open online courses, have garnered worldwide attention as a new platform for learning over the past three years. A growing community of researchers from various disciplines have studied engagement patterns in MOOC (e.g., Kizilcec, Piech, & Schneider, 2013; Clow, 2013), and found out MOOC learners exhibited highly varied ways of interacting with and using the courses they enroll in. Unlike in traditional online learning platforms, many MOOC learners do not consider completing a course their primary goal (Fini, 2009; Belanger & Thornton, 2013).

As such, challenges have been raised on how to define learner success or course effectivenss in the context of MOOCs (e.g., Breslow et al. 2013). This question is pertinent in that traditional metrics assessing the success of online learning may not be relevant to the needs and goals of MOOC learners. Existing research has focused on two dimensions in addressing this question: achievement and persistence. First, some studies of learner success have examined metrics mirroring a traditional classroom, such as grades of quizzes and exams (cf. Belanger & Thornton, 2013; MOOC @ Edinburgh, 2013). Second, studies linking success with persistence focused on learner usage and participation with course components such as videos (e.g., Guo, Kim, & Rubin, 2014) and discussion forums (e.g., Yang, Sinha, David, & Rose, 2013). Nevertheless, attempts made in evaluating MOOCs have been mostly based on online environments prior to the MOOC era.

1.2 Motivation of the study

It has been noted that MOOCs, as a new learning platform, presents learning and educational variables beyond those seen in conventional learning environments (Deboer et al., 2014; Whitmer et al., 2014). The present study proposed to explore one group of such MOOC-specific variables with regarding to students interest scores in course content and course platform inspired by a previous study investigating MOOC student motivation and course completion (Wang & Baker, 2015).

In a previous study on the same data set, it is shown that course completers tend to be more interested in the course content, whereas non-completers tend to be more interested in MOOCs as a type of learning experience (Wang & Baker, 2015) via correlating student answers on ten MOOC-specific motivational items in an early-course survey. Although the previous study focused on "course completion", unarguably one of the most prevalent traditional course metrics; it pointed out a direction in that the study showed that it appears that students who are particularly motivated by the new and unique aspects of MOOCs as a new platform of learning are less likely to complete the course according to the pace set by the instructor. This finding led to the current study to further examine students who have high interests in the course content area as compared to those who are more likely to be drawn to the course due to the novelty of the MOOC platform.

The present study conducted a principal component analysis on the ten survey items address MOOC-specific motivations on both course content and platform. Two principal component scores were extracted and saved as new variables to approximating student interests scores on course content (content-interest score) as well as the learning platform (platform-interest score). Afterwards, two separate multiple regression models with these two variables as dependent variables. These two regression models were applied to study how interests scores on content and platform relate to classic motivational items such as mastery-goal orientation and academic efficacy, and individual variables such as professional background, language, plus a self-rated completion confidence score were included as explanatory variable in both models.

2. Data Source and Participants

2.1 Data sources

The present study was conducted within the context of a MOOC,"Big Data in Education", delivered via Coursera. A survey was distributed to students through the course E-mail messaging system to students who enrolled in this course prior to the course start date.

2.2 Participants

The MOOC had an overall enrollment of about 48,000 students at the time of completion (since that time, over 5,000 more students have enrolled in the course). The pre-course survey received 2,792 responses; among which 38% of the participants were female and 62% of the participants were male. All survey respondents were 18 years old or older, among which 9% were between 18 to 24 years old, 38% were between 25 to 34, 26% were between 35 to 44, 17% were between 45 to 54, 8% were between 55 to 64, and 1% were 65 or older. This indicates a student profile not too dissimilar to graduate student populations taking more traditional online courses.

2.3 Motivation Survey

To measure MOOC learner motivation, the pre-course survey incorporated 3 sets of questions: MOOC-specific motivational items; two PALS (Patterns of Adaptive Learning Survey) sub-scales (Midgley, et al., 2000), Academic Efficacy and Mastery-Goal Orientation; and an item around confidence in course completion.

The MOOC-specific items consisted of 10 questions drawn from previous MOOC research studies (cf. Belanger & Thornton, 2013; MOOC @ Edinburgh, 2013) asking respondents to rate their reasons for enrollment. These 10 items address traits of MOOCs as a novel online learning platform. Specifically, these 10 items included questions on both the learning content and features of MOOCs as a new platform. For example, items such as “Subject relevant to my academic field of study” and “Extending current knowledge of the topic” relates to the content of the course; whereas items like “Course is offered by a prestigious university” and “Curious to take an online course” emphasize features of the MOOC platform. Participants were asked to rate on how important each potential benefit of a MOOC was to them, using a 5-point Likert scale.

Two PALS Survey (Midgley, et al., 2000) scales measuring mastery-goal orientation and academic efficacy were used to study standard motivational constructs. PALS scales have been widely used to investigate the relation between a learning environment and a student’s motivation (cf. Clayton et al., 2010; Meece, Anderman & Anderman, 2006; Ryan & Patrick, 2001). Participants were asked to select a number from 1 to 5 with 1 meaning least relevant and 5 most relevant.

3. Analysis

3.1 Principal Component Analysis

As a follow-up analysis to the previous investigation on course completers and non-completers, a principle component analysis for the ten survey items that are specific to the context of MOOCs and the contemporary societal interest in MOOCs was applied. Two components explaining 46% of the variance were extracted. A promax rotation provided the best-defined component structure.

Component 1 presented higher loadings on items addressing the unique opportunity afforded by the MOOC platform (labeled as “platform” in Table

1), They all addressed the features of MOOCs as a new learning medium. For example, “Course is offered by a prestigious university”, and “Cannot afford to pursue a formal education” address the unique opportunity afforded by the MOOC platform, whereas “Curious to take an online class” and “Geographically isolated from educational institutions” involve features common to all online learning platforms. In comparison, component 2 (labeled as “content”), with higher loadings on three items, connects to respondents’ knowledge or interest in the specific content area of the course.

Table 1. Factor loadings and communalities based on a principle components analysis with promax rotation for 10 items from the learner intention questions of the motivational survey:

	Component 1 (Platform)	Component 2 (Content)	Communality
Think the course will be fun and enjoyable	.351	.010	.126
Subject relevant to my academic field of study	.027	**.690**	.489
Class teaches Skill that will help my job/career	-.026	**.749**	.548
Course is offered by a prestigious university	**.552**	.283	.491
Curious to take an online course	**.619**	-.034	.371
Want a credential to enhance my CV/resume	**.662**	.174	.547
Supplement other college/university class	**.681**	.122	.536
Extending current knowledge of the topic	-.090	**.642**	.381
Geographically isolated from educational institutions	**.811**	-.201	.588
Cannot afford to pursue a formal education	**.769**	-.175	.531

In a previous analysis (Wang & Baker, 2015), independent t tests were conducted on the same group of ten items between course completers and non-completers. The results showed that it was the same group of items highlighted under Component 1 (Platform) showed statistically significant differences.

The consistency on the ten MOOC-specific motivational items between the previous t-test analyses on course completion and the present principal component analysis suggests that the concept of interests in content and platform can potentially apply to not only just a limitedly small subset of course completers but potentially relevant to a broader group of MOOC learners. Since course completers is only a relatively small subset of the group of registered students, it is worthy of further investigation to understand what other features do students who are more likely to care more about the content share.

3.2 Regression Analyses

Dependent variables

As an exploratory effort and to further look into whether other variables from the survey can explain variation on student preference over content or platform, two multiple regression analyses were conducted with the two components scores (Content and Platform) as dependent variables.

Explanatory variables

For both models, five variables from the survey were entered as explanatory variables consisting of three continuous variables measuring self-reported completion confidence score (SRATE), mastery-goal orientation (PAL-MG), and academic efficacy (PAL-AE). Two PALS Survey (Midgley, et al., 2000) scales measuring mastery-goal orientation and academic efficacy were used to study standard motivational constructs. PALS scales have been widely used to investigate the relation between a learning environment and a student’s motivation (cf. Clayton et al., 2010; Meece, Anderman & Anderman, 2006; Ryan & Patrick, 2001). Participants were asked to select a number from 1 to 5 with 1 meaning least relevant and 5 most relevant.

Also included were two categorical variables including whether students are self-described as a native speaker or not (EngN) and whether the professional background fall into the anticipated areas or not (JOBOthers). JOBOthers is a categorical binary variable indicating whether students selected the ten anticipated professional backgrounds identified by the course instructor or chose the "Others" category in the survey.

3.2.1 "Content"– Interest Analysis

A multiple linear regression was conducted to predict student preference on course "content" from the five explanatory variables explained above. From Table 2 below, the overall model statistically significantly predicted "Content", $F(5, 1223) = 67.833$, p < .oo1, R^2=.217.

Table 2. Model Summary "Content"

				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
.466[a]	.217	.214	.88124721	.217	67.833	5	1223	**.000**

a. Predictors: (Constant), EngN, JOBOTHERS, PAL_AE, SRATE, PAL_MG

Professional background in "other" categories (JOBOTHERS)
From Table 3 below, we can see that there is a statistically significantly difference between students who have a professional background in the anticipated areas and those who are not in terms of their "content" scores, *b* = -.225, t (-8.827), p< .001.

Self-rated confidence in course completion (SRATE)
Self-reported confidence scores of course completion from students also significantly predicted the content scores, *b* = .067, t(2.356), p<0.05.

Mastery-goal orientation (PAL_MG)
In addition, scores of mastery-goal orientation also significantly predicted the content scores, *b* = .387, t(12.776), p<.001

Table 3. Coefficients[a] - "Content"

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound
(Constant)	-2.111	.148		-14.238	.000	-2.402	-1.820
JOBOTHERS	**-.496**	**.056**	**-.225**	**-8.827**	**.000****	**-.607**	**-.386**
SRATE	**.030**	**.013**	**.067**	**2.356**	**.019***	**.005**	**.055**
PAL_MG	**.489**	**.038**	**.387**	**12.776**	**.000****	**.414**	**.564**
PAL_AE	.034	.037	.028	.922	.357	-.038	.106
EngN	-.051	.052	-.025	-.990	.323	-.153	.050

a. Dependent Variable: Component Score - Content

3.2.2 "Platform" - Interest Analysis

Similarly, a second multiple linear regression was conducted to predict student preference on course "platform" from the five explanatory variables explained above. From Table 4 below, the overall model statistically significantly predicted "platform", $F(5, 1223) = 65.473$, p < .001, R^2=.211.

Table 4. Model Summary "Platform"

				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
.460[a]	.211	.208	.88278229	.211	65.473	5	1223	**.000**

a. Predictors: (Constant), EngN, JOBOTHERS, PAL_AE, SRATE, PAL_MG

Professional background in "other" categories (JOBOTHERS)
From Table 3 below, we can see that there is a statistically significantly difference between students who have a professional background in the anticipated areas and those who are not in terms of their "platform" scores, *b* = .089, t(3.477), p<0.05.

Self-rated confidence in course completion. (SRATE)
Self-reported confidence scores of course completion from students also significantly predicted the "platform" scores, *b* = .067, t(2.356), p<0.05.
Mastery-goal orientation (PAL_MG)
Similar to the previous "content" model, scores of mastery-goal orientation also significantly predicted the "platform" scores, *b* = .387, t(12.776), p<.001.

Native English Speakers (EngN)
Contrary to the "content" model, whether students self-identified as a native English speaker significantly predicted the "platform" scores, *b* = -.188, t(-7.332), p<.001.

Table 5. Coefficients[a] "Platform"

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B	
Model	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
(Constant)	-1.751	.149		-11.788	.000	-2.042	-1.459
JOBOTHERS	**.196**	**.056**	**.089**	**3.477**	**.001****	**.085**	**.306**
SRATE	**.058**	**.013**	**.130**	**4.544**	**.000****	**.033**	**.083**
PAL_MG	**.434**	**.038**	**.345**	**11.319**	**.000****	**.359**	**.509**
PAL_AE	-.052	.037	-.044	-1.421	.156	-.124	.020
EngN	**-.380**	**.052**	**-.188**	**-7.332**	**.000****	**-.482**	**-.278**

a. Dependent Variable: Component Score - Platform

In both models, scores on academic efficacy did not show statistical significance in predicting "content"-interest scores and "platform"-interest scores.

4. Conclusions and Discussions

The present project included a principal component analysis on ten survey items related to student course enrollment motivation. Two component scores representing "Content" and

"Platform" were extracted. To further investigate how students who favor more on "Content" and those on Platform, two multiple linear regressions were conducted with

4.1 Variables showing statistical significance to both models

Professional background in "other" categories (JOBOTHERS)
This variable was included since about 30% of students selected the "Others" categories rather than the ten anticipated job categories. The content of the course was designed as an advanced level course for researchers and practitioners in learning analytics and educational data mining. It came out as a surprise that many a students revealed in an open-ended question that they are working in industries that are not directly related to this course including Finance, Consulting, etc. Some of the respondents stated that they were currently unemployed.

From the analyses in the previous section, it is shown that whether students coming from a course-related background or not statically significantly influence interest scores in both "content" and "platform" models. Specifically, for the "content" model, students who chose the "other" job category option have significantly lower interest in content compared to those who choose one of the ten anticipated job categories when holding other variables constant; whereas in the "platform" model, students who chose the "other" job category option have significantly higher interest in platform compared to those who choose one of the ten anticipated job categories when holding other variables constant.

The results from both models suggested that students who have a background that is not directly relevant to the course content area are likely to have registered for this MOOC due to interests related to the platform than the content or knowledge area. One possible interpretation is that students who are looking for career transitions might be those who have not yet decided on which new area they will transition to. Rather, they might be taking advantage of the opportunities provided by the MOOC platform to sample different areas of possibilities. This can partially explain the high drop-out rates since learners can be sampling the course at the beginning and decided it was not a good fit. Therefore, the dropping out behavior can not simply to interpreted as a course "failure"; rather it is only an indicator of incompatible background match between the learners and the requirements of taking the course.

Self-rated confidence in course completion. (SRATE)
Self-rated confidence in course completion has shown to influence interests scores on both content and platform. This result is consistent with the hypothesis that at the beginning of a course, students who show interests in either the content area or the learning platform tend to also have higher expectations in completing the course.

Mastery-goal orientation (PAL_MG)
Mastery-goal orientation is another variable that has show to relate to student interests in both course content and platform. It is also anticipated that students who are mastery-goal oriented are more likely to show interest in the content area of the course. Concurrently, the interests in the content areas co-exist with the interests in the course content.

4.2 Variables showing statistical significance unique to the "Platform" model

Native English Speakers (EngN)
Students who self-identified as native English speakers showed less interests in the course platform than non-native speakers. There were no statistically significant difference on interests in course content when comparing native speakers and non-native speakers. One way of interpreting of the finding on the relative higher interests on course platform from non-native speakers is that many non-native speakers reside in less-developed countries with lower level access to educational resources. Therefore, non-native speakers might be more interested in the educational opportunities provided by the MOOC platform.

4.3 Non-significant items to both models

Academic Efficacy (PAL_AE)
Academic efficacy did not show significant influence to neither the content nor the platform models. This finding was moderately unexpected in that self-efficacy is often related to interests in learning. One possible explanation is that items include in this scale might have been overly general without referring to the course content area.

5. Limitations and Future Work

The present study explored concepts of student interests on course content and the learning platform followed by regression analysis on other survey items. The present study represents one-step further in understanding MOOC learners via non-traditional course metrics as compared to course completion and students grades. By examining the survey motivational variables and demographic variables, we increased understanding of MOOC learners who care more about content versus those who care more about platform. Future analyses has been planned to look at not only variables from the course surveys but also those extract from log files to uncover potential relationships between learning behaviors during the course and the their interests in content/platform.

Meanwhile, since the present study was carried out in the context of only one MOOC, which might limit its generalizability, future work should also collect and analyze data from different MOOCs across different disciplinary areas and course platforms to determine whether the findings obtained here are general (Wang, Paquette, & Baker, 2014). For instance, it is reasonable to ask whether results may vary between MOOCs on science subjects and humanities subjects, or between more introductory and more advanced MOOCs. Similarly, national and cultural differences may also play a relevant role, which could be studied by analyzing differences between students in multiple populations taking the same MOOCs, and by comparing MOOCs offered in different languages.

Additionally, since data on MOOC learner information is highly fragmented and diffuse (McAulay, Stewart, & Siemens, 2010), it may be useful to incorporate data beyond the immediate course platform and the pre-designated time frame for a specific course in the efforts of better understanding learner behaviour. For example, data from social networks and student career advancement after taking a MOOC can also be incorporated in future analyses.

6. References

Belanger, Y., & Thornton, J. (2013). Bioelectricity: A Quantitative Approach Duke University's First MOOC.

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*, 13-25.

Clow, D. (2013). MOOCs and the funnel of participation. Paper presented at the LAK'13: 3rd International Conference on Learning Analytics & Knowledge, Leuven, Belgium. Retrieved from Retrieved from http://oro.open.ac.uk/36657/1/DougClow-LAK13-revised-submitted.pdf

Clayton, K., Blumberg, F., & Auld, D. P. (2010). The relationship between motivation learning strategies and choice of environment whether traditional or including an online component. *British Journal of Educational Technology*, *41*(3), 349-364.

DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing "course":

Reconceptualizing educational variables for massive open online courses. *Educational Researcher*, 43(2).

Fini, A. (2009). The technological dimension of a massive open online course: The case of the CCK08 course tools. *The International Review Of Research In Open And Distance Learning*, *10*(5).

Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). 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). ACM.

Guo, P. J., Kim, J., & 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).

McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice.

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*, 487-503.

Midgley, C., Maehr, M. L., Hruda, L., Anderinan, E.M., Anderman, L., Freeman, K. E., et al. (2000).*Manual for the Patterns of Adaptioe Learning Scales (PALS).* Ann Arbor: University of Michigan.

MOOC @ Edinburgh 2013 – Report #1 (2013). MOOC @ Edinburgh 2013 – Report #1. University of Edinburgh, Edinburgh, Scotland. Retrieved from http://www.era.lib.ed.ac.uk/bitstream/1842/6683/1/Edinburgh%20MOOCs%20Report%202013%20%231.pdf

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, *38*(2), 437-460.

Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs? *MERLOT Journal of Online Learning and Teaching*, *11*(1).

Wang, Y., Paquette, L., Baker, R. (2014) A longitudinal study on learner career advancement in MOOCs. *Journal of Learning Analytics, 1 (3),* 203–206.

Whitmer, J., Schiorring, E., & James, P. (2014, March). Patterns of persistence: what engages students in a remedial English writing MOOC?. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (pp. 279-280).

Yang, D., Sinha, T., Adamson, D., & Rosé, C. P. (2013). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Proceedings of the 2013 NIPS Data-Driven Education Workshop*.