Computers & Education 98 (2016) 157–168

Contents lists available at ScienceDirect

Computers & Education

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

Exploring the factors affecting MOOC retention: A survey study

Kate S. Hone [a, *], Ghada R. El Said [b]

[a] Brunel University London, Kingston Lane, Uxbridge, UB8 3PH, UK
[b] Future University in Egypt (FUE), 90th Street, Fifth Settlement, New Cairo, Egypt

ARTICLE INFO

Article history:
Received 21 August 2015
Received in revised form 12 February 2016
Accepted 29 March 2016
Available online 31 March 2016

Keywords:
Massive Open Online Courses (MOOCs)
Distance education
Telelearning

ABSTRACT

Massive Open Online Courses (MOOCs) hold the potential to open up educational opportunities to a global audience. However, evidence suggests that only a small proportion of MOOC participants go on to complete their courses and relatively little is understood about the MOOC design and implementation factors that influence retention. This paper reports a survey study of 379 participants enrolled at university in Cairo who were encouraged to take a MOOC of their own choice as part of their development. 122 participants (32.2%) went onto to complete an entire course. There were no significant differences in completion rates by gender, level of study (undergraduate or postgraduate) or MOOC platform. A post-MOOC survey of students' perceptions found that MOOC Course Content was a significant predictor of MOOC retention, with the relationship mediated by the effect of content on the Perceived Effectiveness of the course. Interaction with the instructor of the MOOC was also found to be significant predictor of MOOC retention. Overall these constructs explained 79% of the variance in MOOC retention.

1. Introduction

Massive Open Online Courses (MOOCs) are a rapidly growing mode of educational provision, holding the potential to open up access to world class teaching and educational resources beyond geographical and social boundaries. The potential benefits are particularly high for those students in developing countries for whom the costs of travel and tuition to attend face-to-face education at top western universities would be prohibitive. Studies where participant demographics are reported, however, suggest that relatively few participants are from developing world regions such as Africa and Asia, with the majority being from North America or Europe (Liyanagunawardena, Adams, & Williams, 2013). In addition, despite the potential and hype associated with MOOCs, retention rates overall are typically very low (figures of 10% retention are widely cited). While research studies are starting to examine the reasons behind the low retention rates, these tend to focus on a single MOOC as a case study (e.g. Greene, Oswald, & Pomerantz, 2015; de Freitas et al., 2015) or look at intention to complete rather than actual behaviour (e.g. Alraimi, Zo, & Ciganek, 2015). Due to this methodological focus, little is known about the experiences of MOOC non-completers (Liyanagunawardena et al., 2013). This study therefore set out to explore the factors which affect actual learner retention in MOOC learning. A survey study was conducted with learners who completed a MOOC in full as well as those who dropped out, thus allowing comparisons to be made between these two sub-groups. The focus of

* Corresponding author.
E-mail address: kate.hone@brunel.ac.uk (K.S. Hone).

http://dx.doi.org/10.1016/j.compedu.2016.03.016

the study was on students resident in Egypt, representing a developing country in North Africa whose population appears under-represented in MOOC participation, but where the potential benefits are high. The learners enrolled in MOOCs across a number of different popular platforms and in a range of topics, allowing exploration of the effect of MOOC features on retention as well as the impact of participant demographics. Learner perceptions of MOOC features were collected using scales developed from the literature on distance and eLearning evaluation. Our main aim was to explore whether experiential variables (i.e. learners' reported experiences with the course) were predictive of the extent to which learners engaged with the course, in order to potentially inform the design of more effective MOOCs.

2. Literature review

MOOCs are the subject of increasing attention, both in the popular press (e.g. Pappano, 2012; Kovanović, Joksimović, Gašević, Siemens, & Hatala, 2015) and as a topic for academic research. An early systematic review of MOOC research was provided by Liyanagunawardena et al. (2013). This review found that much of the reported empirical research followed a case study approach; it also highlighted a number of gaps in the literature including a lack of exploration of MOOC experiences in the developing world and a lack of research considering the perspective of MOOC non-completers. Raffaghelli, Cucchiara, and Persico (2015) similarly found a high reliance on case studies in their review of methodology within the MOOC literature. Ebben and Murphy (2014) provide a systematic review of MOOC scholarship, highlighting two distinct phases of discourse: from an early connectivist focus to a later behavioural focus. In this later phase the growth of learning analytics is highlighted, with the process of delivering MOOCs itself providing a rich source for data mining to better understand learner behaviour.

The specific focus of the current paper is on learner experience with MOOCs and the effect of experience of MOOC characteristics on learner retention. Learner retention is important as a measure of MOOC success since only those learners who persevere with a course have a chance of reaping the intended educational benefits of the learning experience.

2.1. MOOC learner retention

Early reports on learner retention in MOOCs focussed on the observation that completion rates tend to be very low. For example Alraimi et al. (2015) cite a number of sources to conclude that retention rates are on average less than 10%. More recently there has been a small but growing literature looking at the factors which affect learner retention within MOOCs, both in terms of participant characteristics and MOOC characteristics.

A frequent research approach is to use a single MOOC as a case study to look at retention factors. This echoes the wider trend identified by Liyanagunawardena et al. (2013) and Raffaghelli et al. (2015) and also appears to reflect the situation where many MOOCs instructors are also using their own course as an opportunity to gather research data, supporting the data mining trend identified by Ebben and Murphy (2014). In a case study of a single astronomy MOOC, de Freitas et al. (2015) speculate that both challenging assessments and gamification elements positively impact completion, but they do not present data to support this. Greene et al. (2015) conducted a case study looking at learner retention within a single MOOC (on 'Metadata: Organization and Discovering Information'). They collected survey data from participants who started this course and then looked at how this data predicted retention using survival analysis. They found participants with prior experience of MOOCs were less likely to drop out, as were older and more educated participants. Self-rated commitment to completing the course was the most statistically significant predictor of outcome. Greene et al.'s (2015) paper focusses on the characteristics of the participants, rather than of the MOOC.

Some studies are also starting to look at retention factors over a wider sample of MOOCs. Hew (2014) studied three top rated MOOCs across three disciplines, combining participant observation with analysis of reflection data from 965 course participants. From this work they propose five features that promote student engagement: problem-centric learning, instructor accessibility and passion, active learning, peer interaction and using helpful course resources. However, they did not look directly at retention and they did not include any lower ranked MOOCs as controls. Reich (2014) reports a study based on survey and log data from nine HarvardX courses which also showed a relationship between learner intention to complete and actual completion (to earn a certificate) and this was a stronger predictor of outcome than any demographic variable. However, of those who indicated that they intended to complete a course in this study, only 22% did so (compared to 6% who said they only intended to browse). Adamopoulos (2013) analysed a real-world data set of user generated content to model the factors that predicted self-reported course progress. This study included sentiment analysis of 1163 textual comments submitted to an online course review hub (on CourseTalk.org) by 842 students across 133 courses from thirty universities across six providers. The analysis suggested that positive sentiment expressed in relation to the course instructor had the largest positive effect on likelihood of completion; sentiment expressed in relation to assignments and course material also had a positive effect. However, the nature of the data source for the studied variables, being via self-reported via online course reviews, may bias the sample. The scope for further research based on data mining of MOOC data across multiple courses (though only on one platform at present) has been increased by the decision of Harvard and MIT to release a dataset containing de-identified original learning data from 16 HarvardX and MITx courses offered in 2012–13 (MIT News, 2014). However, the utility of such data remains limited by the lack of insight into the experiences of users who do not complete the course. There have also been concerns expressed over the potential research ethics issues of MOOC data mining in cases where insufficient consideration has been given to issues such as informed consent, privacy, anonymity and confidentiality (Marshall, 2014).

Alraimi et al. (2015) conducted a survey study with 316 users of three major US-based MOOC platforms (Coursera, EdX and Udacity) looking at intention to continue using MOOCs. They found that intention was significantly influenced by perceived reputation, perceived openness, perceived usefulness, perceived enjoyment and user satisfaction. While this study had the advantage of looking at experience across multiple MOOC platforms and courses, the focus on intention to use rather than actual completion is a limitation.

The literature on MOOC learner retention has also given some insight into patters of retention and led to a questioning of the most appropriate measures of retention. A 'funnel of participation' (Clow, 2013) is well recognised within MOOC education. de Freitas et al. (2015) report a relatively rapid dropout rate at the start of their Astronomy MOOC which stabilises to an almost negligible rate by the final part of the course. Similarly Greene et al. (2015) found a strong decrease in students (around 40%) between the start of the course and the end of the first unit, a further 25% dropped out between the first and second assessment, but dropouts were increasingly unlikely as participants neared the end of the course. DeBoer, Ho, Stump, and Breslow (2014) argue that educational variables such as enrolment and dropout need to be reconceptualised for MOOCs. Part of the issue here stems from the fact that those who initially sign-up for a MOOC may do so without intending to participate, so including them in the statistics for calculating retention can be seen as misleading.

Overall, while there is a growing literature on MOOC learner retention, existing studies suffer from some limitations. In particular there is a paucity of understanding of how MOOC features contribute to student experience and retention, and a lack of insight into the views of participants who fail to complete their course.

2.2. Dimensions of online learning effectiveness

While the literature specific to MOOCs is fairly nascent, there is an earlier body of work on online and distance learning which clearly has potential relevance to this new form of online provision. Peltier, Drago, and Schibrowsky (2003) reviewed the literature on distance education, virtual communities and teaching effectiveness to propose a model of the effectiveness of the online educational experience. Peltier et al.'s (2003) research supported the relevance of six constructs, three focussing on interpersonal/communication issues (student-to-student interactions, student-instructor interactions and instructor support and mentoring) and three focussing on course design (course content, course structure and information delivery technology). Their study of 299 learners on online MBA courses found that all of these constructs significantly affected ratings of overall perceived effectiveness of the course. A later reanalysis of the data from this study using structural equation modelling further suggested that course content was the most important factor in determining the perceived quality of the online learning experience (Peltier, Schibrowsky, & Drago, 2007).

Eom, Wen, and Ashill (2006) examined the effect of similar constructs to Peltier et al. (2003) (course structure, interaction, and instructor knowledge & facilitation) as well as instructor feedback and two constructs related to the participants themselves (learning style and self-motivation) within an e-learning context. Their research, using structural equation modelling in PLS with data from a sample of 397 online learners, found that all of these constructs predicted user satisfaction. However, only instructor feedback and participant learning style were found to affect learning outcomes; additionally user satisfaction predicted learning outcomes.

Marks, Sibley, and Arbaugh (2005) argue that researchers should study three aspects of online learning: instructor-student interaction; student-student interaction and student-content interaction. In an empirical study they found significant effects of both the instructor-student interaction and student-student interaction on self-rated learning effectiveness, with the instructor-student interaction having twice the effect of student-student interaction. They were unable to demonstrate statistically significant effects of the student-content variables in their model, however these were measured by single-indicator latent variables related to presentation features, some of which were rarely used in the sample of courses (such as use of streaming audio, streaming video) and some of which were ubiquitous (e.g. PowerPoint presentations).

While these studies are interesting, they were conducted in a rather different context to MOOC usage; specifically the participants were enrolled in online programmes and therefore needed to complete the courses for credit. In the next section we therefore propose a model to investigate whether constructs investigated within earlier e-learning studies are relevant to understanding MOOC continuance.

3. Research model

The aim of the current research was to explore the factors which affect MOOC completion/learner retention. We focus predominantly on user perception of MOOC features, rather than individual learner characteristics, since our primary concern is the design and provision of more effective MOOCs for the wide range of learners that might access them, including those from less developed countries such as those who formed the sample for this study. Similar to Marks et al. (2005) we considered three broad categories of experience: experiences with the course instructor; experiences with other learners on the course; and experience with the design features of the course; as together these cover the broad scope of how individual MOOCs will differ from one another. We draw on our review of past literature reported above to propose a number of constructs which we hypothesise may affect learner retention within this emerging context. The choice of constructs was guided by consideration of those features which have either been shown to significantly predict outcomes in the limited research to date which has specifically considered MOOC adoption, or (where MOOC-specific evidence is absent) been shown to be relevant in contexts with analogous characteristics (i.e.

online learning more generally). Given the exploratory stage of this research, an inclusive approach was taken to construct selection to ensure a wide coverage of potentially relevant variables. This resulted in a relatively large number of initial hypotheses (n = 15) which were subsequently refined through examination of the resulting measurement model (see section 4.4).

3.1. Instructor effects

As discussed above Adamopoulos (2013) provides some preliminary evidence of the role of instructors in MOOC retention, with positive review comments about course instructors (from a sample of 1163 comments posted online) correlating with completion. However, as Adamopoulos (2013) used sentiment analysis rather than subjective measurement constructs in his research, we looked to earlier (non-MOOC) e-learning research to identify reliable measures of instructor interaction. As reviewed in section 2.2, researchers looking more generally at e-learning effectiveness had identified several distinct facets of instructor interaction. These include instructor-learner interaction (Peltier et al., 2003), instructor support (Peltier et al., 2003) and instructor feedback (Eom et al., 2006) all of which had been found to predict at least one measure of course effectiveness. Thus Peltier et al.'s (2003) survey study of 299 online MBA course learners found effects of both instructor-learner interaction and instructor support on perceived effectiveness. Eom et al.'s (2006) study of 397 US university students who had studied at least one online course found that instructor feedback affected both learning outcomes and user satisfaction. For the purposes of this research all three constructs were retained to see if these effects also play out in a MOOC context and thus we hypothesise:

H1–3. instructor behaviour (instructor-learner interaction, instructor support and instructor feedback) in MOOCs will have a positive impact on learner retention

Our main outcome variable in this research was learner retention. However, many of the earlier studies on e-learning effectiveness used perceived effectiveness (a measure of satisfaction with the learning environment) as their dependent variable due to study designs which prevented the direct measure of retention (e.g. Peltier et al., 2003). We therefore also included the following hypotheses:

H4–6. instructor behaviour (instructor-learner interaction, instructor support and instructor feedback) in MOOCs will have a positive impact on perceived effectiveness

3.2. Co-learner effects

While our literature review did not identify any MOOC-specific empirical findings for the role of co-learners, discussion boards and support for interaction with other participants represent a key feature of MOOC implementations across all the major platforms, presumably reflecting an underlying design assumption that these are beneficial for learners. Indeed earlier survey research on online learning does support a positive role for learner-learner interaction. Thus Peltier et al.'s (2003) survey of online MBA participants found a significant relationship between student-to-student interactions and perceived effectiveness; similarly Marks et al. (2005) surveyed 659 online MBA learners in the US and found a significant relationship between student-student interactions and perceived learning. We therefore predict that:

H7. Learner-learner interaction in MOOCs will have a positive impact on perceived effectiveness

H8. Learner-learner interaction in MOOCs will have a positive impact on learner retention

3.3. Design and implementation effects

Adamopoulos's (2013) sentiment analysis study of MOOC reviews is suggestive of a positive role of course material within MOOC retention but fails to provide a direct measure of course content that can be applied here. Peltier et al.'s (2007) earlier survey work with online MBA participants also suggests a key role for course content in perceived effectiveness and provides a measure which can be tested in the new MOOC context. We therefore predict:

H9. Course content of MOOCs will have a positive impact on perceived effectiveness

H10. Course content of MOOCs will have a positive impact on learner retention

Peltier et al.'s (2003) and Eom et al. (2006) research on online learning also suggested a role for course structure on perceived effectiveness (see Section 2.2 for a discussion) and we therefore hypothesise:

H11. Course structure of MOOCs will have a positive impact on perceived effectiveness

H12. Course structure of MOOCs will have a positive impact on learner retention

The final design feature that we considered was the effectiveness of the delivery modes used since MOOCs rely on interactive content to support learning. Peltier et al. (2003) argued that for online distance education the effectiveness of electronic materials (in their case delivered via CD-ROM) is likely to affect learner satisfaction. We therefore hypothesise a similar effect would apply within the MOOC context, thus:

H13. Information delivery technology will have a positive impact on perceived effectiveness

H14. Information delivery technology will have a positive impact on learner retention

3.4. Dependent variables

MOOC retention was included as a key dependent variable since it represents the main focus of the current study. For the purposes of our main research model we chose to measure MOOC retention on a scale (rather than a binary consideration of whether participants completed or not) as we recognise that learners may benefit from a MOOC even if they do not complete it to certification and that completion per se is therefore not the best measure of MOOC effectiveness and value.

As discussed above, perceived effectiveness was also included as an outcome measure in our study since it represents the dependent variable in several of the e-learning effectiveness studies which have informed our hypotheses. It is common for perceived effectiveness to be considered in contexts where it was impractical to measure actual behaviour and it is also a measure that has been found to relate to behavioural intention in numerous studies of online behaviour (e.g. Davis, 1989). We therefore finally hypothesise that:

H15. Perceived effectiveness of MOOCs will have a positive impact on learner retention.

4. Research method

4.1. Survey design

A printed questionnaire was developed primarily using scales which have been used in previous research in a learning context. The questionnaire was written in English since English is the teaching language for undergraduates and postgraduates in most of Egyptian higher education institutes, including the one used as a context for this study. Attitude items used a five-point Likert scale (1 = strongly agree; 5 = strongly disagree). The following scales were adapted from Peltier et al. (2003): Perceived effectiveness (3 items), Information Delivery Technology (3 items), Course Content (5 items), Course Structure (3 items), Learner to Learner Interaction (5 items), Instructor to Learner Interaction (4 items), and instructor support (5 items). In addition the Instructor Feedback scale (4 items) from Eom et al. (2006) was used. A self-report measure of learner retention (3 items) was created and was used alongside a categorical measure of whether the MOOC had been completed to earn a credential signifying completion (this was subsequently also used as the end point of the learner retention scale measuring proportion of the MOOC completed). Demographic data was collected on age, gender, education and experience. The full scales can be found in Appendix A. At the end of the questionnaire there was space for participants to provide an open text response as to why they did or did not complete the course.

4.2. Sampling

The sample for this study was drawn from the student population at two higher education institutions in Cairo, Egypt. Convenience sampling was used due to the need to identify a study context in which there was a clear academic rationale to invite students to take part in a MOOC. Two groups of participants were invited to participate. The first was the cohort of students (n = 256) taking a Management Information Systems (MIS) course as an elective as part of their undergraduate studies at a private university in Cario, Egypt; of these 241 agreed to participate in the research. The second was the cohort of students (from a range of backgrounds) undertaking a nine month postgraduate diploma in Software Skills at a government postgraduate institute in Cairo (n = 245), all of whom agreed to participate. In each case the selected course provided a context in which it was meaningful and relevant for students to experience a MOOC as an optional self-learning element within their studies. Thus both courses were already designed to include a self-learning element; in the case of the undergraduate programme, participation was an optional course element for a bonus grade; in the case of the postgraduate programme, participation in self-learning was part of the programme of study, but the inclusion of MOOC study as part of this was optional.

4.3. Procedure

Prior to the start of the self-learning element of their course, students were invited to take part in the study through a face to face invitation from a member of the research team. Students were given a 10 min MOOC orientation session with an overview of different MOOC platforms and an explanation of how MOOCs could be used as a resource for their self-

learning module. Participation was voluntary and no incentives or rewards were offered. Participants were informed of their rights not to participate and to withdraw whenever they wished. It was also made clear to them that the course learning outcomes could be demonstrated via other means than taking a MOOC. Participants who chose to enrol in a MOOC as part of their self-learning were given the opportunity to search for a suitable MOOC (of maximum six weeks duration) and were offered in-class support to complete the MOOC sign-up process. Choice of platform and topic was left to the participant and topic did not have to relate to information technology. Within both courses the self-learning took place over six weeks and at the end, participants were asked to give a class presentation on their newly acquired skill; participants did not need to provide a MOOC certificate of completion in order gain credit/recognition for their work. At the end of the six week period of the study, participants were given class time to complete the survey on their experiences with the MOOC they had selected.

4.4. Data analysis method

Data screening and factor analysis was carried out in SPSS. Partial Least Squares (PLS) was used to perform structural model analysis. PLS was chosen because of the exploratory nature of this study.

4.5. Data screening and measurement model

Out of a total of 486 students who initially agreed to take part, 379 questionnaire responses were collected (a response rate of 78%) with no cases with missing data. Three cases (two male, one female, none of whom completed a MOOC) were identified as showing unengaged responses to the Likert scales (s.d. < 0.55) and were removed from the data set for the factor analysis and structural model analysis, leaving a sample size of 376 for these stages of the analysis. Initial data screening also identified the Course Structure scale as problematic, with 90% of all participants answering 'agree' to all items on the scale, so this construct was excluded from further analysis.

Although initial examination of the retained MOOC design and human factors constructs showed these to be reliable (with all Chronbach's alphas > 0.8), the correlation matrix showed high multicollinearity. Exploratory factor analysis confirmed high cross loadings between many of the constructs. After removal of cross-loading items, a clean two-factor model was obtained consisting of Course Content (retaining three out of four items from the original scale) and Interaction with Instructor (formed by combining two items from the Instructor-to-Learner-Interaction scale with one item from the Instructor Support scale). While the Interaction with Instructor construct was therefore new for this research, it nevertheless showed high face validity and high reliability. The two outcome variables of Perceived Effectiveness and Retention also showed high reliability (see Table 2).

4.6. Revised constructs and model

Following factor analysis the hypothetical model was considerably simplified with only two predictor items retained: Course Content and Interaction with Instructor. The hypotheses were therefore reframed as follows:

H1. Course Content (CC) will have a significant effect on the Perceived Effectiveness of a MOOC

H2. Course Content (CC) will have a significant effect on learner retention within a MOOC

H3. Instructor Interaction (Int) will have a significant effect on the Perceived Effectiveness of a MOOC

H4. Instructor Interaction (Int) will have a significant effect on learner retention within a MOOC

H5. Perceived Effectiveness (PE) of a MOOC will have a significant effect on learner retention within a MOOC

5. Results

5.1. Descriptives

5.1.1. Sample

379 participants completed the survey. 282 respondents (74.4%) had studied via the Coursera MOOC platform, 73 respondents (19.3) via FutureLearn, 21 (5.5%) via Khan Academy and three (0.8%) via EdX. The sample included 160 male participants (42.2%) and 219 female participants (57.8%). The majority (266; 70.2%) were aged 24 or under; 93 participants (24.5%) were aged between 25 and 29 and 20 participants (5.3%) were aged 30 and over. The majority (295; 77.8%) were not employed and the split of Undergraduate and Postgraduate study level was 183 (48.3%) and 196 (51.7%) respectively.

5.1.2. MOOC completion/learner retention

Overall 122 participants (32.2%) indicated that they had completed the MOOC to earn a credential signifying official completion. Measures of time before drop-out, proportion of exercises/assessments completed and proportion of content watched showed very similar patterns of responses with a bimodal distribution; participants were most likely to complete

Table 1
Learner retention (self-report).

		Frequency	Percent
Proportion of assessments completed	None	14	3.7
	Few	90	23.7
	Around half	129	34.0
	Most	24	6.3
	All	122	32.2
Proportion of content watched/read	None	5	1.3
	Few	90	23.7
	Around half	141	37.2
	Most	19	5.0
	All	124	32.7
Proportion of course time completed	First few days	33	8.7
	First few weeks	116	30.6
	Towards the middle	103	27.2
	Towards the end/just before the end[a]	5	1.3
	Completed MOOC	122	32.2

[a] These two response sets in the original data set were merged due to the small numbers across the two categories and for equivalence with the other scales (producing a five point scale with 'completed MOOC' as the positive end point).

around half of the course or to complete the whole course. Once participants had passed the half way point (by any measure) they were very unlikely to drop out before the end. These results are illustrated in Table 1.

5.1.3. Demographic effects on learner retention

Chi Square analysis showed no significant effect of gender on overall completion ($\chi2 = 0.60$, df = 1, $p > 0.05$). Similarly age (categorised as 24 or younger vs. 25 and above) did not show any significant effect on completion ($\chi2 = 2.347$, df = 1, $p > 0.05$) and postgraduates and undergraduates showed no significant difference in likelihood of completion ($\chi2 = 0.739$, df = 1, $p > 0.05$). As no significant differences in retention were observed between the undergraduate and postgraduate groups in this study, the remaining analyses reported here included the sample from both groups.

5.1.4. MOOC platform effects on learner retention

As the proportion of the sample studying on EdX and Khan Academy was very low, the data for MOOC completion was compared for the two main choices of platform, Coursera and FutureLearn. Chi Square analysis showed no significant difference in completion rates across these two platforms ($\chi2 = 1.893$, df = 1, $p > 0.05$).

5.2. Structural model

The results of the PLS analysis of the revised model are presented in Fig. 1 below. The model overall explained 79% of the variance in learner retention and 78% of the variance in Perceived Effectiveness. The significant relationships are summarised below in Table 3.

5.3. Open text responses

In total 124 (32.7%) participants provided an open text comment as to why they did or did not complete the course (though in some cases only one or two word answers were provided). The responses of those who completed a MOOC included a number of positive comments that could be broadly categorised as concerning the content of the course; for example one participant stated that:

> "I successfully completed my course because the content was just right"

Other examples included that the courses provided content they were really interested to learn, provided material they could not find elsewhere, provided tips about soft skills, provided content based on real cases/examples and practical content (such as programming skills) especially where this was supplemented with access to additional resources (such as programming tools). Conversely those who did not complete their course noted issues related to content such as the courses were too complex/in-depth/sophisticated/technical, used language that was too complex, had too many modules and were boring.

While those who completed their MOOC successfully did not mention the role of human (instructor or peer) interaction, this was a major theme that emerged from the comments of those that did not complete. Participants mentioned feeling isolated and alone with poor communication with instructor and peers; they described being unmotivated to continue due to low interaction and poor feedback and communication with both instructors and peers, with no teamwork or group interaction; they complained that the instructor did not praise or motivate them after quizzes; and they stated that instructors did not engage learners in discussion or facilitate brainstorming. For example:

Table 2
Measurement model.

Construct	Items	Item loading	AVE	CR	Crobach's α
Course Content	CC1	0.830	0.691	0.870	0.77
	CC2	0.804			
	CC3	0.859			
Interaction with Instructor	ILI3	0.865	0.810	0.927	0.887
	ILI4	0.941			
	IS4	0.892			
Perceived Effectiveness	PE1	0.922	0.796	0.921	0.872
	PE2	0.867			
	PE3	0.887			
Retention	Time	0.973	0.958	0.986	0.978
	Assessments	0.978			
	Content	0.985			

"I did not continue with the Coursera course as I was demotivated due to poor instructor's feedback"

"The instructor did not praise learners"

"The instructor didn't … engage us in discussions"

Another clear theme from those who completed concerned the packaging of material, especially video, into small chunks which could be flexibly accessed. For example participants praised the use of short videos and course organisation that allowed them to skip some sessions. Conversely, participants who did not complete emphasised the presence of very long videos (of an hour or more) which were both boring and difficult to load due to connectivity issues especially from mobile devices, and which made it difficult to scan for relevant content. For example:

"Contents are overwhelming and not given in small chunks"

"Course was not easy to scan in terms of content"

6. Discussion

6.1. Key findings

The intention of this research was to identify the factors which affect learner retention within MOOCs. The empirical results from a survey of 379 learners showed that MOOC Course Content has a significant effect on retention. This effect is mediated by the effect of Course Content on Perceived Effectiveness which in turn affects retention. Interaction with the instructor was also shown to have a direct effect on retention, such that retention is higher when interaction with the instructor is higher. Overall these constructs explained 79% of the variance in MOOC retention. The relevance of course content has had some support in previous MOOC literature, for example Adamopoulos (2013) found that positive sentiments expressed in relation to course material had a positive relationship with retention. Previous work in e-learning contexts similarly found a significant effect of course content on perceived effectiveness of a course (Peltier et al., 2007). The current work therefore supports previous research, reinforcing the role of course content on both perceived effectiveness and retention. It also builds on previous research by suggesting that the effect of content on retention is mediated by perceived effectiveness, an issue that the methodologies of these previous studies prevented them from directly addressing. The findings in relation to instructor interaction also supports the findings from Adamopoulos (2013) who found that positive sentiment in online reviews in relation to the course instructor had a positive effect on retention. Previous research in online learning, while not looking directly at retention as an outcome (given the different research context), had found course instructor factors had an influence on proxies for effectiveness such as learner satisfaction (e.g. Eom et al., 2006).

It is interesting to note that the completion rate observed in this study was higher than previous studies have typically found, with around a third of participants completing their course to the point of certification. This suggests that using a MOOC as a learning opportunity within the context of a broader, University accredited programme is beneficial in terms of increasing the likelihood of completion. However, the majority of learners still did not complete their MOOC, even in this context, suggesting that there still remains an issue with MOOCs in terms of the extent to which they engender completion among a motivated audience. This corresponds with the previous finding of Reich (2014) which found that only 22% of motivated participants completed a MOOC.

The pattern of drop outs from the MOOC (by all retention measures) showed that drop-out was most likely at or before the mid-point of the course; most who got past the mid-point went on to complete. While similar to the pattern observed in earlier studies (DeFreitas et al., 2015; Greene et al., 2015) we did not see very high attrition rates at the start of the course

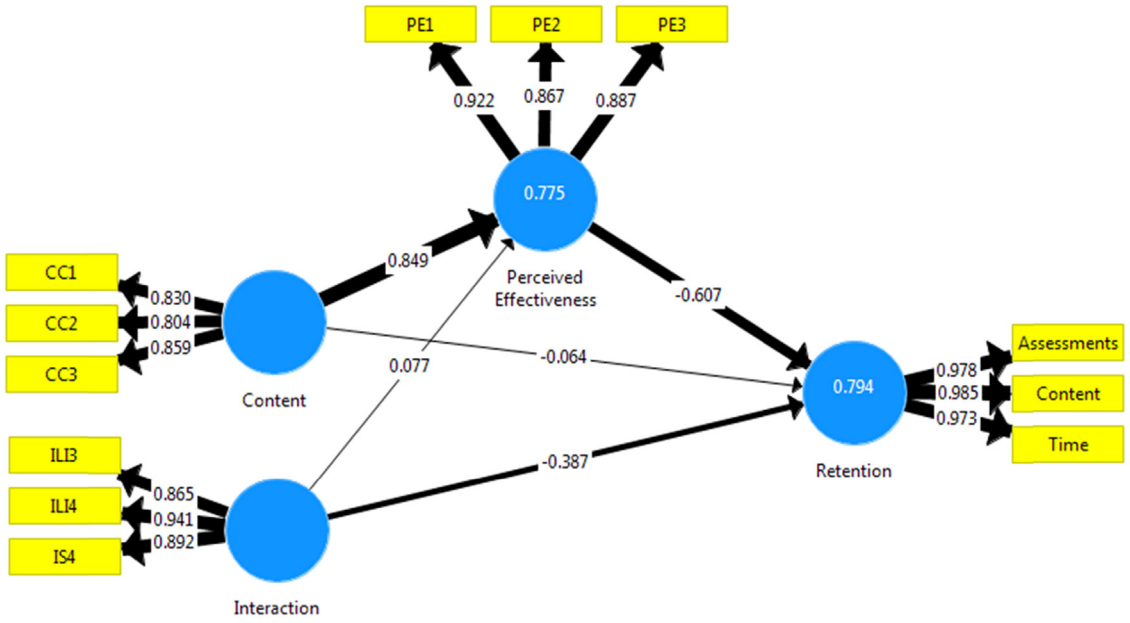

Fig. 1. Structural model PLS results.[1]

Table 3
Results of hypothesis testing.

Hypothesis	Path co-efficient	*t*-value	Results
H1: CC→PE	0.849***	31.248	Supported
H2: CC→Retention	−0.064^{ns}	0.745	Not supported
H3: Int→PE	0.077 ^{ns}	1.916	Not supported
H4: Int→Retention	−0.387***	10.541	Supported
H5: PE→Retention	−0.607***	7.805	Supported

***p < 0.001, ns = non-significant (two-tailed test).

reported in these studies which may relate to our sample being relatively motivated learners. We might infer from the retention patterns observed that participants are demonstrating a form of loss aversion - having invested a significant effort in participating within a MOOC, they will become more reluctant to drop out and therefore lose the chance at accreditation that they have worked towards. However, further work would be needed to explore whether this does indeed explain participants' motivations.

In this study no effects of learner demographics on retention were found. Similarly no differences in completion were observed between the two main MOOC platforms which students chose for their learning (Coursera and Future Learn).

6.2. Methodological implications

A key methodological strength of this research was that the research method was able to capture the views of both students who completed a MOOC and, importantly, those who dropped out. Capturing experience data from this latter group has been particularly problematic in much previous MOOC research which has tended to rely on data captured within MOOCs or associated online forums, where the views of non-completers are under-represented.

This study provides some important methodological contributions for the study of MOOC retention. First we have demonstrated the reliability of a new self -report construct of learner retention in MOOCs and shown the high correlation between time completed, proportion of content viewed and proportion of assessments completed as measures of retention. The research partially supports Perceived Effectiveness as a useful surrogate for measures of learner retention but illustrates

[1] Note that negative relationships are shown with the retention construct in the model – this is because the experience Likert items were coded with strongly disagree as 5, so a high number indicates a more negative experience. Conversely retention was coded such that completion of the course (the most positive outcome) equated to a 5 on the retention scale.

that it does not capture the full picture. Specifically we demonstrated the role of a construct (Instructor Interaction) with a direct effect on retention which was not mediated by Perceived Effectiveness. It is possible that there are other, as yet unexplored, constructs which could similarly affect retention, which studies relying on Perceived Effectiveness measures alone would be unable to identify.

It is interesting to note that we had to remove the Course Structure construct from the analysis as there was very little variance in responses (90% of all respondents selected 'agree' to all items). This suggests that Course Structure may to some extent be a solved problem within the major MOOC platforms (which impose a strong platform-wide organisational framework for the presentation of course material). Course Structure therefore appears not to act as a differentiator of experience as it might if the organisation of content was left to individual course designers.

It is surprising that in this study we failed to extract a separate factor relating to student-student interaction, as this is often thought of as a key component of the MOOC experience. The student-student items were found to frequently cross-load in the factor analysis with items describing instructor-student interaction. Interestingly Marks et al. (2005) quote Mason (1991) as finding that instructors play a major role in directing online discussions. If this were the case in our study it might explain the methodological problem that we encountered (i.e. good student-student interactions might be dependent on the behaviour of the instructor in encouraging these). However, more evidence would be needed to support this interpretation.

6.3. Limitations and future work

While this study was designed to overcome several of the limitations we identified in previous work, as with all research some limitations remain. The research was exploratory in nature given the relatively nascent state of the current literature of MOOCs. This meant that the applicability of the constructs applied from previous e-learning literature was untested and the research model had to be revised during the analysis phase to accommodate the issues with the measurement constructs which were identified. Nevertheless, a clean factor structure was obtained with the questionnaire items and we were able to propose a new measure of Instructor Interaction which has high reliability and high face validity. Further work is needed to confirm the relevance of this construct in other MOOC research contexts. Further work is also recommended to explore whether an appropriate construct to capture learner experience of student-student interaction (distinct from the role of the instructor) can be developed for the MOOC context.

The research context for this study deliberately focused on university learners within a developing world context as an under-researched and under-represented group within MOOC learning. It is promising that a reasonably high proportion of learners in this context went on to complete a MOOC. However, the geographically limited population of the study also represents a limitation; further work would be needed to examine whether the results observed here generalise to MOOC learners in other countries and learning contexts.

While the research here provides a model of MOOC retention which explains a high proportion of the variance, the quantitative approach which was adopted means that the nuances of experience are not captured in this data and the findings therefore don't fully capture the reasons behind participants' judgements about the key constructs in the model. To ameliorate this somewhat, participants were given the opportunity to give a qualitative response to an open-ended question at the end of the survey about their reasons to complete or drop out of the MOOC. The responses to these questions provide support for the quantitative analysis in that positive comments related to MOOC content were frequent amongst those who completed, while negative comments related to both content and instructor interaction were frequent amongst those who did not complete a MOOC. However, the qualitative comments also highlighted further themes which were not fully addressed within the research model; in particular the chunking of material into digestible segments, with flexibility around accessing these appeared to be an important feature of the completed MOOCs. In addition the 'added value' provided by the MOOC content also emerged as potentially important, be this through provision of additional web resources (such as programming tools), through illustrative case studies supported by video examples, through diagrammatic elements or simply through the provision of material that could not easily be found elsewhere. These tentative findings suggest additional factors which could be usefully explored further through additional research with MOOC participants. Further, more in-depth qualitative research is particularly recommended for uncovering emergent issues affecting learner retention.

7. Conclusion

The aim of this research was to examine the effect of experience of MOOC characteristics on learner retention. The findings identified MOOC content, Perceived Effectiveness and Instructor Interaction as having a significant effect on learner retention, together explaining a substantial percentage of the variance in retention. While MOOC content has a major effect on Perceived Effectiveness and subsequently on retention, interaction with the instructor is also important for learner retention. This may represent a limiting factor for the viability of very large scale MOOCs and suggests that MOOC providers need to think carefully about how to provide appropriate human interactive elements to their courses as well as excellent content.

Appendix A. Survey items.

Constructs	Items	Measures	References
Instructor to Learner Interaction (ITI)	ITI1	I felt free to ask questions throughout this course	Peltier et al., 2003
	ITI2	The instructor responded to my questions in a timely manner	
	ITI3	The instructor was easily accessible to me	
	ITI4	I felt free to express and explain my own views throughout this course	
Instructor Support (IS)	IS1	The instructor played an important role in facilitating learning in this course	Peltier et al., 2003
	IS2	The instructor contributed to the discussions in this course	
	IS3	The instructor was actively helpful when students had problems	
	IS4	I have interacted with the instructor in this course	
	IS5	The instructor emphasized relationships between and among topics	
Instructor Feedback (IF)	IF1	The instructor was responsive to student concerns	Eom et al., 2006
	IF2	The instructor provided timely feedback on assignments, exams or projects	
	IF3	The instructor provided helpful timely feedback on assignments, exams or projects	
	IF4	I felt as if the instructor cared about my individual learning on this course	
Learner to Learner Interaction (LLI)	LLI1	The group work contributed significantly to this course	Peltier et al., 2003
	LLI2	Group size was appropriate for course purposes	
	LLI3	Student interaction was an important learning component of this course	
	LLI4	This course provided an opportunity to learn from other students	
	LLI5	I had sufficient opportunity to interact with other students on this course	
Course Content (CC)	CC1	This course effectively challenged me to think	Peltier et al., 2003
	CC2	Course assignments were interesting and stimulating	
	CC3	This course was up-to-date with developments in the field	
	CC4	Student evaluation techniques such as projects, assignments, and exams were related to the learning objectives of this course	
	CC5	This course included applied learning and problem solving	
Course Structure (CS)	CS1	The structure of the modules was well prepared and organized	Peltier et al., 2003
	CS2	Projects/assignments were clearly explained	
	CS3	I understood what was expected of me	
Information Delivery Technology (IT)	IT1	The interactive content of this course was effectively communicated	Adapted from
	IT2	The interactive content of this course included information not covered in printed material of the same course	Peltier et al., 2003
	IT3	The interactive content of this course contributed towards learning	
Perceived Effectiveness (PE)	PE1	I would recommend this course to friends/colleagues	Peltier et al., 2003
	PE2	I have learned a lot in this course	
	PE3	I have enjoyed taking this course	
Learner Retention (LR)	LR1	Did you complete the MOOC to earn a credential signifying official completion? (Yes/No). If no, when did you drop out? (First few days, first few weeks, towards the middle, towards the end/just before the end)	New for this study
	LR2	How many exercises/assessments did you complete in the MOOC? (All, most, around half, a few, none)	
	LR3	How much of the MOOC content do you estimate you watched or read? (All, most, around half, some, none)	

References

Adamopoulos, P. (2013). What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. In *Thirty fourth international conference on information systems, Milan, 2013*.

Alraimi, K. M., Zo, H., & Ciganek, A. P. (2015). Understanding the MOOCs continuance: the role of openness and reputation. *Computers and Education, 80*, 28–38.

Clow, D. (2013). MOOCs and the funnel of participation. In *Proceedings of the third international conference on learning analytics and knowledge (LAK), Leuven, Belgium, April 8–12* (pp. 185–189).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

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), 74–84.

Ebben, M., & Murphy, J. S. (2014). Unpacking MOOC scholarly discourse: a review of nascent MOOC scholarship. *Learning, Media and Technology, 39*(3), 328–345. http://doi.org/10.1080/17439884.2013.878352.

Eom, S. B., Wen, H. J., & Ashill, N. (2006). The determinants of students' perceived learning outcomes and satisfaction in university online education: an empirical investigation. *Decision Sciences Journal of Innovative Education, 4*(2), 215–235.

de Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology, 46*, 455–471. http://dx.doi.org/10.1111/bjet.12268.

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015, May 8). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal (Online First)*. http://doi.org/10.3102/0002831215584621.

Hew, K. F. (2014). Promoting engagement in online courses: what strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology (Online First)*. http://doi.org/10.1111/bjet.12235.

Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2015). What public media reveals about MOOCs: a systematic analysis of news reports. *British Journal of Educational Technology, 46*(3), 510–527. http://doi.org/10.1111/bjet.12277.

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: a systematic study of the published literature 2008–2012. *The International Review of Research in Open and Distance Learning, 14*(3), 202–227.

Marks, R. B., Sibley, S. D., & Arbaugh, J. B. (2005). A structural equation model of predictors for effective online learning. *Journal of Management Education, 29*(4), 531–563.

Marshall, S. (2014). Exploring the ethical implications of MOOCs. *Distance Education, 35*(2), 250–262. http://doi.org/10.1080/01587919.2014.917706.

Mason, R. (1991). Analyzing computer conferencing interactions. *International Journal of Computers in Adult Education and Training, 2*(3), 161–173.

MIT News. (2014, May 30). *MIT and Harvard release de-identified learning data from open online courses*. Retrieved 20 August 2015, from http://newsoffice.mit.edu/2014/mit-and-harvard-release-de-identified-learning-data-open-online-courses.

Pappano, L. (2012). *The year of the MOOC*. New York Times. November 2 http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all&_r=0.

Peltier, J. W., Drago, W., & Schibrowsky, J. A. (2003). Virtual communities and the assessment of online marketing education. *Journal of Marketing Education, 25*(3), 260–276.

Peltier, J. W., Schibrowsky, J. A., & Drago, W. (2007). The interdependence of the factors influencing perceived quality of the online learning experience: a causal model. *Journal of Marketing Education, 29*(2), 140–153.

Raffaghelli, J. E., Cucchiara, S., & Persico, D. (2015). Methodological approaches in MOOC research: retracing the myth of proteus. *British Journal of Educational Technology, 46*(3), 488–509. http://doi.org/10.1111/bjet.12279.

Reich, J. (2014). *MOOC completion and retention in the context of student intent*. EDUCAUSEreview online Available at: Accessed 28.07.15. http://www.educause.edu/ero/article/mooc-completion-and-retention-context-student-intent.