
Article:
Swinnerton, BJ orcid.org/0000-0002-4241-4952, Hotchkiss, S and Morris, NP orcid.org/0000-0003-4448-9381 (2017) Comments in MOOCs: who is doing the talking and does it help? Journal of Computer Assisted Learning, 33 (1). pp. 51-64. ISSN 0266-4909

https://doi.org/10.1111/jcal.12165

Making comments in MOOCs: Who is doing all the talking and does it help?

Bronwen Swinnerton[1], Stephanie Hotchkiss[2], Neil P Morris[1]

[1]Research Centre in Digital Learning, School of Education, University of Leeds, Leeds, United Kingdom

[2]Socio-Technical Centre, Leeds University Business School, University of Leeds, Leeds, United Kingdom

Correspondence to: Dr. Bronwen Swinnerton, School of Education, University of Leeds, Leeds. LS2 9JT, UK. E-mail: b.j.swinnerton@leeds.ac.uk

Abstract

This paper investigates the characteristics and behaviour of learners on nine Massive Open Online Courses (MOOCs) on the FutureLearn platform in 2014 and 2015. FutureLearn's social constructivist approach, which emphasizes learning through social interaction, makes the focus on forum posting within these MOOCs particularly appropriate. This study makes a significant contribution to the research on MOOCs by exploring firstly the extent to which learners on FutureLearn MOOCs make comments, secondly, whether groups of learners who comment to a lesser or greater extent have different demographic characteristics and thirdly, whether commenting activity is associated with MOOC completion. The study uses a large dataset of over 25,000 learners to examine learners who comment versus those who don't, and learners who generate many posts. The results show that learners who are older, work part-time or not at all are more likely to post a comment, and those who are older, work part-time or not at all, are better educated and have prior online experience make the most comments. Making comments is also strongly associated with completing a MOOC, in particular for those who make many comments.

Keywords: MOOCs, social constructivism, commenting, MOOC completion,

Introduction

Massive Open Online Courses (MOOCs) have proliferated dramatically since the advent of the modern form in Canada in 2008, and their popularity globally has risen rapidly. The first MOOCs were connectivist or cMOOCs where learners were encouraged to connect through learning communities using a variety of web tools, to create knowledge. These MOOCs were followed by xMOOCs, which are more didactic in nature, offered on platforms such as Coursera, edX and Udacity. These xMOOCs often comprise a series of video lectures and assessments, with less emphasis on social interaction. By January 2014, the main xMOOC providers had offered almost 700 courses and enrolled more than 8 million users worldwide (Perna et al., 2014), and this number is increasingly rapidly. According to Dhawal Shah, founder of Class Central, a free online MOOC aggregator, the number of registered students added in 2015 is nearly equal to the previous three years combined (Wexler, 2015). In 2013, the UK's Open University launched a new MOOC platform, FutureLearn, which has almost four million registered participants and partnerships with over 80 universities and other organisations from across the world (FutureLearn). FutureLearn differs from both the cMOOC and the xMOOC approach by having a social constructivist pedagogy underpinning its design, and is labelled as a social learning platform, due to its 'water-cooler' style discussion areas provided alongside learning content intended to encourage social interaction (Ferguson & Clow, 2015; Ferguson & Sharples, 2014). A number of studies suggest that social interaction within MOOCs can impact on retention of learners, is associated with learner engagement, and can help prevent feelings of isolation, a perceived problem in online, distance courses.

Social interaction and retention in MOOCs

It is well known that MOOCs have low completion rates, and this has been a continued perceived problem, with reported rates of between 5% and 12% (Jordan 2014; Perna et al., 2014). Consequently, it is not surprising that there have been many studies examining the factors associated with retention in MOOCs (Breslow et al., 2013; Cisel, 2014; Guo & Reinecke, 2014; Kizilcec et al., 2013; Palin, 2014),

some of which have focused on the role of social interaction. Several studies have found that MOOC learners who post in discussion forums are less likely to drop out than those who don't (Balakrishnan & Coetzee, 2013; Breslow et al., 2013; Kizilcec et al., 2013; Woodgate et al., 2015). Balakrishnan & Coetzee (2013), examining data from a Coursera MOOC found that students who posted on a weekly basis were very unlikely to drop out, whilst those who didn't post were more likely to drop out but not as likely as those who didn't read posts. Woodgate et al. (2015) examined data from six Coursera MOOCs offered by the University of Edinburgh in 2013 and 2014 and found that learners who persist and gain a Statement of Accomplishment were more likely to have read a discussion forum post or to have posted themselves than a learner who did not persist, and they were more likely to have done it more often. Breslow et al. (2013) found that although only 3% of learners participated in a MOOC discussion forum, certificate earners used it much more, with 52% of certificate earners participating in the forums.

Engagement

Several studies suggest that social interaction and retention are related as both reflect a level of motivation and engagement by the learner (e.g. Wen et al., 2014; Yang et al., 2013; Yang et al., 2014). Yang et al. (2013) examined data from a Coursera MOOC and found that late joiners found it difficult to integrate into the discussion forum, remaining at the periphery, and tended to post at a lower rate than those who joined earlier. Yang et al. (2014) examining data from three Coursera MOOCs, divided MOOC learners into sub-communities based on the interest and focus of the discussions and found that drop out/retention rate can be predicted by membership of these sub-communities. Those students who express lower levels of motivation and cognitive engagement, and are less comfortable with the course procedures are more likely to drop out. Wen et al. (2014) reporting on further work from this team, found by analysing discussion forum content, that the more personalised the posts by a student, suggestive of higher cognitive engagement, the less likely they were to drop out. Ferguson & Clow (2015) in their work exploring clusters of learners on four FutureLearn MOOCs suggest that those learners who make more comments are generally more extensively engaged with the course, demonstrated by their assessments

being completed on time. Robinson (2011) however, suggests that discussion forums help to engage students, rather than being a reflection of their level of engagement

MOOCs and isolation

Gasevic et al. (2014) report research that suggests that a lack of social interaction can have a negative impact and be one of the factors which leads to online and distance learners feeling isolated and that such feelings may be more of a problem in MOOCs, given their 'massiveness'. (Gasevic et al., 2014; Muilenberg & Berge, 2001) It is thought that for distance learners, becoming part of an online learning community can help alleviate these feelings of isolation, which Rovai (2002) found increases persistence, although Hart (2012) suggests that these feelings of community and support are short-lived if they don't persist from course to course. Much of the research on MOOCs and the extent to which learners comment has been derived from data from xMOOCs such as those developed on the Coursera platform. Within that platform, discussions take place in one or more forums, where learners can continue discussions with the same group of learners throughout the course. The FutureLearn structure separates content and the associated discussions into discrete 'steps', in order that the discussions take place alongside the content. This means that the comments from step 1 are not connected to the comments in step 2 or step 3 etc. The existence of a 'follow' function, however, allows learners to keep track of educators' and other learners' posts. In this context, does a sense of community persist, evidenced by those who interact progressing further in the MOOC?

In their highly influential work on communities of practice, Lave and Wenger (1991) asserted that learning is a social process rather than something that happens just inside one's own head. Reed et al. (2010) use the term 'social learning' when referring to the learning that occurs through social interaction and processes between actors within a social network (Reed et al. 2010). Many others have suggested that collaborative learning is valuable and positively influences student persistence, and attitudes towards learning (Sharples et al., 2014; Stump et al., 2011). The FutureLearn approach has evolved from these ideas of social learning and is informed, in particular, by Laurillard's conversational framework, which

posits that a dialogue must exist between the learner and themselves, between the learner and teacher, and between learners and other learners for learning to take place (Laurillard, 2002). Mike Sharples, Academic Lead for FutureLearn, argues that 'the implementation of conversational learning on FutureLearn has been largely successful' (Sharples, 2016, p.339), stating that 36% of learners engage in social interaction on the platform (Sharples, 2016).

This study examines this approach more closely to assess firstly, whether FutureLearn MOOC learners participate in social learning by making comments and engaging in a dialogue in the discussion areas. Secondly, this study explores who the commenters are and who makes the most comments, an area so far largely neglected when discussing the social interaction within the FutureLearn MOOCs. In doing so we hope to shed some light on the different commenting behaviours of learners within MOOCs. Morris et al. (2015) explored the relationship between the demographic characteristics of several thousand FutureLearn MOOC learners and retention and found that learners who are older, with prior online experience, not working and having a degree or higher are more likely to complete more of a MOOC. This study will explore demographics to discover whether those groups of learners, more likely to complete, are the same as those who comment. Or are those learners who comment a different group with different characteristics? Thirdly, this study will explore the relationship between social interaction and engagement, in the form of completion. Does making comments have an impact on how long learners stay in a MOOC, how far they progress through a MOOC? What about those learners who don't comment? Are they unengaged with the MOOC and as a result drop out early?

Research Questions

This study answers three research questions:

- 1: To what extent do FutureLearn MOOC learners make comments?
- 2: Who are the commenters, and who are the learners who make lots of comments?
- 3: Does making comments, and the number of comments made, have any relationship with completion behaviour?

Method

This study uses two sources of data related to nine FutureLearn MOOCs developed by the University of Leeds. The first source is a pre-course survey relating to demographics, motivation, prior experience etc., which all learners are asked to complete. A second source of data is the learner activity within the MOOC. The two sources of data are combined via a unique learner ID. The data for this study were collected by FutureLearn and supplied to the University of Leeds. FutureLearn users are advised that data collected on the platform may be used for research purposes, and this study was conducted in accordance with the FutureLearn Code of Practice for Research Ethics. The study also gained ethical approval from the University of Leeds Ethics Committee. All data are completely anonymous and individuals cannot be identified by the University of Leeds.

Previous work and preliminary analysis for this study has shown that different MOOCs elicit different learner behaviour and different MOOCs attract different types of learners, sometimes intentionally. For example, the Politics MOOC administered in the run up to the 2015 UK General Election was marketed to AS level students, (a pre-university qualification in the main taken by 17-18 year olds) although other types of learners did take the MOOC. The median age of active learners on this MOOC is 21.5 years, whilst the first run of the World War 1 MOOC appealed disproportionately to older learners, having a median age of 50.5 years. Previous research has found that older learners are more likely to complete more of a MOOC (Morris et al., 2015) and thus by including a range of MOOCs we hope to obtain a broader picture of commenting and completion behavior, not one based on a specific, and possibly unrepresentative, cohort. We have included the nine University of Leeds MOOCs on the FutureLearn platform for which data was available when we commenced this study. In each section we will provide results for the overall dataset from the nine MOOCs, and where individual MOOCs show significant differences we will report those too.Table 1 identifies the nine FutureLearn MOOCs included in this study, and the number of steps in each MOOC which have a discussion area.

Table 1
FutureLearn MOOCs in this study and the number of steps with discussion areas

MOOC	Year of delivery	Abbreviated name	Steps that allow comments (n)
Exploring Anatomy: The Human Abdomen	2014	Anatomy 2014	66
Exploring Anatomy: The Human Abdomen	2015	Anatomy 2015	37
World War 1: Changing Faces of Heroism	2014	WW1 2014	37
World War 1: Changing Faces of Heroism	2015	WW1 2015	36
Physical Theatre	2014	Actor 2014	46
Physical Theatre	2015	Actor 2015	32
Innovation: The Key to Business Success	2014	Innovation 2014	45
Fairness and Nature: When Worlds Collide	2014	WWC 2014	40
Politics	2015	Politics 2015	36

The abbreviated name will be used throughout this paper

Within these nine MOOCs there are six unique MOOCs, three of which have been run twice. However, in all three reruns changes have been made to the content of the MOOC between run 1 and run 2.

Participants

For the purposes of this paper the participants who are included in our dataset are those learners who complete at least two steps in any MOOC. We have categorised these participants as ‘active learners’.

Table 2

Datasets used in this study

Dataset	Number	Research question/analysis
Active learners in nine MOOCs	25316	RQ1(number of learners who comment) RQ3 (relationship between being a commenter and completing)
Active learners who are commenters	8122	RQ1 (number of comments commenters make) RQ3 (relationship between number of comments made and completing)
Active learners who completed a pre-course survey	5037	RQ2 (demographics of commenters and non-commenters)
Active learners who are commenters and completed a pre-course survey	2429	RQ2 (demographics of those who make fewest/most comments)
Superposters	1095	RQ1 (number of learners who make lots of comments) RQ3 (relationship between making lots of comments and completion)
Superposters who completed a pre-course survey	445	RQ2 (demographics of those who make lots of comments)

Table 2 shows the datasets used to answer each of the research questions in this study. The active learners who make at least one comment will be known as 'commenters' whilst those who make no comments will be known as 'non-commenters'. The total dataset contains 25,316 active learners, of which 8,122 are commenters. 5,037 active learners (19.90%) across the nine MOOCs completed a pre-course survey, and of these 2,429 are commenters.

Table 3
Number of commenters and non-commenters, and mean number of steps, by survey respondents and non-survey respondents

	Survey respondents	Survey non-respondents
n=25316	n=5037	n=20279
Commenters n (%)	2429 (48.2)	5693 (28.1)
Non-commenters n (%)	2608 (51.8)	14586 (71.9)
Total steps completed (mean, sd)	27.5 (sd = 19.9)	22.11 (sd = 20.2)

Table 3 shows that survey respondents are more likely to be commenters than those who did not complete a survey. The chi-square statistic (X^2 = 751.820, df=1, $p<0.001$) shows that this difference is significant. Survey respondents also complete a higher mean number of steps than non-survey respondents. An independent t-test shows this difference is significant (X^2=17.199, df=7815.140, $p<0.001$) equal variances not assumed. We recognise this as a limitation of the study but can only access the demographic data of those who provided it in their survey responses, and we must take note that it explains the higher median number of comments in the demographics section than described when discussing the whole sample of active learners in Table 4. This does not affect our analysis of the relationship between commenting and completion of steps as the whole sample of active learners is included irrespective of whether they completed a survey.

For the third research question, which examines completion and its relationship to commenting, learners have been categorized into four groups according to classifications adopted by FutureLearn:

Week one only - learners who only engaged in the first week of the course;

Returning learners - learners who complete the first week, and return to the second week but do not complete the course;

Completers - learners who complete at least 50% of the steps and all of the assessment (required by FutureLearn to be eligible for a certificate);

100% engaged - learners who complete all the steps and assessments.

Data analysis

Data were collated, organized and analysed using Microsoft Excel and SPSS®, to explore overall trends and within MOOC differences. The analysis for this study has used a range of statistical tests. The categorical demographic data has been analysed to compare commenters and non-commenters using the chi-square test to look for group differences. The volume of comments data is highly skewed and therefore non-parametric tests have been used. Difference in volume of comments for groups of learners

has been explored using Mann Whitney U. Where there are multiple groups of learners, for example when the data set is split by employment status, their commenting behaviour has been analysed using Kruskal-Wallis test. Chi-square tests and independent t-tests have been used to assess the representativeness of the sample of survey respondents as a subset of the MOOC learner sample. Statistical significance has been measured at the $p<.05$ level throughout.

Results

RQ1: To what extent do MOOC learners make comments?

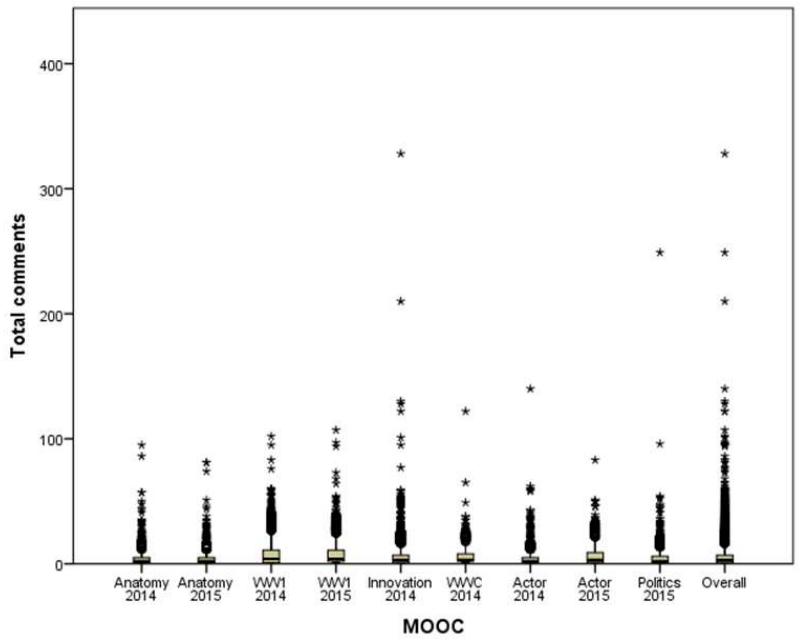
Table 4
Extent to which active learners make comments within the nine MOOCs and overall

MOOC (year)	Active learners (n)	Active learners who are commenters (n)	% of active learners who are commenters	Comments by active learners (n)	Median comments per commenter (n)
Anatomy (2014)	4382	989	22.6	4677	2.0
Anatomy (2015)	4100	766	18.7	3712	2.0
WW1 (2014)	2927	1342	45.9	11332	4.0
WW1 (2015)	1765	769	43.6	6802	4.0
Innovation (2014)	5176	1955	37.8	12510	3.0
WWC (2014)	1256	605	48.2	3782	3.0
Actor (2014)	1436	573	39.9	3171	2.0
Actor (2015)	1525	568	37.2	3895	3.0
Politics (2015)	2749	555	20.2	3453	2.0
Overall	25316	8122	32.1	53334	3.0

The median has been calculated rather than the mean due to the highly skewed nature of the data.

Of the 25,316 active learners, 8,122 (32.1%) make at least one comment. Table 4 shows the proportion of learners who are commenters for each of the nine MOOCs, and shows that it varies between MOOCs. The total number of comments made by active learners across these nine MOOCs is 53,334, and this varies between MOOCs too. Most commenters tend to make a small number of comments; the median number of comments is 3, whilst for individual MOOCs the median number of comments ranges from 2 to 4. These data suggest that even for those third of learners who make at least one comment, most don't make many comments. However, a small number of commenters make lots of comments. (See Figure 1)

Figure 1: Number of comments made by active learners (all active learners) on each of the nine MOOCs and overall



The boxplots in Figure 1 show that all MOOCs in this study and overall show a similar pattern in this respect. Overall, 2,635 commenters (32.4% of all commenters) make one comment only and 1,185 commenters (14.6% of all commenters) make two comments only. There are also commenters who make many comments. The 'superposters' as Huang et al. (2014) call learners who post most frequently, make tens and hundreds of comments (see Figure 1). The top 10% of commenters make 16 comments or more; the top 5%, 23 comments or more, the top 1% 47 comments or more. 11

commenters make over 100 comments and the highest number of comments by one learner on a single MOOC is 328. However, to be more rigorous we have taken a superposter to be the upper outliers in each set of commenters for each MOOC. We have used the definition of an outlier as any value less than the first quartile minus the interquartile range, or greater than the third quartile plus the interquartile range (Horber, 2015). Using this definition, the number of superposters in the whole data set is 1,095 (13.48% of all commenters) who made more than 13 comments.

RQ2: Who are the commenters, and who are the learners who make lots of comments?

Table 5
Proportion of commenters, non-commenters and number of comments by demographic characteristics for the active learner sample of nine MOOCs

n=5037*		Commenter n (%)	Non-commenter n (%)	Comments (n)	Median comments (n)
Gender	Male	878 (37.6)	887 (36.8)	7195	4
	Female	1445 (61.9)	1497 (62.2)	12317	4
Age	Under 18	158 (6.8)	310 (12.9)	628	2
	18-25 years	267 (11.4)	458 (19.1)	1253	2
	26-35 years	369 (15.8)	423 (17.6)	2710	3
	36-45 years	324 (13.9)	314 (13.1)	2329	4
	46-55 years	424 (18.2)	340 (14.2)	4041	4
	56-65 years	467 (20.0)	346 (14.4)	4794	5
	66+ years	325 (13.9)	211 (8.8)	3795	7
Employment	Work FT	840 (36.1)	875 (36.6)	6802	4
	Work PT	422 (18.1)	393 (16.4)	3589	4
	FT education	315 (13.5)	576 (24.1)	1568	3
	Not working	750 (32.2)	547 (22.9)	7550	5
Prior education	Less than high school/sec ed	60 (2.6)	85 (3.6)	393	2
	High school/sec ed	557 (24.0)	698 (29.4)	3392	3
	Univ degree	1026 (44.3)	992 (41.8)	8588	4
	Univ masters	574 (24.8)	482 (20.3)	5902	4
	PhD	100 (4.3)	114 (4.8)	1131	6
Online experience	Prior online	1098 (47.3)	923 (38.9)	10499	4
	No prior online	1223 (52.7)	1447 (61.1)	8894	3
Country	UK	1432 (61.2)	1457 (60.9)	11716	3
	Non-UK	907 (38.8)	935 (39.1)	7869	3

*n varies as not all survey respondents answered every question

Age

The data in Table 5 show that commenters are more likely to be older than non-commenters. The chi-square statistic shows this difference is significant (χ^2=154.07, df=6, p< 0.001). Standardised residuals (in brackets) show that the significant difference is caused by more learners under 18 (4.7) and aged 18-25 (4.7) not commenting and more learners aged 46-55 (2.4), 56-65 (3.3) and 66 + (3.7) being commenters. The data also show that older commenters make more comments. Kruskal Wallis test shows this difference is significant (χ^2=71.38, df=6, p< 0.001).

Employment status

Table 5 also shows that commenters are slightly more likely to work part-time or to not work than to work full time and much less likely to be in full-time education. Chi-square shows this difference is significant (χ^2 =109.125, df=3, p< 0.001). Standardised residuals (in brackets) show that the significant difference is caused by more of those in full-time education (5.9) not commenting and more of those not working (4.4) being commenters. Full-time workers and students who comment also make fewer comments than those working part-time or not at all. Using Kruskal Wallis, pairwise comparisons, the adjusted values show that those in full-time education make significantly fewer comments than all other groups (p< 0.001 for working full time and working part time, p< 0.001 for not working) and those not working make significantly more comments than those who work full time (p< 0.001) .

Prior education

There is very little evidence to suggest that commenters are more likely to be better educated. Within the commenters group there is a slight tendency for those who are more educated to make more comments. Kruskal-Wallis test shows a significant difference, (χ^2= 21.95, df=4, p< 0.001), and pairwise comparisons show that the significant difference is between those with a high school or secondary

education making fewer comments than those with a degree ($p=0.002$), masters ($p<0.001$) or PhD ($p=0.004$).

Prior online experience

Overall, commenters are more likely to have taken an online course previously and this difference is significant ($\chi^2=33.44$, df=1, $p<0.001$). A similar pattern is found for six of the nine MOOCs. However, for three MOOCs (WW1 2015, WWC 2014 and Actor 2014), the proportion of commenters was higher amongst those who had not previously taken a course online, although not significantly so. The data also show that commenters who have taken an online course previously tend to make more comments and this difference is significant (Mann Whitney U = 748,622.50, $p<0.001$).

The analysis also examined gender and country (UK vs. non-UK), but overall no significant differences were found.

Superposters

We have already noted that most commenters don't actually make very many comments, whilst a substantial minority make lots of comments. Do these 'superposters' have similar demographic characteristics to commenters generally? Amongst commenters who completed a pre-course survey there are 445 superposters and 1,984 commenters who are not superposters. Table 6 shows the different demographic characteristics of these two groups.

Table 6
Demographic characteristics of commenters, comparing superposters to non-superposters for the active learner sample of nine MOOCs

n=2429		Superposter (n=445) n (%)	Non-superposter (n=1984) n (%)
Gender	Male	168 (39.0)	710 (37.3)
	Female	261 (60.6)	1184 (62.2
Age	Under 18	8 (1.9)	150 (7.9)
	18-25 years	21 (4.9)	246 (12.9)
	26-35 years	49 (11.4)	320 (16.8)
	36-45 years	57 (13.3)	267 (14.0)
	46-55 years	89 (20.7)	335 (17.6)
	56-65 years	113 (26.3)	354 (18.6)
	66+ years	93 (21.6)	232 (12.2)
Employment	Work FT	143 (33.3)	697 (36.7)
	Work PT	87 (20.2)	335 (17.7)
	FT education	27 (6.3)	288 (15.2)
	Not working	173 (40.2)	577 (30.4)
Prior education	Less than high school/sec ed	5 (1.2)	55 (2.9)
	High school/sec ed	75 (17.5)	482 (25.5)
	Univ degree	190 (44.3)	836 (44.3)
	Univ masters	136 (31.7)	438 (23.2)
	PhD	23 (5.4)	77 (4.1)
Online experience	Prior online	251 (58.9)	847 (44.7)
	No prior online	175 (41.1)	1048(55.3)
Country	UK	248 (57.5)	1184 (62.1)
	Non-UK	183 (42.5)	724 (37.9)

The data in Table 6 show that superposters compared to other commenters are older (χ^2= 79.896, df=6, $p<0.001$), standardized residuals (in brackets) show that those under 18, (-3.9), 18-25 (-4.0) and 26-35 (-2.3) are less likely to be superposters, whilst those aged 56-65 (2.9) and 66+ (4.3) are more likely to be superposters. A significant difference is found for employment status (χ^2= 33,465 df=3, $p<0.001$). standardized residuals (in brackets) show that those not working (2.9) are more likely to be a superposter whilst those in full-time education (-4.1) are less likely to be a superposter. There is a significant difference by prior education (χ^2= 25.073, df=4, $p<0.001$). Standardised residuals (in brackets) show that

more superposters are likely to have a Master's degree (2.9) and be less likely to have secondary education only (-2.8). Superposters are also more likely to have done a course online before (χ^2= 28.228, df=1, p< 0.001). There is no significant difference based on gender or country.

RQ3: Does making comments, and the number of comments made, have any relationship with completion behaviour?

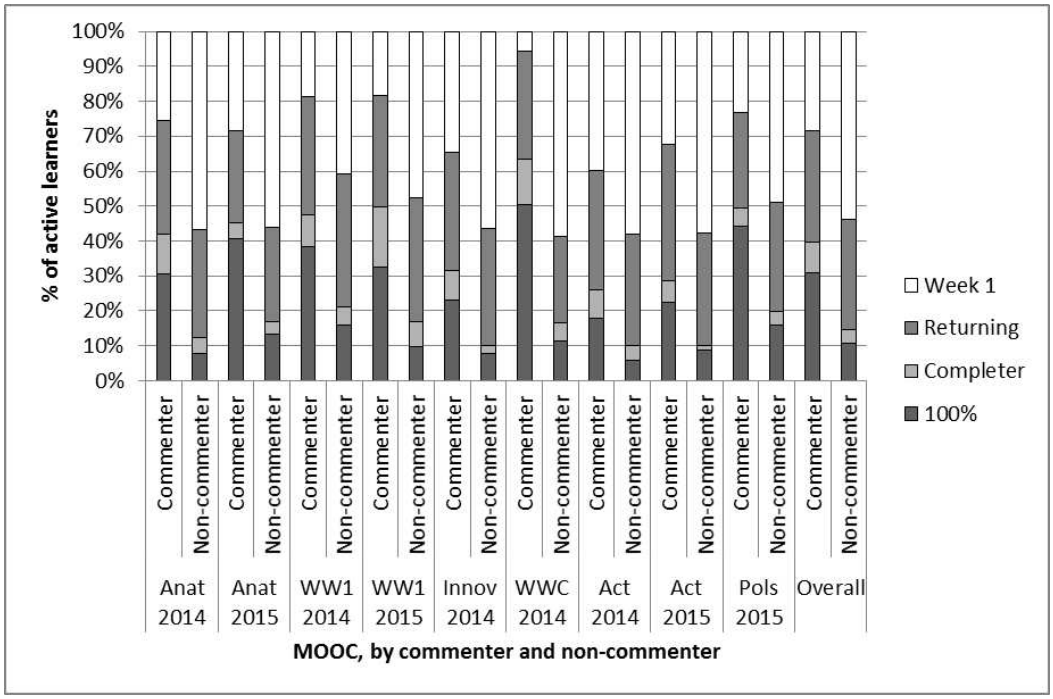
Morris et al. (2015) found that older learners, with a degree or higher, who are not working and have done a course online previously are more likely to complete more of the MOOC. This study has also found that these learners are also more likely to make a comment or to comment more. Given that the tendencies to comment, and to comment more, concur with the tendencies to complete, this raises the question as to what, if any, the relationship is between commenting and completing.

Table 7
Number and (percentage) of commenters and non-commenters in each of the completion groups, overall

n=25316	Commenter	Non-commenter
Overall	8122	17194
Week 1 n (%)	2312 (28.5)	9243 (53.8)
Returning learner n (%)	2578 (31.7)	5428 (31.6)
Completer n (%)	728 (9.0)	660 (3.8)
100% n (%)	2504 (30.8)	1863 (10.8)

The completion groups of week 1, returning learner, completer and 100% are not strictly ordinal as returning and completer do overlap somewhat. However, generally the proportion of the MOOC completed increases through these groups from week 1 to 100%. Figure 2 and the data shown in Table 7, indicate that learners in the higher completion groups are more likely to be commenters and chi-square shows this difference is significant (χ^2= 2315.799, df=3, p< 0.001). This pattern is found for each of the nine MOOCs too (see Figure 2).

Figure 2: Percent of active learners in each completion category for each MOOC, by commenters/non-commenters



It is more difficult to assess the relationship between the number of comments a learner makes and their completion category, because an individual commenter has more opportunity to make comments the further they progress in a MOOC. The data in Table 8 support this showing that the median number of comments made (by commenters) increases as the completion group increases. Kruskal-Wallis Test, (χ^2=1615.295, df=3, $p<0.001$) shows this difference is significant and pairwise comparisons show the difference is significant in this direction between all completion groups.

Table 8
Median number of comments by each completion group

N=8122	Week 1 (n=2312)	Returning learner (n=2578)	Completer (n=728)	100% (n=2504)
Median no. of comments	1	3	5	6

Each MOOC has a different number of steps and thus a different number of steps need to be completed to achieve 100% completion. In order to show the relationship between completion groups and volume of comments made, each MOOC is shown separately in Figure 3. These graphs show that completing more steps is not consistently associated with an increase in comments (and vice versa), until the number of steps completed approaches the total number of steps in the MOOC. As total steps completed approaches 100% the prevalence of those who make lots of comments, the superposters, increases greatly.

Figure 3: Scatterplots showing relationship between number of steps and number of comments per active learner who comments, on each MOOC, by completion group

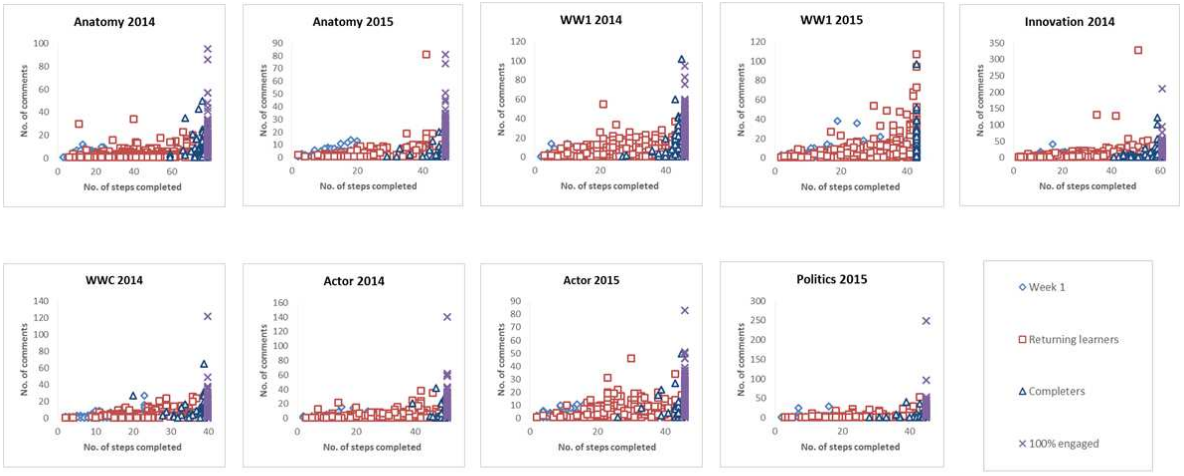


Figure 3 shows that the superposters tend to be in the 100% completion group, more than any other completion group. The proportion of superposters in each MOOC who complete all of the MOOC ranges from 53.6%, to 84.5%, (mean = 70.7, sd = 11.8). These active learners make the most comments and complete every step. However, this is not to suggest that all learners who complete all of the MOOC are making lots of comments. There are a substantial proportion of learners who complete 100% of the MOOC whilst making no comments whatsoever. The

complete dataset of active learners contains 4,367 learners who completed all the steps in their MOOC. Of these 1,863 (42.7%) made no comments at all.

Discussion

The FutureLearn social constructivist approach emphasises the value that social interaction has for the learning experience within their MOOCs and in order to facilitate this most steps within FutureLearn MOOCs have a discussion area. This study has found that despite a strong encouragement for learners to get involved in the online discussions, most learners do not make any comments and of those learners who do post at least one comment, most make only a very small number of comments, with 47.0% of commenters making just one or two comments only. There are however, a small proportion of learners who make a lot of comments.

MOOC learners have a variety of motivations for enrolling on a course (Hew & Cheung, 2014), some of which encourage social interaction. Zheng et al. (2015), use four categories of MOOC motivations in their research, one of which is 'connecting with people', and suggest that some learners enroll on MOOCs to find like-minded people with similar interests that they can connect with, and one could suggest that those who enrol with this goal are more likely to get involved in discussions. However, it is not so easy to predict commenting behaviour when the goal is 'I'd like to explore this subject area for possible career change' or 'This subject is relevant to my academic field of study' (both from the FutureLearn pre-course survey). MOOC learners who do not complete the whole of the course may still achieve their learning goal (Liyanagunawardena et al., 2014; Reich & Ho, 2014). The same is true of commenting; learners can achieve their goals in a MOOC without making any comments at all.

The subject content of a MOOC, and the type of learners it attracts may also affect the extent to which learners get involved in the discussions. For example, both runs of the WW1 MOOC elicited a relatively high number of comments and both also have an older cohort of learners, and we have found that older learners are more likely to comment. Conversely, Politics 2015, which was marketed to and attracted many learners of school age, had the youngest cohort of learners, a very low proportion of

commenters and the lowest median number of comments. Previous research has shown that dropout from MOOCs can often be caused by lack of time (Khalil & Ebner, 2014; Liu et al. 2015; Zhang et al. 2015), and it seems reasonable to suggest that engaging in discussion forums may also be affected by time constraints as those in full time occupations, working full time or in full time education make fewer comments. The 17-18 year olds in the Politics 2015 MOOC may have made fewer comments because they're younger and less confident, or it may be due to a lack of time due to prioritizing their full time studies. Those who are more highly educated may be more confident about the subject matter, or about learning generally and believe they have something to offer to the discussion, which makes such learners more likely to make a comment. Alario-Hoyes et al. (2016) suggest that those learners who make the most comments sometimes take on the role of educators, presumably feeling they are 'expert' enough in the MOOC subject to take on such a role. Learners who have prior experience online are also more likely to make more comments, and this may also reflect a higher level of confidence which derives from being familiar with the environment and having had experience of commenting previously. Milligan et al. (2013) found that cMOOC learners who were active in blogs and on Twitter were very likely to have prior experience of participation in a cMOOC, whilst passive participants were not. They assert that 'It seems clear that learners must learn how to learn in a cMOOC.' (2013: 156).This may well be the case for FutureLearn MOOC learners too. Demographics of learners may go some way to explain some of the different behavior of learners. Further research linking demographics, motivations and activity in MOOCs is needed to shed further light on why some learners comment and why some do not.

In relation to the third research question, we found that learners who comment have similar demographic characteristics to those learners who complete more of the MOOC, and that there is a strong relationship between being a commenter and completing. These findings concur with previous research examining forum activity and retention (Balakrishnan & Coetzee, 2013; Breslow et al., 2013; Kizilcec et al., 2013; Woodgate et al., 2015). However, for most commenters, those who make just a few comments, the volume of comments they make has only a weak relationship with completion. It is the superposters, those active learners who make lots of comments, who are much more likely to complete. Without further

analysis these relationships between commenting and completing cannot be said to be causal. However, it is likely that similar characteristics which result in completion also make it more likely to comment; being used to completing a course, being confident and self-assured, having experience of this type of course previously and having relatively more free time. It is also likely that 'engaged' learners both comment and complete. As Yang et al. (2014) have suggested, those students who express lower levels of engagement are more likely to drop out. Is this the case in these MOOCs? Are our commenting completers more motivated than those who don't comment and don't complete? Or is the support they receive from being socially interactive in the discussion areas keeping them focused and engaged to complete the MOOC? This may well be the case for the superposters who are involved in sustained dialogue throughout the MOOC. However, for those who make just a few comments, and whose interactions cannot really be described as conversations and possibly having a more similar experience to non-commenters than superposters, are they falling away due to the unsustained nature of their social interaction and any support that may provide. For these learners, does a sense of a community exist at all, or does it only exist in those steps in which they comment, making it even more short-lived than Hart suggests? There is a small but significant proportion of those who go on to complete most or all of their MOOC who make no comments, a group who seemingly do not require support from educators or other learners. The less educated, younger learners who are in full-time education might be thought to benefit more from this kind of support and from social interaction than older, more educated learners. But it is the older more educated learners making more comments. However, for most learners it is too simplistic to describe the commenter / non-commenter groups as separate unrelated groups and suggestive of active and passive. This distinction does not take into account the other ways in which social interaction takes place within MOOCs; by reading posts, liking posts and following educators, mentors and other learners. The commenters, and especially those who comment a lot, may well affect the learning experience of the non-commenters. Those learners who choose to not comment may read many of the comments posted, they may like one or more posts, they may follow educators or other learners and by doing so will almost certainly feel more part of a learning community than the label of non-commenters suggests. The mere

existence of the thousands of comments could make those non-commenters feel less isolated and provide them with what Hart (2012) calls 'increased comfort' and a feeling of being part of a learning community. Morris et al. (2005) found statistically significant differences in the amount of time spent in engagement activities between students who withdraw from a course and those who complete, and for them engagement activities include time spent reading as well as responding to posts. Balakrishnan & Coetzee (2013) found that students who did not read forum posts were more likely to drop out as the course progressed, whereas viewing just one thread a week made a student much more likely to stay. They suggest that students can participate actively in the course without contributing to a forum and suggest that not making comments is not the greatest predictor of dropout, not reading them is a more important factor. Younger, less educated learners might well be getting the support they need, and a feeling of being in a social community by doing precisely these things but without actually commenting themselves.

Limitations

Unfortunately, at this time we do not have access to data relating to the extent to which non-commenters (or commenters) read comments made by others, whether they like comments made by others or whether they follow other learners. Data of this nature would provide more evidence about the behaviour of non-commenters, and those who don't comment so much. How passive are they? It is likely that it varies, with some non-commenters being very active within the MOOC and others less so. This is research for the future when this type of data becomes available.

We also found that those learners who completed a pre-course survey are more likely to be commenters and to have completed more of the MOOC than non-respondents. We recognise this is a limitation of the study in respect of the demographic analysis carried out to provide a description of who comments and who is more likely to complete a MOOC, although when comparing one demographic group with another, they are all from the dataset of survey respondents and thus this factor is present for

all in this subset. Also this does not affect the analysis for RQ1 or RQ3 as the full data set is used, not the sub-set of pre-course survey respondents

Conclusion

This study has found that around a third of FutureLearn active learners post at least one comment, of whom most make just a handful of comments. However, a small number of learners make many comments. The study has found that older learners who work part-time or not at all are more likely to be commenters, and those who are older, working part-time or not at all, are better educated and have prior online experience make the most comments, and are very much more likely to be superposters. The study has also found that there is strong relationship between those who complete more steps in a MOOC and being a commenter, and in particular there is a strong relationship between being a superposter and completing all of the steps in a MOOC. This is a first step in understanding the relationship that social interaction within a MOOC has on engagement in the form of completing a MOOC. The limitations of this study related to other kinds of activity within MOOCs, such as reading, liking and following will be able to be overcome when this data becomes available, hopefully in the near future.

References

Alario-Hoyos, C., Munoz-Merinos, P.J., Perez-Sanagustin, M., Delgado Kloos, C. & Parada G, H.A. (2016). Who are the top contributors in a MOOC? Relating participants' performance and contributions. Journal of Computer Assisted Learning. Vol 32, No. 3, pp.232-243.

Balakrishnan, G. & Coetzee, D. (2013). Predicting Student Retention in Massive Open Online Courses using Hidden Markov Models. EECs Department, University of California, Berkeley.

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 and Practice in Assessment. 8, pp.13-25.

Cisel, M. (2014). Analysing completion rates in the First French xMOOC. EMOOCs 2014: European MOOCs Stakeholders Summit. Proceedings. Research Track. pp.26-32.

Ferguson, R. & Clow, D. (2015). Examining Engagement: Analysing Learner Subpopulations in Massive Open Online Courses (MOOCs). In: 5th International Learning Analytics and Knowledge Conference (LAK15), 16–20 March 2015, Poughkeepsie, NY, USA, ACM.

Ferguson, R. & Sharples, M. (2014). Innovative pedagogy at massive scale: teaching and learning in MOOCs. 9th European Conference on Technology Enhanced Learning (EC-TEL 2014): Open Learning and Teaching in Educational Communities, 16-19 September 2014, Graz, Austria, Springer International Publishing, pp.98–111.

Gasevic, D., Kovanovic, V., Joksimovic, S. & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. The International Review of Research in Open and Distance Learning. 15(5), pp.134-176.

Guo, P.J. & Reinecke, K. (2014). Demographic Differences in How Students Navigate Through MOOCS. Proceedings of the first ACM conference on Learning@ scale conference (21-30). ACM

Hart, C. (2012). Factors Associated With Student Persistence in an Online Program of Study: A Review of the Literature. Journal of Interactive Online Learning. 11(1), pp.19-42.

Hew, K.F. & Cheung, W.S. (2014). Students' and instructors use of massive open online courses (MOOCs): Motivations and challenges. Educational Research Review. 12, pp.45-58.

Horber, E. (2015). Outliers in SPSS. http://www.unige.ch/ses/sococ/cl/spss/concepts/outliers.html Accessed September 11, 2015

Huang, J., Dasgupta, A., Ghosh, A., Manning, J. & Sanders. M. (2014). Superposter behaviour in MOOC forums. In Proceedings of the first ACM conference on Learning@ scale conference. pp.117-126. ACM.

Jordan, K. (2014). Initial Trends in Enrolment and Completion of Massive Open Online Courses. The International Review of Research in Open and Distance Learning. 15(1), pp.133-160.

Khalil, H. & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention - A Literature Review. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2014. pp.1236-1244.

Kizilcec, R.F., Piech, C & Schneider, E. (2013). Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses. LAK conference presentation. Accessed 15 December 2014.

Laurillard, D. (2002). Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies. Routledge: London.

Lave, J. and Wenger, E. (1991). Situated learning: legitimate peripheral participation. Cambridge: Cambridge University Press.

Liu, M., Kang, J. & McKelroy, S. (2015). Examining learners' perspectives of taking a MOOC: reasons, excitement, and perceptions of usefulness. Educational Media International. 52(2), pp.129-146.

Liyanagunawardena, T.R., Parslow, P. & Williams, S.R. (2014). Dropout: MOOC Participants' Perspective. EMOOCS 2014: European MOOCs Stakeholders Summit. Proceedings. Research Track. 95-100.

Milligan, C., Littlejohn, A. and Margaryan, A. 2013. Patterns of engagement in Connectivist MOOCs. Journal of Online Learning and Teaching. 9(2), pp.149-159.

Morris, L.V., Finnegan, C. & Wu, S-S. (2005). Tracking student behavior, persistence, and achievement in online courses. The Internet and Higher Education. 8(3), pp.221-231.

Morris, N.P., Hotchkiss, S. & Swinnerton, B. et al. (2015). Can demographic information predict MOOC learner outcomes? http://www.emoocs2015.eu/sites/default/files/Papers.pdf eMOOCs Conference, May 2015

Muilenberg, L.Y. and Berge, Z.L. (2001). Barriers to distance education: A factor-analytic study. The American Journal of Distance Education. 15(2), pp.7-22.

Palin, A. (2014). Moocs: Young students from developing countries are still in the minority. FT.com. March 9. 2014.

Perna, L.W., Ruby, A., Boruch, R.F., N, Scull, J, Ahmad, S & Evans, C. (2014). Moving Through MOOCs: Understanding the Progression of Users in Massive Open Online Courses. Educational Researcher. 43(9), pp.421-432.

Reed, M.S., Evely, A.C., Cundill, G., Fazey, I., Glass., Laing, A., Newig, J., Parrish, B., Prell, C., Raymond, C. & Stringer, L.C. (2010) What is Social Learning?. Ecology and Society. 15(4).

Reich, J. & Ho, A. (2014). The Tricky Task of Figuring Out What Makes A MOOC Successful. The Atlantic. 23 January 2014, . http://www.theatlantic.com/education/archive/2014/01/the-tricky-task-of-figuring-out-what-makes-a-mooc-successful/283274/ Accessed 8 January 2015.

Robinson, J. (2011). Assessing the value of using an online discussion board for engaging students. Journal of Hospitality, Leisure, Sport and Tourism Education. 10(1), pp.13-22.

Rovai, A.P. (2002). Building Sense of Community at a Distance. The International Review of Research in Open and Distributed Learning. 3(1).

Sharples, M., Adams, A., Ferguson, R., Gaved, M., McAndrew, P., Rienties, B., Weller, M. & Whitlock, D. (2014). Innovating Pedagogy 2014: Open University Innovation Report 3. The Open University: UK.

Sharples, M. (2016, April). Effective Pedagogy at Scale: Social Learning and Citizen Inquiry. In Proceedings of the Third (2016) ACM Conference on Learning@ Scale (pp. 339-340). ACM.

Stump, G., Hilpert, J., Husman, J., Chung, W-T. & Kim, W. (2011). Collaborative Learning in Engineering Students: Gender and Achievement. Journal of Engineering Education. 100(3), pp.1-24.

Wen, M., Yang, D. & Rosé, C.P. (2014). Linguistic Reflections of Student Engagement in Massive Open Online Courses. In Proceedings of the International Conference on Weblogs and Social Media.

Wexler, E. (October 19 2015). MOOCs are still rising, at least in numbers. Wired Campus. http://chronicle.com/blogs/wiredcampus/moocs-are-still-rising-at-least-in-numbers/57527 Accessed 28 October 2015.

Woodgate, A., Macleod, H., Scott, A-M. & Haywood, J. (2015). Differences in online study behavior between sub-populations of MOOC learners. Educacion XXI. 18(2) pp.147-163.

Yang, D., Sinha, T., Adamson, D. & Rosé, C.P. (2013). Turn on, Tune in, Drop out: Anticipating student dropout in Massive Open Online Courses. NIPS Data–Driven Education Workshop.

Yang, D., Wen, M., Kumar, A., Xing, E.P. & Rosé, C. P. (2014). Towards an Integration of Text and Graph Clustering Methods as a Lens for Studying Social Interaction in MOOCs. The International Review of Research in Open and Distance Learning.. 15(5), pp.214-234.

Zheng, S, Rosson, M.B., Shih, P.C. and Carroll, J.M. (2015). Understanding Student Motivation, Behaviors and Perceptions in MOOCs. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. ACM, 2015. pp.1882-1895.