Moving through MOOCS: pedagogy, learning design and patterns of engagement

Conference or Workshop Item

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Design for Teaching and Learning in a Networked World	
Series Title		
Chapter Title	Moving Through MOOCS: Pedagogy, Learning Design and Patterns of Engagement	
Copyright Year	2015	
Copyright HolderName	Springer International Publishing Switzerland	
Corresponding Author	Family Name	**Ferguson**
	Particle	
	Given Name	**Rebecca**
	Prefix	
	Suffix	
	Division	
	Organization	The Open University
	Address	Milton Keynes, UK
	Email	Rebecca.Ferguson@open.ac.uk
Author	Family Name	**Clow**
	Particle	
	Given Name	**Doug**
	Prefix	
	Suffix	
	Division	
	Organization	The Open University
	Address	Milton Keynes, UK
	Email	
Author	Family Name	**Beale**
	Particle	
	Given Name	**Russell**
	Prefix	
	Suffix	
	Division	
	Organization	University of Birmingham
	Address	Birmingham, UK
	Email	
Author	Family Name	**Cooper**
	Particle	
	Given Name	**Alison J.**
	Prefix	
	Suffix	
	Division	
	Organization	University of Birmingham
	Address	Birmingham, UK

	Email	
Author	Family Name	**Morris**
	Particle	
	Given Name	**Neil**
	Prefix	
	Suffix	
	Division	
	Organization	University of Leeds
	Address	Leeds, UK
	Email	
Author	Family Name	**Bayne**
	Particle	
	Given Name	**Siân**
	Prefix	
	Suffix	
	Division	
	Organization	University of Edinburgh
	Address	Edinburgh, UK
	Email	
Author	Family Name	**Woodgate**
	Particle	
	Given Name	**Amy**
	Prefix	
	Suffix	
	Division	
	Organization	University of Edinburgh
	Address	Edinburgh, UK
	Email	

Abstract	Massive open online courses (MOOCs) are part of the lifelong learning experience of people worldwide. Many of these learners participate fully. However, the high levels of dropout on most of these courses are a cause for concern. Previous studies have suggested that there are patterns of engagement within MOOCs that vary according to the pedagogy employed. The current paper builds on this work and examines MOOCs from different providers that have been offered on the FutureLearn platform. A cluster analysis of these MOOCs shows that engagement patterns are related to pedagogy and course duration. Learners did not work through a three-week MOOC in the same ways that learners work through the first three weeks of an eight-week MOOC.

Keywords (separated by '-')	Learning analytics - Learner engagement patterns - Learning design - Moocs - Pedagogy

Moving Through MOOCS: Pedagogy, Learning Design and Patterns of Engagement

Rebecca Ferguson[1(✉)], Doug Clow[1], Russell Beale[2], Alison J. Cooper[2], Neil Morris[3], Siân Bayne[4], and Amy Woodgate[4]

[1] The Open University, Milton Keynes, UK
Rebecca.Ferguson@open.ac.uk
[2] University of Birmingham, Birmingham, UK
[3] University of Leeds, Leeds, UK
[4] University of Edinburgh, Edinburgh, UK

AQ1

Abstract. Massive open online courses (MOOCs) are part of the lifelong learning experience of people worldwide. Many of these learners participate fully. However, the high levels of dropout on most of these courses are a cause for concern. Previous studies have suggested that there are patterns of engagement within MOOCs that vary according to the pedagogy employed. The current paper builds on this work and examines MOOCs from different providers that have been offered on the FutureLearn platform. A cluster analysis of these MOOCs shows that engagement patterns are related to pedagogy and course duration. Learners did not work through a three-week MOOC in the same ways that learners work through the first three weeks of an eight-week MOOC.

Keywords: Learning analytics · Learner engagement patterns · Learning design · Moocs · Pedagogy

1 Introduction

The overarching label 'massive open online course' sets out some of the common features of this approach to learning and teaching: size, philosophy, mode of delivery and structure. The generic title does not specify the pedagogy, the theory of learning and teaching, which underpins MOOCs. Nor does it specify the types of learning design – the ways in which courses are planned, sequenced and managed – that work well with this model [4].

The first MOOCs [2] employed a connectivist approach, linking people and information on the basis that 'the connections that enable us to learn more are more important than our current state of knowing' [13]. Connectivist MOOCs were followed by xMOOCs, which typically employ a pedagogy based on content delivery [10].

The FutureLearn MOOC platform takes a different approach, and employs a social constructivist pedagogy that is based on the Conversational Framework [9, 11]. This is a theory of learning effectively through conversations with yourself and others [12]. In order to do this successfully, learners need access to both a shared representation of the

G. Conole et al. (Eds.): EC-TEL 2015, LNCS 9307, pp. 1–15, 2015.
DOI: 10.1007/978-3-319-24258-3_6

subject matter and tools that support reflection, comment and responses. These were incorporated into the design of the FutureLearn platform.

Despite these differences, MOOC providers all share a concern about dropout rates. Almost without exception, there is a large disparity between numbers registering and numbers completing a MOOC [6]. There are positive explanations for this discrepancy and many of those who leave early have gained what they wanted from the course and do not regard themselves as dropouts [1, 3]. Nevertheless, the open nature of MOOCs is not just about making more resources available to more people; it is also about extending opportunities.

An overview of this issue is provided by Jordan [5], and her website offers an opportunity to explore the data in more depth [6]. At the time of writing, 'MOOC Completion Rates: The Data' includes information about more than 180 MOOCs. Only two of these have reported a retention rate of over 40 %, and most have a completion rate of less than 13 %. A possible explanation is that this variation is due to differences in pedagogy and learning design.

Section 2 of this paper looks at past research to investigate patterns of engagement, including a cluster analysis of Coursera MOOCs that identified four engagement patterns, and subsequent work on Open University FutureLearn MOOCs, which showed that patterns of engagement are affected by pedagogy. Sections 3 and 4 introduce the method used to expand this study, using data from four universities. Section 5 reports the first research phase, indicating that learning design influences learner engagement. Section 6 covers the second research phase, which identifies elements of design that affect patterns of engagement, and Sect. 7 deals with the third phase, which identifies patterns of engagement in MOOCs with different learning designs. Sections 8 and 9 discuss these findings and set out the implications for research and practice.

2 Using Analytics to Investigate Engagement

Learning analytics are concerned with the use of trace data relating to learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs [14]. They therefore provide ways of investigating patterns of engagement on MOOCs and, potentially, highlighting pedagogic approaches and learning designs that prove successful in retaining learners.

Kop and her colleagues have published several studies of connectivist MOOCS, including an analysis of patterns of engagement on PLENK2010 [8]. Their work pointed to two patterns of engagement, the *Visible Contributors* and the *Consumers*.

Kizilcec and his colleagues classified engagement on three MOOCs on the Coursera platform ('the Coursera study') [7] in terms of video lectures and assessments, and identified four patterns of engagement:

- *Sampling*: learners explored some course videos.
- *Auditing*: learners watched most videos, but completed assessments rarely, if at all
- *Disengaging*: learners reduced their engagement over time.
- *Completing*: learners completed most assessments

They noted that ‘MOOC designers can apply this simple and scalable categorization to target interventions and develop adaptive course features’.

2.1 Engagement on Social Constructivist MOOCs

In order to explore whether these patterns could indeed be used by MOOC designers on any platform, a study of four MOOCs run by The Open University on the FutureLearn platform (‘the OU study’) [4] aimed to replicate the method employed on the Coursera study [7]. Learners’ activity in each week was assigned to one of four categories: *on track* if they submitted assessment in the week it was set, *behind* if they completed an assessment after the week in which it was set, *auditing* if they engaged with content but not with the assessment, and *out* if they did not participate in a course week. Once engagement profiles had been created for learners, the *k*-means clustering algorithm was used to partition learners into a small number of groups. This produced two clusters that resembled the *Completing* and *Sampling* groups from the Coursera study, but neither an *Auditing* nor a *Disengaging* group [4, 7].

The original methodology prioritized engagement with content and assessment. However, the FutureLearn MOOCs laid stress on the social construction of knowledge, including discussion alongside each content step. The OU study therefore went on to create engagement profiles for learners that reflected engagement with content, with assessment and with discussion [4]. The method used to do this was employed in the current study, and is described in Sect. 4.1.

This generated seven clusters. *Samplers* visited a course briefly. *Strong Starters* left after the first week’s assessment. *Returners* completed assessments in the first two weeks, then left. *Mid-way Dropouts* completed 3–4 assessments before leaving. The *Nearly There* cluster completed most assessments but left early. *Late Completers* completed most assessments but were either late in submitting these or missed some. *Keen Completers* engaged actively throughout. Typically, learners posted an average of around one comment for each week of engagement with the course [4].

MOOCs in the OU study were mainly eight-week courses with an assessment point at or near the end of each week. OU MOOC2 ran for a shorter period of time, and here the *Midway Dropouts* cluster was replaced by another cluster that fell between the *Samplers* and the *Strong Starters*. OU MOOC3, on the other hand, ran for eight weeks, but only included three assessments. In this case, the *Returners* and the *Midway Dropouts* were replaced with a cluster of *Samplers Who Comment*, and by a much smaller cluster of those whose engagement was concentrated on the final week.

The OU study showed that patterns of engagement are influenced by differences in pedagogy and course duration. This was an important point in terms of practice. If engagement patterns on different MOOCs are essentially similar once variations in pedagogy have been taken into account, then reducing course duration should reduce dropout. If MOOC learners find it difficult to complete assessment after the first week of a course, MOOCs could be structured to take this into account. However, if duration influences engagement patterns, then these approaches are less likely to succeed.

The research reported in this paper builds on the Coursera study and the OU study and investigates patterns of learner engagement on FutureLearn MOOCs from four universities. Each of these MOOCs has a different and carefully considered learning design, containing many elements. In this study, we have focused on two of these – length of course and distribution of assessment – as these can be easily identified within sets of activity data. The study asks whether the engagement patterns identified on the four Open University MOOCs are found on MOOCs delivered on the same pattern by different universities, and whether engagement patterns are influenced by changes in learning design.

3 FutureLearn Multi-University Dataset

To investigate these questions, we used data supplied by FutureLearn, a company owned by The Open University that has developed a social-constructivist platform for delivering free online courses, which currently has over 1.5 million registered learners. Course weeks are divided into steps. Except for assessments, each of these is associated with a free-flowing discussion. 'Typically, the discussion associated with any step on a FutureLearn MOOC will attract hundreds or thousands of contributions, with 'Like' and 'Follow' options providing ways of navigating these' [4].

Each partner institution has access to its own course data. Here, we focus on five MOOCs from four institutions. Together, these five sets of data include information about the activity of all 32,942 learners who engaged actively with these courses.

3.1 MOOCs in the Cross-University Dataset

LongMOOC1 was an eight-week course from the OU, including assessment in each week, focused on the hard sciences. It is included here because it was the MOOC for which the eight OU study cluster descriptions were first developed. The dataset includes activity data from 5,069 learners who were active on this MOOC.

LongMOOC2 was a seven-week course from the University of Edinburgh, including assessment each week, focused on the hard sciences. It was included to test whether the clusters generated by the OU study could be replicated on courses at other institutions. In all, 10,136 learners were active on this MOOC.

TalkMOOC3 was a six-week politics course from the University of Edinburgh. This was included because it had a distinctive learning design and pedagogy. The course provided a period of structured discussion, enabling participants to develop an informed position on an issue, so the course included no assessment. The dataset includes activity data from 6,141 learners who were active on this MOOC.

ShortMOOC4 was a three-week course from the University of Birmingham with a focus on the life/medical sciences. This was included because it was a short course, offering a broad introduction to concepts that could be explored in more depth in subsequent courses. In all, 6,839 learners were active on this MOOC.

ShortMOOC5 was a three-week course from the University of Leeds with a focus on the life/medical sciences. This was included alongside ShortMOOC 4 to test whether three-week courses showed consistent patterns of engagement. The dataset includes activity data from 4,756 learners who were active on this MOOC.

4 Clustering

4.1 Replicating the Method

The current study began by replicating the method used in the OU study. We made use of anonymised individual-level activity data. The data processing that took place before we received the data meant they were partially aggregated, and so we had access to the data and time of a learner's first visit to a content step, but we did not have access to the date and time of any subsequent visits. Graphical inspection of these data did not reveal any major differences that could be attributed to this process.

Although there is no compulsion to keep to the course schedule, an overview of data showed that activity followed a weekly cycle, spiking after the weekly course email was sent out on a Monday [4]. This weekly pattern meant that this study could follow previous studies and divide the data into weekly segments.

For each course week, we assigned learners an activity score of 1 if they viewed content, 2 if they posted a comment, 4 if they submitted their assessment in a subsequent week, and 8 if they submitted it early or on time. Individuals' activity scores were summed in order to give an activity score for each week. For example, if a learner visited content, commented and submitted on time in weeks 1–5 visited content and did the assessment late in week 6, visited content in week 7 and did not participate in week 8, their engagement profile would be [11, 11, 11, 11, 11, 5, 1, 0].

Possible weekly activity scores:

1 = Visited content only
2 = Posted comment but visited no new content
3 = Visited content and posted comment
4 = Submitted the assessment late
5 = Visited content and submitted assessment late
6 = Posted late assessment, saw no new content
7 = Visited content, posted, late assessment
8 = Submitted assessment early /on time
9 = Visited content, assessment early /on time
10 = Posted, assessment early /on time, no new content
11 = Visited, posted, assessment early /on time

Using these engagement profiles, we applied the *k*-means clustering algorithm to split the learners into a small number of groups. To cluster engagement patterns, we used the L_1 norm as the basis for one-dimensional k-means clustering, thus minimising the sum of the differences between individual patterns in each cluster. As *k*-means has random aspects, we repeated the clustering 100 times, selecting the solution with the highest likelihood.

Based on the number of weeks in a course, we used the k-means algorithm to extract clusters directly from the profiles, as a 3-, 6-, 7- or 8-dimensional vector for each learner, in order to allow for the possibility of clustering by time of activity.

A feature of k-means clustering is that it will always generate k clusters. We trialed the use of values for k ranging from 3 to 8, and compared the silhouette widths, in order to identify the best-fit value. This provided a graphical representation of how well each object lay within its cluster. We also used a scree plot to visualize the total within-groups sum of squares – an indication of how closely the clusters group together, and used 'kinks' in the slope to suggest the best choice of k.

4.2 The Three Phases of the Study

In **Phase One** of the study, we looked in detail at the clusters for which $k = 7$ provided the best fit. This value for k had also proved to be the best fit in the OU study. The Phase One dataset therefore included LongMOOC1 and LongMOOC2 and used the cluster descriptions that had been developed during the OU study. **Phase Two** used the sets of data for TalkMOOC3, ShortMOOC4 and ShortMOOC5 and explored why a value of 7 for k was not a good fit in these cases. **Phase Three** of the analysis used the most suitable value for k for TalkMOOC3, ShortMOOC4 and ShortMOOC5. Table 1 shows the values of k used for each phase of the study.

Table 1. Values for k used in the phases of this study.

MOOC	Phase 1: best-fit value for k aligned with OU study	Phase 2: testing $k = 7$ where this was not the best fit	Phase 3: Most suitable value for k in each set of data
LongMOOC1	7	-	-
LongMOOC2	7	-	-
TalkMOOC3	-	7	3
ShortMOOC4	-	7	4
ShortMOOC5	-	7	5

5 Analysis of Clusters: Phase One

LongMOOC1 and LongMOOC2 were both courses in the hard sciences, running for 7 or 8 weeks on a platform that linked discussion with comment throughout. In both cases, as on the OU study, seven clusters was the best fit for the data.

This section therefore describes the seven clusters outlined in the OU study, noting where LongMOOC2 differs from the data considered in that study. Where no differences are noted, the description holds true for the four MOOCs in the OU study (which included LongMOOC1) and for LongMOOC2. Typical engagement profiles are based on the average weekly activity scores in LongMOOC1.

Cluster I: Samplers is the largest cluster of learners (37 %–56 %). Its members, on average, visit about eight steps, with around one in ten visiting just one step. It includes a large number of latecomers – around a fifth of the people in this cluster were not active during Week 1. A few members of the cluster complete the Week 1 assessment, but do not go on to do so in subsequent weeks. No more than 18 % of cluster members post a comment and typically less than 15 %. A typical engagement profile for this cluster is: [1, 0, 0, 0, 0, 0, 0, 0].

Cluster II: Strong Starters is a smaller cluster, accounting for 8 %–14 % of the entire cohort. All members of this cluster submit the assessment in Week 1 but there is then a sharp drop-off in activity and members of the cluster typically do not complete another assessment. On average, over a third post a comment, with cluster members posting fewer than four comments each. On LongMOOC2, only 17 % of the cluster posted a comment and, overall, learners in all clusters on this course posted less frequently than on the OU MOOCs. This may be because the OU courses had no requirements for entry, other than an interest in the subject, while LongMOOC2 required 'a basic level of mathematical skills, at the level of a final-year school pupil'. This could have meant that it was more difficult to contribute to the discussion on this MOOC. A typical engagement profile for this cluster is: [9, 1, 0, 0, 0, 0, 0, 0].

Cluster III: Returners, 6 %–8 % of the cohort, all complete the assessment in Week 2, although up to 10 % may have missed it out in Week 1. On average, they visit fewer than half the steps on the course – for example, the mean number visited by this group on LongMOOC2 was 42 of 110 steps. After submitting their Week 2 assessment, members of this cluster usually leave. Overall, they post less than one comment each. A typical engagement profile for this cluster is: [9, 9, 0, 0, 0, 0, 0, 0].

Cluster IV: Midway Dropouts is typically a small cluster, making up just 4 %–6 % of the overall cohort. Learners in this cluster visit about half the course steps and submit assessment in three or four weeks then leave the course halfway through its run. In the OU study, just under half this cluster posted a comment. In LongMOOC2, this cohort posted less than on the OU study and only 29 % of this cluster contributed a comment. A typical engagement profile for this cluster is: [9, 9, 9, 4, 1, 1, 0, 0].

Cluster V: Nearly There is a variable cluster in percentage terms (5 %–19 %), and accounts for just 3 % of the LongMOOC2 cohort. These learners engage for longer than those in the clusters described above, submitting work in half or more of the course weeks and visiting around 80 % of the steps. They leave without completing the entire course and few, if any, submit work in the final week. A typical engagement profile for this cluster is: [11, 11, 9, 11, 9, 9, 8, 0].

Cluster VI: Late Completers Both this cluster and the following cluster are similar to the *Completing* cluster identified in the *Coursera* study. *Late Completers* make up 6 %–8 % of the cohort. They view most steps on the course. Each week, more than 94 % of the cluster completes the assessment. They submit the final assessment and most others, but either do so late or miss at least one of them. Less than half of them post a comment. A typical engagement profile for this cluster is: [5, 5, 5, 5, 5, 9, 9, 9].

Cluster VII: Keen Completers, 7 %–23 % of the cohort, can be regarded as model students. They visit more than 90 % of the course steps, submit all work on time and engage throughout. Typically, at least two-thirds of them comment at least once, posting about twice a week. Even on LongMOOC2, where engagement with comments was lower than on the courses included in the OU study, 58 % of this cluster posted a comment. A typical engagement profile for this cluster is: [11, 11, 9, 9, 11, 11, 9, 9].

6 Analysis of Clusters: Phase Two

In the case of TalkMOOC3, ShortMOOC4 and ShortMOOC5, seven clusters did not prove to be the best fit for the data, and so these MOOCS did not form part of the Phase One dataset. However, the *k*-means approach enabled us to segment the data to investigate why seven was not an appropriate value for *k* in these cases.

6.1 TalkMOOC3: Seven Clusters

In the case of TalkMOOC3, $k = 7$ did not generate clusters that were clearly equivalent to those found on other MOOCs. This was because TalkMOOC3 differed from others in that did not include any assessment. This affected the activity scores that could be obtained in any one week. As the description of possible weekly activity scores outlined above shows, weekly activity scores on most MOOCs considered here could range from 0–11. However, on TalkMOOC3 the range was between 0 (no activity) and 3 (viewed content and posted comment). This limited range effectively increased the importance of comments at the expense of content. Splitting this dataset into seven therefore produced seven clusters that were heavily dependent on small variations in posting behaviour and, of course, not at all influenced by assessment.

6.2 ShortMOOC4 and ShortMOOC5: Seven Clusters

ShortMOOC4 and ShortMOOC5 included assessment, but differed from other MOOCs studied in that they were much shorter: both ran for just three weeks.

When these two courses were segmented into seven, the resulting clusters included five that were the same as those found in Phase One. These clusters met the criteria described above for *Samplers*, *Strong Starters*, *Nearly There*, *Late Completers* and *Keen Completers*.

Three clusters were not found in these two MOOCs. *Returners* (learners who drop out in week 2 of an eight-week course) and *Midway Dropouts* (learners who drop out halfway through the course) were not represented, presumably because, on a three-week course, these clusters would be very similar to *Nearly There* (learners who drop out in the penultimate week). These two MOOCs also included no cluster of *Late Completers*, presumably because there were only three opportunities for late submission, rather than eight.

Other clusters were found on these MOOCs. Two, the *Saggers* and the *Improvers*, were found on both. *Surgers* and *Weak Starters* were found on only one MOOC.

Cluster a: Surgers concentrate their effort after the first week of a three-week course. On average, they visit more than two-thirds of the steps, engaging to some extent with each week. They do little in Week 1 other than submit their assessment late, engage more in Week 2, but still submit their assessment late (working on it in Week 3), and engage but do not submit in Week 3. On average, they post one or two comments. A typical engagement profile for this cluster is: [4, 6, 2].

Cluster b: Improvers fall behind in Week 1, submitting their first assessment late. They engage more in Week 2 and by Week 3 they are on schedule and submit their assessment on time. They view the majority of steps on the course and typically post more than one comment. A typical engagement profile for this cluster is: [5, 6, 9].

Cluster c: Saggers was only found in ShortMOOC4. The members of this cluster start strongly. They submit on time in the first week, they visit most steps on the course and they post an average of just over one comment a week. They engage less than Keen Completers, submitting late in Week 2, then engaging more in Week 3 and submitting their final assessment on time. Their engagement profile is: [10, 5, 8].

Cluster d: Weak Starters only appears in ShortMOOC5, engaging for longer than *Samplers* and less actively than *Strong Starters*. They engage in Week 1 and some submit the assessment. In Week 2 there is a low level of engagement, and few return in Week 3. On average, members of this cluster post one comment. A typical engagement profile for this cluster is: [4, 1, 0].

7 Analysis of Clusters: Phase Three

In the third and final phase of the study, we looked at the best-fit number of clusters for TalkMOOC3 ($k = 3$), for ShortMOOC4 ($k = 4$) and for ShortMOOC5 ($k = 5$).

7.1 TalkMOOC3: Three Clusters

As noted above, the lack of assessment on TalkMOOC3 meant that weekly activity scores ranged only from 0–3.

In this case, the method of assigning activity scores described above meant that higher scores were given to those who commented (2) and those who commented and viewed content (3). In the other MOOCs examined here, commenting had a less significant effect because it could only contribute 1 to a possible activity score of 11.

As discussions were linked with content, an activity score of 2 was rare because it was only possible for learners who did not engage with the current week’s activities, but did return to a previous week to comment. The effective options for a weekly activity score were therefore 0, 1 or 3.

Cluster 1/3: Quiet is the largest cluster, containing around two-thirds of the cohort. On average, members of this cluster visit a quarter of the course steps. None of them contribute a comment during the first week of the course, and only 7 % post at any point. This cluster contains many late arrivals. Over a third arrive after Week 1, and over 9 %

only engage with the second half of the course. A typical engagement profile for this cluster is [1, 0, 0, 0, 0, 0].

Cluster 2/3: Week 1 Contributors Although number and frequency of comments are among the defining elements of many of the MOOC clusters identified in this study, it is only TalkMOOC3 that contains clusters in which every member comments at least once. In this cluster, which makes up 19 % of the cohort, every member posts a comment during the first week of the course, although half of them do not comment again. On average, this cluster visits 38 % of the course steps. A typical engagement profile for this cluster is [3, 1, 1, 0, 0, 0].

Cluster 3/3: Consistent Engagers makes up 11 % of the cohort. They have a mean activity score at least slightly above 1 in each week of the course, indicating that they engage throughout. On average, they visit 82 % of the course steps. All of them contribute at least one comment during the course, 95 % of them contribute more than three comments and 7 % contribute more than 100 comments each. A typical engagement profile for this cluster is [3, 3, 3, 3, 1, 1].

7.2 ShortMOOC4: Four Clusters

Cluster 1/4: Very Weak Starters is the largest group on this MOOC, accounting for 35 % of the cohort. Its profile is closest to that of the *Weak Starters* who were found in ShortMOOC5 when it was split into seven clusters. Overall, the *Very Weak Starters* show low levels of engagement, visiting 20 % of steps on average, commenting less than once per learner, and no individual achieves an activity score higher than 5 in Week 1 of the course. This indicates that none of these learners finishes their assessment on time in Week 1, although 18 % of the group completes this assessment late. A fifth of the cluster do not engage at all during the first week of the course. The engagement profile for this cluster is: [2, 1, 0].

Cluster 2/1: Strong Starters (Truncated) can be regarded as a version of the *Strong Starters* cluster described above. This is one of the smaller clusters, accounting in this case for 17 % of the entire cohort. Members of this cluster submit the assessment in Week 1, although 6 % submit late. There is then a drop-off in activity and members of the cluster typically do not complete another assessment. In the case of ShortMOOC4, almost half the group posts a comment and, overall, cluster members post just over one comment each. The engagement profile for this cluster is very similar to the beginning of the *Strong Starters*' profile: [10, 1, 0].

Cluster 3/4: Returners (Truncated) can be regarded as a version of the *Returners* cluster described above. These learners all complete the assessment in Week 2, although a few (17) miss it out in Week 1. On average, this cluster visits 81 % of the steps. This is a higher percentage than that visited by the *Returners* described above but, because this was a short course, it only represents 36 steps. After submitting their Week 2 assessment, members of this cluster engage in very little activity. On an eight-week course, their activity after Week 2 is spread thinly across the final six weeks. On this three-week course, all additional activity is concentrated in Week 3, which gives them a slightly higher average activity score in Week 3, as the engagement profile shows. The amount of people commenting was high across this MOOC (47 %) and was also high in this cluster, with half the group posting at least one comment. The engagement profile for this cluster is very similar to the beginning of the *Returners*' profile: [9, 9, 2].

Cluster 4/4: Keen Completers (Truncated) can be regarded as a version of the *Keen Completers* cluster described above. The description of the group is the same in both cases but the engagement profile shows that, overall, members of this group were less keen, because their average activity score in any week was never higher than 9. A typical engagement profile for this cluster is: [9, 9, 9].

7.3 ShortMOOC5: Five Clusters

Cluster I/5: Samplers (Truncated) can be regarded as a version of the *Keen Completers* cluster described above. The description of the group is the same in both cases. Once again, a large number of people do not participate during Week 1, with more than 25 % not registering any activity during that time. The typical engagement profile for this cluster is [1, 0, 0], but it should be noted that many learners did not participate in Week 1 although they visited in Weeks 2 or 3.

Cluster 2/5: Strong Starters (Truncated) is another instance of the *Strong Starters (Truncated)* cluster described in the analysis of ShortMOOC4. A small cluster, in which everyone submits the Week 1 assessment, but there is then a sharp decline in engagement. Engagement profile: [9, 1, 0].

Cluster 3/5: Returners (Truncated) is another instance of the *Returners (Truncated)* cluster described in the analysis of ShortMOOC4. These learners all complete the assessment in Week 2, although a few (20) miss it out in Week 1. In this case, average activity levels rise slightly in Week 2. This is not clear from the typical engagement profile of 8, 8, 2, but can be seen when the average engagement score for each week is shown to one decimal place: 7.8, 8.1, 1.6.

Cluster 4/5: Improvers is another instance of the *Improvers* cluster that was found when the three-week MOOCs were segmented into seven clusters. Activity in this cluster rises each week, and the final assessment is typically submitted on time. The engagement profile remains: [5, 6, 9].

Cluster 5/5: Keen Completers (Truncated) is another instance of the cluster found in ShortMOOC4. Engagement remains steady throughout the three weeks, with an average activity score of 9 or 10.

8 Discussion

This study investigated whether the engagement patterns identified in the OU study [4] were only found on MOOCs developed by The Open University, or whether they were also apparent in MOOCs from other universities. It also investigated whether engagement patterns are influenced by changes in learning design – in this case assessment and course length. This is an important issue, because the original, and highly cited, study of Coursera MOOCs suggested that MOOC designers would be able to apply the four patterns they had identified 'to target interventions and develop adaptive course features' [7]. The OU study [4] showed that MOOC designers should not do this without taking variations in pedagogy into account. The current study shows that they should also take account of design variations.

The seven clusters identified in the OU study: *Samplers*, *Strong Starter*s, *Returners*, *Midway Dropouts*, *Nearly There*, *Late Completers* and *Keen Completers* were found here on MOOCs that employed similar assessment patterns to those from the OU and that also ran for seven or eight week. The same set of clusters was not found on three-week MOOCs, or on a MOOC that did not include assessment.

In the case of TalkMOOC3, the omission of assessment from the course can be seen not only as a shift in learning design, but also as a change in the underlying pedagogy. The original Coursera study clustered learners on the basis of their engagement with two key pedagogic elements: content and assessment. The OU study found that a different set of engagement patterns was seen when the key pedagogic elements were discussion, content and assessment. In the case of ShortMOOC4, the key pedagogic elements were discussion and content, and so a new set of clusters emerged.

These variations related to pedagogy highlight that the results of a cluster analysis are dependent on the variables that are selected as significant by researchers. A *k*-means analysis will produce *k* clusters for any value of *k*, but these will only be meaningful if priority is given to elements of the data that are significant in the context.

In the case of ShortMOOC4, where engagement with discussion was a key element of the course, the cluster descriptions set out here can be only partial. This is because the activity data included information about how many comments individuals posted, and when they posted them. However, there were no data available about ways in which people engaged with the discussion by reading or liking comments. Adding these elements in future studies would make it easier to identify and characterise engagement patterns in this type of course.

ShortMOOC4 and ShortMOOC5 both employed similar learning designs in that both ran for just three weeks. In the case of ShortMOOC5, this aspect of the learning design was, in part, intended to encourage completion by making the *Disengagement* identified in the Coursera study more difficult to achieve.

To some extent, this was the case. ShortMOOC4 included learners who displayed truncated versions of the engagement patterns of *Strong Starters*, *Returners* and *Keen Completers*. ShortMOOC5 included truncated forms of the engagement patterns of *Samplers*, *Strong Starters*, *Returners* and *Keen Completers*.

However, in both cases the *Late Completers* cluster was missing. In addition, there was a subtle shift in the engagement pattern of the *Keen Completer*s. The overall activity levels of learners in this cluster were lower each week than they were on eight-week courses. On average, *Keen Completer*s on the eight-week MOOCs engaged fully in most weeks, posting comments, visiting content and completing assessment on time. In the three-week MOOCs, the average *Keen Completer (Truncated)* did not do all those things in any week.

The clusters also suggest that participants in three-week MOOC tended to concentrate their activity in one or two weeks, and these did not necessarily include the first week. *Strong Starters* and *Weak Starters* concentrated on Week 1, but a notable feature of the *Sampler* cluster was the high percentage (on all the MOOCs examined here and in the OU study) that did not engage in the first week.

ShortMOOC5 included an *Improvers* cluster, which was seen for $k = 5$ and $k = 7$. The *Improvers* increased their activity over time. In Weeks 1 and 2 their average activity scores indicate that most of them submitted their assessment late, while they completed their Week 3 assessment on time. This suggests that they completed all the course assessment in Weeks 2 and/or 3.

When ShortMOOC4 and ShortMOOC5 were segmented into seven, both contained a *Surgers* cluster. In this case, cluster members typically did not engage during Week 1, submitted their Week 1 assessment late, engaged more in Week 2 but again submitted their assessment late, then did not submit their Week 3 assessment. Activity here is concentrated at the end of the course.

The study also included a cluster of *Saggers*, found on ShortMOOC4 when it was segmented into seven, who engaged more in Week3 and Week 1 than in Week 2.

It seems clear that many learners do not approach a three-week course in the same way as an eight-week course. There are limited opportunities to get far ahead of, or behind, the cohort, and so it is possible to dip in at different points without losing the sense of working on the course as part of a cohort.

Although these courses remained open to learners after their end-dates, and most FutureLearn assessment does not have to be submitted by a set date, none of the typical engagement profiles ends with a Fig. 4, which would have indicated that a significant engagement pattern involved completing the course after its end date. AQ2

9 Conclusion

This study has extended the findings of the OU study using new sets of data. The recommendations of that study can therefore be restated with more confidence – with the added proviso that they only apply in the case of seven- to eight-week MOOCs that support engagement with content, assessment and discussion.

The OU study suggested strategies for intervention and improvement:

- providing previews of course material
- setting up discussion steps for latecomers
- encouraging late arrivals to register for another course or for a later presentation
- providing bridges between course weeks, stressing links between those weeks [4].

In the case of course length, there are many good pedagogic reasons for creating short courses. However, educators should not assume that shortening the length of a course will necessarily increase learner engagement and course completion. This study indicates that may not be the case because the patterns of engagement shift.

In the case of pedagogy, it is clear from the Coursera study, the OU study and this study that changes to the basic pedagogic elements of a course are associated with shifts in patterns of engagement. The analysis of TalkMOOC3 gives some indication of the patterns of engagement on a course focused on content and discussion, but a lack of data limited that analysis and thus its implications are unclear.

As shifts in pedagogic approach can change the elements of a course that can be regarded as key, it is important that any future analysis of patterns of engagement is based on the elements that are most important to the pedagogy of the course under examination. Where the focus is heavily on discussion elements, it is necessary to have access to data about the reading and liking of contributions, as well as when and by whom they were written.

It is clear that changes to some elements of learning design can change learners' patterns of engagement with a MOOC. Future work should therefore investigate other elements of learning design in order to identify which of these are associated with desirable patterns of engagement.

Our understanding of patterns of engagement in MOOCs is developing rapidly. It initially seemed clear that these patterns held steady across MOOCs; we now know that those patterns are influenced by both learning design and pedagogy. Cutting a MOOC short in order to promote engagement originally seemed to a promising plan, but this study shows that such an approach is not a panacea. These insights can help us to produce learning designs that will support desirable patterns of engagement.

Acknowledgments. We thank FutureLearn, The Open University, the University of Birmingham, the University of Edinburgh and the University of Leeds for allowing us access to the data used in this study. We also thank the FutureLearn Academic Network for establishing connections between researchers with an interest in this area.

References

1. Clow, D.: MOOCs and the funnel of participation. In: Learning Analytics and Knowledge 2013 (LAK 13), pp. 185–189. ACM, New York (2013)
2. Downes, S.: Connectivism and connective knowledge (2012). http://www.downes.ca/files/books/Connective_Knowledge-19May2012.pdf
3. Downes, S.: Like reading a newspaper (2014). http://halfanhour.blogspot.co.uk/2014/03/like-reading-newspaper.html
4. Ferguson, R., Clow, D.: Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs). In: Learning Analytics and Knowledge 2015 (LAK 15), pp. 51–58. ACM, New York (2015)
5. Jordan, K.: Initial trends in enrolment and completion of massive open online courses. IRRODL **15**(1), 133–160 (2014)
6. Jordan, K.: MOOC completion rates: the data. http://www.katyjordan.com/MOOCproject.html
7. Kizilcec, R., Piech, C., Schneider, E.: Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In: Learning Analytics and Knowledge 2013 (LAK 13), pp. 170–179. ACM, New York (2013)
8. Kop, R., Fournier, H., Sui Fai, J.M.: A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. IRRODL **12**(7), 74–93 (2011)

9. Laurillard, D.: Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies. RoutledgeFalmer, New York (2002)
10. Littlejohn, A.: Understanding Massive Open Online Courses. CEMCA, New Delhi (2013)
11. Pask, G.: Conversation Theory: Applications in Education and Epistemology. Elsevier, New York (1976)
12. Ferguson, R., Sharples, M.: Innovative pedagogy at massive scale: teaching and learning in MOOCs. In: de Freitas, S., Rensing, C., Ley, T., Muñoz-Merino, P.J. (eds.) EC-TEL 2014. LNCS, vol. 8719, pp. 98–111. Springer, Heidelberg (2014)
13. Siemens, G.: Connectivism: a learning theory for the digital age (2004). http://www.elearnspace.org/Articles/connectivism.htm
14. SoLAR: Open learning analytics: an integrated & modularized platform (2011). White Paper, http://solaresearch.org/OpenLearningAnalytics.pdf

Author Query Form

Book ID :	**370743_1_En**
Chapter No.:	**6**

Springer
the language of science

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author
During the process of typesetting your chapter, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query Refs.	Details Required	Author's Response
AQ1	Per Springer style, both city and country names must be present in the affiliations. Accordingly, we have inserted the city name "Edinburgh" in affiliation "4". Please check and confirm if the inserted city name "Edinburgh" is correct. If not, please provide us with the correct city name.	
AQ2	Figure 4 is cited in the text, but the source is missing. Kindly provide the source or delete the citation.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	,
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	