DROPOUT RATES OF MASSIVE OPEN ONLINE COURSES: BEHAVIOURAL PATTERNS

D.F.O.Onah¹, J.Sinclair¹, R.Boyatt¹
¹The University of Warwick (UNITED KINGDOM)

Abstract

Massive open online courses (MOOCs) have received wide publicity and many institutions have invested considerable effort in developing, promoting and delivering such courses. However, there are still many unresolved questions relating to MOOCs and their effectiveness. One of the major recurring issues raised in both academic literature and the popular press is the consistently high dropout rate of MOOC learners. Although many thousands of participants enrol on these courses, the completion rate for most courses is below 13%. This paper investigates MOOC attrition from several different perspectives. Firstly, we review existing literature relating to MOOC dropout rates, bringing together existing findings on completion rates and analyses of several specific courses, which identify factors that correlate to likelihood of dropout. We provide a meta-analysis of the basic figures on overall dropout rates previously collected to identify relationships between course factors and dropout rates. In addition, the literature is reviewed from a qualitative perspective drawing together findings on the reasons for dropout and methods suggested for resolving or reducing the dropout rate. Secondly, using themes emerging from the initial investigation, we provide a preliminary analysis of data gathered from a Computing MOOC run by the University of Warwick, UK and presented using a Moodle platform. Different aspects of students’ demographic data are examined to see if relationships to persistence exist. An important feature of this course is that it has been run in two different parallel modes (“traditional” MOOC mode with peer support, and “supported” mode with real time, tutored programming labs). This allows direct comparison between the dropout figures for the two different modes. Qualitative information from student evaluations is also considered. Finally, we discuss our findings relating MOOC dropout rates, considering what factors are within the control of a MOOC provider and suggesting the most promising avenues for improvement. Our results indicate that many participants who may be classed as dropouts (for example, because they do not complete the necessary components to gain a certificate) are still participating in the course in their own preferred way (either at a slower pace or with selective engagement). This suggests that the structure of “a course” may not be helpful to all participants and supporting different patterns of engagement and presentation of material may be beneficial.

Keywords: MOOC, massive open online course, dropout rate, reasons for dropout, participants’ engagement, motivation, behavioural patterns, Moodle.

1 INTRODUCTION

Massive open online courses (MOOCs) are a relatively new model for the delivery of online learning to students. As “massive” and “open”, these courses are intended to be accessible to many more learners than would be possible through conventional teaching. They are often free of charge and participation need not be limited by the geographical location of the learners [1, 2]. MOOCs have received a good deal of publicity and media attention in the past two years, with many highly respected Higher Education institutions joining the race to develop and deliver courses in a wide range of subjects [3]. Claims have been made that MOOCs will be able to solve many educational problems by providing free to access, cutting edge courses, thus reducing the cost of university level education and enabling learners in developing countries [3]. There has also been speculation on their potential as a “disruptive” educational platform, which will force re-thinking and re-structuring of existing educational models [4].
Despite the great enthusiasm for and rapid growth of MOOC courses and platforms, there has also been rising concern over a number of MOOC aspects. One feature in particular that is difficult to ignore is that these massive courses also have massive dropout rates [5,8]. As noted by Kolowick: “massive open online courses have gained renown among academics for their impressive enrolment figures and, conversely, their unimpressive completion rates” [7]. Few MOOCs have a percentage completion, which reaches double figures [6]. Of the millions of learners who have already participated in MOOCs, the vast majority do not get to the stage of obtaining a certificate of completion. This is seen by some as “an indictment of the MOOC format” [7] and Devlin has noted commentators in this camp “hinting that therein lies a seed of the MOOC’s eventual demise” [8]. However, the counter-argument holds that completion rate statistics should not be viewed in this way. If even a small percentage of a very large group completes, the actual number of successful students will still be far greater than would otherwise have been possible. A number of authors point out that the bare figures do not provide a realistic view of the situation and suggest the need for a new metric. It is claimed that courses on the Coursera platform have a 45% completion rate if only those students who reach at least the first assessment are included [8]. This rises to 70% for students on the “Signature Track” scheme (for which a $50 payment is required). Devlin also argues that a fairer comparison would be with the numbers who apply for entry to traditional university courses [8].

With no agreed MOOC metric, compiling and comparing statistics can be problematic. In her compilation of MOOC dropout rates, Jordan notes 13 separate criteria being used in published results on MOOC completion [6]. The most commonly used measure is obtaining a course certificate.

Although it may be unfair to dismiss MOOCs on the basis of over-simplistic completion statistics, it would also be a mistake to accept the current situation as satisfactory. It is important to understand the factors leading to attrition in order to identify those which are preventable or can be reduced. This paper reviews issues relating to MOOC dropout, considering published data on MOOC completion and discussing factors implicated in previous studies as being related to attrition. In particular, we focus on three MOOCs for which published evaluations exist. Many of the factors suggested in the literature remain untested since few studies have yet been done to discover underlying causes or to evaluate the effect of possible interventions. We report initial results from the “Computing for Teachers” (CfT) MOOC developed by the Department of Computer Science at the University of Warwick. The MOOC was conducted in two simultaneous modes: a “traditional” mode (with support from peers and via forums) and a “supported” mode (with group and individual support from experienced tutors. This allows direct comparison of engagement and attainment data for the two groups of learners.

2 LITERATURE SURVEY

Although published research and analysis relating to MOOCs was noticeably absent at the time of their initial expansion, there is now a steadily growing body of relevant literature [9]. This section focuses on work relating to dropout rates and looks at three aspects. Firstly we set the scene by considering three separate case studies. Secondly, the general picture of known completion rates is presented and some observations from a brief meta-analysis are added. The main part of the survey considers issues associated with MOOC attrition, including suggested causative factors and indicators linked to prediction of dropout.

2.1 MOOC Dropout and Completion: Existing Evaluations

Large amounts of data are collected by the major MOOC platforms, but access to this is not generally available. Several evaluations have been published by specific institutions providing data and analysis on courses they have delivered. These provide a valuable source of information on a variety of aspects relating to learner background, engagement and attainment. We focus on data relating to participation and dropout rates.

The University of Edinburgh launched six MOOCs on the Coursera platform in January 2013 [10,11,12]. The information in this paragraph is obtained from a published report on these MOOCs [11]. The six “short, fully-online” courses ran for 5-7 weeks and attracted a total initial enrolment of 309,628 learners. Six different course structures were developed and, in addition to the usual features of the Coursera platform, new methods of content delivery and collaborative working were introduced. Evaluation of the Edinburgh MOOCs revealed that 123,816 of those enrolled (about 40%) accessed the course sites during the first week (‘active learners’), of whom 90,120 (about 29%) engaged with the course content. Over the duration of the course, the number of active participants rose to 165,158
(53%). As a gauge of persistence, 36,266 learners (nearly 12%) engaged with week 5 assessments. This represented 29% of initial active learners, although interestingly there was a large variation across the six courses ranging from 7% to 59%. Obtaining a statement of accomplishment required attainment of a certain percentage in the assessment (the specific level varying between courses) and this was achieved by 34,850 people (roughly 11% of those who enrolled). The report [11] provides more demographic data and analysis, but engagement and dropout rates are not investigated further with respect to these.

A further case study is available from Duke University which ran a Bioelectricity MOOC in 2012 [13]. In this evaluation, figures are presented in a different way to the previous study, so direct comparison is hampered. However, 12,175 registrations were made of which 7,761 students watched at least one video. This figure, representing around 64% of enrolments, might be compared to the Edinburgh figure of 53% for those who were active at any point during the duration of the course. Statistics on resource access (such as video viewings) give one measure of participation, but as students may access each resource many times, it does not show how many participants are still active at any stage. Quiz submission is perhaps a more useful metric and in the Duke MOOC, 3,200 students (26% of enrolments) attempted at least one quiz in the first week. This might be compared to the 29% of Edinburgh MOOC students who engaged in week 1. The statement of accomplishment for this course was again based on reaching a certain level of achievement in the quizzes, and 313 participants (2.6%) attained this level. This is on the low side even for MOOC completion and learner feedback suggested three specific reasons for failure to complete [13]. These are addressed further below.

A third useful evaluation is available for the UK Open University’s Open Learning Design Studio (OLDS) MOOC [14]. This was a smaller course, with 2420 registrations. Nearly half of these (1197) accessed at least one key course page in the first week. The report provides a rich analysis of user perspectives, participation and interaction. The course itself was experimental and designed to promote social learning rather than simply present course materials. Participants were asked to suggest criteria of success and to set their own learning goals. In this type of course it is very difficult to provide a simplistic “completion” measure. The report [14] refers to “approximately 30 active contributors and at least 30-60 other participants”. Only 22 learners completed the post course survey but of these, only half felt they had achieved their learning objectives.

These three published case studies provide interesting information on a variety of aspects including engagement and dropout. However, the different measures which are gathered, the varying ways in which statistics are presented and the different perspectives on “participation” and “success” within the courses themselves make it difficult to provide a direct comparison between them. Of course, the general trajectory is clear: many enrol; fewer start out; a small minority complete.

### 2.2 MOOC Completion Data

Jordan [6] provides a compilation of available data on MOOC completion. This is an on-going initiative which provides a useful resource for basic comparisons. Currently (May 2014) 169 courses are represented, and completion rates may be viewed according to factors such as platform, institution and length. The graphical representation of this data illustrates a number of relationships: shorter courses have higher completion rates; small courses (with up to 200 enrolments) are much more likely to have a completion rate of over 20% than larger courses; MOOCs rely on peer grading only have often had very low completion rates.

Courses from 13 different platforms are currently represented in [6], with only three of these contributing more than 10 courses. Further analysis of the data shows that, of the 61 courses hosted by Coursera, the average completion rate was just over 6%. The Open2Study courses, of which there are 64, are all very short (4 week) and are automatically graded. The average completion rate for these is just under 30%. The EdX courses included (19 in total) were generally longer in duration, with only one being less than 10 weeks, but all were automatically graded. These had average completion of around 8%.

Another interesting comparison can be made between two different presentations of the same course using different platforms [6, 15]. Circuits and Electronics 6.002x was offered by MITx in March 2012 and by edX in September 2012. The first run had 154,763 registered participants of whom 7,157 completed the course (4.6%). The later edX delivery had 46,000 enrolments and 3,008 completions (6.5%). The dropout rate observed on the course is therefore broadly similar across the two platforms.

Although the collected data builds a useful background picture of MOOC completion, it does not evaluate or even suggest the underlying factors and features which may contribute to learners’
decisions to continue in a course. The following section examines possible contributing factors identified in the literature.

2.3 Reasons for Dropout

Although a number of reasons for student dropout have been suggested, there has been little research to assess how far these influence MOOC learners in practice or to identify which are within the sphere of influence of MOOC developers.

No real intention to complete

A number of authors have noted that reasons for participation given by users often include motivations such as “out of curiosity” and “to learn more about MOOCs” rather than to learn the subject itself [7,10,11]. It is therefore suggested that many enrolments are from people who do not intend to participate fully, including professionals who want to gain understanding of the format in order to produce their own courses [14]. Casual, recreational learners may not wish to invest effort into attempting assessments which are generally used as evidence of completion [16]. Lack of prerequisites and open entry encourage casual enrolment. Grover et al [16] view this broad range of background, intention and participation as “a by-product of the open access nature of the courses and the novelty of the medium”. If users do not really intend to complete, it is argued that they should not be included in statistics which may then be used as an indictment of the course [17]. A better measure might well be whether those who register achieve their own learning outcomes, but, as evidenced by the evaluation of the OLDS MOOC [14] is very difficult to capture and assess.

Lack of time

Students who fully intend to complete the course may fail to do so because they are unable to devote the necessary time to study [4,13]. This has been noted even in courses where participants have a high level of motivation to complete [9]. Personal circumstances may be to blame, but in some cases the workload of the course may be too high. Diversity of learner background means that the current “one size fits all” MOOC format does little to adapt to individual needs. Learning materials which are appropriate for some may take others much more (or less) time to master.

Course difficulty and lack of support

Related to the previous point is the level of difficulty of a course and the lack of necessary background. Insufficient mathematical skills are noted in relation to the Duke Bioelectricity course [13]. As one respondent in Mackness et al’s survey said: “The reason I stopped is because I cannot understand the issues being discussed any more” [18]. Student blogs often refer to the inadequacy of peer support and lack of instructors when topics become difficult.

Lack of digital skills or learning skills

Online learning generally requires a high degree of autonomy and depends on users being able to work with the technologies and formats used. Even those who are familiar with using a range of everyday technologies may be uncomfortable when new systems must be quickly mastered. Conole [4] points to learners’ confusion and frustration as a reason for high dropout rates. Evaluation of the Duke Biochemistry MOOC [13] notes that students were unable to make the transition from theoretical learning to practical application required for the assessments.

Bad experiences

Some MOOC participants have pointed to a variety of bad experiences as being a barrier to continued participation. These include: inappropriate behaviour of peers in forums; lack of focus and coordination in forums; depletion of study groups due to attrition; poor quality and incorrect learning materials; technical problems in the MOOC platform [9,18,19,20,21].

Expectations

Students may enrol with little understanding of what the course requires and may have unrealistic expectations either of the course or of their own ability to undertake it.

Starting late

Late starters on a course may find it very difficult to catch up and outcomes are likely to much lower for this group of students [19]. It is not simply a matter of catching up with learning materials. Support groups and learning networks will already have formed and newcomers may struggle to fit in to the existing structure. Students who join after community discussion is already well developed are often
unable to orientate themselves in the forums [22].

**Peer review**

Some authors have noted that courses relying on peer grading often have much lower completion rates than others [5,6]. Peer grading may well require more work on the students’ part. It has also been suggested that some students are unhappy with the concept of peer review and that training is often lacking [16,23,24]. Other participants have been disheartened by bad practice discovered through peer review, for example, by unhelpful or dismissive comments on their work, lack of response or discovery of plagiarism in peers’ work.

### 2.4 Predicting Dropout

Observation of users’ behaviour and participation has been used to identify a number of indicators which can provide predictors of future dropout. This can be a useful tool and allows the possibility of timely intervention. Little work exists on the mechanics and nature of interventions on a personal level that could be both effective and realistically implementable within a MOOC format, particularly models where tutor involvement is impossible. However, feedback to the tutor on general progress can allow real-time adjustment of difficulty or additional materials to be provided [20,21,23].

Activity in discussion forums is often used as a measure of MOOC users’ participation [20]. The discussion within forums develops into multiple and complex conversations. Analysis can provide information not just on initiation of and response to themes within the discussion but also gives insight into social networks and the trajectories of participation associated with different groups [19,25].

Most MOOCs gather large amounts of data about users and their activity within the course [20,23]. Data analytics can discover patterns within the data which often leads to surprisingly accurate prediction of future success. Some studies have identified patterns of activity which deviate from that expected by the course developers and which might be associated with positive engagement [14]. For example deviating from a task or choosing to pursue a different activity to that suggested may indicate that the participant is directing their own learning and making choices appropriate to their needs. However, some deviations from expectations (such as failure to interact or to submit attempt at activities at key stages) can be strong indicators of future dropout. In such cases, strategic intervention would be appropriate [21].

As discussed above, existing literature suggests a number of factors which may influence MOOC attrition rates. Published course evaluations have also included background information taken from pre-course surveys which helps to establish the demographics of those who enrol. However, there is little current research to link this to the students who drop out or to determine the factors that are most likely to influence users and for which effective counter-measures might be developed within the MOOC format. The following section presents initial results from an experimental MOOC in which the effect of tutor support on student engagement and achievement can be directly assessed and evaluated.

### 3 METHODOLOGY

The Computing For Teachers (CfT) MOOC was designed and produced by the Department of Computer Science at the University of Warwick. The purpose of the CfT MOOC was to assist school teachers in the delivery of the increased computing content being introduced to U.K. schools. The course was delivered in two different modes; “traditional” MOOC mode with peer support and automated assessment options, and a paid “supported” mode which also included on-line tutor support and tutor supported forums and resources.

#### 3.1 Basic Enrolment Figures

Over 550 teachers registered for the CfT MOOC with 30 of those registering for the tutor supported mode, with all others on the 'peer support' model. For evaluation purposes, a snapshot of data was taken on the 18th March 2014 and used to investigate participants’ engagement levels at that point and compare activity on the different modes of delivery. A basic comparison was completed to show the differences in participation rates of the two modes. Table 1 contains information on the number of registrations and the overall participation within the course. The data shows 72 participants never accessed the course after registration. The majority of participants, about 480 at time of analysis, participated in at least one element of the course.
### Table 1. Computing for Teachers MOOC Enrolment

<table>
<thead>
<tr>
<th>Participants</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrolled</td>
<td>552</td>
<td></td>
</tr>
<tr>
<td>Accessed material</td>
<td>480</td>
<td>87%</td>
</tr>
<tr>
<td>Never accessed any material</td>
<td>72</td>
<td>13%</td>
</tr>
</tbody>
</table>

#### 3.2 CfT Participant’s Basic Quiz Data

During the MOOC, quizzes were conducted which focused on the concepts of computing and programming taught in each of the sessions. Summary data from these quizzes was used to facilitate the tutor’s understanding of the students’ knowledge of the course. Figure 1 and Table 2 below show the various participation levels of the students involved in the quizzes. As the sessions progressed we observed the number of participants gradually declining. For example, 134 participants in the programming quiz declining to 19 participants (at time of analysis) in the fifth session. We have noticed the effect of a significant participants taking a very long time to complete quizzes and work through material sometimes many weeks after the initial burst of activity.

### Table 2. Record of participants in quizzes

<table>
<thead>
<tr>
<th>Sessions</th>
<th>Quizzes</th>
<th>No. Of Participants</th>
<th>Average Grade/10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session1</td>
<td>Concepts</td>
<td>134</td>
<td>7.76</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>125</td>
<td>7.84</td>
</tr>
<tr>
<td>Session2</td>
<td>Concepts</td>
<td>123</td>
<td>9.06</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>105</td>
<td>7.61</td>
</tr>
<tr>
<td>Session3</td>
<td>Concepts</td>
<td>86</td>
<td>7.79</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>47</td>
<td>8.87</td>
</tr>
<tr>
<td>Session4</td>
<td>Concepts</td>
<td>45</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>32</td>
<td>5.07</td>
</tr>
<tr>
<td>Session5</td>
<td>Concepts</td>
<td>39</td>
<td>7.59</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>19</td>
<td>8.84</td>
</tr>
</tbody>
</table>

#### Fig.1 Percentage of participation in the quizzes

Figure 2 displays an analysis of average quiz scores during the first five sessions of the MOOC, with some noticeable spikes in average grades, e.g. an average of 9.06 in ‘Computing Concepts’ for
season 2, and also some average scores indicating participants experiencing significant difficulty, e.g. ‘Python Programming’ quiz in session 4 with an average score of 5.07.

![Fig.2 Average grades of the quizzes scores](image)

### 3.3 Comparison of Traditional and Tutor Supported Groups

The traditional mode comprises registered learners interacting in peer-to-peer support group. This mode allows student to receive support from and provide support to fellow learners on the course through interactions in discussion forums. The tutor supported version allows access to additional resources and direct support from experienced teaching staff. As the tutor supported version required an upfront financial commitment we expected that all the participants on the mode of delivery would motivate them to engage regularly with the material. However to our surprise some participants who paid for support never participated (or in other cases have very low participant levels). Financial investment did not guarantee participation in the course. At time of writing, the course is on-going we expect that this may change for some participants as they engage with the material in ‘bulk’ rather than in an on-going basis. We have anecdotal evidence that teachers want to consume the material when they have a block of time available (perhaps in school holidays or on training days). Table 3 and Fig.3 contains some low participation in Python programming from the tutor support column in (session 5) with only 5 participants while the traditional mode has 14 in the same ‘Python programming’ quiz out of 19 attempted. This analysis reveals low participation rate in session 5 Python programming of 19 (3.4%) participants out of the number of enrolled participants of 552, in traditional mode 14 (2.7%) participants out of 522 and the tutor-supported mode of 5 (16.7%) out of 30.

<table>
<thead>
<tr>
<th>Sessions</th>
<th>Quizzes</th>
<th>No. Of Participants</th>
<th>Traditional Participants</th>
<th>Tutor Support Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session1</td>
<td>Concepts</td>
<td>134</td>
<td>115</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>125</td>
<td>110</td>
<td>15</td>
</tr>
<tr>
<td>Session2</td>
<td>Concepts</td>
<td>123</td>
<td>107</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>105</td>
<td>91</td>
<td>14</td>
</tr>
<tr>
<td>Session3</td>
<td>Concepts</td>
<td>86</td>
<td>76</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>47</td>
<td>38</td>
<td>9</td>
</tr>
<tr>
<td>Session4</td>
<td>Concepts</td>
<td>45</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>32</td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>Session5</td>
<td>Concepts</td>
<td>39</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>19</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>
4 DISCUSSIONS AND FURTHER WORK

Understanding the reasons behind dropout rates in MOOCs and identifying areas in which these can be improved is an important goal for MOOC development. Many widely-quoted dropout rates are calculated from baseline numbers which include registrations by people who never engage with the course or who engage in their own way but without completing assessments. Despite this, it is clear that many of those who do wish to follow and complete a course are hindered by factors such as level of difficulty, timing and lack of digital and learning skills. These problems become even more acute when MOOCs are proposed as a replacement for traditional teaching (rather than just free, spare time activities) and particularly when they are suggested as the means to teach large remedial classes.

One factor which has been suggested as influencing dropout is lack of support. This paper reports initial results from a MOOC with two modes: one “traditional” MOOC mode with minimal tutor support, the other with full, small group tutor support. Although numbers are fairly small, particularly for the supported version, the results show a higher percentage of supported students completing each quiz. The surprising feature is that most of the supported students did not make use of the real-time tutorial sessions or the tutor-monitored forum. Improvement in performance cannot therefore be put down to tutoring. It may be that this group of people were the more highly motivated from the start, leading them to enrol for the mode they thought would give them the best chance of success but then also providing them with the impetus to study. It should also be noted that although the main MOOC was free, the tutoring element incurred a cost of £100. Again, it might be argued that payment demonstrates commitment and those who have made that commitment are more likely to persevere.

Another interesting point emerging from the course is that many participants are a long way behind in the course but are still working steadily and would be happy to continue at their own pace. The MOOC participants are teachers whose busy schedule makes regular study difficult but they are keen to make bursts of progress when possible. This brings in to question the usefulness of the “Course” element of the MOOC. What might be intended as useful structure and a schedule which allows learners to collaborate may also be seen as inflexible and unhelpful for learners who cannot, for whatever reason, work to the same timetable. Many MOOCs provide a very fixed and inflexible format in terms of timing, structure and learning materials. Adaptively and choices within the structure would give users more control in reaching their learning objectives. These are areas for future work.

5 ACKNOWLEDGMENT

The author wishes to acknowledge Mr Adakole. S. Onah for all his supports financially in his research, his family members and his friends for their moral supports.
REFERENCES


[12] Coursera, Available at: https://www.coursera.org/edinburgh [Accessed: 10/01/14].


worked Learning Conference, University of Lancaster, United Kingdom, pp. 266-275.


