J Comput High Educ (2017) 29:147–159
DOI 10.1007/s12528-016-9129-6

Supporting, categorising and visualising diverse learner behaviour on MOOCs with modular design and micro-learning

Matthew Leach[1] · **Syed Munib Hadi**[1]

Published online: 16 November 2016

Abstract Much is made of the diversity for MOOCs learners—with their varied motivations and interests; yet MOOCs are often run and judged on the assumption that learners would progress through the course in its entirety, to completion. This paper presents an analysis of three recently delivered MOOCs that were designed to support a broader set of learner goals. A modular design was used, where each part included well defined learning outcomes and assessment criteria, and where completion was rewarded with digital badges. The paper proposes a new categorisation of learner achievement and methods of visualising learner behaviour that compliment this more open design. Results show that this approach recognises micro-learning that is missed if only completion rates are considered.

Keywords Modular MOOCs · Digital badges · Completion rates

Overview

This article begins with a summary of literature that is relevant to defining the success of learners on MOOCs (Massive Open Online Courses), and how a set of three MOOCs were designed to allow both greater autonomy and more accurate recognition of achievement. The research methodology and methods of data collection are through which this design is assessed are then described. A new system of categorising learner achievement in MOOCs is proposed, and used to frame the detailed analysis of the data itself that follows. Finally the key conclusions

✉ Matthew Leach
m.leach@derby.ac.uk

Syed Munib Hadi
s.hadi@derby.ac.uk

[1] University of Derby, Academic Innovation Hub, innovationhub@derby.ac.uk, Derby, UK

that can be drawn from the data are discussed and potential future opportunities for research are highlighted.

Background

MOOCs arguable reached mainstream recognition in 2011 when the Introduction to Artificial Intelligence course, run by Sebastian Thrun and Peter Norvig at Stanford University, reached 160,000 enrolments (Rodriguez 2012). The availability of MOOCs has since grown rapidly, with over 4200 having been released up to the end of 2015 (Shah 2015).

The Academic Innovation Hub, at the University of Derby, developed and delivered three MOOCs in 2015/2016. The first titled "Digital.Me: Managing your Digital Self" (henceforth referred to in the short form Digital.Me), the second titled "Bridging the Dementia Divide: Supporting People Living with Dementia" (henceforth referred to as Dementia), and a third titled "Innovating in Operations Management" (henceforth referred to as Operations Management). All three MOOCs were delivered through the Canvas.Net platform, and ran for eight weeks—with 6 weeks of content/activities, a week for an optional peer review assignment, and a final week contingency to complete any unfinished units. These MOOCs reused some existing materials, but for the most part were designed specifically to be run as MOOCs, and consisted of newly commissioned text and video content.

Two main classifications of the design philosophy of MOOCs have come into use: xMOOCs—following a similar content-driven approach to standard online and on-campus courses; and cMOOCs—courses built around connectivist principles and collaboration between learners (Kennedy 2014). However, alternative classifications have been suggested, with Clark (2013) providing a set of eight distinct classifications at one extreme, and Conole (2013) suggesting the removal of classifications and instead defining MOOCs individually in ten dimensions. For this article it is sufficient to say that although discussions formed the core activity within the three University of Derby MOOCs, the volume of provided content and structure to discussions would place them in the xMOOC category.

MOOCs differ from traditional Higher Education courses primarily due to their open nature—with the courses officially available to anyone able to obtain an email address, such as over the age of 13 under EU law (European Commission 2012). In order to support the widest possible range of learners, while still facilitating an individualised and active approach (Jasnani 2013), the decision was taken to make all of the content in these courses available from the start. This would allow learners to adopt their own strategy, such as following the course synchronously with the provided support, complete at their own rate, join late and catch-up, or just access materials that they were interested in.

There is a growing consensus that completion of the whole course is not the best criteria for analysing MOOCs (Ho et al. 2014; Hayes 2015); but instead a more granular approach should be followed (Lackner et al. 2015).

All three MOOCs consisted of six units with a digital badge available for completing each one. Each unit was self-contained, with its own identified learning outcomes verified through a UK Higher Education validation process (Robertshaw et al. 2015). The specific requirements for each unit varied but typically included accessing all provided materials and engaging in a discussion; with some units requiring completion of a multiple choice quiz. The aim of using badges was to give learners recognition for achieving the learning outcomes in each unit they chose to study. In this way we are moving into the realm of micro-learning—a concept that has emerged to not only mean smaller portions of learning, but also the flexibility for learners to choose what and when to learn (Jomah et al. 2016).

Research methodology

The primary goal of the MOOCs was to deliver high quality learning experiences for the enrolled learners. The factors affecting this experience were to be analysed; but as complete courses individual features could not be tested in isolation. For this reason, a design-based research methodology was chosen. Applying the same course design structure to three MOOCs with different subjects and target audiences allows a better analysis of how that design affected learner behaviour.

Learners completed three optional surveys during each MOOC. The 'welcome survey' that collected intention and demographic data, and the midway survey that sought their opinions on the course, were both standard components in all Canvas.Net courses. In addition, learners were asked to complete a course review survey at the end. This article focuses solely on the quantitative data generated from the multiple choice questions in these surveys.

The Canvas.Net platform, as with many web-based applications, automatically collected a vast amount of data—ranging from high-level learner submissions of assessments and discussion posts, down to the lowest level of times and locations for individual mouse clicks. Portions of this data are available through the web interface for those with the appropriate level of permission (e.g. teachers and support staff); such as grades attained in assessments. For this paper the lower-level data was obtained through the available API. Software was written to make repeated calls to the API, and then to collate and structure the data.

The key data harvested was the dates on which learners completed units, the contributions that they made to discussions, dates of enrolment and last access, and the total amount of time spent on the course. In some cases behaviour was inferred from missing data—for example a learner with a blank value for their last login date was known to have never logged into the course.

Learners could continue to enrol on the MOOCs up to one week before the end. This provided some challenges in terms of assessing retention and unit completions while the course was running—with learners both joining and leaving the courses daily (where leaving results in much of their data being removed). Learners also continue to have limited access to the MOOC once it has completed, which further corrupts the data. For these reasons the required data was collected in totality the day after each MOOC closed.

Definitions of learner participation

The literature provides a number of different categorisations for learners. Of particular relevance to this article are the more concrete definitions that allow direct comparisons to be made between MOOCs. These typically focus on the end state of the learner after they have ceased interacting with the MOOC, and several are summarised in Table 1.

Our use of these categories, and the relation to other authors' definitions is discussed below:

Enrolled Generally the most widely advertised metric is the total number of individuals who have gone through the registration process, to be eligible to access materials on the MOOC. When Future Learn (2015) announced the largest MOOC that had ever run, with 440,000 learners, the figure was for enrolments.

Active Learners in this category are those who are recorded are recorded as having accessed the materials, or at least logged into the course, once it has started. The comparison in Table 1 shows that the categorisation used in this article is less granular in differentiating active learners. Both Nelson (2014) and Ho et al. (2014, 2015) differentiate based on repeated activity, while Huin et al. (2016) differentiate based on learner intention. Neither of those aspects are the focus of this article.

Engaged Learners in this category are those that have actively contributed to the course, through submissions to quizzes, discussions, assignments etc. It should be noted that this metric does not necessarily indicate an increased level of activity per se. A learner may fully engage with retrieving materials and undertaking their own learning, but choose not to demonstrate any of the activity through contributions. In this manner their level of learning is unchallenged and un-evaluated, but not necessarily at a lower level. The use of engaged in this article matches Future Learn's social category (as reported by Nelson 2014) because contributions to discussions was fundamental to engagement with our MOOCs. There were only a few areas (some quizzes) where learners could register engagement that was not

Table 1 Comparison of learner categorisation

This article	Jordan (2014)	Nelson (2014)	Ho et al. (2014)	Ho et al. (2015)	Huin et al. (2016)
Enrolled	Enrolled	Joiners	Only registered	Enrolment	Registrants
					No-shows
Active	Active	Learners	Only viewed	Participant	Participants
		Active	Only explored	Explorer	Learners
		Returning			Observers
Engaged		Social			Active learners
Achieved					
Completed	Completed	Fully participating	Certified	Certified	Complete

social. The MOOCs did include some interactive materials, statistics on their use was not individually recorded.

Achieved This category is fundamental to the design of the presented MOOCs. The modular design and awarding of badges allowed learners to gain recognised achievement, without needing to complete the entire course.

Completed A final metric considered in this paper is the number of completions. This is the number of learners who can be considered to have completed the entire course.

These can be viewed as a progression of nested categories, e.g. a learner must be active before they can be classed as engaged. However, learners are only assigned to one category for the purpose of reporting; which will be the category that takes them closest to completion. This continues the assumption that completion of the entire course is the ideal situation.

Results

Overview of learner participation

Figure 1 applies the proposed categorisation of learners to the three MOOCs that were delivered. Through analysing the data all learners could be easily categorised, and there were no overlaps between categories.

Evidence of different learner goals

A total of 735 learners completed the welcome survey for the dementia MOOC, which represented 35.5% of active learners. The majority of the respondents (477) said that they expected to fully engage with the course; which leaves 258 who from the outset did not intend to fully complete (35.10% of respondents). For Digital.Me 966 learners completed the welcome survey (68.75% of active learners), and 445 (46.07% of respondents) indicated that they did not intend to fully engage. For

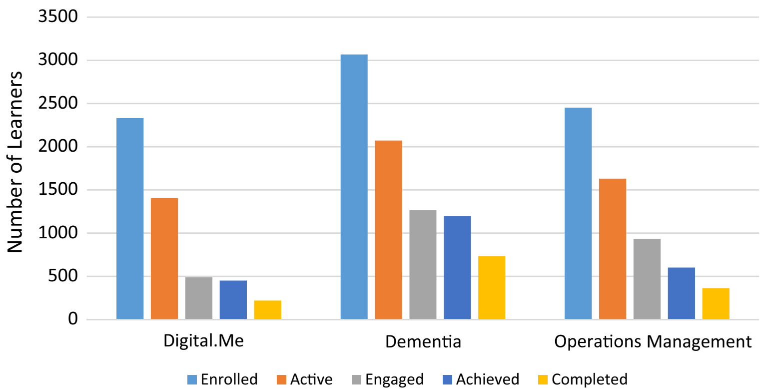

Fig. 1 Overview of MOOCs using proposed metrics

Operations Management 907 learners completed the welcome survey (55.61% of active learners), with 331 (36.94% of respondents) indicating that they did not intend to fully engage. Alternative intentions were to go through all materials but not engage in assessment, to 'check the course out' but not engage in assessment, or to only partially interact with sections of the course that they were interested in.

Although unit 1 was considered the first on the course, some learners on all three MOOCs took the opportunity to start with a different unit. The figures were 7.08% of learners on Digital.Me, 6.17% of learners on Dementia, and 5.15% on Operations Management. There were also groups of learners that completed at least one unit, but never completed unit 1. These figures were 2.65% of learners on Digital.Me, 2.08% Dementia, but only 0.33% on Operations Management.

Completion of units

A substantial number of learners registered but never accessed the course—31.61% for Digital.Me, 26.29% for Dementia and 33.51% for Operations Management. The figures are consistent with a general trend in MOOCs (Onah et al. 2014). For this reason the initial analysis was performed against active learners—those that logged into the course at least once. In Digital.me 15.59% of active learners (9.40% of those enrolled) completed the course; 35.47% (23.96%) completed the Dementia course; 22.32% (14.84%) completed the Operations Management course—where course completion meant meeting the requirements for all six units of a course.

Figure 2 shows the percentage of active learners that completed each unit of the course. Note that the figures shown for completing unit 6 is different to the figures reported above for completing the course (all six units). Since all content was open from the beginning, there were learners who chose to complete unit 6 that did not complete all of the other five. This theme is explored in greater depth elsewhere in this article.

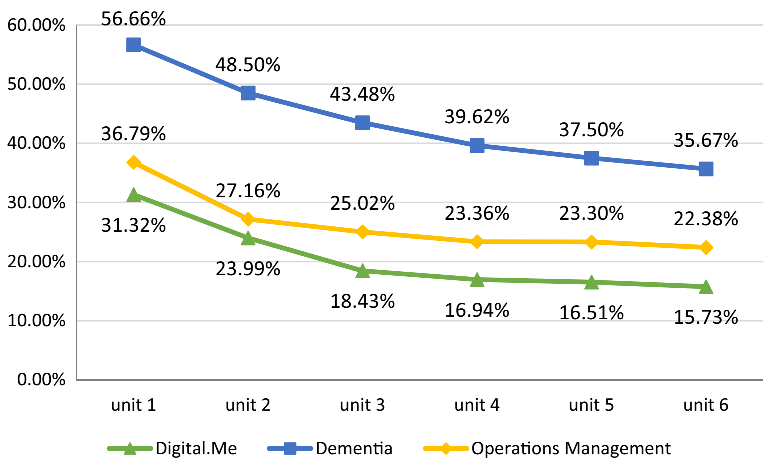

Fig. 2 Learners completing units as a percentage of active learners

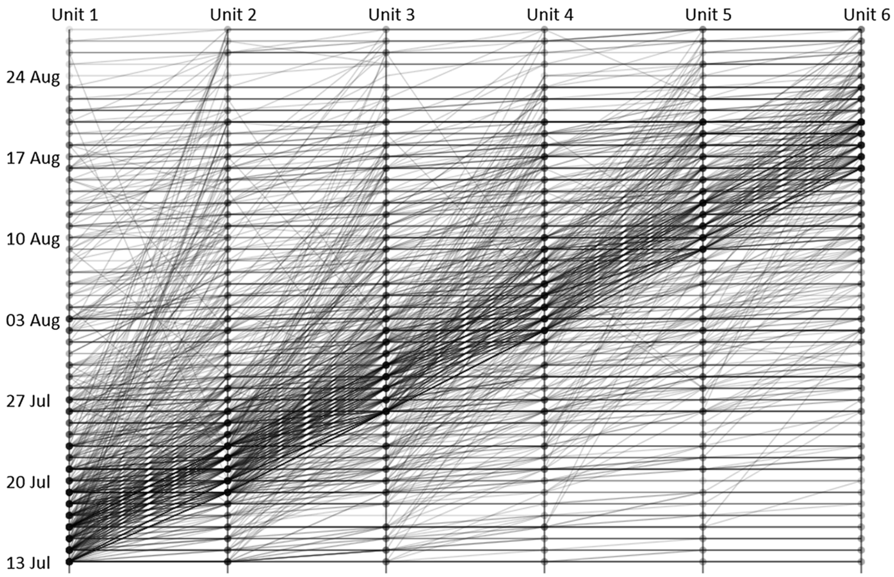

Fig. 3 Pattern of unit completion for dementia

The percentage of badges awarded, relative to the total number that could have been awarded to active learners, provides another metric for judging learner activity. For example, in Dementia 2072 learners were active, and so across all 6 units a maximum of 12,432 badges could have been awarded. In total 5417 badges were actually awarded on the Dementia course, which translates to 43.57% of the potential maximum. For Digital.me 20.49% of available badges were awarded to active learners, and for Operations Management the figure was 36.20%. In all cases these figures are higher than the course completion percentages, but are still legitimate measures of learning that has taken place.

The pattern through which learners completed units also provides interesting insights into their behaviour. Figures 3 and 4 show this pattern of completion—with each learner represented by a single line that shows on what date they completed each unit.[1]

There is a strong diagonal feature which represents a large portion of the learners following the weekly pace of each course—in line with the provided support. It is important to note that the only horizontal lines in the figures come from the data itself, and indicate that two units were completed on the same date. In some cases a learner may only have completed two units on the same date, but in both Dementia and Digital.Me around 10% of learners completed the entire course in one day. For Operations Management the figure was slightly higher at 15%.

Also note that lines going diagonally downwards (higher on left than the right) represent units being completed out of sequence. This variation in order is better

[1] The tools used to create the visualisations in Figs. 3, 4, 5, 6 and 7 have been made available for other researchers that may find them of use at https://github.com/Thoughtfulmonkey/deby-mooc-data-vis.

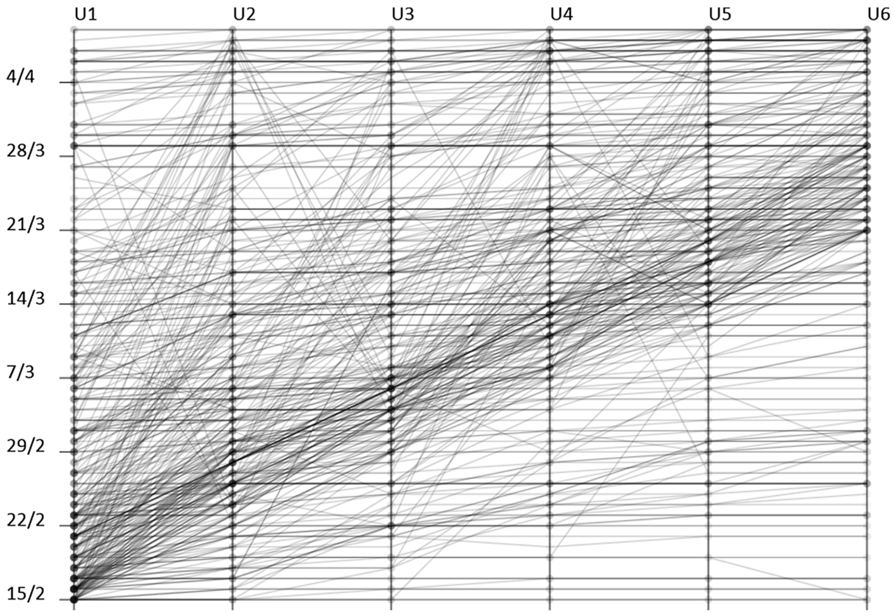

Fig. 4 Pattern of unit completion for operations management

visualised in a state diagram, as recommended by Coffrin et al. (2014) and shown in Figs. 5, 6 and 7.

Here each circle is a unit—with size denoting number of completions. A line entering from above indicates the number of learners that started with that unit, and the line below is the number of learners for which that was their last unit completed. Curved lines above represent learners choosing which unit to move onto next, while curved lines below represent back-tracking to complete an earlier unit. Again this

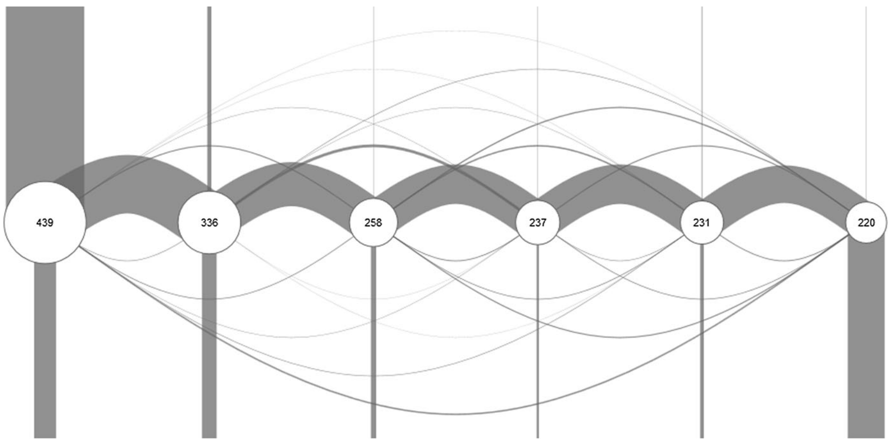

Fig. 5 State diagram representing the order in which units were completed—Digital.Me

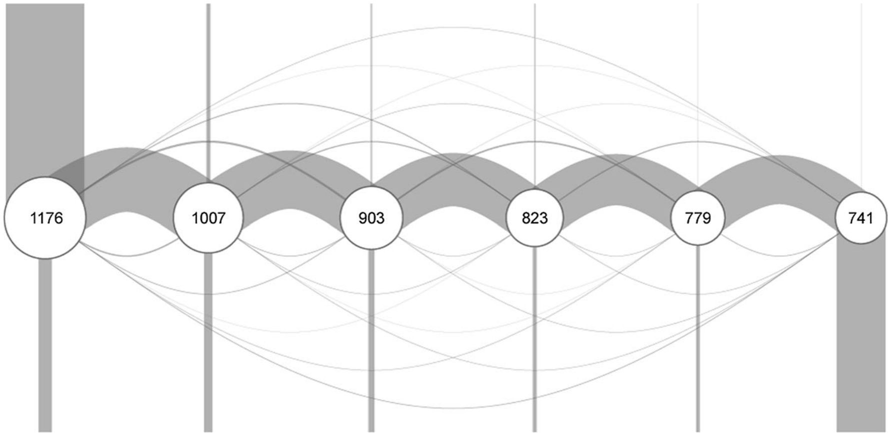

Fig. 6 State diagram representing the order in which units were completed—dementia

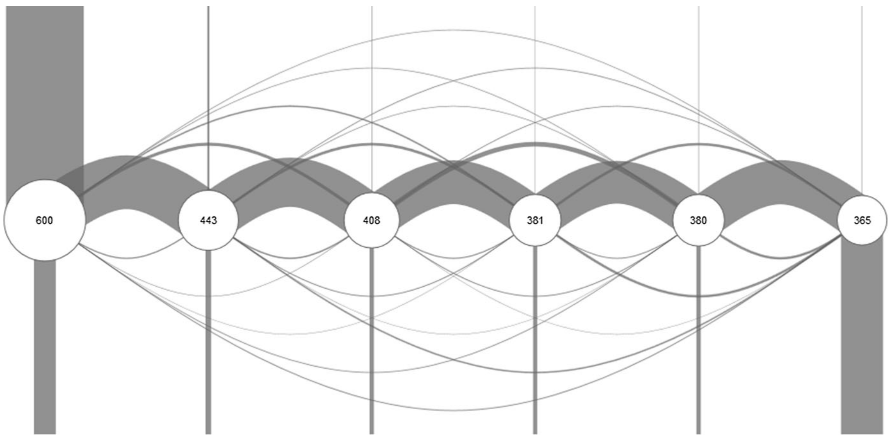

Fig. 7 State diagram representing the order in which units were completed—operations management

demonstrates a core body of learners moving linearly through the course, but also some learners taking different approaches.

Although most learners completed each course in a much shorter time than the six scheduled weeks, this did not translate into spending less time on the course. Figure 8 compares the gap in days between the first and sixth units completed, against total time logged on Dementia. There is only a small positive correlation between these datasets of 0.12 (and 0.08 for Digital.Me), and we can see that completing all six units within 3 days compares well to spreading activity over 29 days.

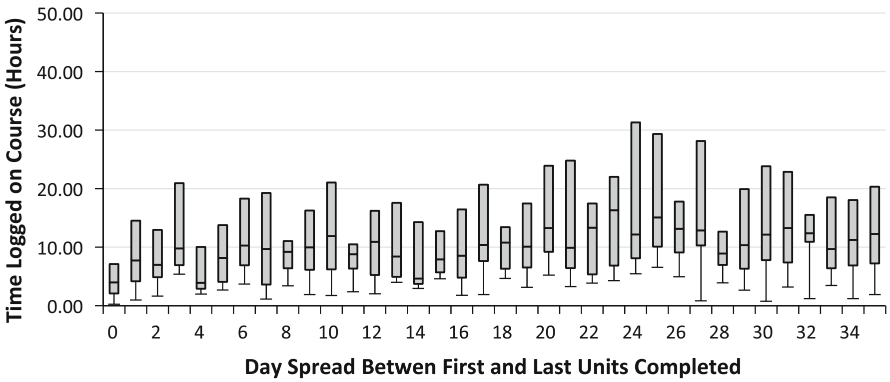

Fig. 8 Distribution of time logged on course against gap in days (up to 35) between completing first and sixth units, for learners that completed the course. Top whiskers not shown

Modular flow

Evidence for learner activity within specific units comes from dates when those units were completed, but also more tellingly from the dates of posts in the discussions. There were on average twice as many posts in the supported unit than any of the others, with 2.02 times as many in Digital.Me, 2.22 times as many in Dementia, and 2.45 times as many in Operations Management. This indicates that in each week of the MOOCs more activity took place in the week that was currently being supported, than in any of the other weeks.

Although the increased discussion activity represents a large number of learners that were synchronised with the current week (e.g. participating in Unit 2 during week 2), it should be noted that combined together more activity was taking place in unsupported units than supported ones. That is to say that although in week 3 more discussions were taking place in unit 3 than any of the other units, combined together the discussions in units 1, 2, 4, 5 and 6 were higher. In Digital.Me an average of 65.8% of overall posts were made in an unsupported week, for Dementia the figure was 58.4%, and for Operations Management the figure was 59.8%.

Conclusions and recommendations

By awarding badges for achievements on each unit, many more learners gained recognition of their achievements than if there had just been a completion certificate. In Digital.me 451 learners gained badges (compared to 209 being rewarded by the completion certificate); in Dementia 1198 learners gained badges (compared with 527 certificates); in Operations Management 602 learners gained badges (compared with 364 certificates).

The higher percentage of learners who's only recognition came from badges, compared to those completing the whole course, proves that engagement takes place in MOOCs that is not reflected in overall completion rates. This suggests that when measuring learning success or completion in MOOCs, this micro learning needs to be taken into account. However, designing modular MOOCs should not be just about providing recognition for individual segments of the course. Instead these modules should also have clear learning outcomes and assessment criteria to maintain academic integrity.

The second aim of the open, modular approach was to provide learners with more freedom in how they approached the course, and this is certainly evident in the data. The presented graphs show learners taking non-standard routes through the courses, and the discussion data shows the majority of activity taking place outside of what would be considered the supported schedule.

The diverse range of abilities, intentions and behaviours make identifying any patterns almost impossible. There were no correlations between data sets such as amount of time logged, duration to complete the course, amount of contributions to discussions, and the quality of contributions. Learners joined at different times and completed different units, with the majority of activity taking place outside of the supported week. If the aim is to provide beneficial materials to as wide an audience as possible then a certain amount of chaos may have to be embraced.

Further research

There is the potential to go further—using a purely modular structure without sequence or hierarchy, where learners are not just allowed but actively encouraged to choose their own sequence.

The considered metrics are useful for making direct comparisons between MOOCs, but lack a detailed classification that attempts to establish motivations and background experiences. A system for this is presented by Huin et al. (2016); but also by Khalil et al. (2016) who have used clustering techniques to group learners into four distinct categories: *drop-outs*; *gamblers*, who attempt to complete with minimal effort; *sociable students*, who had high levels of engagement but low scores; and *perfect students*, who had high levels of engagement and success. These are categories that seem, superficially at least, to be evident in our data, and warrants further investigation.

Bali (2014) and Ho et al. (2014) suggest placing less emphasis on course completion and completion rates. Given this, an area for further research is the exploration of minimal criteria to "complete" a course, but with more engaging and challenging optional stretch activities. When scores do not contribute to the final grade then more accurate assessments of student progress might be discerned.

There is significant focus on what constitutes success on the part of the learners; but the proposed categorisation may also allow the success of the course to be evaluated. Drop-offs from one category to another may highlight issues with aspects such as course design or materials. If a learner becomes active but not engaged, were the discussion topics not enticing enough? If a learner becomes engaged but does not achieve anything, did the materials not adequately prepare them for a quiz? As such the transitions between categories could be a fruitful area of research.

As part of the welcome survey learners are asked to indicate their intention (how much they intend to engage), as well as various demographic information. Segmenting the data along these lines may reveal different behaviours and is work that will be undertaken in the near future.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The authors declare that the research complied with ethical standards.

References

Bali, M. (2014). MOOC pedagogy: Gleaning good practice from existing MOOCs. *MERLOT Journal of Online Learning and Teaching, 10*(1), 44–56. Retrieved from http://jolt.merlot.org/vol10no1/bali_0314.pdf.

Clark, D. (2013). MOOCs: Taxonomy of 8 types of MOOC. Donald Clark Paln B. Retrieved 9 May 2016 from http://donaldclarkplanb.blogspot.co.uk/2013/04/moocs-taxonomy-of-8-types-of-mooc.html.

Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOC. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge (LAK '14)* (pp. 83–92). New York, NY: ACM.

Conole, G. (2013). MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs. *RED. Revista de Educación a Distancia*, *39*. Retrieved 14 May 2016, from http://www.um.es/ead/red/39/conole.pdf.

European Commission. (2012). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Retrieved 5 October 2015, from http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf.

Future Learn. (2015). [Updated] FutureLearn delivers the largest MOOC ever as more than 440,000 learners convene for English language learning. Retrieved May 12, 2016, from https://about.futurelearn.com/press-releases/futurelearn-delivers-the-largest-mooc-ever-as-nearly-400000-learners-convene-for-english-language-learning/.

Hayes, S. (2015). MOOCs and quality: A review of the recent literature. Retrieved 5 October 2015, from http://www.qaa.ac.uk/en/Publications/Documents/MOOCs-and-Quality-Literature-Review-15.pdf.

Ho, A. D., Reich, J., Nesterko, S., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The first year of open online courses. Retrieved 22 September 2015, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2381263.

Ho, A. D., Chuang, I., Reich, J., Coleman, C., Whitehill, J., Northcutt, C., Williams, J. J., Hansen, J., Lopez, G., & Petersen, R. (2015). HarvardX and MITx: Two years of open online courses (HarvardX Working Paper No. 10).

Huin, L., Bergheaud, Y., Caron, P. A., Codina, A., & Disson E. (2016). Measuring completion and dropout in MOOCs: A learner-centered model. In *Proceedings of the European Stakeholder Summit on Experiences and Best Practices in and Around MOOCs (EMOOCS 2016)*, 22–24 February, 2016. Graz, Austria.

Jasnani, P. (2013). Designing MOOCs: A white paper on instructional design for MOOCs. Retrieved 5 October 2015, from http://www.tatainteractive.com/pdf/Designing_MOOCs-A_White_Paper_on_ID_for_MOOCs.pdf.

Jomah, O., Masoud, A. K., Kishore, X. P., & Sagaya, A. (2016). Micro learning: A modernized education system. *Broad Research in Artificial Intelligence and Neroscience, 7*(1), 103–110. Retrieved 10 September 2016 from https://www.edusoft.ro/brain/index.php/brain/article/viewFile/582/627.

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), 133–160.

Kennedy, J. (2014). Characteristics of massive open online courses (MOOCs): A research review, 2009–2012. *Journal of Interactive Online Learning, 13*(1). Retrieved May 10, 2016, http://www.openeducationeuropa.eu/sites/default/files/asset/Characteristics-massive-online-courses-MOOCs.pdf.

Khalil, M., Kastl, C., & Ebner, M. (2016). Portraying MOOCs learners: A clustering experience using learning analytics. In *Proceedings of the European Stakeholder Summit on Experiences and Best Practices in and Around MOOCs (EMOOCS 2016)*, 22–24 February, 2016. Graz, Austria.

Lackner, E., Ebner, M., & Khalil, M. (2015). MOOCs as granular systems: design patterns to foster participant activity. Retrieved 10 September 2015, from http://www.researchgate.net/publication/277890739_MOOCs_as_granular_systems_design_patterns_to_foster_participant_activity.

Nelson, S. (2014). Updated numbers from our platform. Future learn blog. Retrieved 11 May, 2016, from https://about.futurelearn.com/blog/updated-numbers/.

Onah, D. F. O., Sinclair, J., & Boyatt, R. (2014) Dropout rates of massive open online Courses: Behavioural patterns. In *Proceedings of the 6th International Conference on Education and New Learning Technologies (EDULEARN14)*, 7–9 July, 2014. Barcelona, Spain.

Robertshaw, D., Owen, E., & Hadi, S. M. (2015). An approach to quality assurance of MOOCs: Bringing MOOCs into mainstream quality processes. Presented at Learning with MOOCs II. 2–3 October, 2015.

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance and E-Learning*. Retrieved from http://www.eurodl.org/?article=516.

Shah, D. (2015). By the numbers: MOOCS in 2015. *Class Central*. Retrieved 13 September 2016 from https://www.class-central.com/report/moocs-2015-stats/.

Matthew Leach is a postdoctoral researcher in the University of Derby's Academic Innovation Hub. His interests are in new approaches to education and training, including MOOCs and mixed realities.

Syed Munib Hadi is the current head of the University of Derby's Academic Innovation Hub. His key areas of interest are improving retention through use of technology and learner Analytic, innovation in learning design and pedagogy and the role innovation and technology in changing educational delivery models.