Comparison of Live, Late and Archived Mode Learner Behavior in an Advanced Engineering MOOC

Kerrie Douglas[1], Harsh Aggarwal[2], Taylor V. Williams[1], Yichen Fan[2], Peter Bermel[3]

[1] School of Engineering Education, Purdue University, West Lafayette, IN 47907
[2] School of Industrial Engineering, Purdue University, West Lafayette, IN 47907
[3] School of Electrical & Computer Engineering, Purdue University, West Lafayette, IN 47907

***Abstract*—Advanced engineering Massive Open Online Courses (MOOCs) aim to teach highly technical content to learners around the world at a much greater scale than the traditional classroom. Unlike the latter setting, it is common for many MOOC learners to initiate learning after the official start date, or even after the course ends. This archive-mode learning allows students an open-ended period for course completion. However, little research on MOOCs focuses on how such learners utilize these resources. To address this gap, the present study seeks to compare live, late, and archive learners' intentions, interests, and behaviors in accessing the course materials in a highly technical engineering MOOC. Utilizing the Kruskal-Wallis test, we find that live learners desire to achieve a passing grade at the onset of the course, while archive learners are less concerned about grades. Also, learners who work full-time tend to enroll late, while full-time students tend to enroll as live or archive learners. Most importantly, while similar classes of behaviors are found in all groups, a significantly greater proportion of archive-mode learners complete the course than live or especially late-enrolling learners. These results suggest that relaxing or removing time constraints may significantly improve learning outcomes in MOOCs.**

Keywords—MOOCs, highly technical courses, intentions, effect size analysis

I. Introduction

There is no 'one-size-fits-all' high quality learning experience. Hence, the US National Academies of Engineering has posed a Grand Challenge to Advance Personalize Learning. In practice, this challenges engineers to design learning environments that account for learners' motivations, interests, preferred style of learning, cultural background, age, and other factors [1]. While MOOCs, *Massive Open Online Courses*, have certainly increased availability of learning content for learners world-wide, the fundamental aim has been to offer courses at a much greater scale than the traditional classroom, i.e. "massive". The content of courses offered through MOOC platforms range from courses taken for pure leisure (e.g., a book club [2]) to courses for personal finance, health and wellness education (e.g., caring for a loved one with Alzheimer's [3]) to a wide range of undergraduate-level technical courses (e.g., algorithms and data structure [4], writing composition), and even to highly specialized topics for graduate or professional learners (e.g., the physics of conductive polymers [5]).

By virtue of the scale of the courses, MOOC learners are quite heterogeneous, with still largely unknown learning needs [6], [7]. Individual learners may also have their own varied needs and intentions, depending on the topic of the course. For example, the same person could desire social interaction when participating in a MOOC on caring for someone with Alzheimer's, but prefer to learn independently in MOOCs focused on engineering research. The level of individual differences and learning style preferences is largely unknown in the MOOC environment. MOOCs arguably provide much flexibility for learners (e.g., anytime learning), yet the format of the courses are essentially the same regardless of learner needs or course topic: there are videos, assignments, assessments, and discussion forums assembled in a fixed order.

While the format of the courses have changed very little since MOOCs were first offered, most MOOC platforms are moving away from being completely free and open, to "freemium" models, where learners or their affiliated employers pay for access to premium content [8]. New content continues to be offered through MOOC platforms, particularly targeting professional development and continuing education opportunities for employees of large corporations. Although MOOCs originally aimed to provide free courses to anyone, the provider platforms have sought to partner more closely with industries as a potential viable revenue stream [8].

As a result, MOOC platforms have understandably begun to target professional learners, with US companies spending on average $1,273 per employee on education in 2016; a total of $164.2 billion as of 2013 [9]. Science and engineering-related content is a popular topic for such training that is often delivered either fully online or in a hybrid format. Finding ways to tailor content delivery to specific engineering learners has potential to make the learning process much more efficient and effective. For instance, in the context of MOOC platforms, a single course could have multiple branches specifically designed with particular learners in mind. Recent research on adaptive learning in MOOCs has shown that it is possible to adapt course delivery based on learning style [10].

To innovate more personalized online learning experiences targeted specifically for advanced engineering learners, there is a fundamental need to better understand learners that seek out advanced engineering content. The purpose of this paper is to understand similarities and differences in learner needs and their usage of a highly-specialized engineering MOOC, based on whether they enrolled at the beginning of a course, joined after the course started, or utilized the materials fully in the archived-mode. The research questions are: (1) What is the significance of when a learner joins an advanced engineering MOOC in terms of their intentions and interests in the course? (2) What is

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

the significance of when a learner joins an advanced engineering MOOC in terms of their usage of course materials?

II. Literature Review

In US-based MOOC providers, such as edX, Coursera, and Udacity, learners can enroll at any point in the course, and even view and interact with all course materials in an archived mode, after it ends. Learners have a great deal of flexibility in terms of choice; the ability to choose which resources to utilize and which to skip, and when to start and disengage from a MOOC. Researchers have consistently found essentially 4-5 groups of learners in terms of their usage: those that fully engage with all materials, those that utilize most materials (without assessments), those that begin fully engaging, but then drop out after one or two weeks; and those that sample the material with sporadic bursts of activity [11], [12].

While some have claimed that low completion rates are an artifact of learners not intending to utilize MOOCs as a full course, cluster research and survey research regarding MOOC learners' intention demonstrate that the majority of learners who initiate activity, actually begin with an intention of fully utilizing the learning materials [7].

Research by Douglas *et al*. evaluated learners' intrinsic and extrinsic motivations, and found that MOOC learners generally exhibit overall high levels of innate desire to engage, but vary in their extrinsic motivation, such as supervisor request and position requirement [7]. Students generally reported the highest levels of extrinsic motivation when compared to workers and unemployed individuals. However, the research was focused on live learners in two courses, one of which was the same course of our study. More research is needed on how to help learners achieve their own learning goals in open courses and what course characteristics encourage ongoing engagement.

Researchers have been finding features that describe learners who display evidence of learning through performance on assessments. Activity alone is a large prediction of how well a learner is expected to perform on assessments [13]. In addition to actually utilizing materials, social interaction has been found to be a very important variable in learner achievement [14], [15]. Yet, many learners join MOOCs long after the course starts, significantly limiting interactions with anyone associated with the course. To what extent these learners who join long after official start date utilize materials is particularly of interest, considering that social networks have been found to be the most predictive variable in detecting early drop-out in a large online engineering course [16]. On the other hand, when enrolling several weeks into a course with thousands of other learners, one might not have an expectation of ever interacting with the actual instructor, even if intended to utilize all other non-social resources.

Campbell and colleagues examined learners' intention and behavior in two undergraduate-level courses on statistics and programing [17]. The researchers created two groups of comparison: learners who participated in the live-mode of the courses and those that participated in an archived-mode. The researchers found that learners in both groups were similar in terms of usage of course materials (videos, assignments, assessments, and discussion forum). Previous studies have viewed late enrollers as a problem to be minimized [18], [19]. Archive-mode (self-paced) learners have been viewed as having greater potential, but viewed as requiring further regulation to be successful [20].

Considering that working, professional engineers may have a need for information 'just-in-time,' it is important for researchers to understand their needs, whether or not they participated in the live-mode of the course. While previous research has found similarity between the groups when the courses were introductory-level, it is important to examine whether the trend is the same in courses with advanced engineering topics. Furthermore, with flexibility of enrolling at any point in a course, creating groups of "live" and "archive" alone may not be sufficient to capture types of learners. A learner could enroll one-week before the "archive" mode and then continue utilizing the material into the archive-mode. To understand more about the fundamental intentions of learners and their behavior in advanced engineering open courses, we begin first examining when learners begin engagement in a course and then empirically determine meaningful groups to compare intention and behavior. As people continue to enroll in the course as late as two years after the course has ended, it is important to examine how learner needs vary based on whether they participate in the course during the live offering or once the materials are simply published and available online, with no personal interaction.

III. Methods

A. Data Collection and Cleansing

The pre-class survey was administered as part of the first week's course material in nano520X, Fundamentals of Nanoelectronics Part A, offered on the edX platform for 9 weeks in the spring of 2015. The survey included questions on both user background and intention for taking the course. After the nine-week course ended, survey responses from archive mode learners were continuously collected until February 11, 2017. During this time, a total of 9412 students enrolled in the course.

The survey data was cleaned by removing incomplete survey responses in a two-step process. In the first step, responses, which had at least 50% of the survey items answered, were considered complete and retained for the second step. For the second step, completed surveys were checked for 100% completion of the survey items related to this study's research question, or were eliminated from analysis. Less than 2% of the responses that passed step one were removed in the second step.

The behavioral data was collected as module access data associated with enrolled learners in the Nanoelectronics course. We identified 320 course materials (i.e., modules) made available throughout the course, including lectures (video and PDF formats), homework, tutorials, quizzes, and exams. In our subsequent learner behavior analysis, we identified whether each learner accessed each item, or not. We then employed a k-means clustering algorithm to group users with similar behavior. The number of clusters used for each group of learners (live, late, and archive) was chosen through elbow plots, following the procedure described in Ref. [11].

B. Identifying Learner Groups

To understand the intentions of different types of learners, learners were divided into three groups based on their survey completion date: live learners, late enrollers, and archive learners. The archive learner group consists those who complete the survey after course end date, May 26, 2015.

Those who completed the survey before the course end date were initially identifies as live users. However, when we plotted the survey completion date (Fig. 1 and Fig. 2), most learners completed the survey within 10 days of the course start date, March 26, 2015. A total of 1272 learners completed the survey (out of 9412 possible respondents), and live learners completed approximately 34%. Approximately 55% learners completed the survey during the first two weeks of the live course. After that, the number of learners who took the survey decreased. Starting from April 15, 2015, the daily number of people who completed the survey was less than 10. Therefore, a third group, late enrollers, was identified.

Late enrollers are those who completed the survey after April 15 and before the course end date, May 26, 2015. Although late enrollers joined the class when the class was in live-mode, and the professor was involved in the discussion board, the course had started three weeks prior, and it is doubtful if the learners could catch up and fully engage in the course (as shown in the behavioral data). Often, many course assessments and concepts (which could be foundational for later parts of the course) were already completed before the late enrollers joined the course.

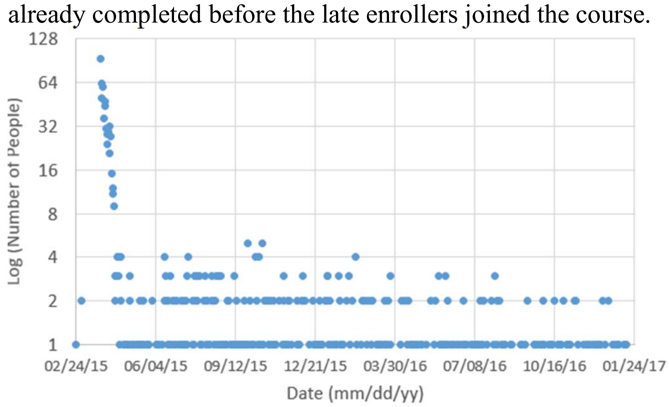

Fig. 1. Survey Completion Date for All Learners

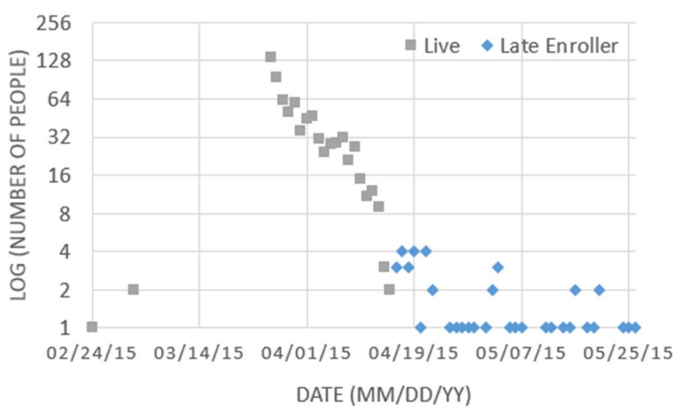

Fig. 2. Survey Completion Date for Live and Late Enrolling Learners

C. The Kruskal–Wallis Test

To determine if live, late enrolling, and archive mode learners were statistically significantly different in terms of their intentions, interests, and behaviors in accessing the course material, we used the Kruskal–Wallis test. We initially conducted a one-way ANOVA test, and the result showed that data is not normally distributed, which violates a fundamental assumption of the one-way ANOVA test. Hence, a non-parametric test, Kruskal-Wallis test, was used instead to analyze our data, as discussed in the next section [21].

IV. Results

A. Are live, late enrolling, and archive mode learners significantly different in terms of their intentions, interests, and behaviors in accessing the course material?

Out of the total 1272 learners analyzed in this study, 773 were live learners, 49 were late enroller learners, and 450 were archive mode learners. Table 1 lists all the questions used to detect the differences in intentions, interests, and behaviors of the learners in this study.

The non-parametric Kruskal-Wallis analysis of variance shows that there is a significant difference in the learning goals among the learner groups ($p < 0.001$). Live learners hope to achieve a passing grade, while late enrollers and archive learners are not concerned with grade outcome.

There was no significant difference among reasons for taking the course ($p = 0.065$), what they hope to gain ($p = 0.136$), and planned time commitment ($p = 0.355$). Most people enrolled in this course because of their interest in nanoelectronics. They hoped to become well acquainted with foundational principles of nanoelectronics. Their planned time commitment was 3-6 hours per week.

TABLE I. Survey Question Description

Question Description		
Questions		***Options***
1	Which of the following best describes the reason you are taking this course?	1. I am curious about nanoelectronics. 2. I am very interested in nanoelectronics. 3. I am familiar with other courses offered by Professor Datta. 4. This course was recommended to me by someone I know.
2	Which of the following best describes what you hope to gain from the Nanoelectronics course?	1. I want to learn a broad overview of what Nanoelectronics is. 2. I want to become well acquainted with foundational principles of nanoelectronics. 3. I want to be able to apply material from this course in future projects related to nano electronics.
3	Which of the following best describes you learning goals for this course?	1. I wish to achieve a high final grade 2. I wish to achieve a passing grade 3. I am not concerned with grade outcome
4	How much time do you plan to dedicate each week on this course?	1. 1-3 hours 2. 3-6 hours 3. 6-9 hours 4. 9-12 hours 5. 12-15 hours 6. uncertain

B. What are the demographic characteristics of live, late enrolling, and archive mode learners?

Fig. 3 indicates that full-time workers are less likely to learn information about relevant courses on a timely basis, compared to full-time students as the number of people not working is much higher in live-mode than the archive mode. We conducted a Kruskal-Wallis test on employment status, and results showed that there is a significant difference (p <0.001) among the employment status of the three groups of learners (see Fig. 3). Late enrollers appeared to be full-time workers, while live and archive learners were mostly full-time students.

All three groups appear similar in educational backgrounds as can be seen from Fig. 4. Approximately 67% of learners in each group hold a bachelor's degree or a master's degree. Approximately 67% of learners in each group report having prior related coursework experience (Fig. 5). Approximately 60% of learners in each group report taking more than two semesters of college-level calculus (Fig. 6). For live and late enrollers, approximately 50% of learners heard about the concept of nanoelectronics from the edX website (Fig. 7). However, for archive learners, more learners (~67%) knew of the concept from other sources.

Finally, the behaviors of the learners in each segment were studied via a k-means cluster analysis, described in Methods. The live learners are clustered in Fig. 8; the late-enrolling learners are clustered in Fig. 9; and the archive-mode learners are clustered in Fig. 10.

One key result is that the elbow plot analysis predicts four clusters for each learner group, pointing to fundamental similarities between learners in this course. The four clusters are fully-engaged learners, who participate in all course materials; consistent viewers, who review most content but do not complete most assessments; one-week engaged learners, who terminate substantial involvement after the first week; and sporadic users, who only participate in a small fraction of the course materials. While abandonment rates are not directly measured, sporadic user rates may be viewed as a suitable proxy.

A second key result is that the proportion of users in each cluster vary with the time of enrollment. In particular, we observe about ⅛ of the live-mode learners are fully engaged; whereas virtually none of the late-enrollment learners are fully engaged (although ⅕ are often engaged); and ¼ of the archive-mode learners are fully engaged. On the other end of the spectrum, about ¾ of live learners are sporadic; about ⅔ of late learners are sporadic; and less than ½ of archive-mode learners are sporadic participants.

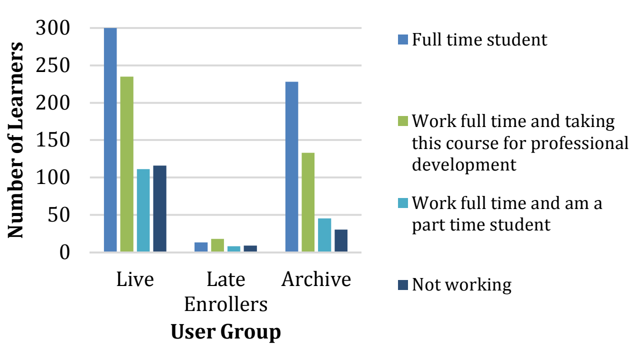

Fig. 3. Employment Status, sorted by Learner Groups

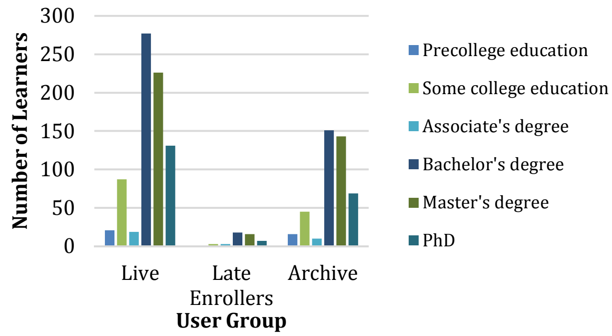

Fig. 4. Level of Education, sorted by Learner Groups

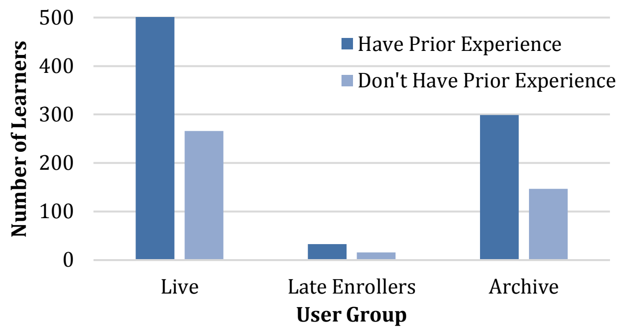

Fig. 5. Prior Experience on Nanoelectronics, sorted by Learner Groups

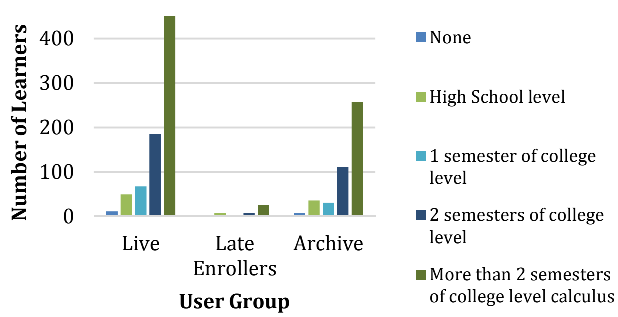

Fig. 6. Level of Math, sorted by Learner Groups

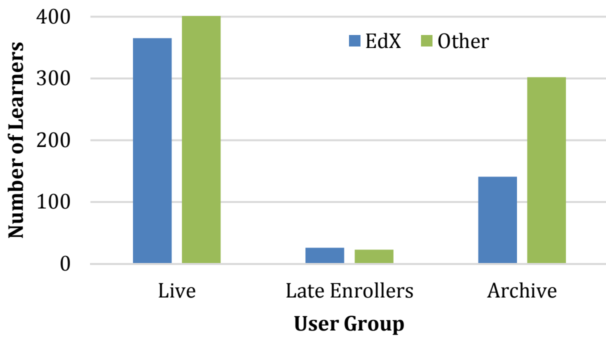

Fig. 7. Source of the Concepts of taking a course on Nanoelectronics, as sorted by Learner Groups

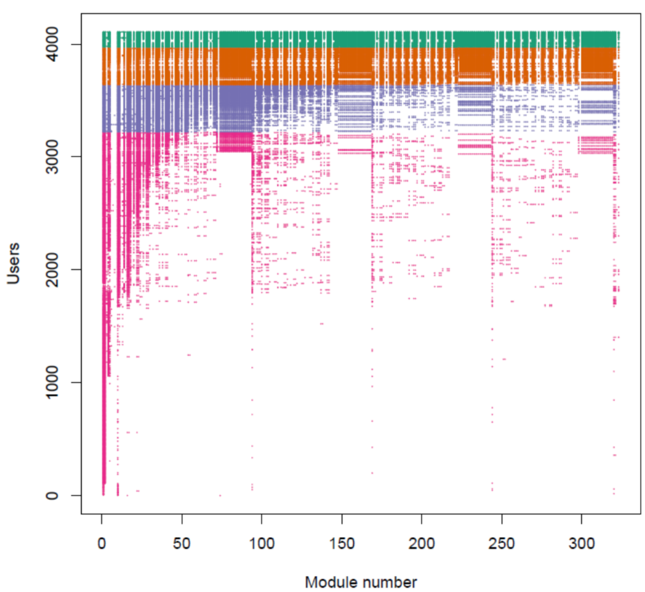

Fig. 8. Clustered behavioral results for all live learners in the studied MOOC, derived from clickstream data. Highly-engaged learners are in the top cluster (about 1/8 of the total number of students), with sporadic learners clustered toward the bottom (about ¾ of total students).

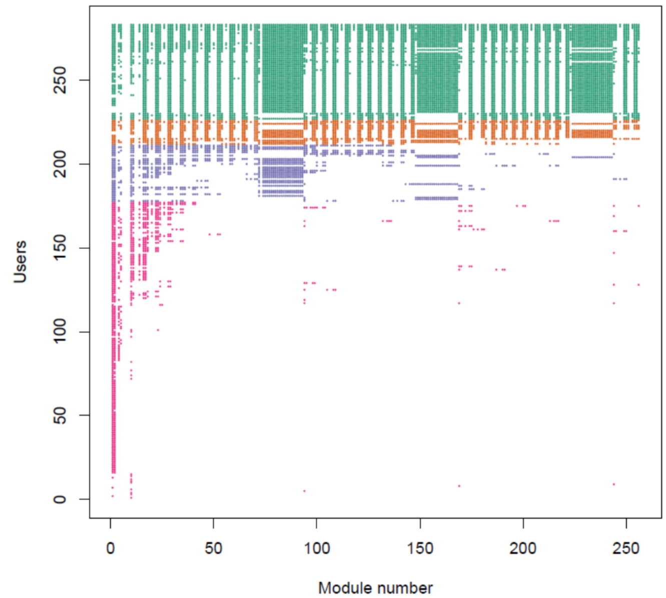

Fig. 9. Clustered behavioral results for all late-enrolling learners in the studied MOOC. A distinctive cluster of assessment-focused learners emerges at the top (approximately 1/5 of the total), with sporadic learners clustered toward the bottom. Virtually none of these students are deemed to be highly engaged, because of their consistently large fraction of missed content.

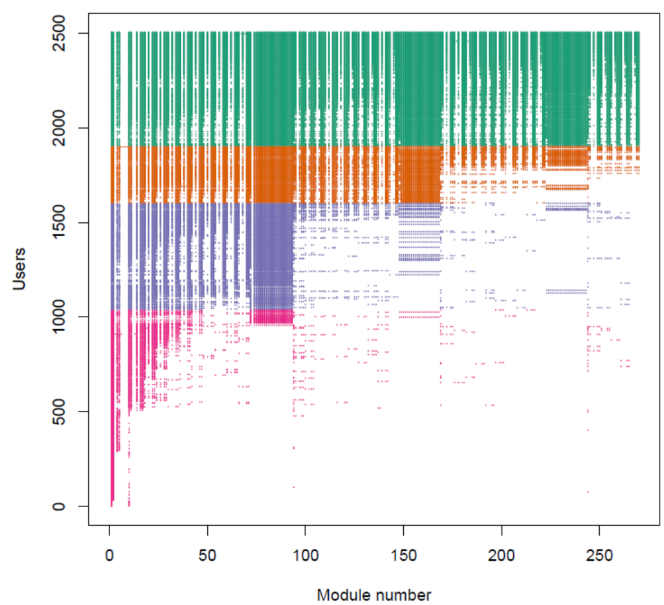

Fig. 10. Clustered behavioral results for all archive-mode learners in the studied MOOC. Highly-engaged learners are in the top cluster, with sporadic learners clustered toward the bottom. A larger fraction of learners (about ¼) are highly engaged, compared to the other groups. Correspondingly, a smaller fraction of learners (about ½) are only sporadically engaged.

V. Discussion

A. Intentions, Interests, and Behavior in Accessing Course Materials

While previous researchers found little difference in learners' interests and behaviors based on time of enrollment [16], these results suggest that live learners in advanced engineering courses have different learning goals compared to late enroller and archive mode learners. Most live learners hope to achieve a passing grade in the course, while late enrollers and archive mode learners were not concerned about the grade. This difference in goals is possibly because live learners have more intention of participating in the course while the teaching staff is still involved. They are active in learning and would like to achieve a passing grade outcome. Late enroller and archive mode learners did not have the resource of teaching staff; therefore, they may not have been concerned about the grading outcome.

Our behavioral clustering on live, late, and archive learners (Figs. 8, 9, and 10, respectively) reveals two crucial results. First, it becomes clear similar classes of behavior are found in all cases, as in prior work [9]. The four classes of behavior found in all cases are highly-engaged learners, consistent learners, one-week engaged learners, and sporadic learners. Second, it appears that a significantly greater proportion of archive-mode learners complete the course, compared to live or especially late-enrolling learners. In other words, the fraction of highly engaged learners increases as one goes from late enrollment to live-mode to archive-mode. The implication of this finding is that time constraints create a major penalty in terms of full participation by students, and that relaxing or removing these restrictions may substantially increase the fraction of learners able to fully engage and learn the desired knowledge from the corresponding MOOC.

B. Demographic Characteristics

The number of unemployed people drops significantly in the archive mode group as compared to live and late enrollers This phenomenon might occur because this course, Fundamentals of Nanoelectronics Part A, provides little opportunity for social interaction in the archive mode. The archive-mode course could still be attractive to full-time students who have a study focus in the field of nanoelectronics and professional employees who work in this field, who already have adequate alternatives for socialization. It may be appropriate to develop additional survey questions to probe this finding and understand different types of learners concerning their employment status.

VI. Conclusions

In this study, we found that live learners hoped to achieve a passing grade in an advanced engineering MOOC, while late enrollers and archive learners were less concerned about their grade outcomes. All three groups of learners had similar reasons for taking the course, planned time commitment, and educational backgrounds. As for the employment status, late enrollers were mostly full-time workers, while live and archive learners tended to be full-time students. In terms of behavior, it was clear that late enrollers were at a significant disadvantage in terms of their ability to complete course content. However, archive-mode learners were in fact significantly more likely to complete courses than live-mode learners.

The significant gap between stated motivations and behavior suggests that structural factors created by the platform may play a key role in determining individual behaviors, and that one of the key constraints affecting many learners is time. Relaxing this constraint appears to be highly conducive to greater engagement by individual learners in the courses.

Further study is recommended to analyze motivations and intentions among full-time students, full-time workers, and unemployed individuals, concerning the three groups of learners to better understand the relationship between employment status and learner motivation. Such information could be used to aid MOOC developers in designing targeted courses for learners at different levels of their career.

Acknowledgment

This work was made possible by a grant from the National Science Foundation (NSF DGE-1544259). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] G. Christensen, A. Steinmetz, B. Alcorn, A. Bennett, D. Woods, and E. J. Emanuel, "The MOOC Phenomenon: Who Takes Massive Open Online Courses and Why?," *SSRN Electron. J.*, Nov. 2013.

[2] Maggie Sokolik, "'The Jungle' by Sinclair: BerkeleyX Book Club," *edX*, 2015. [Online]. Available: https://www.edx.org/course/jungle-sinclair-berkeleyx-book-club-uc-berkeleyx-colwri3-7x. [Accessed: 01-May-2018].

[3] Jolene Brackey, "Helping people with Alzheimer's disease - Online Course," *FutureLearn*, 2018. [Online]. Available: https://www.futurelearn.com/courses/creating-moments-of-joy. [Accessed: 01-May-2018].

[4] K. Wayne and R. Sedgewick, "Algorithms, Part I - Online Course," *Coursera*, 2018. [Online]. Available: https://www.coursera.org/learn/algorithms-part1. [Accessed: 01-May-2018].

[5] B. Boudouris and V. Agarkar, "The Physics of Electronic Polymers (PEP) - Online Course," *edX*, 2016. [Online]. Available: https://www.edx.org/course/physics-electronic-polymers-pep-purduex-nano600x. [Accessed: 01-May-2018].

[6] D. Engle, C. Mankoff, and J. Carbrey, "Coursera's Introductory Human Physiology Course: Factors that Characterize Successful Completion of a MOOC," *Int. Rev. Res. Open Distrib. Learn.*, vol. 16, no. 2, Apr.

2015.
[7] K.A. Douglas, B.P. Mihalec-Adkins, N. Hicks, N. Diefes-Dux, P. Bermel, and K. Madhavan. "Learners in advanced nanotechnology MOOCs: Understanding their intention and motivation." In *123rd Annual Conference of the American Society of Engineering Education*, pp. 26-29. 2016.
[8] Y. Jia, Z. Song, X. Bai, and W. Xu, “Towards Economic Models for MOOC Pricing Strategy Design,” Springer, Cham, 2017, pp. 387–398.
[9] ATD Research, *2017 State of the Industry Report*. Alexandria, VA, 2017.
[10] Sonwalkar, Nishikant. "The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I." In *MOOCs Forum*, vol. 1, no. P, pp. 22-29. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: Mary Ann Liebert, Inc., 2013.
[11] R. F. Kizilcec, C. Piech, and E. Schneider, “Deconstructing disengagement,” in *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK ’13*, 2013, p. 170.
[12] K.A. Douglas, P. Bermel, Md M. Alam, and K. Madhavan. "Big data characterization of learner behaviour in a highly technical MOOC engineering course." *Journal of Learning Analytics* 3, no. 3 (2016): 170-192.
[13] J. Reich, “Education research. Rebooting MOOC research.,” *Science*, vol. 347, no. 6217, pp. 34–5, Jan. 2015.
[14] J. Richardson and K. Swan, “Examining Social Presence in Online Courses in Relation to Students’ Perceived Learning and Satisfaction,” *Online Learn.*, vol. 7, no. 1, 2003.
[15] C. Baker, “The Impact of Instructor Immediacy and Presence for Online Student Affective Learning, Cognition, and Motivation,” *J. Educ. Online*, vol. 7, no. 1, 2010.
[16] W. Chen, C. G. Brinton, D. Cao, A. Mason-singh, C. Lu, and M. Chiang, “Early Detection Prediction of Learning Outcomes in Online Short-Courses via Learning Behaviors,” *IEEE Trans. Learn. Technol.*, pp. 1–1, 2018.
[17] J. Campbell, A. Gibbs, H. Najafi, and C. Severinski, “A Comparison of Learner Intent and Behaviour in Live and Archived MOOCs,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 15, no. 5, Nov. 2014.
[18] J. Bailey, C. Schneider, and T. Vander Ark, *Navigating The Digital Shift: Implementation Strategies for Blended and Online Learning*. Digital Learning Now!, 2013.
[19] P. Tompkins, “Late registration and student success in on-campus and online classes,” Old Dominion University, 2013.
[20] J. M. Lim, “Predicting successful completion using student delay indicators in undergraduate self-paced online courses,” *Distance Educ.*, vol. 37, no. 3, pp. 317–332, Sep. 2016.
[21] T. Van Hecke, “Power study of anova versus Kruskal-Wallis test,” *J. Stat. Manag. Syst.*, vol. 15, no. 2–3, pp. 241–247, May 2012.