Beyond the Means – Visualizing Learner Activity and Outcomes for Online Instructors

Mr. Taylor V. Williams, Purdue University-Main Campus, West Lafayette (College of Engineering)

Taylor Williams is a Ph.D. student in Purdue’s school of engineering education. He is currently on an academic leave from his role as an instructor of engineering at Harding University. While at Harding he taught undergraduate biomedical, computer, and first-year engineering. Taylor also spent time working in industry as a systems engineer. Taylor received his master's in biomedical engineering from Tufts University and his bachelor's in computer engineering and mathematics from Harding University. His primary research interest is in how to use machine learning in fully online and hybrid educational environments to understand students and improve their learning.

Dr. Kerrie A. Douglas, Purdue University-Main Campus, West Lafayette (College of Engineering)

Dr. Douglas is an Assistant Professor in the Purdue School of Engineering Education. Her research is focused on improving methods of assessment in large learning environments to foster high-quality learning opportunities. Additionally, she studies techniques to validate findings from machine-generated educational data.

Prof. Peter Bermel, Purdue University-Main Campus, West Lafayette (College of Engineering)

DR. PETER BERMEL is an assistant professor of Electrical and Computer Engineering at Purdue University. His research focuses on improving the performance of photovoltaic, thermophotovoltaic, and nonlinear systems using the principles of nanophotonics. Key enabling techniques for his work include electromagnetic and electronic theory, modeling, simulation, fabrication, and characterization.

Dr. Bermel is widely-published in both scientific peer-reviewed journals and publications geared towards the general public. His work, which has been cited over 5500 times, for an h-index value of 28, includes the following topics: * Understanding and optimizing the detailed mechanisms of light trapping in thin-film photovoltaics * Fabricating and characterizing 3D inverse opal photonic crystals made from silicon for photovoltaics, and comparing to theoretical predictions * Explaining key physical effects influencing selective thermal emitters in order to achieve high performance thermophotovoltaic systems

Hillary Elizabeth Merzdorf, Purdue University-Main Campus, West Lafayette (College of Engineering)

College of Engineering
Beyond the Means – Visualizing Learner Activity and Outcomes for Online Instructors

It is now seven years since The New York Times declared 2012 "the year of the MOOC" [1] when the idea of online education through open courses held great promise. MOOCs were anticipated to make advanced learning available to anyone with access to the internet and interest in the content they offered. While they have not disrupted education to the extent predicted, institutions have increasingly sought to integrate MOOCs into formal degree, certificate, and professional development programs [2]. Corporations have begun to utilize MOOC platforms to reach large workforces, and professional learners have turned to MOOCs to increase their on-the-job skills and continuing educational needs. Shifts in online learning from open and free access to formal acknowledgements (e.g., certifications) have in no way reduced the fundamental challenges of evaluation for MOOC providers. MOOCs require an investment of time and resources from stakeholder groups, thus, there remains a need to know how to interpret outcomes from a MOOC and level-set expectations for all vested parties.

Due to the wide range of subject material, international offerings, and low- to no-cost enrollment, MOOCs attract a highly diverse set of learners [3]. At the same time, instructors often do not know what to expect from their students, nor how to interpret the patterns they might see. Previous research has found that instructors are especially interested in the outcomes of the relatively few learners who fully utilize the course, but it is difficult to understand who those learners are and how to help them best [4]. Also, instructors are interested in whether the larger groups of learners achieve their intended learning goals [5]. However, without understanding more about their learners—specifically, what they hope to gain from the course and why they only utilize part of the course—instructors and course designers must primarily rely on their intuition and direct feedback to guide their decisions. Ultimately, instructors of these online courses have relatively little information to guide their responses to students or to perform continuous improvement in their courses.

MOOC learners engage with the course in many dissimilar ways with the result that instructors cannot reasonably assume that most learners will—or even intend to—fully utilize the course. Previous research has found patterns of learner course utilization behavior that hold across platforms and courses, yet there is little to help instructors know which learners they should focus on and which of the course outcomes are experienced by the target groups [6], [7], [8]. If instructors are to be able to provide their learners with the guidance required to meet their personal course goals, they must have a way to identify differences in course interactions among learners. When learner subgroups and interaction patterns are visible, instructors may act upon that information more readily.

This paper is intended to share the progress we have made in answering the question "In what ways do different types of learners interact differently with advanced STEM MOOCs?" We ultimately intend to establish benchmarks for this MOOC subcategory, because these graduate-level STEM MOOCs may attract a different population of learners than general MOOCs. If so, then presenting instructors with information about their course alongside benchmarks derived from comparable courses will be more informative and useful. Instructors can then begin to
evaluate their course against similar courses. Once we understand how these courses compare to each other, we can begin to develop and refine standards and best practices for developing and delivering these advanced STEM MOOCs. These MOOC standards will help instructors more effectively meet the needs and desires of the various learner types taking online courses.

**Literature review**

The diversity and size of enrollment in MOOCs make it impossible to understand learners individually as a traditional instructor would. Research into MOOC effectiveness and outcomes has largely been driven by aggregate results and generalizations about learners. Learners frequently report taking courses for lifelong learning and personal interest, as well as connecting with people who share their interest in a topic, and these goals are predictive of overall course engagement [9]. Factors such as in-course discussion with peers or instructors have also been studied as predicting learner performance [10]. Behavioral engagement predictors of academic performance also typically remain generalized, isolated from other metrics, looking at broad statistics of enrollment, assignment submission, or quiz completion without regard for subgroups [11]. While these methods are useful for comprehensively investigating the activity of learners, the ultimate goal of studying engagement is early identification of learners at risk and finding ways to help them continue [12]. As such, aggregate results are often challenging for instructors to apply for course improvement.

An informative approach to understanding learning is through modeling learner interactions with course material. Machine learning methods can predict engagement and retention based on course activity to provide interventions as early as possible [13]. However, analyzing MOOC data requires additional considerations regarding use and interpretation. Qu and Chen [14] emphasize the power of visual analytics for illustrating the activity of thousands of MOOC learners while recognizing that analytics serve a variety of user groups who may not be familiar with data interpretation. A key goal of applying learning analytics to inform pedagogical interventions is enabling the agency of learners through goal-setting and reflection [15]. To plan effective interventions, the instructor needs to know where learners are in the course and with what they are struggling. Having rich, contextualized behavioral data readily available enables instructors to make these decisions.

Personalization is an important aspect of online education generally and MOOCs specifically, as individual learners have unique motivations and goals. Bonk et al. [16] found that MOOC instructors are generally interested in ways to personalize their courses, but that their personalization tools and approaches are limited to automated grading and feedback, as well as feedback from peers or the learning system. There is a need for real-time personalization approaches for instructors that are straightforward and scalable, which give instructors options for immediate intervention. In the future, this personalization could be driven by artificial intelligence (AI) advancements based on detailed learning analytics. These AI advancements will help instructors tailor their courses and allow learners to practice skills and receive personalized feedback [17]. Personalization can become an achievable goal when instructors can easily and flexibly interact with course data, perhaps with guidance from virtual agents one day, but prepared today with the tools for visualizing and measuring learning.
Another goal for integrating MOOC analytics with instruction is the development of course benchmarks. In contrast with the standards and success measures of traditional postsecondary education, no such benchmarks exist for STEM MOOCs. Metrics of quality, value, and positive outcomes are the beginnings of benchmarks which serve as standards of comparison across MOOCs. These indicators primarily assess the structure, design, and instructor support within the course [18]. High-quality courses will foster interactions among learners and instructors, but instructors may underestimate their potential to shape learners’ experiences through course design [19]. Our goal of developing benchmarks depends on first understanding learner activity and how it relates to the overall goals of instructors, MOOC providers, and institutions.

To address this need to inform online instructors about their learners, tools are needed to visualize groups of learners rather than only showing just aggregate information about the whole course population. Creating benchmarks for each learner group across a single course category (e.g., advanced STEM MOOCs) rather than all types of MOOCs generally should establish more reasonable expectations about the level of learner interaction with each course. There are at least two potential benefits to this approach. First, instructor access to benchmarks will let them see what learner behaviors are typical and will enable them to compare their course to see if they are ahead or behind of expectations. This benchmarking will enable instructors to consider distinct groups of learners and to target their interventions in a meaningful way that prior data demonstrate a reasonable likelihood of success [20, 21]. Second, visualizing the data broken down by groups will allow us to evaluate variations in MOOC quality; particularly those resulting in above-average engagement for certain learner groups, independent of the specific learners taking a given course. This result will in turn help inform future course design, building expectations for learner behavior into the course structure.

Study design
To answer in what ways learners interact with advanced STEM MOOCs, we analyze learner usage patterns across nine advanced STEM MOOCs offered by nanoHUB, a National Science Foundation supported project [22]. In this paper, we present early findings based on an analysis of three of these courses. nanoHUB is an online platform dedicated to “computational nanotechnology research, education, and collaboration” [22] and partners with the MOOC hosting platform edX to deliver nanoHUB courses online.

Data
The individuals we include in this analysis are those we refer to as “live-mode learners.” Consistent with previous research [23], we define these live-mode learners as those individuals whose first recorded interaction with any course material occurred during the 14 days following the official course start date. The learners who are excluded from this analysis are “late-mode learners” (who began after the first two weeks but before the course officially ended) and “archived-mode learners” (who began the course after the official end of the course). We will examine the late and archive learners separately since [23] found meaningful differences between learners in these three time-based groups.

The structure of edX courses and their content are arranged hierarchically in a tree structure. In that tree structure, each piece of course content can have parent and child relationships with other
content items. For instance, a title page for a major unit in the course is known in edX's nomenclature as a chapter. All content items for a unit exist as children beneath the unit item. The other content types included in this paper are sequential, video, and problem. A sequential is a topic within the course, and these are also essentially the web pages to which a learner can navigate to watch a video, read the topic text, or post on the topic discussion board. Two of the content types which can live on a sequential page are videos and problems. A video is typically a lecture or some other piece of media content. A problem is an interactive multiple-choice question, typically part of a quiz or a unit exam. Figure 1 shows the general tree structure for an edX course. There are four additional content types shown in the figure which are not logged in the data set we analyzed: course, vertical, HTML, and discussion. Further details regarding all course content item types and their organization are available on edX’s documentation website.

Figure 1. (left) The hierarchical data structure of an edX course. (right) Examples of content names at each of the hierarchy levels that more readily capture the course characteristics for interpretation. Colors match those used for content elsewhere in the paper.

Methods

We analyzed the raw data collected from the edX courses in several steps. First, we considered the structure of our primary data file. This data file is a record of when each learner's first and last interaction with each course content item occurred. Using these course content access times, we can then filter the learner population down to only live-mode learners by removing learners whose first recorded event occurred outside the first 14 days of the course. We also recorded when each live-mode learner’s final interaction event occurred.

Since the primary goal of this effort is to see how different types of learners interact with the course, we must first categorize each learner. To assign each learner to a learner type category, we utilize a tool we previously developed that identifies learner clusters based on the number of content items each learner accessed [24, 25]. The clustering is performed using k-means clustering. When considering the recommended number of clusters across all nine advanced STEM MOOCs we have available, we found that the gap statistic method consistently recommended either four or five clusters. Since our next goal after the work presented here is to create benchmarks for these advanced STEM MOOCs, we opted to consistently use five learner clusters for all courses to facilitate cross-course learner comparisons. The nonparametric Mann-Whitney test confirmed that each of the five clusters is statistically distinct from each other cluster in all nine courses.

To help with the analysis of these learner clusters we developed multiple interactive visualizations in Tableau Desktop (version 2019.1).

Results

Results from three of the nine nanoHUB courses for which we have data are presented in the following pages: nano535x, Principles of Electronic Biosensors; nano540x, Nanophotonic Modeling; and nano600x, Physics of Polymers. These courses matriculated between 384 and 1218 live-mode learners, with the numbers generally increasing for longer courses. Further details about these courses are available in Table 1.

Table 1. Information about the courses included in the analysis. Live-mode learners are those individuals whose first recorded interaction event with course content occurred during the first two weeks following the course's official start date.

<table>
<thead>
<tr>
<th>Course</th>
<th>Length</th>
<th>Live-mode learner count</th>
</tr>
</thead>
<tbody>
<tr>
<td>nano540x: Nanophotonic Modeling</td>
<td>8 weeks (10/2/2016 – 11/30/2016)</td>
<td>704</td>
</tr>
<tr>
<td>nano600x: Physics of Polymers</td>
<td>12 weeks (5/1/2017 – 7/25/2017)</td>
<td>384</td>
</tr>
</tbody>
</table>

We performed clustering on the data set referenced above. Similar to prior work [24, 26, 27], we found five distinct clusters: (C1) fully-engaged learners who participate in all course content, (C2) consistently-engaged learners who participate in all course content through the majority of the course but dropout in the final course units, (C3) two-week engaged learners who engaged fully through the content intended for the first two weeks of the course, (C4) one-week engaged learners who access most of the course content intended for the first week, and (C5) sporadic learners who either only access selected items or leave after accessing the first few content items. When this clustering analysis was performed on all nine available courses, we found that the learner percentages in each cluster remained stable (see Figure 2).
Figure 2. Users cluster percentages remain reasonably consistent across all nine available nanoHUB advanced STEM courses. The colors are consistent with those used for clusters throughout the paper. Circle areas are proportional.

Several hundred content items typically compose each course, primarily consisting of problems, followed by webpages and videos (respectively 66%, 17%, and 14% of the logged content items in nano535x). The composition percentages for nano540x and nano600x closely align with nano535x.

Figure 3. Course content accessed by nano540x’s 704 learners, colored according to the content type. Each learner’s activity is shown on a separate row. The course content item numbers (x-axis) are a sequential ordering of the course content as specified in the course design. The items that each live-mode learner accessed with are marked. The learners are arranged along the y-axis in increasing number of content items accessed by that learner.

Figure 4 shows the number of still-active live-mode learners in nano540x as time progresses. As previously mentioned, live-mode learners are defined as those whose first interaction with course content occurs within the 14 days following the official start date. The left edge of the graph corresponds with the official start date of each course. The official end date of each course is shown with a black dotted line. Each seven days until the end date is shown with a dashed vertical white line (with every fourth week shown as a solid line).
Cluster C1—the top layer in nano540x’s timeline, seen in Figure 4—shows a group of learners who stopped interacting with the course almost precisely when the course ended. This behavior is in contrast with nano540x's C3 and C4 which each also began with about 100 learners. Although more of C3 and C4’s learners departed early in the course than did C1’s, many more of C3 and C4's learners continued to access the course well past the course's end. Considering that the learners in C1 were “fully engaged” with the course material, we observe that these learners appear to treat the course traditionally (that is, using the course as designed and only participating while the course is active). Cluster C4, however, might be using this MOOC more like reference material—only accessing a subset of the course material, but continuing to return for many months to refer to this information. Similar patterns were found in all three courses analyzed: nano535x, nano540x, and nano600x.

Figure 4. A timeline for nano540x’s active live-mode learners, colored according to the learners’ cluster (the most active learners are in cluster C1). Learners are dropped from the timeline after their last recorded interaction with any content item. The five individual clusters are shown in the inset graphs.

We also performed an analysis of each cluster’s course content access rates for each course. A representative set of these visualizations can be seen in Figure 5(a), the data were plotted using box-and-whisker plots, with boxes indicating the median and interquartile range (IQR) of accesses for each specific content type (and whiskers showing 1.5 times the IQR). Each cluster’s access rate was noticeably different: while highly-engaged learners typically accessed all course content at a high rate, sporadic learners accessed at much lower rates; in fact, many of these learners primarily accessed the table of contents, with relatively infrequent access of in-depth material like video lectures or exam problems. Furthermore, it was found that the three most highly-engaged clusters tended to access most assessments, with some slight drop-off in latter weeks, but the engagement with videos was not as robust. Similar patterns were found to hold across all three courses analyzed.
Figure 5. Learner behavior in nano540x shown in three figures per cluster (see C1 for labels). In (a) and (b), color indicates the course content type. Sub-figure (a) shows the percentage of each content type accessed by that cluster. Sub-figure (b) shows the content items each learner accessed—each user’s activity is shown on a separate row. Sub-figure (c) shows a timeline of still-active learners. Observations with each are included cluster. Written access percentages show the cluster’s mean values.
As seen in Figure 6, only a small fraction of the learners (around 5%) in each course ultimately achieved a passing grade. However, when examining only the learners in C1, we find that about 70% of these learners passed these three courses.

<table>
<thead>
<tr>
<th></th>
<th>nano535x</th>
<th>nano540x</th>
<th>nano600x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster C1 pass rate</td>
<td>73.0%</td>
<td>67.1%</td>
<td>64.0%</td>
</tr>
<tr>
<td>Cluster C2 pass rate</td>
<td>1.2%</td>
<td>7.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>OVERALL pass rate</td>
<td>6.1%</td>
<td>7.1%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

Figure 6. Most nanoHUB courses on edX require users to earn 70% of available points to pass. Users in C1 have about a 70% pass rate while overall pass rates are well under 10%. The area of each circle is proportional to the associated learner count, which is shown within the circle.

Discussion

The benefit of these visualizations broken down by cluster group is their potential to give instructors a more detailed view of their course’s usage. This view of their learners will allow those instructors to draw more precise and useful insights than if they only examined the combined learner population. These visualizations allow us to highlight how the different clusters of learners interact with the different types of course materials differently. These methods are potentially interesting and useful for instructors as they work towards better personalization and flexible presentation of content in their courses.

Regarding curriculum design, our approaches to visualizing learner outcomes could help instructors better predict where and when learners will be accessing types of materials, by knowing at what points within a course they begin to reduce their access. Interventions around instructor-learner interactions can be informed by our research, as they allow instructors to triangulate various learner activities and draw conclusions about their patterns of access and engagement. For current MOOCs, our research can help illuminate interventions that the data show will not work, so as not to waste time with interventions and course modifications already known to be ineffectual.

Limitations

One limitation of this research is that we used a proxy for engagement, namely when a learner accessed a particular course content item. Thus, we considered that a learner "engaged" with a given item if a record existed in the data. However, this access log will not be a good estimation of engagement for some learners; for example, a learner who quickly clicks through the entire
course content sequence will be indistinguishable from a learner who thoughtfully engages with, and possibly even revisits, every course content item.

Another limitation stems from so many of the total count of content items being problem items, resulting in the clustering analysis weighing problem access more heavily than it probably ought when distinguishing between groups. A related data issue on the edX platform is that all assessment problems that exist on a quiz or exam are simultaneously logged as having been accessed once any of them are accessed. This issue means that more assessment problems will have been recorded than likely were engaged with. The resulting categorization of learners into clusters may be less precise due to the higher weight given to the problems when modeling learner activity.

Future research
As rapid data visualization of learner outcomes and course behavior becomes more tightly integrated with MOOC instruction and intervention during a live mode offering, future research may investigate more detailed questions around the time of access. For example, more insight is needed into how long learners spend watching course videos and the order in which learners access the course content. Collecting additional details associated with the access patterns and time spent in the course could be valuable for developing responsive interventions for struggling learners in each category. Now that we have these interactive visualizations available for three courses, future research will include generating the visualizations for the six remaining courses we have available. We will then work on generalizing what is common among these courses, then using these findings to develop and refine the benchmarks discussed at the beginning of this paper. Additional future work may include examining case studies from each of the clusters. When conducting these case studies, additional information about the learners can be extracted from the pre- and post-course surveys completed by some learners. By focusing on these learners who complete these surveys, we could gain more understanding of the characteristics of learners at each level of engagement and provide instructors with additional insight.

Conclusions
This work sought to identify distinct access pattern groups for MOOC learners based on their level of activity. We used a previously published clustering technique to find distinct groups of learners in a variety of advanced STEM MOOC courses. Our preliminary results show that access patterns differ greatly among groups, but within groups, those patterns are more generalizable across courses. For example, it was found that learners in the three most highly-engaged clusters typically accessed most assessments but did not access video content as robustly. Also, it was found that the least engaged (so-called sporadic) learners typically spend most of their time accessing high-level pages like the table of contents, rather than in-depth information. It appears that these learner patterns are consistent between courses for each of the five cluster groups. These findings demonstrate the potential usefulness of considering the heterogeneous composition of the learner population within advanced STEM MOOCs. In future work, we will consider generalizability across a broader range of courses to help develop and refine benchmarks that can aid instructors in targeting interventions for MOOC learners.
Acknowledgments

This work was made possible by the U.S. National Science Foundation (NSF) (EEC #1227110 and PRIME #1544259). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

The authors would like to thank the reviewers who made this a much stronger paper.

References


