

# Massive Open Online Courses (MOOCs): Participant Activity, Demographics, and Satisfaction

Sara Shrader, Maryalice Wu, Dawn Owens, Kathleen Santa Ana  
*University of Illinois @ Urbana-Champaign*

## Abstract

This paper examines activity patterns, participant demographics, and levels of satisfaction in multiple MOOC offerings at the University of Illinois at Urbana-Champaign from August 2012–December 2013. Using the following guiding questions: “Who are MOOC participants, how do they participate, and were they able to get what they wanted out of the course?” we have uncovered unique patterns of engagement that correlate with certain demographic characteristics. Our analysis employed both qualitative and quantitative methods, and serves as a model for further studies seeking to uncover the significance of participant activity within MOOCs.

## Introduction

As enrollments soar and universities scramble to develop and deliver educational content to the masses, Massive Open Online Courses (MOOCs) have captured the attention and imagination of millions of people from around the globe. MOOCs have been designed so that anyone with an Internet connection can sign up for courses at little or no cost, thus eliminating traditional course-bound penalties for inactivity or incompleteness. The University of Illinois at Urbana-Champaign (Illinois) was swift in joining the MOOC movement, partnering with Coursera in the summer of 2012. Initially, Illinois offered six MOOCs in topics that ranged from an introductory level course in *Sustainability* to a highly technical graduate level course in computer science known as *VLSI CAD* (Very-Large-Scale Integration Computer-Aided Design). This whirlwind of activity has prompted various stakeholders in the field of higher education to confront many overarching questions about the nature of MOOCs in order to better understand the significance of this emergent educational phenomenon.

Thus far a number of researchers have pondered key questions concerning who MOOC participants are and how they engage within MOOC platforms (Breslow, Pritchard, DeBoer, Stump, Ho,

& Seaton, 2013; Christensen, Steinmetz, Alcorn, Bennett, Woods, & Emanuel, 2013; Jordan, 2014; Wulf, Blohm, Brenner, & Leimeister, 2014). Our aim is to expand upon this endeavor in two ways. First, by presenting our approach to analyzing data associated with the first six MOOCs offered at Illinois in order to contribute to the growing scholarly discussion of how better to understand MOOCs and those who take them; second, to highlight the various patterns of activity displayed by our MOOC participants. Our focus on activity patterns is not to predict levels of course completion or incompleteness, but instead to learn more about participant satisfaction based on differing activity patterns. Using clickstream and survey data, we correlate demographic characteristics and levels of satisfaction to five separate clusters of participant activity that emerged from our analysis.

## Literature Review

The MOOC phenomenon calls forth an opportunity and need to examine new modalities of education, as well as to utilize different methods for analyzing the enormous amounts of data associated with MOOCs. Given the vastly different parameters that situate MOOCs from traditional modes of education—a context that is indeed both massive and open—researchers are confronted with the daunting task of conceptualizing ways to study MOOCs that allows for research replication, but maintain a readiness to discover promising approaches to gathering and analyzing data. Complicating the study of MOOCs even further is the differentiated nature of MOOCs where factors such as course design and discipline-specific attributes necessitate distinctive course goals and learning objectives, thus making cross-comparisons of data difficult.

Moreover, the ease of registering for MOOCs negates traditional course-bound rules such as formal drops and withdrawals, allowing MOOC participants to come and go as they please, to complete assignments or not. At the time of this study, MOOCs were mostly free of cost, offering a measure of autonomy for participants distinct from that of traditional degree-seeking students, as they are truly free to learn as they see fit because there are minimal financial and educational repercussions. Since involvement in MOOCs is solely at the discretion of participants and, for the most part, without penalty, the critical evaluation of traditional educational metrics is necessary in order to better contextualize the study of MOOCs (DeBoer, Ho, Stump, & Breslow, 2014).

As such, early rounds of MOOC research have called into question the efficacy of MOOC data, based on traditional metrics such as total enrollment numbers in determining completion rates (Chafkin 2013; Marcus 2013; Parr 2013; Pretz 2014). While MOOCs often enjoy a high level of initial enrollment, research shows that on average fewer than 10% of participants earn course completion certificates (Breslow et al., 2013; Ho, Reich, Nesterko, Seaton, Mullaney, Waldo, & Chuang, 2014; Koller, Ng, Chuong, & Chen, 2013). Concerns about low completion have been met with considerable critique (Carey, 2013; DeBoer et al., 2013; Ho et al., 2013; Vu & Fadde, 2014; Kizilcec, Piech, & Schneider, 2013), confirming the insufficiency of using total enrollment figures in determining completion rates. In turn, some researchers have examined intent for taking courses, which according to Koller et al. (2013) varies greatly, thus drawing some important distinctions about who registers for MOOCs and why.

While some authors have highlighted the potential of MOOCs as an effective means of democratizing higher education by increasing the access of non-traditional participants (Lewin, 2012; Wulf et al., 2014), demographic research has shown that MOOC users comprise a fairly homogeneous population. Typically, MOOC participants are young, well-educated males, living in developed countries, and have obtained higher levels of formal education (Christensen et al., 2013). Efforts have been made by institutions to target MOOCs toward certain populations (Mangelsdorf, Russell, Jorn, & Morrill, 2015), while MOOC service providers like Coursera have developed on-demand platforms to accommodate more flexibility in course offerings (Larson, 2014).

Other available MOOC research thus far has focused on identifying patterns of participant behaviors for both descriptive and predictive purposes. In particular, Kizilcec et al. (2013) used a longitudinal approach to track individual MOOC participants in three computer science courses. The authors identified four types of engagement, including completing, auditing, disengaging, and sampling. Similarly, Breslow et al. (2013) utilized predictive modeling in order to determine which course activities led to greater course success as defined by total number of points earned and persistence within a course.

Given the state of MOOC research, our paper seeks to examine the following questions:

1. How do our MOOC participants engage with the course material?
2. What are the demographic profiles of our MOOC participants?
3. Were participants able to get what they wanted out of the course?

## Methods

In this paper we engaged in a retrospective analysis of the MOOCs offered at Illinois from August 2012–December 2013. Table 1 shows the differences in course length, course design, and course content.

Table 1: Summary of Course Length and Activities

Course		Duration	Activities	Enrollment
Introduction to Sustainability		8 weeks (3 offerings)	16 quizzes 46 video lectures 8 forum activities 1 final exam	91,325
Microeconomics Principles		8 weeks (4 offerings) 4 weeks (1 offering) 16 weeks (1 offering)	8 quizzes 89 video lectures 8 forum activities 8 project milestones	104,887
Introductory Chemistry – Part I	Organic	8 weeks	8 quizzes in Coursera 8 quizzes in an external tool 100 video lectures	30,854
Intermediate Chemistry – Part I	Organic	8 weeks	8 quizzes in Coursera 8 quizzes in an external tool 130 video lectures	14,434
VLSI CAD: Logic to Layout		10 weeks	8 quizzes 65 video lectures 4 programming activities 1 final exam	21,854
Heterogeneous Programming	Parallel	8 weeks	7 quizzes 46 video lectures 7 programming activities	36,908

Over the span of two years, we gathered data both from within and outside of the Coursera platform. External to the Coursera platform is survey data that we gathered by administering two surveys to our course participants. At the beginning of each course, we sent out a questionnaire to all enrollees with questions about demographics, reasons for enrolling in the course, and what participants hoped to get out of the course. At the end of the course, a second questionnaire was sent, asking participants about their experience and satisfaction with the course (we did not collect survey data for the first offering of

*Introduction to Sustainability or Heterogeneous Parallel Programming*). Coursera also collected demographic data on course participants, and we merged that data into our own dataset to fill any necessary gaps.

Along with survey data, we collected the clickstream and event data recorded by the Coursera platform. By far the most granular, the clickstream data includes every click, of every participant, within a course site. Examples of clickstream data include forum and wiki views, as well as video views. The clickstream data allows us to know specifically which videos the participants watched, how many times they watched videos, and whether the videos were downloaded or streamed. If the videos were streamed, we can assess when participants paused, stopped, and/or restarted the videos. Clickstream data also contains information about the time and date of each click, the type of device and browser that was used to access various pages in the course site, and provides the IP address of every participant. The second source of data, the event data, contains information about video lectures viewed, quizzes taken, assignments submitted, dates of enrollment, the content of forum posts, quiz scores, and certificates earned.

As mentioned earlier, MOOC enrollment numbers cannot be meaningfully used as a central point of analysis in courses where participants are not held financially or academically accountable. The need to eschew total enrollment numbers in favor of using subtler points of analysis is continually reinforced in our approach to understanding the data. Take for example the following data (Table 2) from our first offering of Microeconomic Principles:

Table 2: Microeconomics Principles Enrollment Summary

Course	Total enrollment	Enrolled by the last day of the course	Enrolled after the last day of the course	Enrolled but never active	Enrolled and active during course period	Received a Certificate of Completion
Microeconomics Principles	50,676	50,375	301	28,107	22,569	2,233

By using “total enrollments” (50,676) as the denominator for determining the percentage of participants who completed the course by earning a “certificate of completion” (2,233), the course completion rate would be a mere 4.4%. However, if we chose to use participants who were “active during the course period” (22,569) as our denominator, the completion rate doubles to a completion rate of 9.9%.

We can fine-tune our analysis even further by making some decisions about what counts as course “activity.” If we defined activity broadly—clicking one time in the course site during the official start and end dates of a course—our active participant number is 22,569. However, by adjusting our definition of activity to include more nuanced time spent within a course, our results change yet again. Table 3 differentiates between activity for single day, for more than one day, and activity based on sustained, weekly clicks within a course:

Table 3: Microeconomics Principles Enrollment by Different Levels of Activity

Course	Active during the course period	Active for only a single day	Active for more than one day	Active every week of the course
Microeconomics Principles	22,569	8,019	14,550	2,261

By simply using “active for more than one day” as our denominator, completion rates jump to 15.3%. Conversely, by using “active every week of the course” as our denominator, completion rates balloon to 98.8%. Consequently, in order to paint a more nuanced picture of MOOCs, we opted to employ the more conservative “active for more than one day” denominator as the foci of our analysis in hopes of encapsulating a broad array of MOOC activity.

Likewise, we defined activity as a multifaceted action that includes various combinations of typical online course behaviors such as watching lectures, taking quizzes, submitting assignments, and participating in forums. Our understanding of activity is limited to the internal and external Coursera data outlined above, though presumably course participants may be engaged in relevant educational activity outside of the Coursera platform, including in-person study groups and social media gatherings. Specifically, we employed the following combinations of participant activity:

- only watching lectures
- only taking quizzes
- only engaging in forums
- watching lectures and taking quizzes, but not engaging in forums
- watching lectures and engaging in forums, but not taking quizzes
- taking quizzes and engaging in forums, but not watching lectures
- displaying activity in all three areas: watching lectures, taking quizzes, and engaging in forums

As such, our rationale for defining a MOOC participant in this manner—as someone who displays activity (as we have defined it) within a course for *more than one day*—is significant because final data reports will vary dramatically (as noted above) depending on the chosen denominator used to define participation.

Along with the need to better contextualize total enrollment data, researchers also need to clarify how they treat those who drop out of a course. In a traditional course structure, individuals who formally withdraw from a course tend to be classified as such, thus creating a separate category of analysis in regards to retention statistics. Yet, the concept of dropping out within the context of MOOCs makes little sense given the flexible signup process. Therefore, due to the conceptual problems involved with course withdrawals, our data includes individuals who discontinued their activity while remaining technically enrolled in the course, as well as those who intentionally un-enrolled themselves from a course.

Finally, some limitations must be acknowledged. Although we had 100% reporting on participant activity via web server logs, we had low response rates on the survey portions of our research where we ascertain participant demographic characteristics and course satisfaction. The survey response rates ranged from 3.1% (*Intermediate Organic Chemistry*) to 8.7% (*Introduction to Sustainability*) with an overall response rate of 5.3%. The response rates for people who are “Active for more than one day” ranged from 11.4% (*Introductory Organic Chemistry*) to 20.1% (*Introduction to Sustainability*) with an overall response rate of 14.0%. A low response rate is only a problem if it creates non-response bias (i.e. non-respondents are like respondents on variables of interest). To help ameliorate this bias, we calculated and employed non-response weights that use available information on participant activity levels regarding the true composition of the surveyed population in order to adjust the results to compensate for survey non-respondents. (Mandell, 1974) After adjusting for non-response bias according to activity patterns, we get a more accurate representation of the demographic distribution of MOOC participants.

After gathering and processing the clickstream, event, and survey data from each of our courses, we merged the data in order to uncover different patterns of behavior exhibited by MOOC participants by accounting for as much variation as possible.

## Results

The following section illustrates the various activity patterns displayed by our course participants. Our analysis of the clickstream and event data reveals that participants engage with courses in traditional and non-traditional ways, with some individuals opting to participate in only one type of course activity, while others participate in multiple types of course activity.

### Traditional and Non-Traditional Course Activity

Traditional course activity often consists of content being delivered through a lecture format, followed by some form of assessment that attempts to measure learning outcomes through an accumulation of points. Typically, in *traditional online courses*, videos are used to deliver content while computer and human-graded tests and quizzes serve as a central conduit of course assessment.

Progressing through a traditional online course, therefore, requires that participants watch a series of videos and complete various assessments over time. The process is structured in such a way that achieving pre-determined, instructor-created goals becomes the pathway for success. As a general convention for most educational schemas, completion or non-completion of the assessment pathway becomes the trigger that determines success. It is then *the quality of completion*, usually determined by a grading scale, which distinguishes high achievers from their less successful peers.

In contrast, non-traditional course activity takes a different form, where the concept of a structured pathway is muddled, and participants determine their own goals for success. Depending on an individual's pre-determined goals, success can mean watching only a handful of videos in order to learn a single concept, or success can mean watching some videos and taking multiple assessments. Since the openness of MOOCs allows for individuals to engage how they see fit, non-traditional course engagement necessarily becomes a valid path for success. Measuring success, however, becomes difficult since traditional course metrics such as completion do not hold sway outside of a pre-determined assessment pathway. Moreover, understanding success as a quality metric becomes difficult because there is no grading scale or other measure by which to determine achievement.

**Participation patterns.** Using “active for more than one day” as our denominator, activity within a course varies by course discipline, as does the kind of activity displayed by participants. However, there are also some consistent trends. Overall, we found two patterns of participant behavior, which include “single-activity” participants and “multi-activity” participants. Single-activity denotes participation in only one of the activities found within a course, while multi-activity encompasses participation in more than one course activity. Figure 1 and Table 4 illustrate the diverse activity types displayed by our Coursera participants.

Combining all courses, a significant percentage of participants only watched video lectures, accounting for nearly 29% (28.9%) of activity in all courses, though there are some clear outliers. Half of all participants in *Intermediate Organic Chemistry* only watched video lectures (49.4%), while 20.6% of participants in *Introduction to Sustainability* only watched videos with no other activity. Conversely, those who only submitted quizzes account for a fraction of participants within all courses (1.4%), with the highest percentage found in *Intermediate Organic Chemistry* at a mere 3.0%. Not surprisingly, forum only participants were also quite small (2.6%), with *Introductory Organic Chemistry* having the highest rates of forum participation (5.1%), and *Microeconomic Principles* having the lowest rates (1.6%).

Figure 1: Activity Type by Course

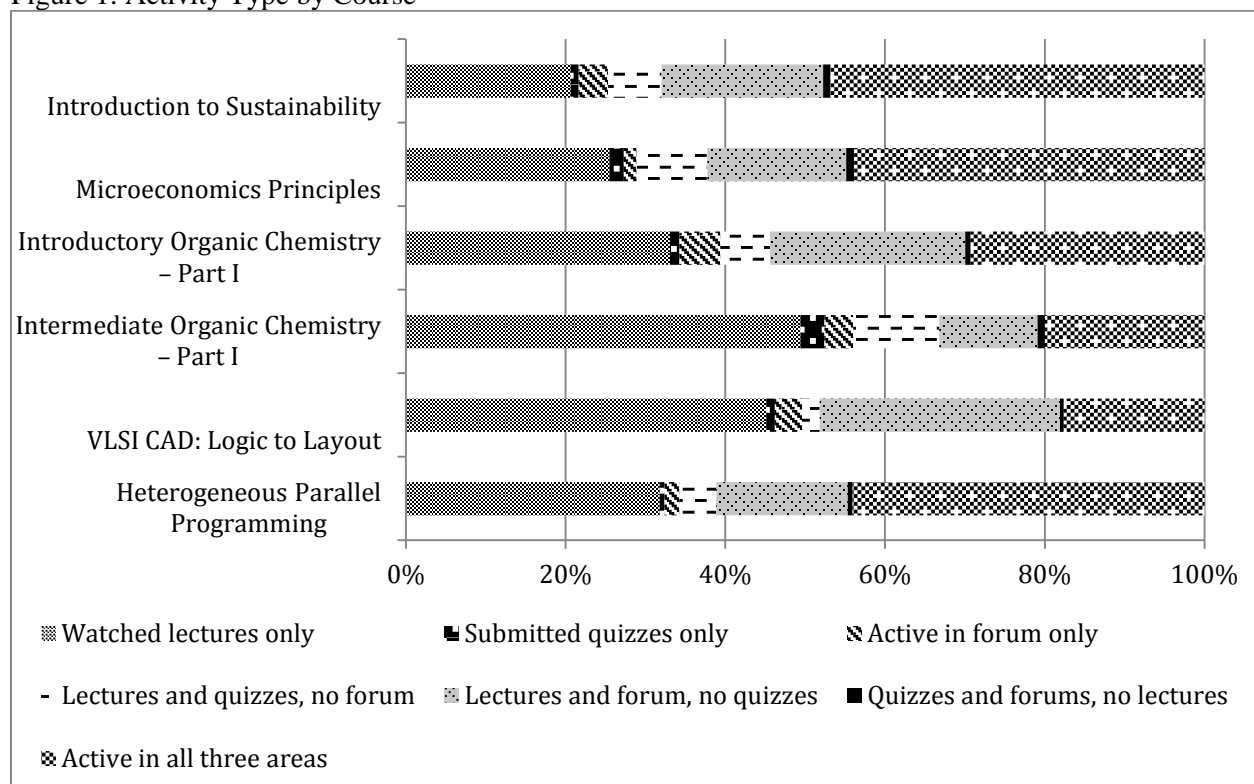


Table 4: Activity Type by Course

Course	Watched lectures only	Submitted quizzes only	Active in forum only	Lectures and quizzes, no forum	Lectures and forum, no quizzes	Quizzes and forums, no lectures	Active in all three areas
	%	%	%	%	%	%	%
Introduction to Sustainability	20.6	1.0	3.7	6.7	20.2	0.9	46.8
Microeconomics Principles	25.5	1.8	1.6	8.9	17.4	1.0	43.9
Introductory Organic Chemistry – Part I	33.0	1.2	5.1	6.2	24.4	0.7	29.2
Intermediate Organic Chemistry – Part I	49.4	3.0	3.6	10.8	12.3	0.9	20.0
VLSI CAD: Logic to Layout	45.1	1.1	3.4	2.2	30.1	0.5	17.6
Heterogeneous Parallel Programming	31.8	0.6	1.9	4.6	16.5	0.6	44.1
Overall	28.9	1.4	2.6	6.9	19.4	0.8	40.0

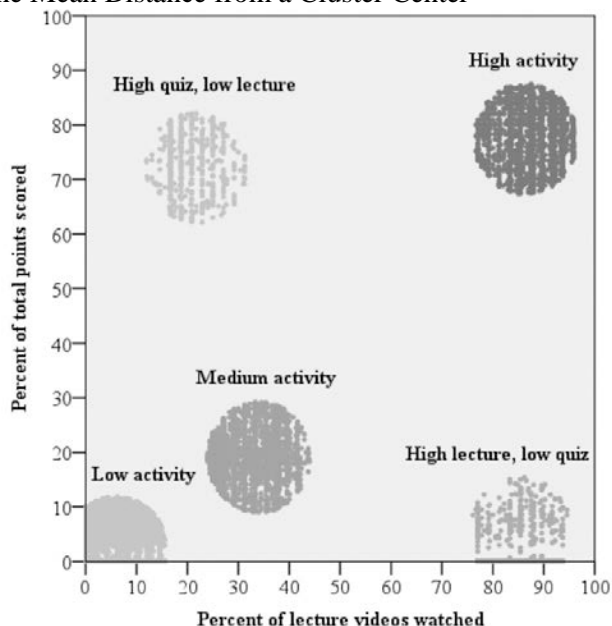
While more research is needed to understand why certain disciplines prompted stronger single-activity participation than others, it is clear that single-activity is a component of MOOC participation. This indicates that some participants did not use MOOCs as an all-encompassing, educative experience, but instead relied on discrete avenues for engaging with the course content. The openness of MOOCs necessarily invites this type of single-path participation, since participants are free to choose, in some respects, how they want to take a course. This type of participation closely mirrors the freedom of choice one engages in when using the Internet, where web users dictate how, when, and where they find and share information. In this way, single-activity participants used courses much like they would use the Internet by searching through a course in order to find a single activity of interest.

Engaging in multiple activities, however, accounts for higher percentages of participation for nearly all courses. On average for all courses, 40% of all participants engaged in all activity types (lectures, forums, quizzes). At 46.8%, participants in *Introduction to Sustainability* were the most multi-active, while *VLSI CAD: Logic to Layout* attracted the fewest multi-active participants (17.6%). “Dual-activity” participants, those who participated in a combination of two activity types, were most active by watching lectures and working in the forums and not taking quizzes (19.4%).

Due to the various combinations of single, dual, and multi-activity type, we applied *k*-means clustering analysis in order to better identify and understand patterns of participation. *K*-means clustering is an iterative technique used to group observations around the nearest means of two or more criterion variables. In our case we used “percent of videos watched” and “percent of total points scored” as criteria for clustering. Distance from the cluster means was calculated using simple Euclidean distance.

Using “active for more than one day” as our denominator, and modeling our two criterion variables of “percent of lecture videos watched” and “percent of total points scored” the following clusters highlight unique patterns of engagement among participants (forum participation was not included in this cluster analysis because it did not reveal any significant trends). These patterns suggest both traditional and non-traditional trajectories of course activity. It is important to note that our clustering model does not illustrate activity patterns on a time-scale; depicted in Figure 2 is the final outcome of participant activity at the end of a course.

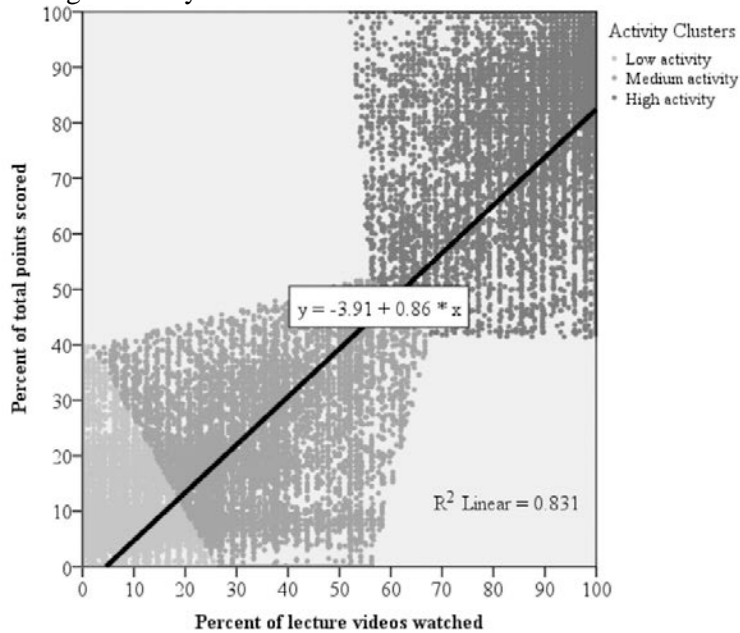
Figure 2: Illinois Coursera Participants with the Mean Distance from a Cluster Center



In particular, our analysis shows that the “high” “medium” and “low” clusters emerged as candidates for proceeding along a traditional course trajectory, albeit at quite different levels of engagement within a course. Figure 3, also using *k*-means clustering, demonstrates a strong linear relationship ( $R^2=0.831$ ) between those who watched videos and those who took quizzes, suggesting that these three groups (high, medium, and low activity) are actually one group on a single continuum.



Figure 3: Illinois Coursera Participants within the Low, Medium, and High Activity Clusters



For example, individuals in the “high activity” cluster demonstrated a strong level of engagement indicative of the traditional expectations that correspond to a traditional course setting. Participants who fell within this cluster watched a large percentage of videos (86% average), and earned a high percentage of quiz points (77% average). It is important to note, however, that engagement along a traditional path indicates little regarding the quality of such participation, and further research is needed to articulate the value of learning within a MOOC framework, relative to other

traditional and non-traditional modes of learning. However, by traditional standards, participants in this cluster demonstrated success since 75% earned a certificate of completion.

In contrast, the “low activity” cluster includes individuals who engaged in a small number of course activities. The limited participation within this cluster, however, makes it difficult to extract any meaningful analysis without having access to more information. Perhaps the most intriguing group on the continuum is the “medium activity” cluster, since these individuals persisted longer than the “low activity” cluster, but fell short of the “high activity” cluster. Overall, 16% of participants in this cluster engaged in 25% of the activities within a course, while 62% completed half of the activities in a course. For the most part, those in the “medium activity” cluster have yet to differentiate what type of course participant they want to become. Although they are poised to travel down the “high activity” path, they are also in a position to go the non-traditional route of the “high quiz, low lecture” participants. Or, they could splinter off into the opposite direction, becoming part of the “high lecture, low quiz” group.

On the other hand, the “high quiz, low lecture” participants as well as the “high lecture, low quiz” participants appeared to be on a non-traditional course trajectory. For the “high lecture, low quiz” group watching videos is more important than completing quizzes, which suggests that these participants are not concerned with earning points and completing the course in the traditional sense. Likewise, the “high quiz, low lecture” group also engaged in the course the way they wanted to, using assessment as the primary means to interact with the content. Overall, individuals who chose non-traditional paths for course participation eschewed an educational experience based on achieving the pre-determined goals laid out for them. Instead, these participants sought an individualized path for themselves in terms of navigating the course.

It is important to note, however, that 54% of participants in the “high quiz, low lecture” group earned certificates of completion in the course. Our courses were primarily designed (with some exceptions) to reward participants who achieved a certain number of quiz points. This mirrors a common, and arguably traditional, educational practice of utilizing test scores as a primary means to award educational credentials. Ironically, in approaching the course in a non-traditional way, the activity of the

“high quiz, low lecture” group reinforced an often critiqued practice of using tests as a primary means to incentivize learning.

**Demographic patterns.** In order to better understand the various profiles of participants who engaged with MOOCs through patterns of both traditional and non-traditional course activity, we ran individual logistic models for each of the five clusters with specific demographic characteristics pulled from our survey data and weighted by participant activity. Weights specific to each activity cluster within each course were constructed using the following formula:

$wt_n = (pp_n/ps_n)$  where:

$wt_n$  = non-response weight for cluster n

$pp_n$  = proportion of all participants in the course who were in activity cluster n

$ps_n$  = proportion of all survey respondents who were in activity cluster n

After examining sex, age, education level, and employment status, the demographic composition of each cluster revealed some surprising insights, as well as some expected trends.

The only demographic characteristic that remained equal across all activity clusters is sex, where female and male participants are equally likely to be represented in each group. This finding is significant because some MOOC research has focused on the gender gap in certain MOOCs, illustrating the difference between female and male enrollments (Christensen, et. al, 2013). This descriptive enrollment gap can be seen in some of our courses as well. For example, although our overall course enrollments were split between males at 55% and females at 45%, *Introduction to Sustainability* and *VLSI CAD* did not follow this trend. The *VLSI CAD* population is highly skewed toward males (87.1%), whereas *Sustainability* is the only course that enrolled a majority of female participants (61.1%). Highlighting descriptive enrollment trends without additional analysis, however, paints a narrow and somewhat distorted view of what is really happening in MOOCs. Although females and males enrolled in different courses at different rates, their likelihood of course engagement at various levels, once enrolled in a course, is equal. So although a gender gap may exist upon enrollment, that gap closes once participants begin to engage in course activities.

**Low activity.** Predictably, the “low activity” cluster included the highest percentage of all age groups, education levels, and employment types. The high percentage of participants in this cluster was not surprising given other studies that have confirmed pervasive levels of low activity within MOOCs (Ho et al., 2014; MOOCs @ Edinburgh, 2013). However, participants who are 24 and younger are twice as likely to be in the “low activity” group compared to all other ages. This indicated that younger people tend to do less than their more mature counterparts.

Further research is needed to determine the factors that cause younger participants to disengage in a more pronounced manner, although some hypotheses can be made. For example, one can hypothesize that because younger participants are closer to the educational experience they are not looking to complete the course in the traditional sense. Students may be simply exploring different academic options in a quest to learn more about their own educational and professional interests. Whatever the reasons may be, once researchers have a more advanced understanding about this “low activity” cluster, decisions can be made on whether and how to adjust courses in order to foster sustained participant engagement for a larger number of registrants.

**High activity.** In contrast, participants in the “high activity” cluster revealed some stark differences in terms of age from the “low activity” cluster. For example, participants between the ages of 26-30 are 1.5 times more likely to be part of the “high” cluster than those between the ages of 18-24. Moreover, participants who are 30 and older are approximately twice as likely to be in this cluster than the 18-24

year olds. At 60 years old, however, participants are 4 times as likely to belong in the “high activity” cluster group than the younger 18-24 year olds. This indicates that older adults are more likely to persist (and even complete in the traditional sense) in courses than younger participants.

One possibility for such a wide age gap in the “high activity” cluster may be due to employment status. When looking at survey respondents who identified as either employed, unemployed, student, or retired, those who reported to be students were half as likely to be in the “high activity” cluster. The other three employment categories—employed, unemployed, and retired—are equally likely to be in this cluster. Students may be less likely to be in this cluster because they are busily working toward their degrees, leaving little time to focus on a MOOC. Or, the subject matter may be beyond a student’s ability, leaving most students unable to keep up with the course material.

Moreover, participants with earned Masters or PhDs are approximately twice as likely to be in the “high activity” cluster, compared to those who have only completed high school or less. Presumably there are numerous reasons why individuals with advanced degrees are more likely to exhibit strong levels of engagement within a framework of traditional course activity. Perhaps their educational experiences have endowed them with the skills and cultural capital needed to persist in what is essentially a self-motivating environment. Or perhaps the sheer number of years that they attended school contributed to a trained mind regarding how one “takes a course.” In any case, there is something to be said for the way in which established educational credentials provide a means for MOOC participants to maintain a high level of activity and achievement atypical of the less credentialed.

**Medium activity.** As discussed earlier, the “medium activity” cluster may be the most puzzling group to understand since individuals in this cluster did not display a clear path of activity. There were no statistical differences in this group in terms of sex, age, or education. The only interesting statistical difference is that unemployed participants are 1.4 times as likely to be in this cluster than their employed counterparts. Further research on the psychology of unemployment may help to explain this statistical anomaly. In terms of our analysis, however, this “medium activity” cluster continues to remain a mystery since it is unclear who these participants are, and why they displayed an inconspicuous level of course activity.

**High lecture, low quiz.** In terms of age, this group included the most striking statistical likelihoods. Overall, participants who are 50 years and older are four times more likely to be in this group than those ages 18-24. What’s interesting to note is that starting at the age of 30, the likelihood of being in this group increases. This suggests that with increased age, individuals are more likely to prefer watching videos rather than taking quizzes or engaging with MOOCs through a traditional pattern. A similar pattern follows for educational level. Participants with a Masters and Bachelors are twice as likely to be in this cluster than those without a college degree. Moreover, participants who hold PhDs are over three times more likely to be in this cluster than those with lower educational credentials.

Finally, similar to the “high activity” cluster, the employment status of participants in the “high lecture, low quiz” group is composed mostly of non-students. In fact, students are half as likely to belong to this cluster, while retired participants are over twice as likely to be in this cluster. This reinforces a trend that seems to apply to older, well-educated, retired participants—not only is it highly likely that they will participate in courses through a non-traditional trajectory, but it is also highly likely that they will watch lectures without feeling a need to test their knowledge through predetermined assessments (such as quizzes).

**High quiz, low lecture.** Of all the clusters, the “high quiz, low lecture” group included the only equal distribution of all demographics—sex, age, education, and employment. Put another way, there is absolutely no difference regarding any of the four demographics in this cluster. This result was somewhat

unexpected since it would be reasonable to hypothesize that this cluster would be composed heavily of students. It is quite conceivable to imagine that students would use MOOCs as a means to complement their current on-campus courses by taking advantage of free assessment activities. According to our analysis, however, this does not appear to be the case. Additional research is needed to determine what type of participants are more inclined to fall within the “high quiz, low lecture” cluster.

**Participant satisfaction.** As mentioned earlier, we administered two surveys at the beginning and end of each course. In the first survey, we asked two open-ended questions: “*What are your reasons for taking this course? What do you hope to get out of it?*” The function of asking these open-ended questions was twofold. First, we wanted to better understand the intent of Coursera participants, but also we hoped to discover more nuanced information about our participants that we could not predict fully by merely asking close-ended questions. While devising our course surveys, MOOCs were in their infancy (arguably, they still are) and little was known about how MOOCs would fit within the established system of higher education. As such, our research team was careful to construct survey instruments that allowed our research process to be informed by the survey data that we received, rather than imposing our preconceived notions about MOOCs onto the research process.

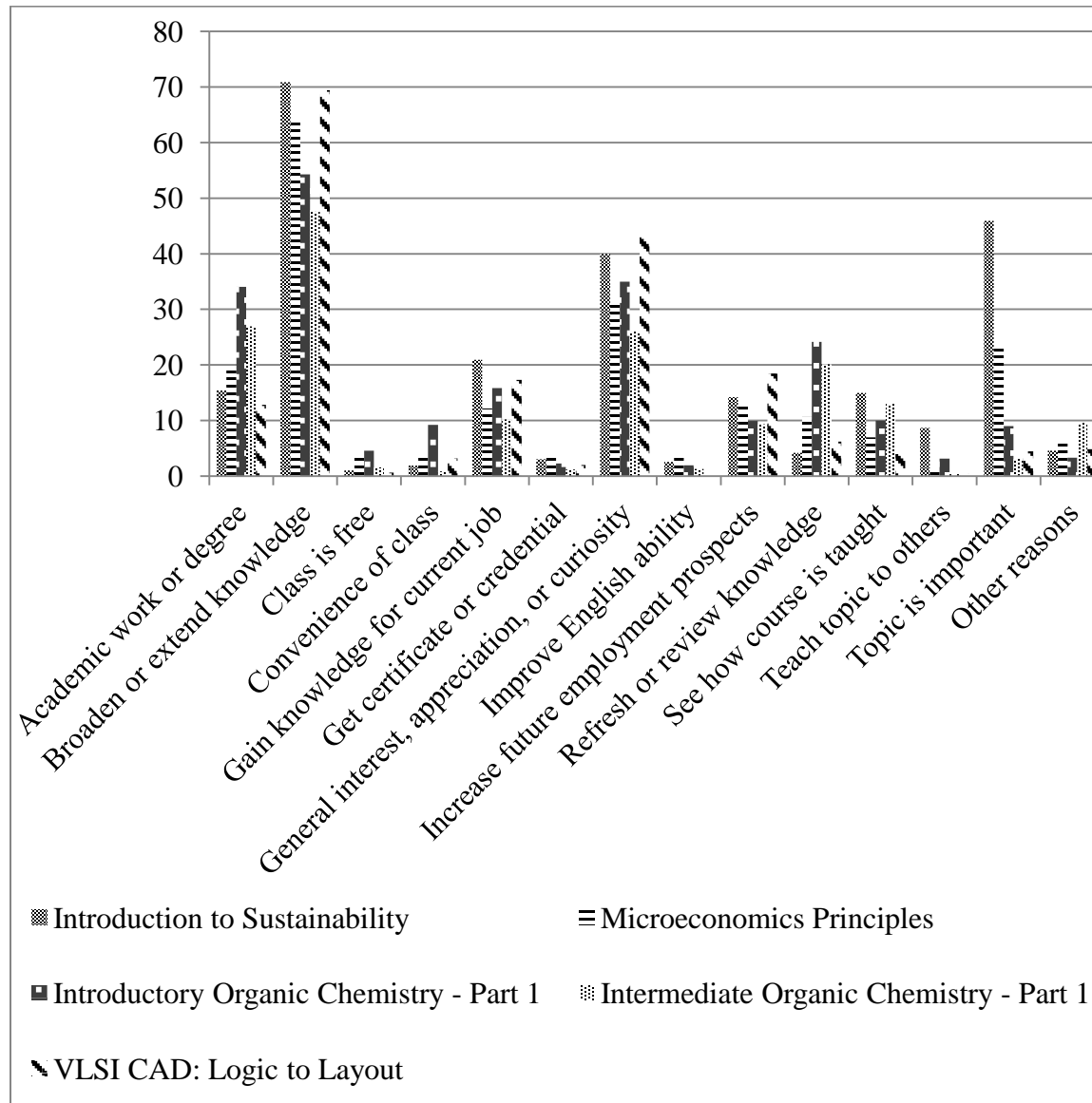
In order to analyze the open-ended survey questions, a coding process was established through thematic analysis resulting in the creation of a qualitative codebook. The codebook was then applied to 6,866 responses and coded by two people independently. All codes were then compared and analyzed for differences; all differences were reconciled between coders. This process resulted in 15,406 segments coded within the 6,866 responses. As a result, we uncovered some unique findings (Figure 4, Tables 5 & 6) with regard to the variety and novelty of responses offered by our course participants.

In particular, it was eye-opening to learn that some participants (albeit a small average, 2.8%) signed up for Coursera courses in order to improve their English ability. In designing our courses for this new, massive audience, we could not have imagined individuals would take courses for this reason. Further, it was surprising to learn that nearly 12% (11.7%) of participants signed-up for our MOOCs in order to see how the courses were taught. Presumably these participants were not interested in learning the content per se, but were simply curious about the buzz surrounding MOOCs. Another explanation is that they were interested in either designing their own MOOCs, or they hoped to get ideas for their own teaching.

Another important and surprising finding was that earning a certificate or credential was not an important reason for most individuals taking our Coursera courses (3.3%). Instead, the majority of participants offered reasons that are best described as falling under the umbrella of lifelong learning. In particular, a large percentage of survey respondents (35.6%) mentioned their general curiosity or interest in the topic, as well as their desire to broaden or extend their knowledge (65.6%).

In the second survey we asked the question: “*How much were you able to get what you wanted out of the course?*” Answer categories ranged on a five-point scale from: (1) *not at all* to (5) *to the fullest extent*. Through a comparison of means with an ANOVA test and a post hoc test of Games-Howell, three response patterns emerged. Not surprisingly, participants in the “low activity” and “medium activity” clusters reported the lowest levels of satisfaction at 3.04/5 and 3.08/5 respectively. These two groups were statistically different from the other three clusters, which reported a nearly one-point higher level of satisfaction. For example, participants in the “high activity” cluster reported significantly high levels of getting what they wanted out of the course (3.98/5). The “high quiz, low lecture” cluster was not far behind, reporting a level of satisfaction at 3.82/5. Participants in the “high lecture, low quiz” cluster indicated a satisfaction level of 3.44/5.

Figure 4: Reasons for Taking the Course



It is significant to note that all of the cluster groups reported above-average levels of getting what they wanted out of the course. Even the “low” and “medium” groups, which were somewhat of an enigma in terms of course activity and demographic composition, revealed a strong level of satisfaction with their course experience. The other three groups, which represented both traditional and non-traditional course trajectories, were highly satisfied with their course experience as well. Further, although none of the participants in the “high lecture, low quiz” cluster completed the course in the traditional sense (since receiving a certificate of completion in all of our MOOCs required that participants score a pre-determined number of points on course assessments), participants in this cluster felt that they were able to get what they wanted out of the course.

Table 5: Reasons for Taking the Course

Reason	Introduction to Sustainability		Microeconomics Principles		Introductory Organic Chemistry – Part I		Intermediate Organic Chemistry – Part I		VLSI CAD: Logic to Layout	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Academic work or degree	142	15.4	213	19.5	97	34.0	16	26.9	27	12.9
Broaden or extend knowledge	654	70.9	702	64.4	154	54.3	28	47.6	144	69.5
Class is free	9	1.0	38	3.5	13	4.6	1	1.6	1	0.7
Convenience of the class	18	1.9	37	3.4	26	9.2	1	0.9	7	3.2
Gain knowledge for current job	193	20.9	133	12.2	45	15.8	6	10.2	36	17.3
Get a certificate or credential	28	3.1	45	4.1	6	2.3	1	1.2	4	2.0
General interest, appreciation, or curiosity	373	40.4	334	30.6	101	35.7	16	26.2	90	43.3
Improve English ability	23	2.5	42	3.9	6	2.0	1	1.3	1	0.1
Increase employment prospects	131	14.2	138	12.7	28	10.0	6	9.4	38	18.5
Refresh or review knowledge	39	4.2	116	10.7	69	24.2	12	20.1	13	6.2
Teach the topic to others	80	8.7	11	1.0	9	3.2	1	0.4	1	0.4
See how course is taught	137	14.8	119	6.9	29	10.1	8	13.3	9	4.2
Topic is important	424	45.9	251	23.1	26	9.0	2	3.1	9	4.5
Other reasons	43	4.6	70	6.4	9	3.3	6	9.6	10	4.9

Table 6: Reasons for Taking the Course  
(Summary for All Courses)

Reason	Total	
	<i>n</i>	%
Academic work or degree	494	19.3
Broaden or extend knowledge	1682	65.6
Class is free	62	2.4
Convenience of the class	88	3.4
Gain knowledge for current job	413	16.1
Get a certificate or credential	84	3.3
General interest, appreciation, or curiosity	913	35.6
Improve English ability	72	2.8
Increase employment prospects	341	13.3
Refresh or review knowledge	248	9.7
See how course is taught	300	11.7
Teach the topic to others	102	4.0
Topic is important	712	27.8
Other reasons	138	5.4

## Discussion and Further Considerations

The differentiated nature and duration of participant activity within MOOCs signifies a paradigm shift in how researchers are able to appropriate traditional metrics for understanding educational data. The porous structure of MOOCs complicates long-standing and universally accepted definitions associated most often with traditional course activity, minimizing the usefulness of benchmarks like completion or retention. In part, the varied demographic profiles of MOOC participants, as well as their intentions for taking courses, contribute to the enigmatic development of MOOCs as a new educational phenomenon.

For these reasons, our research approach has been to utilize different data sources (clickstream, event, surveys) in order to highlight the nuances of MOOCs to account for as much variation as possible. We identified various demographic characteristics of participants, and correlated those characteristics with the activity patterns displayed within our MOOCs. As a result, we uncovered activity patterns that reveal both traditional and non-traditional engagement within a course. The demographic characteristics of participants who engaged in the different course pathways varied, but overall satisfaction was consistently high regardless of one's chosen path.

The implications of our findings are important for a number of reasons. First, our results confirm the need to reconceptualize certain educational variables with regards to MOOCs. As discussed earlier, research has already indicated that using traditional educational metrics for understanding MOOCs is a fruitless endeavor (DeBoer et al., 2014). MOOCs are continually evolving as distinct organisms that do not necessarily fit within the archetype of traditional education. As such, the uncritical application of terms such as completion does little to further our understanding of MOOCs. According to our analysis, non-completers are not necessarily unsuccessful, nor are they necessarily non-active participants. While the traditional pathway may be preferred and even encouraged by a MOOC instructor, the non-traditional pathways elicit merit since these participants are reporting high levels of satisfaction.

In contrast, however, there is also a need to reaffirm some existing conceptions of traditional education with regard to MOOCs. Although MOOCs function within a space that is not fully aligned with traditional educational schemas, our findings show that nonetheless many participants choose to engage in MOOCs in a traditional manner. The “high activity” cluster clearly illustrates this phenomenon. Likewise, the “medium activity” cluster indicates that many MOOC participants instinctively attempt courses along a traditional trajectory. This instinct to “take a course” in a traditional way is, for many individuals, a natural pathway for engaging in sustained learning.

Part of the value in MOOC research, then, lies in further articulating notions of success in an online environment that continues, in some respects, to mirror components of a traditional online education, but has yet to fully develop a distinct identity of its own. MOOCs have rapidly evolved to meet the institutional needs of higher education as well as the populations served by these institutions. For many MOOC participants, earning a certificate of completion does not necessarily align with their intentions for taking a course, and it is important to further understand those specific MOOC populations who have redefined the course experience to fit their needs.

In rethinking MOOCs, new questions about the purpose and value of MOOCs as an educational model must be examined. For example, what utility should MOOCs serve for those single-activity individuals who are only concerned with watching videos or taking quizzes? If institutions view the value of MOOCs as updated versions of extension programs, or as an effective means to market the educational brand of an institution, then what should be done in order to meet these unique needs? On the other hand, if the goal is to provide pathways for those who have historically been marginalized from receiving a quality education, then what type of educational experience should be constructed? In asking (and attempting to answer) these types of questions, we need to be cognizant of whether or not these queries are mutually exclusive, and thus possibly contributing to a false narrative about MOOCs. In better understanding not only the MOOC phenomenon, but also continually examining the participant experience, researchers can effectively shape MOOC policy and practice as institutional goals and values continue to take shape.

## Acknowledgements

We would like to thank Deanna Raineri, Rashid Robinson, Jane Blanken-Webb, Jake Astin, and Julian Martinez-Moreno for their contributions in the conceptualization of this paper. We’d also like to thank Jose Vazquez Cognet, Brian Ross, and Jason Mock who read earlier drafts, as well as the anonymous reviewers of *Online Learning* for their comments and recommendations.



## References

- Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A., & Seaton, D., (2013). Studying Learning in the Worldwide Classroom: Research into edX's First MOOC. *Research & Practice in Assessment*, 8, 13-25.
- Carey, K. (2013, December 12). Pay no attention to supposedly low MOOC completion rates [web log post]. Retrieved from <http://www.edcentral.org/pay-attention-supposedly-low-mooc-completion-rates>
- Chafkin, M. (December 2013/January 2014). Udacity's Sebastian Thrun, Godfather of free online education changes course. *Fast Company*, 181. Retrieved from <http://www.fastcompany.com>
- Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. (2013). The MOOC phenomenon: Who takes massive open online courses and why? *Social Science Research Network*, Retrieved from <http://dx.doi.org/10.2139/ssrn.2350964>
- DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing "Course": Reconceptualizing Educational Variables for Massive Open Online Courses. *Educational Researcher*, 20(10), 1-11.
- Ho, A. D., Reich, B. J. F., Nesterko, S. O., Seaton, D. T., Mullaney, T. P., Waldo, J. H., & Chuang, I. (2014). HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013. *Social Science Research Network*, Retrieved from <http://dx.doi.org/10.2139/ssrn.2381263>
- Jordan, K. (2014). Initial Trends in Enrollment and Completion of Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, 15(1), 133-160.
- Kizilcec, R., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *LAK '13: Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium*, 170-179. doi 10.1145/2460296.2460330
- Koller, D., Ng, A., Do, C., & Chen, Z. (2013, June). Retention and Intention in Massive Open Online Courses: In Depth. *Educause Review*. Retrieved from <http://www.educause.edu>
- Larson, S. (2014, September). Online Education Gets Fast-Tracked with Coursera Classes On-Demand. *readwrite*. Retrieved from <http://readwrite.com>
- Lewin, T. (2012, March 24). Instruction for masses knocks down campus walls. *The New York Times*. Retrieved from <http://www.nytimes.com>
- Mandell, L. (1974). When to weight: Determining nonresponse bias in survey data. *The Public Opinion Quarterly*, 38(2), 247-252.
- Mangelsdorf, S., Russell, J., Jorn, L., & Morrill, J. (2015, March). MOOCs for Wisconsin and the World. *Educause Review*. Retrieved from: <http://www.educause.edu>

Marcus, J. (2013, September). All hail MOOCs! Just don't ask if they actually work. *Time*. Retrieved from <http://www.time.com>

MOOCs @ Edinburgh. (2013). *MOOCs @ Edinburgh Report #1*. Edinburgh, Scotland: University of Edinburgh. Retrieved from <http://hdl.handle.net/1842/6683>

Parr, C. (2013, May 9). Mooc completion rates 'below 7%'. *The Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk>

Pretz, K. (2014, 3 February). Low completion rates for MOOCs [web log post]. Retrieved from <http://theinstitute.ieee.org/ieee-roundup/opinions/ieee-roundup/low-completion-rates-for-moocs>

Vu, P., & Fadde, P. (2014). Rings of engagement: A model for MOOC enrollment. *Journal of Global Literacies, Technologies, and Emerging Pedagogies*, 2(3), 240-247.

Wulf, J., Blohm, I., Brenner, W. & Leimeister, J.M. (2014). Massive open online courses. *Business Information System and Engineering* 6(2), 111-114. doi: 10.1007/s12599-014-0313-9