Behaviour & Information Technology

ISSN: 0144-929X (Print) 1362-3001 (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20

Foreseeing the endgame: who are the students who take the final exam at the beginning of a MOOC?

Chen Chen, Gerhard Sonnert, Philip M. Sadler & David J. Malan

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2019.1711452

Foreseeing the endgame: who are the students who take the final exam at the beginning of a MOOC?

Chen Chen [a], Gerhard Sonnert[a], Philip M. Sadler[a] and David J. Malan[b]

[a]Harvard Smithsonian Center for Astrophysics, Science Education Department, Cambridge, MA, USA; [b]Harvard John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA, USA

ABSTRACT

Massive open online courses (MOOCs) show highly irregular participation behaviour among users. In this study, using data from Computer Science 50x of HarvardX, we investigated one extreme, yet common strategy to foresee the endgame: taking the final problem set at the beginning of the course. We found such a strategy to be the only dominant trajectory alternative to following the sequence prescribed by the syllabus. Whereas all students who took and passed the final problem set at the beginning of the course subsequently completed the course, those who took and failed the final problem set at the beginning of the course finished the fewest number of milestones, even fewer than those who never attempted the final problem set. Moreover, students with a lower prior programming proficiency were more likely than better prepared students both to take the final problem set early and to fail it. This study revealed the disconcerting phenomenon that many students dropped out of a MOOC because, apparently, their confidence was crushed even before they learned any course content. The study suggests that future MOOC practices and policies should offer informative and constructive syllabi to accommodate students' need for previewing the endgame.

ARTICLE HISTORY
Received 13 March 2019
Accepted 30 December 2019

KEYWORDS
Behaviour modelling; distance learning; e-learning; learner – centred design; sequential pattern; visualisation

1. Introduction

Massive open online courses (MOOCs) have been expanding rapidly in terms of the number of courses provided, number of universities collaborating in them, and number of students enrolling (Shah 2018). By providing high quality instructional material from prestigious institutions, with a low or no fee and low barriers to admission, MOOCs have attracted learners with diverse backgrounds (Koutropoulos et al. 2012; Sharples et al. 2016). A great number of MOOC practitioners and researchers have anticipated that MOOCs would revolutionise and democratise higher education (Belanger and Thornton 2013; Haggard 2013; Jacobs 2013; Rice 2013) in the service of underprivileged populations (Dillahunt, Wang, and Teasley 2014). However, high dropout rates (Jordan 2015; Rovai 2003) and irregular learning paces and trajectories (Fini 2009; Maldonado-Mahauad et al. 2018) have dimmed the prospects of MOOCs in the eyes of some critics (Pope 2014; Zemsky 2014). This study investigates one of the extremely irregular MOOC learning trajectories–namely, to jump to the final problem set at the initial stage of the course – and its effect on MOOC completion.

1.1. Typologies of learning trajectory

Most MOOC research coarsely dichotomised students into those who completed the MOOC and those who did not and then predicted full completion as the benchmark of success (He et al. 2015; Jiang et al. 2014; Kloft et al. 2014; Li, Wang, and Tan 2018; Peng and Aggarwal 2015). Some research considered the extent of completion, using the users' performance, active duration, and dropout timestamps (Greene, Oswald, and Pomerantz 2015; Wang et al. 2018; Wen, Yang, and Rosé 2014; Yang et al. 2013). Such studies often adopted the technique of survival analysis, which is suitable to answer questions about the engagement and persistence of the students. Nevertheless, as Kizilcec, Piech, and Schneider (2013, 1) argued, the 'monolithic view of so-called "noncompleters" obscured the many reasons that a learner might disengage from a MOOC'. By identifying a small, yet meaningful set of patterns of engagement and disengagement, Kizilcec, Piech, and Schneider (2013) demonstrated four prototypical trajectories of engagement (completing, auditing, disengaging, and sampling) based on participants' interaction with video lectures and assessments. Using similar methods, DeBoer et al. (2014) showed the course trajectories (the number, order, and timing of each unit)

CONTACT Chen Chen chen.chen@cfa.harvard.edu D306, 60 Garden St., Cambridge, MA 02138, USA

among MOOCs users to be strikingly asynchronous. Coffrin et al. (2014) further visualised the transition between course milestones and compared two MOOC curriculum structures: one course, which had a rolling deadline (similar to traditional courses) and constrained the students to follow the course sequence, and another course, which had an open curriculum structure (similar to a gigantic dropbox) and enabled students to choose their own learning trajectories.

1.2. The double-edged sword of openness

The majority of MOOCs have adopted the open curriculum structure. It is easy for MOOC providers to manage and it gives students a great degree of freedom to learn or sample according to their own preferences. It has been shown that students in open curriculum structures are more likely to transit forward and backward and make full use of course materials, compared with students in a sequential structure (Coffrin et al. 2014). Furthermore, it has been found that the perception of autonomy in learning decreases the perceived boredom and increases enjoyment (Buhr, Daniels, and Goegan 2019). However, Mukala, Buijs, and Van Der Aalst (2015) found that students who follow a regular sequence were more likely to achieve higher grades. As an irregular learning path is often associated with poor time management and lower motivation, scholars have recently proposed to enforce regularity by restricting the visibility of later chapters until learners have completed the early chapters (Kim et al. 2017; Zheng et al. 2015).

When learners are given the freedom to design their own learning progression, that progression reflects not only self-regulation or motivation, but also the learners' learning or course-taking strategy. One common strategy that students adopt is to access a preview of the endgame, i.e. to skip to the exams or later chapters before coming back to the earlier sessions and quizzes (Coffrin et al. 2014; Guo and Reinecke 2014). From a human-centered design perspective, the design of an interactive system should capitalise on the users' or potential users' needs and address their needs at every stage of the design process (Kotamraju and van der Geest 2012; Kujala 2008). Thus, it is necessary to closely examine learners' preview behaviour to inform MOOC design.

1.3. Theoretical framework

1.3.1. Learning progression

Traditionally, learning progression theories have suggested that learning should be sequenced in a hierarchical progression from elementary to complete concepts (Wilson and Cole 1991), because such a progression would keep the transitions smooth and the cognitive load low (Larkin and Chabay 1989; Steinberg 2008). In a modification of a strict sequential progression theory, others have proposed, and shown the necessity, to present a complete and global overview of the end product (or endgame) before zooming into the elementary units (Reigeluth 1999), which gives meaning to the elementary units and prevents students from dwelling on disconnected or simplified contexts (Chen, Schneps, and Sonnert 2016; Ingham and Gilbert 1991; Muller, Sharma, and Reimann 2008).

Regular classroom pedagogy can be viewed as an instantiation of the hierarchical progression, with the instructor enforcing the learning sequence he or she deems the most appropriate. In a typical MOOC setting, by contrast, students are afforded the freedom to deviate from the designed progression. By previewing the global structure and context of the course, learners may develop a deeper understanding and appreciation of elementary units.

1.3.2. Expectancy-value theory

Students' choice of a learning sequence, particularly the preview behaviour, can be interpreted through the lens of expectancy-value theory. Expectancy-value theory was originally developed by Atkinson (1957) to model a person's achievement motivation as a function of expectancy of success and task value. In Atkinson's definition, expectancy was the proportion of individuals who had succeeded at the task in the past. In the following modifications of the theory, other scholars noted that past success in the task was primarily a reflection of the task difficulty, and task difficulty became one of the major factors that was thought to influence a learner's expectancy of success. Other major factors driving expectancy include self-concept (general belief about one's own competence) and self-efficacy (domain-specific belief about one's own competence) (e.g. Eccles and Wigfield 1995).

Eccles et al. (1983) formally introduced the expectancy-value theory to the study of education. In our context, expectancy-value theory (Wigfield and Eccles 2000) would conceptualise student behaviour as driven by two main factors: (a) the students' estimate of the probability that a specific behaviour will be successful (expectancy), which was influenced by the combination of perceived task difficulty, self-concept and self-efficacy; and (b) the value that students assign to this behaviour (value). While expectancy-value theory distinguishes four subcategories of value, the two most salient subcategories in the MOOC context are intrinsic value (interest in, and enjoyment of, the MOOC participation in itself) and utility value (the usefulness of the MOOC participation for external purposes). In addition to the motivating values, a third important subcategory of value is cost. Cost is the time to be invested and obstacles to be overcome to achieve the desired goals (in the MOOC context, the monetary cost is extremely low).

The intrinsic and utility subcategories of value in expectancy-value theory dovetail nicely with self-determination theory, one of the most frequently cited learning motivation theories in the MOOC literature. Self-determination theory also posits that learners can be motivated by external and intrinsic values (Ryan and Deci 2000).

Previewing the endgame creates crucial information in terms of expectancy. A preview of the course content may greatly help a learner to assess the course difficulty, his/her existing competency in the subject area, and to estimate the learning curve and the chances of mastering the skills (expectancy of success). In terms of value, it may also help the learner to understand better the content to be learned and to estimate the time and psychological cost (e.g. the amount of mental effort and frustration) for completing the course and thus to calculate if the completion of the course and mastery of the skills would be worth the efforts and pains necessary (De Barba, Kennedy, and Ainley 2016; Wigfield and Cambria 2010). Passing the final pset is a very strong indication that the student can master all psets and succeed in obtaining a course certificate. Failing the final pset suggests that a substantial amount of work and effort would be needed, and that success will be uncertain. The feedback from taking the final pset serves to help the learner estimate their expectancy (task difficulty, self-efficacy) and the task value (content to learn and cost). Obtaining such a robust expectancy and value estimate may then influence the learner's course persistence behaviour.

1.4. Hypothesis

Based on the above theoretical discussion, we hypothesise:

> H1. Learners who have a weaker background knowledge of (or preparedness in) the course content have a stronger need for self-assessment; thus, they are more likely to adopt the endgame previewing strategy.

> H2. Learners who jump to the final problem set at the beginning of the course have higher chances of dropout if they fail the test, compared with those who jump to the final problem set and pass the test (who have higher chances of completing the course).

The MOOC setting is a strategic research site for testing the predictions of expectancy-value theory because MOOCs have (a) large samples with the opportunity to engage in irregular learning trajectories, including the preview behaviour, and (b) low stakes in exams and low cost for dropout so that dropout events can be easily observed.

2. Data and analysis

2.1. Sample

For this study, we used data about students' characteristics, activities, and performance in the MOOC CH50x taught in 2014 on the HarvardX platform, a member of the EdX collaboration. The course contained eight units, each unit contained one problem set (hereafter, the *pset*). Each time a student submitted a test, he/she was considered to have survived one milestone (regardless if the student passed or failed the pset, because we are interested in whether a student was active by a milestone). Students were allowed to freely choose the order of tests; they could, of course, also drop out of the class any time. In this analysis, we broadly defined course completion as the submission of any seven of the eight pset tests, regardless of the final pset, because meeting such a criterion would indicate that a student had actively engaged with most of the course content. The final pset required participants to complete a JavaScript template so that, for a given city name or zipcode, a corresponding marker would be pinned upon a Google Map. The marker would have to be expandable to present the most notable information for its corresponding area (e.g. historic landmarks). For anyone who is new to computer programming, it would be impossible to pass this pset without learning prior chapters in the course.

The number of individuals pre-registering for CS50x on HarvardX was 6,143,535; however, only 28,350 of those who pre-registered came back to the course, and 20,134 of these finished the pre-survey, which was a prerequisite for gaining access to the course material. Among those who finished the pre-survey, only 9899 participated in any course work. In this study, we considered these as formal enrollees and applied statistical analysis only to these. We obtained the electronic file providing information about the participants' survey, pre-test, number of days active in the course, as well as problem set submission and outcome from the MOOC instructor.

2.2. Data

The dependent variable was the type of milestone transition trajectory that the participants adopted. For each pset (which we regarded as milestones in the course sequence), we knew whether it was submitted (submission; yes = 1; no = 0), whether it was passed (outcome; yes = 1; no = 0), and the sequence rank in which the participant submitted the particular pset (pset-submission-ranking; starting from 1 for the pset submitted first). As will be discussed in detail in the Results section, we created four categories of trajectories: (a) *no-final*: those who did not submit the final pset at all (final-pset-submission = 0); (b) *final-early-fail*: those who submitted the final pset at the beginning of the course before they submitted any other pset and failed the final pset (final-pset-submission = 1, final-pset-submission-ranking = 1, final-pset-outcome = 0); (c) *final-early-pass*: those who submitted the final pset at the beginning of the course before they submitted any other pset and passed the final pset (final-pset-submission = 1, final-pset-submission-ranking = 1, final-pset-outcome = 1); and (d) *final-later*: those who submitted the final pset, but not as the very first submission (final-pset-submission = 1, final-pset-submission-ranking ≠ 1).

Predictors included in this model were: gender, age, education, current school status, importance of certification, interests in participating in forum, English fluency, number of MOOCs completed previously, availability of help from others, and prior programming proficiency. Table 1 shows the details of the variables.

In addition to the item that asked participants to self-report prior programming proficiency, we used a pre-computational thinking test to measure participants' computer programming preparedness. We drew on several types and sources of questions to create this pre-test. From the University of Kent Computer Programming Aptitude Test (https://www.kent.ac.uk/ces/tests/computer-test.html), we took questions on logical thinking, pattern recognition, and ability to follow complex procedures, with the authors' kind permission. From Tukiainen and Mönkkönen (2002), we adapted questions targeting mathematical and logical reasoning and pattern recognition. From sample AP Computer Science Exam questions released by the College Board, we adapted questions on programming in Java. Inspired by the American Computer Science League (ACLS) contests, we also adapted questions on calculating the values of recursive functions. In the case of questions adapted from Tukiainen and Mönkkönen (2002), the AP Computer Science Exam, and the ACLS, we modified the numerical values, item format (all our questions were multiple-choice), or programming language. In this way, we generated a preliminary pre-test of 31 questions and evaluated it by administering it to 911 Amazon Mechanical Turk (AMT) participants. Based on classical test theory and item response theory analyses, we identified the top 11 questions, which explained 83.8% of the variance in the total pre-test scores. We used these 11 questions as the pre-test given to CS50x students. The mathematical reasoning, pattern recognition, and following complex procedures questions from the University of Kent Computer Programming Aptitude Test and those based on Tukiainen and Mönkkönen (2002) were most predictive and hence heavily represented in the pre-test (6 items from Kent; 4 items adapted from Tukiainen and Mönkkönen [2002]). One item was a modified AP Computer Science Exam question.

2.3. Analysis

We first carried out a descriptive analysis, focusing on the subgroups that submitted the final pset at different milestones during the course. Second, because different participants chose different orders of psets, we visualised the course trajectories to show the most common strategies that the participants adopted. Lastly, we applied multinomial logistic regression to predict final pset status using a list of variables.

3. Results

3.1. Pre-survey variables

Among the analytic sample of 9899 participants, 82% were male; 18% were female. The average age was 29.44

Table 1. Wording and coding for the variables measured in the survey.

Variable name	Wording	Coding
Education	What is your highest level of education?	0 = elementary school; 1 = middle school; 2 = high school; 3 = associate degree; 4 = bachelor degree; 5 = masters or professional degree; 6 = doctorate
Current school status	Are you currently enrolled in school?	1 = currently in school; 0 = currently not in school
Importance of certification	How important is the certificate that you receive from HarvardXupon successful completion of CS50x?	0 = not important; 1 = slightly important; 2 = somewhat important; 3 = very important; 4 = extremely important
Interests in forum	How do you intend to participate in the forums	0 = I will contribute to discussion threads frequently; 1 = I will contribute to discussion threads occasionally; 2 = I will view discussion threads, but will not contribute; 3 = I will not visit the discussion forums.
Number of MOOCs completed previously	How many online courses have you completed?	0–12 or more (12-or-more was the maximum one could choose)
English fluency	How fluent are you in English?	0 = basic; 1 = weak; 2 = intermediate; 3 = proficient; 4 = fluent
Availability of help from others	Do you have anybody who will be available to help with studying and homework in this course?	0 = no; 1 = yes
prior programming proficiency	What is your proficiency in any computer programming language?	0 = never used; 1 = basic; 2 = intermediate; 3 = expert

years ($SD = 10.16$, ranging from 10 to 69); 42% were living in a country outside of the United States; 48% could speak more than one language; 56% had a college degree as their highest educational level; 3% had an advanced degree; 36.53% of the enrollees were concurrently going to school. The enrollees spent 7.95 h/week, on average, playing digital games. 67.6% had some computer programming experience prior to the MOOC. On average, enrollees (including those with no prior knowledge) had some experience (more than nothing) in roughly three kinds of programming languages. Forty-three percent of the enrollees rated their proficiency with computer programming to be 'intermediate' or 'expert' (rating from a range of 0 'never', 1 'basic', 2 'intermediate', or 3 'expert', with $M = 1.20$, $SD = 1.14$); 48.5% answered that they did not have friends or family members who could give them programming help; 68.2% predicted that they were very likely or extremely likely to finish the course in order to attain a certificate. On average, enrollees had registered in 1.9 MOOCs and had completed 1.1 MOOCs prior to this MOOC. The average pre-computational skill test score was 0.85 (9.40 out of 11 questions, $SD = 0.15$).

3.2. MOOC participation variables

On average, participants finished 3.2 ($SD = 2.4$) out of the eight problem sets. If (as previously mentioned) we adopt the relaxed definition of completion as finishing at least seven out of the eight psets, 15.7% of the analytic sample could be considered as completers. For the whole sample, the average time span from the first to the last active date was 48.2 days (ranging from 1 to 398 days, $SD = 74.83$). Those who did not participate in the final pset (hereafter the '*no-final*' group) comprised 88.6% ($n = 8769$) of the sample; 11.4% ($n = 1130$) of the sample participated in the final (hereafter the '*final-taker*' group). Among the *final-takers*, 82.4% ($n = 931$) took the final pset in the very beginning of course; that is, they took the final pset right after taking the pre-survey, before engaging with any other course material (hereafter the '*final-early*' group); 17.6% ($n = 199$) of the *final-takers* took the final pset at a later milestone (hereafter the '*final-later*' group). On average, participants in the *final-later* group took the final pset at the 6.8th milestone, which means that most of them attempted the final pset after finishing most of the psets.

Among the *final-early* participants, 91.5% ($n = 852$) failed the final pset (hereafter the '*final-early-fail*' group), and only 8.5% ($n = 79$) passed the final pset (hereafter the '*final-early-pass*' group), confirming that the final pset was very difficult and those who passed it possessed a strong programming competency. If we keep the same definition of completion as previously mentioned, 14.25% of *no-final*, 82.41% of *final-later*, 100% of *final-early-pass,* and 6.9% of *final-early-fail* had completion. The *final-early-fail* group had the lowest completion rate, and a chi-square test showed that it was statistically significantly worse than that of the no-final group ($\chi^2 = 36.89$, $p < 0.001$).

Table 2 presents the descriptive statistics of participation and pre-survey variables broken down by each group.

3.3. Milestone trajectories

Participants were allowed to freely choose the sequence of milestones and drop out at any time (there were more than a hundred thousand selectable sequences that a participant could have taken). The most informative way to illustrate the pattern behind the irregular sequences was to plot the trajectories and summarise the most common paths. Figure 1 is an arc diagram that presents all of the paths that had been taken by at least 25 participants.

The width of the lines corresponds to the number of the participants in each path. Each node on the horizontal axis corresponds to a specific milestone (P1 to P8 are psets, Pre is pre-survey). The size of the nodes indicates the retained sample at the corresponding milestone.

Table 2. Descriptive statistics of participation and pre-survey variables by final exam status groups.

	Final-early fail	Final-early pass	Final-later	No-final	Full-sample
N	852	79	199	8769	9899
Completion	4%	100%	77%	14%	15%
Male	82%	78%	86%	82%	82%
Has programming experience	62%	62%	69%	68%	68%
In school	38%	46%	44%	36%	37%
No help from others	46%	39%	49%	49%	49%
College degree or above	54%	53%	53%	56%	56%
Certification important	70%	77%	74%	68%	68%
Age	29.60 (10.25)	27.66 (11.47)	29.01 (11.74)	29.46 (10.10)	29.44 (10.16)
Programming proficiency	1.03 (1.12)	1.06 (1.15)	1.20 (1.16)	1.22 (1.15)	1.20 (1.15)
MOOCs completed	0.89 (2.03)	0.45 (1.05)	1.05 (1.91)	1.07 (2.27)	1.05 (2.24)
MOOCs registered	1.77 (2.81)	0.97 (1.73)	1.83 (2.54)	1.94 (2.96)	1.91 (2.94)
Number of days active	33.74 (56.13)	140.76 (93.49)	147.65 (105.48)	46.46 (73.22)	48.15 (74.83)
Num of Pset finished	2.58 (1.92)	8.73 (0.47)	7.96 (1.61)	3.07 (2.32)	3.21 (2.42)

Note: Standard deviation is in the parenthesis.

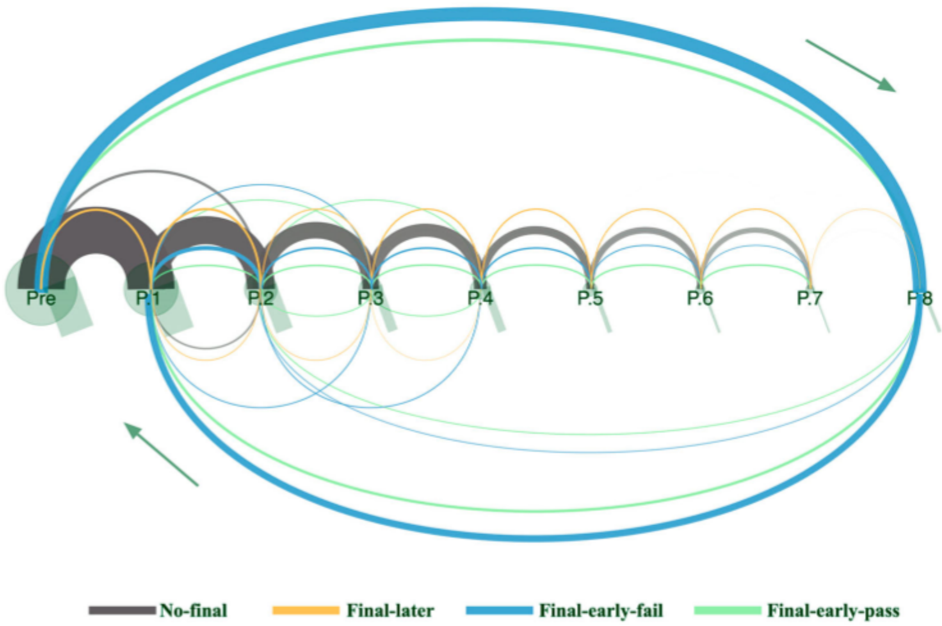

Figure 1. Arc diagram for popular pathways ($n > 25$).

There is a downward-going path at each node, which indicates the participants who dropped out at the milestone. The arcs above the horizontal axis are forward-going trajectories, and the arcs below it are backward-going trajectories. Each group is identified by colour, as shown in the legend.

We can read Figure 1 in combination with Table 1. About 10% of the whole sample were '*final-early*' takers; 8.5% of the *final-early* takers passed, and 91.5% failed the final pset.100% of the '*final-early-pass*' group persisted to the end of the course, whereas only 4% of the '*final-early-fail*' group persisted to the end. 24.5% of the *final-early-fail* group dropped out immediately after trying the final pset, without taking any pset, and 37.4% of the remaining participants in the *final-early-fail* group dropped after finishing only one pset. 90% of the sample did not jump to the final early on; 97.7% of those did not make it to the final. Nevertheless, 14% of the *no-final* group finished enough psets to be considered having completed the course.

In general, there were three common types of trajectories: (a) sequential participants: the paths that followed the sequence designated by course syllabus; (b) semi-sequential participants: the paths that misplaced the sequence of one milestone, but returned to the sequence afterwards (e.g. P1 P3 P2 P4 P5); and (c) participants in the final-early group who took the final pset in the beginning of the course (right after the pre-survey).

Figure 1 clearly shows that, although *final-early* takers only represented about 10% of the sample, this was the only dominant trajectory among all trajectories alternative to the sequential one.

3.4. Predicting final taking status

To predict final pset taking status among the four possible groups, we used multinomial logistic regression, designating the *no-final* group as the reference. All continuous predictors, including those coded in Likert scales, were standardised to z-scores before being entered into the model. Table 3 presents the model estimation. Each cell contains the coefficient (change of log odds of being in one category versus the *no-final* category associated with one-unit change in the predictor) with the standard error in the parenthesis.

The coefficients from each column in Table 2 should be interpreted in similar fashion to the coefficients from binary logistic regressions, except that the reference group in this context was *final taking status* = *no-final*. For example, for the *final-early-fail* column, we could write our model equation as:

$$\ln\left(\frac{P(\text{status} = \mathit{final\_early\_fail})}{P(\text{status} = \mathit{no\_final})}\right)$$
$$= (-2.484) + 0.097\,\text{male} + 0.064\,\text{age} + \ldots$$
$$+ (-0.205)\mathit{pre\_computational\_thinking}$$

Focusing on coefficients that were statistically significant, a one standard deviation increase in the variable *prior programming proficiency* was associated with the decrease in the log odds of being in *final-early-fail* versus *no-final* status in the amount of 0.138. Exponentiating this coefficient yielded the relative risk (also referred to as odds ratio) of 0.87 for being in *final-early-fail versus no-final* status for a one standard deviation increase in

Table 3. A multinomial model predicting final exam status relative to the no-final category.

	Final-early-fail		Final-early-pass		Final-later	
(Intercept)	−2.484	(0.132)	−5.206	(0.452)	−4.351	(0.283)
Male vs. Female	0.097	(0.112)	−0.161	(0.347)	0.174	(0.234)
Age	0.064	(0.050)	0.147	(0.163)	0.048	(0.100)
Education level	0.008	(0.048)	−0.067	(0.152)	0.055	(0.095)
Currently in school	0.115	(0.106)	0.324	(0.334)	0.314	(0.203)
Certificate is important	0.083	(0.094)	0.594	(0.343)	0.345	(0.198)
Interest in forum participation	0.004	(0.043)	−0.138	(0.141)	−0.147	(0.087)
English fluency	0.014	(0.044)	−0.304	(0.104)	−0.119	(0.075)
Numbers of MOOCs completed	−0.070	(0.047)	−0.633	(0.310)	−0.017	(0.085)
No help from others	−0.148	(0.085)	−0.399	(0.275)	0.043	(0.167)
Prior programming proficiency	**−0.138*** **	(0.043)	−0.024	(0.138)	0.011	(0.084)
Pre-computational thinking	**−0.205*** **	(0.037)	−0.106	(0.126)	**−0.141*** **	(0.076)

Notes: ***$p < 0.001$, after Bonferroni correction. The coefficients are log-odds, with standard errors in parentheses. $N = 8960$; pseudo-$R^2 = 0.265$.

prior programming proficiency. Controlling all other covariates at the mean, we calculated that, as *prior programming proficiency* increased by one standard deviation, the marginal decrease in probability of being in the *final-early-fail* group was 1.02%. In the meantime, the changes in the log odds of being in *final-early-pass*, or in *final-later*, versus *no-final* were not associated with changes in *prior programming proficiency*. In other words, *prior programming proficiency* predicted if one belonged to the *final-early-fail* group vis-à-vis the *no-final* group, and participants who had lower prior programming proficiency were more likely to take the final pset early and fail, as opposed to not taking the final at all.

Similarly, we concluded that participants with higher *pre-computational thinking* scores were less likely to be in the *final-early-fail* (log odds = −0.205; relative risk (odds ratio) = 0.815; marginal probability = −1.47%), and *final-later* (log odds = −0.141; relative risk = 0.868; marginal probability = −0.27%) versus the *no-final* group. The *pre-computational thinking* score was not associated with the likelihood of being in the *final-early-pass* group, as opposed to the *no-final* group. Conversely, when we used the *final-early-fail* group as the reference category (in a post-hoc test), we concluded that *pre-computational thinking* scores were positively associated with the likelihood of being in the *final-early-pass* or *no-final* groups. It was to be expected that the *final-early-pass* group had higher pre-computational thinking scores than did the *final-early-fail* group. *Pre-computational thinking* had the largest effect size on group allocation for *final-early-fail* vis-à-vis *no-final.* Figure 2 illustrates both the effect of *pre-computational thinking* (x-axis) and *prior programming proficiency levels* (colours: orange for never used, green for basic, blue for intermediate, purple for expert) on the probability of group allocation for all three final-taker groups (line type: solid line for *final-early-fail* [all sub-groups], dashed line for *final-later*, dotted line for *final-early-pass*). This figure shows that *pre-computational thinking* had effects on *final-early-fail* and *final-later* groups, with the largest effect size on the *final-early-fail* group. It also shows that an effect of *prior programming proficiency levels* occurred only within the *final-early-fail* group, but not for the other two groups (Figure 2 actually plots the *final-later* and *final-early-pass* groups by programming proficiency levels as well; however, the coloured levels are almost indistinguishable for each of these two groups because the effects of programming proficiency were non-significant and minimal).

4. Discussion

The key message from this study is straightforward: In CS50x, most participants who took the final pset chose to take it in the beginning of the MOOC. Such a strategy was the only dominant strategy alternative to following the sequence on the syllabus. Among participants who took the final pset at the beginning, if they passed the test up front, 100% of them would complete the full course; if they failed the exam, their completion rate was the lowest, even worse than that of participants who never attempted the final pset.

4.1. Theoretical implications

We set out in this study with two major theoretical frameworks: the learning progression theories, and the expectancy-value theory. Our finding supported the traditional learning progression theory that suggests learning should follow a stepwise sequence. Our finding did not support the modification of learning progression theory that suggests a global overview of the course may help learners learn more effectively. However, it is noteworthy that taking the final pset may not constitute a global overview, as the final pset is still focused on a separate unit (although it requires the synthesis of multiple prior acquired skills) and cannot reveal the global context of the course structure. Thus, our finding may not be suited to disproof the global overview component of the learning progression theory.

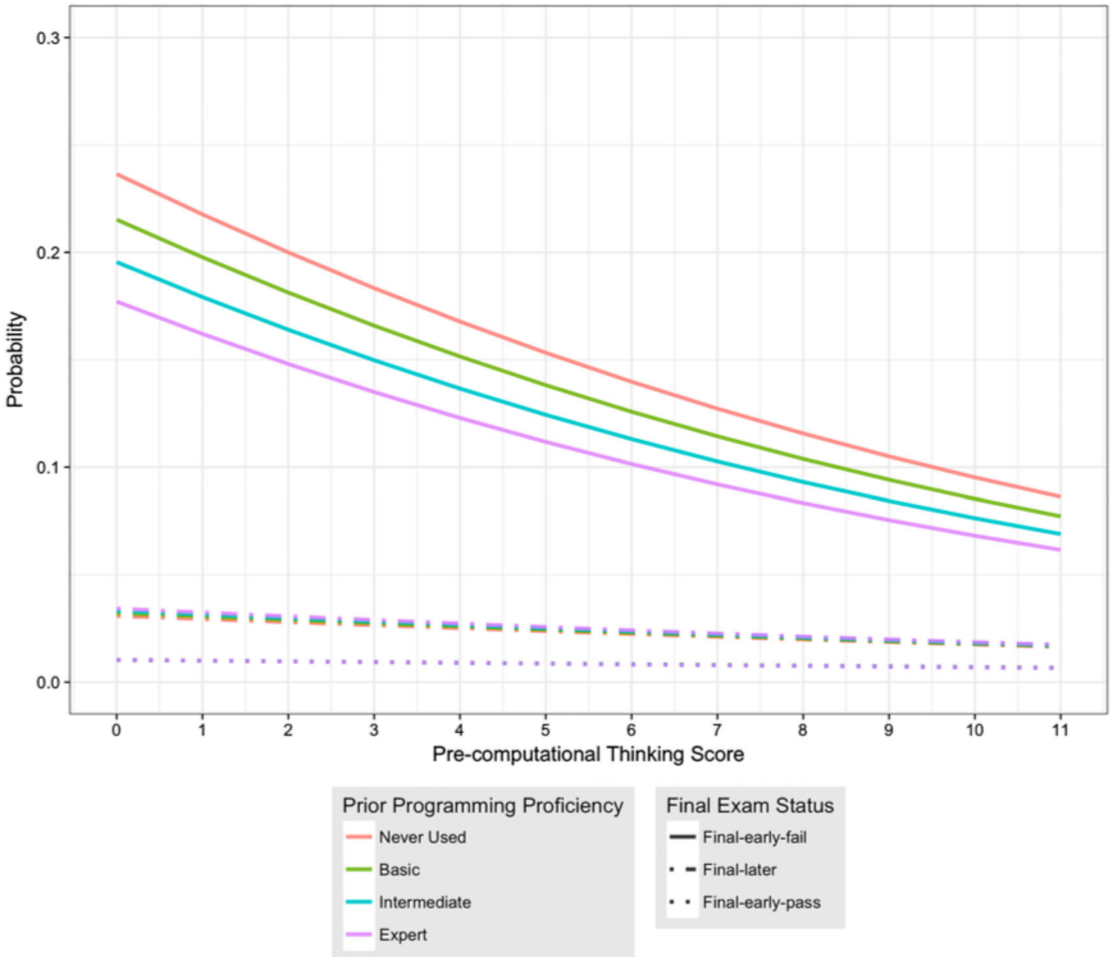

Figure 2. Relationship between Pre-computational thinking scores and the predicted probability of the three final-taker groups, by prior programming proficiency.

From an expectancy-value theory standpoint, our finding suggested that there was a strategic expectancy and value estimation effect of previewing the final-stage course work at the beginning of the course. The preview provided critical expectancy and cost information to the learner that then was used to drive his/her subsequent course taking decisions. This result corroborated our hypothesis that participants who adopted the endgame-previewing strategy used the results of that preview to decide on their course taking behaviour. Those who succeeded in the endgame were motivated to complete the course; those who failed the endgame would drop out quickly.

In addition, we found a relative lack of prior preparedness to predict previewing and failing the final pset. Specifically, learners with lower self-reported programming proficiency and lower pre-computational thinking skills were more likely to be final-early-failures; they took the final in the beginning, failed it, and completed fewer milestones than average before dropping out.

In light of the (counterintuitive) finding that so many novices challenged themselves with the final pset at the beginning of the course, we suggest an add-on to the expectancy-value theory framework – that novices had higher needs to estimate their expectancy and to calculate their cost. This is likely due to novices having less prior experience to assess their own competency – they were more likely to adopt the endgame-previewing strategy to self-diagnose and get a better expectancy estimate.

4.2. Function of the final exam

It is apparent that the majority of the final-takers in this study did not treat the final pset as a capstone project in final stages of the course, but rather for a self-pre-screening in the initial stage. Those who passed the exam before learning the course content received assurance that the course difficulty was within their comfort zone. These participants proceeded to finish all other milestones. We were surprised at the 100% completion rate for final-early-pass participants. It was possible that these participants came with advanced programming skills that enabled them to finish the course effortlessly. It was also possible that a positive experience with passing the final pset early in the sequence gave them the self-efficacy to persist even if they encountered difficulties later. Dropping out after passing the final pset may have appeared to them as a wasted opportunity. It is noteworthy that the average time that these participants spent on the course was 140 days, shorter than that of other participants who completed the course, suggesting that the final-early-pass participants had above-average aptitude in computer programming or eagerness to receive the certificate; however, 140 days was not so short a time as to give reason to believe that the course was so easy for them, on average, that they could complete it within a week or a month.

By contrast, those who failed the final pset upfront had the poorest performance in terms of completion. We speculate that failing the final pset led to multiple reactions by the participants. First, it revealed to the participants their lack of programming background and skills. Those who were shopping for a course within their comfort zone might lose interest in this course. Second, participants who took an early peek at the final pset might be overwhelmed by its difficulty. We analogize this to playing computer games. It may be tempting to take on the big boss from the endgame or challenge the professional level in the first trial of a game. Getting destroyed by the big boss is usually a discouraging, if not humiliating, experience. It reveals the daunting gap an amateur has to fill (with time and practice) to achieve an expert's level of mastery. Such a gap appears even more insurmountable if the amateur does not know the existence or amount of the equipment (e.g. weapons, gadgets, upgrades) that must be acquired to be simply on a par with the big boss. Similarly, failing the final pset at the beginning of the course may exaggerate the difficulty of the course in the mind of the test taker. In particular, when persons have not yet seen the course content, they do not know the existence of the tools they will need to acquire to accomplish the tasks, which may layer despair upon self-doubt. Indeed, previous studies have shown self-efficacy to be a strong predictor of MOOC participation (Bates and Khasawneh 2007; Pellas 2014). Third, those who were defeated by the final pset in the beginning and did proceed to the initial sessions in the course would soon realise that the endgame was extraordinarily more difficult than the basics they learned in the beginning (e.g. realising that understanding and creating a loop using Scratch was far from sufficient to manage a website). Projecting from the learning pace in the initial sessions, they might estimate the level of the endgame to be unreachable, which might explain why final-early-fail participants had in average only 2.58 psets completed, the fewest among all groups.

4.3. Target audience

Since their inception, MOOCs have been projected as the great democratizer in education. The grand goal of the MOOC movement is to make higher education accessible to a population that does not have adequate educational resources and background. For introductory level MOOCs in particular, the target audiences are students at the novice and intermediate levels. However, our analysis revealed the disconcerting reality that MOOC participants self-diagnose at the beginning of the course, using the final pset, its most difficult test, to determine their course participation.

As noted in the literature review, motivation has been considered a key predictor of MOOC persistence. We, too, included two motivation measures in our model – importance of certificate and interests in participating in the forum. However, in a multinomial model that predicted subgroup allocation, we did not find the motivation factors to distinguish between the subgroups. The only predictors that effectively distinguished subgroup allocation were measures of preexisting programming skills (programming proficiency and pre-computational thinking). Therefore, we cannot conclude that final-early-pass learners had a different motivation from that of final-early-fail learners before they enrolled in the course. Rather, our findings suggest that motivation, if it indeed played a role in course persistence for final early takers, was likely to be upgraded or downgraded after learners tested themselves in the final pset.

Those advanced students who passed the final pset at the outset, a subgroup that arguably least needs the course, finished the rest of the course to attain the certificate; those novice students who failed the exam – the subgroup that the introductory level courses were designed for – were the first to drop out. It is noteworthy from Table 1 that both final-early-pass and final-early-fail participants had self-described their prior programming experience as well as their proficiency to be statistically significantly lower than the grand average. This might appear counterintuitive because the final-early-pass participants turned out to be highly proficient. We speculate that self-reporting one's prior programming proficiency was a rather subjective measure. It might not accurately reflect one's true skill, but the assuredness (confidence) in one's proficiency; and participants who lacked confidence might have been more likely to go to the endgame to self-diagnose. Those who passed would have gained assurance, and those who failed would have lost confidence or reinforced their self-doubt.

This voluntary self-screening and self-selection tendency among MOOC participants counteracts the MOOC providers' aim to reach out to underserved populations by designing beginner-friendly courses and disseminating them for free (or a low fee). This finding echoes previous studies that have shown that the participants who benefit the most from MOOCs tend to come from more resourceful and better educated populations (Stich and Reeves 2017) and have warned that 'MOOCs … can exacerbate rather than reduce disparities' (Hansen and Reich 2015, 1). We believe this is a serious challenge for the MOOC movement. By way of comparison, to drop out because of gradual interest loss, fatigue, or time conflicts, is less problematic than to drop out because of an attitude of self-defeatism before learning the content.

4.4. Policy implications

We argue that the root of the aforementioned problem is that (a) MOOC students have the need to preview the endgame before taking the course when (b) there is no proper channel for previewing, thereby (c) forcing students to misuse the final pset to get a glimpse of the endgame. This suggests that a forced sequence, or a semi-sequenced approach (e.g. allowing freedom to choose the unit sequence, but disallowing viewing the final pset before finishing all units) could prevent students from taking the self-screening and self-defeating glimpse. This would constitute a simple solution. Indeed, after the initial excitement over the open structure of MOOCs (Anderson 2008; Hales 2000), scholars have recently revisited the benefit of infinite freedom and proposed restricting accessibility and limiting repeatability in MOOC designs in order to ameliorate the dropout problem (Kim et al. 2017; Zheng et al. 2015).

If, as argued by prior MOOC studies, the irregular learning progression can be fully explained by the lack of self-regulation (Cho and Shen 2013; Littlejohn et al. 2016; Martinez-Lopez et al. 2017; Pellas 2014), time management (Lee 2018; Lin, Lin, and Hung 2015), or goal-orientation (Maldonado-Mahauad et al. 2018), we anticipate that constraining students' focus by limiting access would help students persist in the course. Yet, findings from our study strongly suggest that at least some of the irregularity can be explained by students' intentional choice of a particular strategy for participation: students choose to preview the endgame, based on which they will decide their degree of engagement. Thus, we argue that simply prohibiting taking the final early does not accommodate students' need for previewing the endgame. In response to such a student need, we propose that a better solution resides in designing a constructive preview of the endgame. This would constitute a more complicated – yet ultimately perhaps more successful – solution. The goal of the constructive preview is to provide the big picture of what a student will achieve by the end of the course, display the content and difficulty progression accompanied by the tools that students will acquire, and provide self-diagnosis tools or exercises for students to evaluate their current level of understanding and skills in computer programming to help students make study plans. Kizilcec, Davis, and Cohen (2017) have shown affirmation to be a powerful tool to counteract the social identity threat and to close the achievement gaps in MOOC performance. Similarly, the constructive preview aims to present a manageable learning progression to reassure all levels of students. In the meantime, it needs to reduce the number of avenues that could lead to a defeat of students' self-confidence.

Twenty-five years ago, Matejka and Kurke (1994) wrote in 'Designing a great syllabus' that a syllabus is 'the cognitive map of the destination … on an intellectual journey' (117). This metaphor still holds in the MOOC age, but the design of the syllabus as a cognitive map of the destination should evolve as consequences, such as those found in this study, become known. In this instance, an improved syllabus should incorporate a human-centered approach of presenting the endgame, taking advantage of media technology, and adapting to the changes in teacher-student interaction. Nevertheless, how the effects of the simple solution (preventing the preview of the final pset) and the more complicated solution (improving the preview of the cognitive map) compare with each other in terms of student behaviour remains an intriguing empirical question to be investigated in future research.

Finally, to protect participants' confidence from being crushed by the final pset, one may be tempted to consider deliberately assigning simple questions or challenges for those looking at the preview of the final pset. However, this would be deceitful and lead the participant to forming a mis-estimation of the difficulty and mis-anticipation of the end goal of the course. Thus, we suggest, if confidence-boosting tests were needed, they should be separated from the final pset (or any pset), and only be allocated in the constructive preview to reaffirm participants' current competency without mis-representing the end game.

In short, MOOC providers should consider (a) constraining the visibility of the later psets of the course, (b) providing an early assessment for diagnostic and reaffirming purposes that is unrelated to the end product, and (c) providing a preview of the end product through a constructive syllabus that not only previews the end game alone, but also introduces the stepwise skill growth in the learning trajectory.

4.5. Limitations

The obvious limitation of this study was self-selection bias in group allocation: participants freely chose if, and when, they would take the final pset; the MOOC providers and the researchers could not randomly assign exam taking strategies. Thus, we could not make causal claims, such as taking the final early and failing the final would lead to lower completion. The main message that this study intends to deliver was the typology of MOOC users: to present the common trajectories and describe their background and course completion characteristics. Future studies should carefully examine the effects of innovative designs and applications of aforementioned constructive previews.

The discussion of this article did not focus on those who did not take the final, although they constituted the majority of the sample. There was not any evidence that this subgroup intentionally avoided the final pset. Predominantly, the no-final participants followed the regular sequence but dropped out early. In other words, no-final participants revealed the issues of dropout in general, not the issue of final pset avoidance.

This study showed a simple yet salient pattern among MOOC learners; it did not disentangle the intricate time series of learners' participation history. Prior studies have adopted survival analysis (e.g. Chen et al. 2019) or Hidden Markov models (e.g. Balakrishnan and Coetzee 2013) to predict learners' dropout or sampling behaviour. Future study should consider the final pset as an important transition state, rather than simply an endpoint. If we keep tracking learners' motivation at each milestone (rather than measuring it once at the beginning of the MOOC), we can adopt Hidden Markov models to study the interaction/synchronization between the fluctuating motivation and the irregular backward and forward sampling behaviours.

5. Conclusion

In summary, taking the final pset in the beginning of the CS50x MOOC was the only dominant strategy alternative to following the sequence on the syllabus. 100% of those who passed the final pset early would complete the full course. In the meantime, those who failed the final early had the lowest completion rate, even worse than that of participants who never attempted the final pset. The results from this study suggest that MOOC students, especially students who were unsure about their prior proficiency, have the need to preview the endgame at the beginning of the course. A negative feedback from the previewing experience tends to discourage students from persistence. Future MOOC studies and policies should take a human-centered design perspective and consider constraining the visibility of the final pset at the beginning of the course as well as exploring constructive strategies to accommodate students' need for previewing the endgame.

Acknowledgements

Any opinions, findings, and conclusions in this article are the authors' and do not necessarily reflect the views of the National Science Foundation. We thank Alaalden Ibrahim and John Murray for processing the MOOC data. We also thank those that gave us additional support and direction: Charles Alcock, Lori Breslow, Andrew Ho, Annie Valva, Rob Lue, and Wendy Berland.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by National Science Foundation [grant number 1352696].

ORCID

Chen Chen http://orcid.org/0000-0002-6065-8889

References

Anderson, T., ed. 2008. *The Theory and Practice of Online Learning*. Edmonton, AB: Athabasca University Press.

Atkinson, J. W. 1957. "Motivational Determinants of Risk-taking Behavior." *Psychological Review* 64 (6p1): 359–372.

Balakrishnan, G., and D. Coetzee. 2013. "Predicting Student Retention in Massive Open Online Courses Using Hidden Markov Models." *Electrical Engineering and Computer Sciences University of California at Berkeley* 53: 57–58.

Bates, R., and S. Khasawneh. 2007. "Self-efficacy and College Students' Perceptions and Use of Online Learning Systems." *Computers in Human Behavior* 23 (1): 175–191.

Belanger, Y., and J. Thornton. 2013. Bioelectricity: A Quantitative Approach Duke University's First MOOC.

Buhr, E. E., L. M. Daniels, and L. D. Goegan. 2019. "Cognitive Appraisals Mediate Relationships Between Two Basic Psychological Needs and Emotions in a Massive Open Online Course." *Computers in Human Behavior* 96: 85–94.

Chen, C., M. H. Schneps, and G. Sonnert. 2016. "Order Matters: Sequencing Scale-realistic Versus Simplified Models to Improve Science Learning." *Journal of Science Education and Technology* 25 (5): 806–823.

Chen, C., G. Sonnert, P. M. Sadler, D. Sasselov, and C. Fredericks. 2019. "The Impact of Student Misconceptions on Student Persistence in a MOOC." *Journal of Research in Science Teaching*. doi:10.1002/tea.21616.

Cho, M. H., and D. Shen. 2013. "Self-regulation in Online Learning." *Distance Education* 34 (3): 290–301.

Coffrin, C., L. Corrin, P. de Barba, and G. Kennedy. 2014. "Visualizing Patterns of Student Engagement and Performance in MOOCs." In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, 83–92. ACM.

De Barba, P. G., G. E. Kennedy, and M. D. Ainley. 2016. "The Role of Students' Motivation and Participation in Predicting Performance in a MOOC." *Journal of Computer Assisted Learning* 32 (3): 218–231.

DeBoer, J., A. D. Ho, G. S. Stump, and L. Breslow. 2014. "Changing 'Course' Reconceptualizing Educational Variables for Massive Open Online Courses." *Educational Researcher* 43 (2): 74–84.

Dillahunt, T. R., B. Z. Wang, and S. Teasley. 2014. "Democratizing Higher Education: Exploring MOOC Use among Those Who Cannot Afford a Formal Education." *The International Review of Research in Open and Distributed Learning* 15 (5): 177–196.

Eccles, J., T. F. Adler, R. Futterman, S. B. Goff, C. M. Kaczala, J. Meece, and C. Midgley. 1983. "Expectancies, Values and Academic Behaviors." In *Achievement and Achievement Motives*, edited by J. T. Spence. San Francisco, CA: W. H. Freeman.

Eccles, J. S., and A. Wigfield. 1995. "In the Mind of the Actor: The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs." *Personality and Social Psychology Bulletin* 21 (3): 215–225.

Fini, A. 2009. "The Technological Dimension of a Massive Open Online Course: The Case of the CCK08 Course Tools." *The International Review of Research in Open and Distributed Learning* 10 (5): 1–26.

Greene, J. A., C. A. Oswald, and J. Pomerantz. 2015. "Predictors of Retention and Achievement in a Massive Open Online Course." *American Educational Research Journal* 52 (5): 925–955.

Guo, P. J., and K. Reinecke. 2014. "Demographic Differences in How Students Navigate through MOOCs." In *Proceedings of the First ACM Conference on Learning@ Scale Conference*, 21–30. ACM, March.

Haggard, S. 2013. "The Maturing of the MOOC." Department for Business Innovation & Skills. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/240193/13-1173-maturing-of-the-mooc.pdf.

Hales, R. S. 2000. "Expectations exploded." In *Issues in Web-Based Pedagogy: A Critical Primer*, edited by R. A. Cole. London: Greenwood Publishing Group.

Hansen, J. D., and J. Reich. 2015. "Democratizing Education? Examining Access and Usage Patterns in Massive Open Online Courses." *Science* 350 (6265): 1245–1248.

He, J., J. Bailey, B. I. Rubinstein, and R. Zhang. 2015. "Identifying At-Risk Students in Massive Open Online Courses." In *AAAI*, 1749–1755, January.

Ingham, A. M., and J. K. Gilbert. 1991. "The Use of Analogue Models by Students of Chemistry at Higher Education Level." *International Journal of Science Education* 13 (2): 193–202.

Jacobs, A. J. 2013. "Two Cheers for Web U." *New York Times* 162 (56113): 1–7.

Jiang, S., A. Williams, K. Schenke, M. Warschauer, and D. O'dowd. 2014. "Predicting MOOC Performance with Week 1 Behavior." In *Educational Data Mining 2014*, July.

Jordan, K. 2015. "Massive Open Online Course Completion Rates Revisited: Assessment, Length and Attrition." *The International Review of Research in Open and Distributed Learning* 16 (3): 341–358.

Kim, T. D., M. Y. Yang, J. Bae, B. A. Min, I. Lee, and J. Kim. 2017. "Escape From Infinite Freedom: Effects of Constraining User Freedom on the Prevention of Dropout in an Online Learning Context." *Computers in Human Behavior* 66: 217–231.

Kizilcec, R. F., G. M. Davis, and G. L. Cohen. 2017. "Towards Equal Opportunities in MOOCs: Affirmation Reduces Gender & Social-class Achievement Gaps in China." In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*, 121–130. ACM, April.

Kizilcec, R. F., C. Piech, and E. Schneider. 2013. "Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses." In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 170–179. ACM, April.

Kloft, M., F. Stiehler, Z. Zheng, and N. Pinkwart. 2014. "Predicting MOOC Dropout Over Weeks Using Machine Learning Methods." In *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs*, 60–65.

Kotamraju, N. P., and T. M. van der Geest. 2012. "The Tension Between User-Centred Design and E-Government Services." *Behaviour & Information Technology* 31 (3): 261–273.

Koutropoulos, A., M. S. Gallagher, S. C. Abajian, I. de Waard, R. J. Hogue, N. O. Keskin, and C. O. Rodriguez. 2012. "Emotive Vocabulary in MOOCs: Context & Participant Retention." *European Journal of Open, Distance and E-Learning* 1 (1): 1–23.

Kujala, S. 2008. "Effective User Involvement in Product Development by Improving the Analysis of User Needs." *Behaviour & Information Technology* 27 (6): 457–473.

Larkin, J., and R. W Chabay. 1989. "Research on Teaching Scientific Thinking: Implications for Computer-Based Instruction." In *Toward the Thinking Curriculum: Current Cognitive Research*, edited by L. Resnick and L. Klopfer, 150–172. Alexandria, VA: Association for Supervision and Curriculum Development.

Lee, Y. 2018. "Effect of Uninterrupted Time-on-Task on Students' Success in Massive Open Online Courses (MOOCs)." *Computers in Human Behavior* 86: 174–180.

Li, B., X. Wang, and S. C. Tan. 2018. "What Makes MOOC Users Persist in Completing MOOCs? A Perspective From Network Externalities and Human Factors." *Computers in Human Behavior* 85: 385–395.

Lin, Y. L., H. W. Lin, and T. T. Hung. 2015. "Value Hierarchy for Massive Open Online Courses." *Computers in Human Behavior* 53: 408–418.

Littlejohn, A., N. Hood, C. Milligan, and P. Mustain. 2016. "Learning in MOOCs: Motivations and Self-Regulated Learning in MOOCs." *The Internet and Higher Education* 29: 40–48.

Maldonado-Mahauad, J., M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, and J. Munoz-Gama. 2018. "Mining Theory-Based Patterns From Big Data: Identifying Self-Regulated Learning Strategies in Massive Open Online Courses." *Computers in Human Behavior* 80: 179–196.

Martinez-Lopez, R., C. Yot, I. Tuovila, and V. H. Perera-Rodríguez. 2017. "Online Self-Regulated Learning Questionnaire in a Russian MOOC." *Computers in Human Behavior* 75: 966–974.

Matejka, K., and L. B. Kurke. 1994. "Designing a Great Syllabus." *College Teaching* 42 (3): 115–117.

Mukala, P., J. C. A. M. Buijs, and W. M. P. Van Der Aalst. 2015. "Uncovering Learning Patterns in a Mooc Through Conformance Alignments." BPM Reports 1509.

Muller, D. A., M. D. Sharma, and P. Reimann. 2008. "Raising Cognitive Load with Linear Multimedia to Promote Conceptual Change." *Science Education* 92 (2): 278–296.

Pellas, N. 2014. "The Influence of Computer Self-Efficacy, Metacognitive Self-Regulation and Self-Esteem on Student Engagement in Online Learning Programs: Evidence From the Virtual World of Second Life." *Computers in Human Behavior* 35: 157–170.

Peng, D., and G. Aggarwal. 2015. Modeling MOOC Dropouts. http://cs229.stanford.edu/proj2015/235_report.pdf.

Pope, J. 2014. What are MOOCs Good For? The MIT Technology. http://www.technologyreview.com/review/533406/what-are-moocs- good-for/.

Reigeluth, C. M. 1999. "The Elaboration Theory: Guidance for Scope and Sequence Decisions." *Instructional Design Theories and Models: A New Paradigm of Instructional Theory* 2: 425–453.

Rice, J. 2013. "What I Learned in MOOC." *College Composition and Communication* 64 (4): 695–703.

Rovai, A. P. 2003. "In Search of Higher Persistence Rates in Distance Education Online Programs." *The Internet and Higher Education* 6 (1): 1–16.

Ryan, R. M., and E. L. Deci. 2000. "Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." *American Psychologist* 55: 68–78.

Shah, D. 2018. "A Product at Every Price: A Review of MOOC Stats and Trends in 2017." Class Central.

Sharples, M., R. de Roock, R. Ferguson, M. Gaved, C. Herodotou, E. Koh, and M. Weller. 2016. Innovating Pedagogy 2016: Open University Innovation Report 5.

Steinberg, M. S. 2008. "Target Model Sequence and Critical Learning Pathway for an Electricity Curriculum Based on Model Evolution." In *Model Based Learning and Instruction in Science*, edited by J. Clement and M. A. Rea-Ramirez, 79–102. Dordrecht: Springer.

Stich, A. E., and T. D. Reeves. 2017. "Massive Open Online Courses and Underserved Students in the United States." *The Internet and Higher Education* 32: 58–71.

Tukiainen, M., and E. Mönkkönen. 2002. "Programming Aptitude Testing as a Prediction of Learning to Program." In *Proceedings – Psychology of Programming Interest Group 14*, edited by J. Kuljis, L. Baldwin, and R. Scoble, 45–57. http://www.ppig.org/sites/ppig.org/files/2002-PPIG-14th-tukiainen.pdf.

Wang, W., L. Guo, L. He, and Y. J. Wu. 2018. "Effects of Social-Interactive Engagement on the Dropout Ratio in Online Learning: Insights From MOOC." *Behaviour & Information Technology* 38 (6): 621–636.

Wen, M., D. Yang, and C. P. Rosé. 2014. "Sentiment Analysis in MOOC Discussion Forums: What Does It Tell Us?" Paper Presented at the 7th International Conference on Educational Data Mining, London.

Wigfield, A., and J. Cambria. 2010. "Expectancy-value Theory: Retrospective and Prospective." *Advances in Motivation and Achievement* 16: 35–70.

Wigfield, A., and J. S. Eccles. 2000. "Expectancy–Value Theory of Achievement Motivation." *Contemporary Educational Psychology* 25: 68–81.

Wilson, B., and P. Cole. 1991. "A Review of Cognitive Teaching Models." *Educational Technology Research and Development* 39 (4): 47–64.

Yang, D., T. Sinha, D. Adamson, and C. P. Rosé. 2013. "Turn On, Tune in, Drop Out: Anticipating Student Dropouts in Massive Open Online Courses." In *Proceedings of the 2013 NIPS Data-driven Education Workshop*, Vol. 11, 14, December.

Zemsky, R. 2014. "With a MOOC MOOC Here and a MOOC MOOC There, Here a MOOC, There a MOOC, Everywhere a MOOC MOOC." *The Journal of General Education* 63 (4): 237–243.

Zheng, S., M. B. Rosson, P. C. Shih, and J. M. Carroll. 2015. "Understanding Student Motivation, Behaviors and Perceptions in MOOCs." In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1882–1895. ACM, February.