Motivation as a Lens to Understand Online Learners: Towards Data-Driven Design with the OLEI Scale

RENE F. KIZILCEC, Stanford University
EMILY SCHNEIDER, Stanford University

Open online learning environments attract an audience with diverse motivations who interact with structured courses in a number of ways. To systematically describe the motivations of these learners, we developed the Online Learning Enrollment Intentions (OLEI) scale, a 13-item questionnaire derived from open-ended responses to capture learners' authentic perspectives. Though motivations varied across courses, we found each motivation to predict key behavioral outcomes for learners ($N = 71,475$ across 14 courses).

From learners motivational and behavioral patterns, we infer a variety of needs they seek to gratify by engaging with the courses, such as meeting new people and learning English. To meet these needs, we propose multiple design directions, including virtual social spaces outside any particular course, improved support for local groups of learners, and modularization to promote accessibility and organization of course content. Motivations thus provide a lens for understanding online learners and designing online courses to better support their needs.

Categories and Subject Descriptors: K.3.1 [**Computer Uses in Education**]: Distance learning, Collaborative learning

General Terms: Human Factors, Design, Measurement

Additional Key Words and Phrases: Motivation, open online learning environments, Online Learning Enrollment Intentions, OLEI scale, Massive Open Online Courses, MOOCs, online course design, success metrics

ACM Reference Format:
René F. Kizilcec and Emily Schneider, in press. Motivation as a Lens to Understand Online Learners: Towards Data-Driven Design with the OLEI Scale. *ACM Trans. Comput.-Hum. Interact.* 22, 2, Article 0 (in press), 24 pages.
DOI:http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

Open online learning environments are a novel convergence of traditional course structures with the scale, norms, and features of online media. Resembling typical in-person courses, these virtual classrooms present learners with linear pathways through educational resources, complete with assignments to provide feedback on progress and the opportunity to be recognized for academic achievement by earning a passing grade. At the same time, open online learning environments can be considered repositories of freely accessible, interactive resources and the site of temporary communities. As one of many options on the Web for finding information, socializing, or collaborating, these environments are as amenable to casual engagement with content as they are to the focused, ongoing activity characteristic of a student in a traditional course. Accounting

Author's addresses: R. F. Kizilcec, Department of Communication, Stanford University; E. Schneider, Graduate School of Education, Stanford University.

DOI:http://dx.doi.org/10.1145/0000000.0000000

for these distinct perspectives and motivations is a central challenge for designers of open online learning environments.

The most prototypical examples of the current generation of open online learning environments are massive open online courses (MOOCs), which take different forms but tend to include video lectures, texts, discussion forums, and various forms of assessment. Learner behavior varies substantially in MOOCs, as many learners appear to pick and choose elements of the learning environment that are of interest and utility to them [Kizilcec et al. 2013; DeBoer et al. 2014; Seaton et al. 2014]. Many learners interact with these courses in ways that would not be considered "successful" with respect to instructor-defined criteria of success. In contrast, these learners' behavior would be considered normal, or even successful, in the context of engagement in online media, where user-driven behaviors are welcome and encouraged. To illustrate, consider a learner who finishes all lectures in a course but skips assignments due to time constraints. In the current generation of MOOCs, this learner would be very unlikely to receive a certificate or another form of acknowledgement. Yet selectively engaging with available resources on social media platforms, such as YouTube or Tumblr is very common and socially acceptable. The design and usage patterns of current MOOCs are strongly influenced by the top-down authority of traditional educational settings and at the same time by the user-focused values that pervade online media production.

The choices that learners make in a learning environment can be thought of as expressions of learners' own motivations for engaging in the environment. While this perspective has been circulating throughout the research community (e.g., DeBoer et al. 2014), there has not been a systematic approach to identifying learners' motivations or how they relate to subsequent behaviors. A measure that captures learners' motivations can provide a lens for understanding learner behavior, which can in turn inform the design of open online learning environments to serve a multiplicity of goals, defined by both the learners and the instructor.

To this end, we developed the Online Learning Enrollment Intentions (OLEI) scale. A standardized survey item of reasons for enrolling that were iteratively developed based on thousands of open response answers from surveys in 14 MOOCs on various topics (Section 3). The descriptive statistics generated by the OLEI scale provide a rich picture of MOOC participants and of the variety of motivations across courses (Section 4). The relationship between enrollment intentions and course behaviors both confirms intuitions and provides novel insights into how learners' initial motivations shape subsequent actions in MOOCs (Section 5). We propose future design directions based on our interpretation of these motivations in terms of the learners' needs. Finally, we discuss the challenges in balancing the top-down authority of traditional educational settings with the user-focused values of online media production.

2. RELATED WORK

Recent work on MOOCs suggests that learners engage in a wide range of behaviors, which appear to reflect differences in motivation. Guo and Reinicke [2014] explored resource navigation strategies within four MOOCs on technical topics with a particular focus on learners who earned certificates. They found that certificate-earners navigated opportunistically through the MOOC, jumping backwards from assessments to related resources and watching, on average, 80% of the available lecture videos. Guo and Reinicke also found differences in activity based on demographic characteristics, such as age and country of origin, and called for further work to investigate MOOC learners' motivations and their influence on navigation behavior. Kizilcec, Piech, and Schneider [2013] examined learner trajectories through three computer science MOOCs and found multiple prototypical patterns of engagement. Auditing learners, who primarily watched videos throughout the course, reported similarly high levels of satisfaction as Completing learners, who took a more traditional path of watching videos and completing assessments. Completing learners were, however, more likely to say that they had enrolled in the course for purposes of enhancing their resume or because they wanted the fun and challenge of taking the course. Wilkowski and her colleagues [2014] asked learners to self-report course-specific goals for a skill-based MOOC about using Google Maps. Variations in learners' patterns of activities (watching videos and taking assignments) appeared to reflect their goals, though learners with more time-intensive goals completed activities and met their goals at lower rates than those with easier-to-achieve goals.

Throughout these studies, different behaviors appear to be an expression of different motivations, but very litte is known about the nature of these motivations and their distribution across learners, as well as their relationship to individual differences in learners and course content. A standardized measure is needed to gain insights that can inform technology design recommendations that are relevant across courses. With this measure we can address three research questions:

RQ1. What motivates learners in MOOCs?
RQ2. How pronounced are individual differences in motivations between demographics or across courses?
RQ3. Which motivations are predictive of behaviors in MOOCs and how predictive are they?

To address RQ1 and RQ2, we require an appropriate framework for characterizing motivations. Research on motivation in physically co-present classrooms and in small-scale, "traditional" online courses shows that students are driven both by the desire to master content because it is interesting and relevant, and by the desire to demonstrate competency to earn external recognition [Pintrich 2003; Ryan and Deci 2000]. Environmental factors, such as classroom pedagogical strategies, interact with these academic and social motivations to influence learning and engagement, as do individual student characteristics. Demographics, prior knowledge, self-perceptions, and self-regulation—the ability to plan, monitor and control learning behaviors—have all been shown to have salient influence on achievement [Bransford et al. 2000; Tinto 1997; Schunk and Zimmerman 2012]. This literature highlights the complexity of the phenomenon and provides a number of proven strategies for supporting students that can be adapted to the online learning setting. However, the instruments developed to measure student motivation do not easily transfer, because they were designed under the assumption that students are driven by achievement as defined by the teacher and the school system (e.g., Pintrich et al. 1993). This assumption is unlikely to be justified in open online learning environments, which suggests a need for developing a new instrument for measuring learner motivations in these environments.

An alternative framework for motivation prevalent in media studies, "uses and gratification" scholarship, analyzes people's engagement with openly available media such as radio and television. Uses and gratification theory [Blumler and Katz 1974] aims to understand the needs of users and the ways that a technology gratifies those needs. Insights about user needs can be used to develop more targeted design strategies [Stafford 2005]. This work primarily relies on self-reported preferences among media users and we follow this model in developing a self-report measure for enrollment intentions. We also draw on the uses and gratification framework in interpreting learners' observed behavior as gratifying the needs expressed by their enrollment intentions.

To address RQ3, we define behavioral outcomes that reflect variation in activity among MOOC participants. The discourse around multiple definitions of success in open online learning environments has echoes in the community college literature.

Community colleges are the closest brick-and-mortar analogue to open learning environments, as they are designed to support social mobility for any individual regardless of their prior educational experience and have the lowest prices of any higher educational institution in the United States. Large-scale surveys of community college students indicate that they enroll for a variety of reasons, with about half of students prioritizing personal interest, the development of job skills, or both [Horn et al. 2006]. Eighty percent of these students intend to complete a degree, but these intentions shift over the course of enrollment in college, and many students leave and return throughout their degree-seeking career. To account for the multiple pathways that students take towards degree completion, researchers suggest using “milestone” markers as well as final rates of degree earning to characterize success [Goldrick-Rab 2010]. In a college career, such milestones include number of credits earned and re-enrolling in subsequent semesters. We draw on this insight in the current work by defining milestones based on different levels of engagement with aspects of a single course: lecture videos, forums, and assessments.

3. DEVELOPING THE ONLINE LEARNING ENROLLMENT INTENTIONS SCALE

Instructors, designers, and researchers have been interested in capturing learners' reasons for enrolling in online courses, particularly since MOOCs have attracted a large number of people from diverse backgrounds. Survey questions used in prior work and commonly found in course surveys have asked learners to report their reasons for enrolling using open response or multiple choice questions with response options that were not formally tested.

While open response questions provide a rich source of nuanced information, it is challenging to adequately analyze and categorize textual data at a large scale. Moreover, the unstructured nature of open responses complexifies standardized comparisons across learning environments. In contrast, multiple choice items are simpler to analyze, but may fail to capture learners' true motivations if certain response options are missing or if response options are phrased in terms that do not map onto learners' mental model of their motivations. Moreover, any survey question should be designed with careful attention to known survey biases (for a recent review with applications in HCI, see [Müller et al. 2014]). For multiple choice items, the choice of response options and selection constraints (e.g., ‘select all that apply') are critical determinants of response quality.

In light of the absence of an adequate measure of learner motivations in the context of open online courses, we developed the Online Learning Enrollment Intentions (OLEI) scale. The development of the scale involved iterative refinement of response options which were derived from open response answers.

3.1. Development of Response Options

The process for developing response options was iterative and involved several rounds of pretesting, evaluating, and refining the scale, following best practices for the construction of a new instrument [Colton and Covert 2007]. The choice of response options is critical for the validity with which a question can measure a certain construct. Response options should be mutually exclusive and collectively exhaustive, which means that options should not overlap and all possible responses should be covered. While the latter condition is difficult to satisfy, we set out to approach an exhaustive list by systematically analyzing responses across multiple courses to the open-ended question “Why did you enroll in this course?”.[1] A clear advantage of this approach over a

[1]The enrollment intention question specifies a fixed point in time and the concrete act of enrolling in the course, which leaves less room for interpretation and thus reduces unexplained variation in responses. This

top-down strategy for arriving at a list of options is that the phrasing of the resulting response options is closer to how learners express their reasons.

An iterative process of response option development was based on open response answers from learners in three different MOOCs (on topics in Political Science, Computer Science, and Economics). To efficiently categorize large numbers of responses with multiple raters, the process was crowdsourced using Amazon Mechanical Turk (MTurk), where paid workers manually coded random samples of open responses. Prior work suggests that crowdsourcing survey research on MTurk is a viable alternative to recruiting from a university participant pool [Behrend et al. 2011]. The MTurk coders used a preliminary codebook that had been developed by two independent volunteer coders, using learners' open response texts and a previously developed set of reasons provided in course surveys of a major MOOC platform. For each open response, MTurk coders were instructed to select all appropriate reason from the codebook. An "other" option was provided and coders were strongly encouraged to choose this option if some aspect of the response was not reflected in the existing response options. There was also a "spam" option to capture meaningless responses.

In the first iteration, 300 randomly chosen responses were each coded by four MTurk coders. Each option's frequency and intercoder reliability, as well as the correlations between response options, were evaluated. We closely examined responses that received codes which occured infrequently or which exhibited low intercoder reliability. We also examined all responses that were coded "other" by more than one coder. These analyses highlighted gaps in the codebook, as well as categories that did not align with participants' characterizations of their enrollment intentions. For instance, the option "learn about a familiar topic" was originally included as a coarse proxy for prior knowledge, but very few learners wrote about their motivations in a way that clearly indicated their previous experience with the topic. Based on these analyses, the codebook was modified and applied by MTurk coders to a different random sample of responses (200 responses coded by two coders). A third iteration followed the same refinement procedure (300 responses coded by three coders). The final product is the survey item shown in Table I below. Additional details on the iterative development process, including options that were added and removed, are provided in Section A of the electronic appendix.

3.2. Question Design

The previous sections detailed the method for determining the set of response options that learners choose from to report their motivations. Developing a set of response options that approaches a mutually exclusive and collectively exhaustive list was the first critical step toward a good measure of learner motivations. The second step concerns the design of the survey question; in particular, the choice of selection constraints and the presentation of response options.

3.2.1. Selection Constraints. Selection constraints like 'select one' or 'select three' place an arbitrary limit on the number of reasons respondents can report and coerces them to select a certain number. This tends to induce satisficing behavior, as respondents who would otherwise select a different number of options become less invested in making an effort to respond accurately [Krosnick 1999]. Although 'select all that apply' prompts encourage respondents to decide on the number of options to select, their unguided nature does not require respondents to consider each answer option in turn.

---

is preferred to a broad question about motivations which are not bound by time and subject to change. The measurement of motivations by asking about enrollment intentions is thus expected to more consistently reflect learner motivations at a given point in time in the context of open learning environments.

Table I. Online Learning Enrollment Intentions (OLEI) Scale

Why did you enroll in this course?	Applies	Does not apply
General interest in topic	◯	◯
Relevant to job	◯	◯
Relevant to school or degree program	◯	◯
Relevant to academic research	◯	◯
For personal growth and enrichment	◯	◯
For career change	◯	◯
For fun and challenge	◯	◯
To meet new people	◯	◯
To experience an online course	◯	◯
To earn a certificate/statement of accomplishment	◯	◯
Course offered by prestigious university/professor	◯	◯
To take with colleagues/friends	◯	◯
To improve my English skills	◯	◯

After selecting a few options, a respondent might feel that they selected "enough". As a result, leaving an option unselected does not have a clear and consistent interpretation.

A question design that avoids these issues asks respondents to consider each response option in turn and choose whether it applies to them or not. This recasts each response option into a statement with a dichotomous scale.[2] The response options were labeled "Applies" and "Does not apply", instead of the more commonly found "True/False" or "Yes/No" labels, in an attempt to reduce the likelihood of inducing acquiescence bias—respondents' tendency to agree with questions independent of their content [Krosnick 1999].[3]

3.2.2. Response Option Ordering. It is generally recommended to present non-ordinal response options in random order, as presentation order can influence the respondent's choice (order effects). Simple randomization, however, can be problematic in cases where a more specific question preceeds a more general question. This can bias responses to the general question, a phenomenon known as the subtraction effect [Tourangeau et al. 1991]. In the OLEI scale, the first response option, "General interest in topic", is more general than the other options and should therefore be placed first; if possible, the remaining items should be presented in random order to counterbalance order effects.

4. EXPLORING LEARNER MOTIVATIONS USING THE OLEI SCALE

In this section, we use the OLEI scale to address what motivates learners in MOOCs (RQ1) and the extent to which their motivations vary with demographic characteristics and across courses (RQ2). To further understand learner motivations, we examine which motivations tend to be expressed in pairs and whether an underlying structure in motivations emerges.

[2]Truly construct-specific response options would require asking respondents to rate each motivation individually using different response scales with labels that reflect the underlying construct of each motivation. This is, however, not feasible in this context and would drastically increase the length of the scale. A smaller step toward construct-specific response options could use a 5-point unipolar scale from 'does not describe me at all' to 'describes me a great deal' and rephrase the question to 'To what extent do the following statements describe your reasons for enrolling in this course?' Nevertheless, in the interest of reducing the cognitive burden on respondents and simplifying the scale design as well as the collected data, we opted for the dichotomous scale.

[3]As ordering does not influence the more fundamental tendency to be agreeable, acquiescence bias is not mitigated by randomizing the order of response options.

Table II. Summary information on courses and surveys used in the analysis.

ID	Topic	Date	Platform	Enrolled	Surveyed	Female	Age[a]	College[b]
C1	Biology	Winter 2014	OpenEdX	9,931	3,294	50%	33 (26, 46)	78%
C2	Computer Science	Fall 2013	Coursera	n.a.	2,618	14%	30 (25, 39)	79%
C3	Computer Science	Winter 2014	OpenEdX	28,278	11,933	25%	31 (26, 39)	83%
C4	Computer Science	Winter 2014	OpenEdX	20,720	4,941	16%	30 (25, 38)	73%
C5	Material Science	Fall 2013	OpenEdX	9,979	1,985	16%	31 (26, 43)	84%
C6	Mathematics	Winter 2014	OpenEdX	3,194	2,600	10%	29 (25, 35)	93%
C7	Mathematics	Winter 2014	Coursera	44,112	2,875	34%	34 (26, 47)	76%
C8	Medicine	Winter 2014	OpenEdX	4,411	2,788	91%	30 (25, 38)	83%
C9	Political Science	Fall 2013	Coursera	n.a.	1,506	46%	30 (26, 41)	83%
C10	Quantum Physics	Fall 2013	OpenEdX	7,812	1,503	15%	30 (25, 44)	81%
C11	Sociology	Fall 2013	Coursera	60,884	8,100	46%	34 (28, 44)	n.a.
C12	Sociology	Spring 2014	Coursera	34,963	5,637	43%	34 (27, 43)	90%
C13	Statistics	Winter 2014	OpenEdX	33,362	10,817	22%	33 (28, 42)	96%
C14	Writing	Fall 2013	OpenEdX	37,709	10,878	n.a.	n.a.	n.a.

Note: [a]median (lower, upper quartile) percentage of female learners, self-reported on survey; [b]proportion with college degree or above; n.a. data not available

4.1. Method

The OLEI scale was included in an optional course survey in 14 MOOCs offered by Stanford University through the Coursera and OpenEdX platforms. 71,475 responses were collected between Fall 2013 and Spring 2014. The surveys were distributed in the first few weeks of the course in order to sample from an active learner population. Links to the survey were sent out via email and announced on the course website.

Table II provides information on the courses, the number of enrolled learners, the number who responded to the optional course survey, and demographic information. The number of enrolled learners is the number of learners who signed up for each course either on Coursera or OpenEdX by creating an account and clicking the “enroll” button on the site. There was large variation in the gender balance of survey respondents across courses (e.g., 91% women in C8, but only 10% women in C6). We expect these gender differences in survey respondents are representative of a skewed gender distribution in overall course enrollment. Many courses were in subjects where men tend to be disproportionately represented (e.g., course C10 is on Quantum Physics, a stereotypically male-dominated field) [Foundation 2013], whereas one course appeared to be targeted at women (course C8 included “Women’s Health” in its official title).

Learners chose an average of 6.3 out of the 13 enrollment intentions on the OLEI scale and those few who selected all or none of the items were excluded in the analysis (including them yielded quantitatively similar and qualitatively equivalent results). In a final validation step to check if the OLEI scale actually included the most common enrollment intentions, learners were given the opportunity to report any other reasons for enrolling in the 14 course surveys. Although 7% of learners responded to the question, responses were either repetitions or alternative phrasings of reasons on the OLEI scale, or comments specific to a course (e.g., “to learn R”). Hence, no additional modifications to the OLEI scale were made.

4.2. Overall Distribution of Motivations

To investigate what motivates learners in MOOCs (RQ1), we compared the percentages of learners in each course who reported each enrollment intention (Table III). For each enrollment intention, we computed the median and interquartile range (IQR) across courses. We report medians and IQRs instead of means and standard deviations because the data had a skewed distribution. The median can be interpreted to reflect the popularity of an enrollment intention in a typical course, given that the intention is more popular in half of the other courses and less popular in the other half. In comput-

Table III. Distribution of learner motivations across courses. Percentage of learners who selected each enrollment intention in each course and the median and interquartile range across courses.

Enrollment Intentions	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	Median	IQR
General interest	3	94	90	87	2	91	91	95	87	2	3	90	93	9	89	85
Job relevant	61	47	78	65	46	56	41	57	40	66	22	75	78	21	56	22
School relevant	69	36	31	34	55	47	32	39	33	58	66	25	33	40	38	18
Research relevant	69	34	27	26	55	65	32	35	36	53	60	22	48	15	35	24
Growth/enrichment	4	90	90	87	4	85	91	94	86	3	3	91	89	5	86	85
Career change	74	20	44	33	55	21	18	27	25	70	56	36	35	65	36	30
Fun/challenge	9	80	70	68	18	72	80	75	59	10	26	65	73	38	66	40
Meet new people	66	13	23	25	57	12	14	45	24	69	68	25	13	73	25	43
Experience online	42	36	47	47	40	29	43	54	46	46	42	46	34	50	44	6
Earn certificate	46	49	56	56	40	38	32	66	55	45	38	56	44	39	45	16
Prestigious uni/prof	45	62	58	56	32	61	48	59	64	35	31	59	61	41	57	17
Take with others	70	10	15	16	69	16	9	32	14	74	78	12	20	72	18	55
Improve English	62	24	28	27	57	19	27	17	39	64	57	35	18	14	28	26

ing these summary statistics, each course was given equal weight in order to discount for variation in the number of learners in each course.

Two insights into learner motivation emerged from the overall distribution of motivations and across different courses. First, we gain a novel perspective on learners' motivations for taking MOOCs using a formally constructed instrument.

Earning a certificate of completion is the goal that is commonly emphasized in current MOOCs and expected to motivate many learners. And yet, in a typical course, fewer than half of the learners (45%) indicated an intention to earn a certificate.

Aside from credentialing, most MOOCs contain a discussion forum which allows learners to interact with each other to some extent. A quarter of learners expressed an intention to meet new people in the course. Another social motivation reported by close to 20% of learners in a typical course is to take the course with friends or colleagues. The prevalence of social reasons to enroll is noteworthy, in particular as early MOOCs failed to provide the tools to enable rich social interactions between learners.

Another set of motivations was related to academic circumstances or ambitions. In a typical course, 38% of learners were motivated to enroll because the course was relevant to school or their degree program, and 35% because of relevance for their academic research.

Aside from academic reasons, learners can be motivated by vocational reasons. Over half of learners in a typical course (56%) reported being motivated due to relevance to their job and over a third (36%) were motivated by aspirations to change careers.

Given the global audience of these courses with many learners for whom English is not their first language, learners can be motivated to improve their language skills by taking the course. Over a quarter of learners in a typical course (28%) were motivated to improve their English.

The current generation of MOOCs is frequently offered by professors from prestigious institutions. Over half of the learners in a typical course (57%) reported being motivated by the prestige of the instructor or institution.

MOOCs received a considerable amount of media attention which certainly sparked interest in experiencing an online course firsthand. Almost half of the learners in a typical course (44%) were motivated by their curiosity to experience an online course.

A final set of motivations reflects different aspects of interest, including general interest in the topic, a desire for personal growth and enrichment—common characteristics of lifelong learners—and the motivation to have fun and feel challenged. Two thirds of learners in a typical course (66%) were motivated by an expectation to have fun and be challenged. Almost 9 out of 10 learners reported general interest in the topic (89%) and a desire for growth and enrichment (86%) as a motivating force, though there was considerable variability across courses, as discussed next.

The second insight was that different courses appeared to attract learners with different motivations; or alternatively, properties of the course (e.g., course promotion) induced different motivations in learners. For instance, many learners enrolled in a course on mathematics for fun and to be challenged (80%), while fewer learners indicated this intention for enrolling in a course on quantum physics (10%). Moreover, the variation in the percentage of learners who stated a particular intention differed substantially across courses, which is reflected in differences in IQRs between enrollment intentions. For example, intentions to experience an online course or earn a certificate were consistently reported by close to half of the learners with interquartile ranges of 6 and 16% points, respectively. In contrast, the variation in general interest in the topic or seeking growth and enrichment intentions was substantially larger with interquartile ranges of 85% points.

4.3. Demographic Trends in Motivation

Following the investigation of overall trends in motivations, we examined individual differences in learner motivations by gender, education, and age (RQ2). When viewed across all 14 courses, gender differences in motivation were small—under 5% point difference between the proportion of females and males selecting any given enrollment intention (Table V in the electronic appendix). There was, however, some variability across courses. For instance, in the medicine and political science courses, women were more likely to choose most enrollment intentions, suggesting that men chose somewhat fewer enrollment intentions than women did, $t(60595)$ = 6.13, $p \ll .001$, Cohen's d = 0.05.

Individual differences in motivations by learners' level of education were more pronounced than gender differences (Table VI in the electronic appendix). Learners who held a college degree or a more advanced degree (*more schooling*) more frequently reported enrolling due to relevance to their job than those with some college or less (*less schooling*; 11.2% point difference). However, learners with less schooling more frequently reported enrolling due to relevance to school (14.9% pt.), to experience an online course (9.5% pt.), improve their English language skills (9.1% pt.), and to earn a certificate (8.2% pt). There were notable exceptions to these trends, such as the material science and quantum physics courses in which learners with more schooling more frequently enrolled due to school relevance than those with less schooling (16% pt. and 21% pt., respectively). The observed differences were likely an artifact of the way the population was split, such that current college students were compared to learners with bachelor and master degrees who were more likely to be employed.

As indicated in the summary statistics in Table II, the age distribution across the 14 courses was very similar. This trend extended to individual differences in motivations by age, which were very minor, except for the fact that learners who reported enrolling due to relevance to school or their degree program were three years younger on average than those who did not report this intention.

4.4. Motivation Structure

The previous two sections addressed research questions about what motivates learners and the presence of notable individual differences. In this section, we investigate the structure of these motivations; specifically, relationships between motivations and whether the 13 enrollment intentions are reducible to a smaller number of key motivators. Table IV provides a correlation matrix between enrollment intentions across the

Table IV. Correlations between enrollment intentions based on a linear combination of each course's correlation matrix. Absolute correlations greater than .10, .20, and .30 are color-coded.

Enrollment Intentions	1	2	3	4	5	6	7	8	9	10	11	12	13
1 General interest	1.00	0.03	-0.03	-0.03	0.20	0.01	0.21	0.02	0.05	0.03	0.06	0.01	-0.01
2 Job relevant		1.00	0.11	0.15	0.02	0.06	-0.04	0.02	0.00	0.08	0.06	0.06	-0.01
3 School relevant			1.00	0.43	-0.03	0.06	-0.06	0.07	0.07	0.15	0.09	0.08	0.11
4 Research relevant				1.00	-0.04	0.04	-0.09	0.09	0.06	0.10	0.09	0.08	0.11
5 Growth/enrichment					1.00	0.05	0.23	0.06	0.09	0.09	0.11	0.03	0.04
6 Career change						1.00	0.02	0.16	0.07	0.19	0.11	0.06	0.11
7 Fun/challenge							1.00	0.08	0.13	0.05	0.09	0.06	0.01
8 Meet new people								1.00	0.22	0.21	0.17	0.22	0.21
9 Experience online									1.00	0.18	0.16	0.15	0.23
10 Earn certificate										1.00	0.28	0.11	0.15
11 Prestigious uni/prof											1.00	0.09	0.14
12 Take with others												1.00	0.10
13 Improve English													1.00

14 courses[4]. Absolute correlations below .2 are commonly labeled as "weak" and ones between .2 and .4 as "moderate".

The correlation matrix revealed a moderate positive correlation between relevance to research and relevance to school and one's degree program (r = .43), which could reflect the enrollment intentions of graduate students in particular. The prospect of earning a certificate and the prestige of the university and professor were also positively correlated reasons (r = .28). Moreover, enrollment for general interest in the topic, for personal growth, and for fun and challenge were positively correlated with each other (between r = .20 and r = .23). Learners who intended to meet new people were also motivated to experience an online course, earn a certificate, take the course with friends and colleagues, and improve their English (between r = .21 and r = .22). All correlations were precisely estimated and significantly different from zero, $p \ll .001$.

Most enrollment intentions were only weakly associated with other enrollment intentions, which suggests that there is not a strong basis for an underlying structure in learner motivations. It is also a positive indicator that the response options on the OLEI scale were largely mutually exclusive. We used factor analysis and principal component analysis to evaluate the feasibility of reducing the space of 13 enrollment intentions to a smaller set. A scree plot analysis was performed to determine the number of factors or principal components (PC) that would be optimal given the covariance structure of the data. If the 13 enrollment intentions loaded on a small number of factors, it would suggest that there is a small set of key motivations that is representative of the whole, and thus, the dimensionality of subsequent prediction tasks could be reduced.

The scree plot analysis suggested that the optimal number of factors is six. A factor analysis of the combined correlation matrix with six factors accounted for only 30% of the variance in enrollment intentions. This suggested that the six-factor structure was not an adequate representation of the data. A principal component analysis lent further support to the finding that the space of enrollment intentions was not easily reducible. The first principal component accounted for 21% of variance, and it took 4 (8) PCs to account for over 50% (80%) of variance. This provided strong evidence against reducing the dimensionality of the enrollment intention space a priori.

The resistance to reduction demonstrated by this data highlights the complexity of learner motivations. Enrollment intentions were mostly orthogonal with a couple

[4]Instead of a weighted correlation of all responses across all courses, the combined correlation matrix is an average of 14 correlation matrices, one for each course. This choice was informed by the observation that the average proportions for each reason vary substantially between courses and thus the average across courses is not a desirable normalizing factor for interpreting correlations.

exceptions. Complex systems have emergent properties that arise from the interaction of a multitude of dimensions, including interactions that are longitudinal and dynamic [Mitchell 2009]. In MOOCs, one emergent property is likely to be the extent to which experiences with course content and other individuals in the course meet the needs of learners who were originally motivated to enroll. How do the observable behaviors of learners reflect this complex interplay of intentions and experiences? We employed a predictive model to investigate the extent to which motivations are predictive of behaviors in MOOCs (RQ3).

5. PREDICTING LEARNER BEHAVIORS WITH ENROLLMENT INTENTIONS

The previous sections shed light on what motivates learners in online courses and the structure of their motivations. In this section, we address the research question about which enrollment intentions are predictive of certain behaviors and how predictive they are (RQ3). Although learners' enrollment intentions are certainly not fixed, they reflect motivations at a moment in time, and one would expect this snapshot to be informative of subsequent behaviors in the learning environment. For a number of enrollment intentions, we can construct concrete hypotheses with respect to the behaviors we expect learners with these motivations to engage in.

The first two hypotheses are based on direct implications of learner motivations for observable activities in the courses. Learners who intend to earn a certificate are expected to engage more in activities required to pass the class than learners who do not express this intention. Learners who intend to meet new people are expected to engage more on the discussion forum than those who do not express this intention.

H1. The intention to earn a certificate is predictive of taking more assignments, watching more lectures, and an increased likelihood of earning a certificate.
H2. The intention to meet new people is predictive of posting more on the discussion forum.

A third hypothesis is derived from a large literature on human learning as an inherently social process. Peer learning across a variety of settings is beneficial for deeper understanding and continued engagement [Springer et al. 1999; Johnson et al. 2000; Darling-Hammond et al. 2008]. People make sense of the world by developing shared perspectives and helping one another solve problems [Vygotsky 1986]. While the platforms in this study have limited support for collaboration, learners may also surround themselves with peers offline and experience the benefits of social learning outside of the MOOC. In addition to providing extra support for one another, these learners may encourage each other to keep on schedule with the course. Accordingly, we would expect learners who report taking the course with friends or colleagues to engage more with the learning materials.

H3. The intention to take the course with colleagues or friends is predictive of taking more assignments and watching more lectures.

For the remainder of the motivations we do not have theory-driven hypotheses. Instead, we formulate a number of research questions about the ways certain motivations shape learner behavior, following the same thematic grouping used in the description of results in Section 4. First, it is unclear how learners with enrollment intentions related to academic motivations, namely relevance to school and degree program, and relevance to academic research, use online courses differently than other learners without these intentions. We thus pose the following research question:

RQ3a. How do academic motivations predict learner behavior?

Second, it is unclear how learners with vocational motivations, namely relevance to their job and to change careers, would use online courses differently than those without these intentions. Therefore, we pose the following research question:

RQ3b. How do vocational motivations predict learner behavior?

Third, it is unclear how learners motivated by the prospect of improving their English language skills interact with online courses that were not designed for language learning. Accordingly, we pose the following research question:

RQ3c. How does the motivation to improve one's English predict learner behavior?

Finally, the set of interest-related motivations—general interest in the topic, growth and enrichment, and for fun and challenge—may or may not be predictive of more different behaviors in the course. We thus pose the following research question:

RQ3d. How do interest-related motivations predict learner behavior?

The following investigation of these hypotheses and research questions sheds light on whether learner motivation measured through enrollment intentions can serve as a lens for understanding learners in a way that empowers system designers or course instructors to adapt to learners' needs.

5.1. Method

We investigated the relation between enrollment intentions and learner behavior in ten MOOCs using the OLEI scale to capture enrollment intentions. The ten courses are a subset of those described in the previous section, because some courses were ongoing at the time of analysis and some course surveys were anonymized to the extent that they could not be linked with behavioral course data.

We selected a set of behavioral measures which capture learners' progress in the course, their general performance, and social engagement on discussion forums. In order to generalize across courses and platforms, which varied in terms of course length, requirements, and norms of engagement, we considered relative progress milestones, such as watching more than half of the video lectures in the course. Progress in the course was quantified by three milestones for the proportion of watched video lectures and the proportion of attempted assignments: the learner attempted over 10% (50%, 80%) of assignments available in the course (excluding in-video quizzes); the learner attempted over 10% (50%, 80%) of lecture videos available in the course. The binning was done prior to analysis and was data-independent to avoid introducing biases. The reason for discretizing the continuous variables into binary indicators of cumulative levels (i.e. over 10, 50, and 80%) was that the distribution of the continuous variables was multi-modal. This form of cumulatively discretizing a non-standard distribution provides a valid alternative to parametric modeling strategies, which require additional assumptions [Angrist and Pischke 2008].

A robust measure of performance was particularly difficult to find given the large variation in assessment types and grading procedures. Instead, as a measure of satisfactory performance on the assessment tasks we used whether the learner earned a certificate of completion. Social engagement was quantified by two measures of activity on discussion forums, and a measure of endorsement by the learner community: the learner authored one or more posts/comments on the discussion forum; the learner authored over half as many posts/comments as the most prolific forum posters in the course; and the learner received one or more net votes on the forum.

Among numerous candidate methods for analyzing these data, we opted for a bootstrapped linear regression with main effects for each enrollment intention and each course. This choice was motivated by the specific research questions and properties of the data. In each model, a binary behavioral outcome was predicted by binary variables for each enrollment intention and each course (no intercept was included). Variation between courses was not an object of interest for the present inquiry, but needed to be accounted for, as observations were not independently distributed, but clustered within courses. Courses were modeled as fixed effects as it was computationally not feasible to model them as a random effect. The previous section reported correlations between enrollment intentions, which were mostly low ($|r| < .2$) and in a few cases moderate in magnitude ($.2 < |r| < .4$). This suggested that issues of collinearity would be minor at most.

The linear probability model included main effects for enrollment intentions and courses. No interactions were considered for reasons of parsimony and to ensure interpretability. The model was specified as

$$y_i = \alpha_1 why_{1i} + ... + \alpha_{13} why_{13i} + \beta_1 course_{1i} + ... + \beta_{10} course_{10i} + \epsilon_i$$

where, for each learner i, y_i is a binary outcome indicator (0 or 1), why_{1i} to why_{13i} are binary indicator for whether i selected each enrollment intention, and $course_{1i}$ to $course_{10i}$ are binary indicators for each course. The values depecited in Figure 1 correspond to the coefficients on the enrollment intentions ($\alpha_1, ..., \alpha_{13}$).

A linear probability model instead of logistic regression was employed because the performance of linear fixed-effects models under a range of misspecifications is better understood than for logistic regression, which suffers from aggregation bias if individual differences remain unmodeled [Angrist and Pischke 2008]. In order to ensure the stability of our estimates (i.e. robustness to slight changes in the data) and to select important predictors, the model was fit 10,000 times to bootstrapped data (sampling from the original data with replacement). The final estimate of each predictor's association with each behavioral outcome was the average of 10,000 bootstrapped regression coefficients. The 0.5th and 99.5th percentile of the bootstrapped coefficients provided empirical 99% confidence intervals. Given a sample of close to 44,000 cases, predictors that were significant at 99% were considered important independent of their effect size.

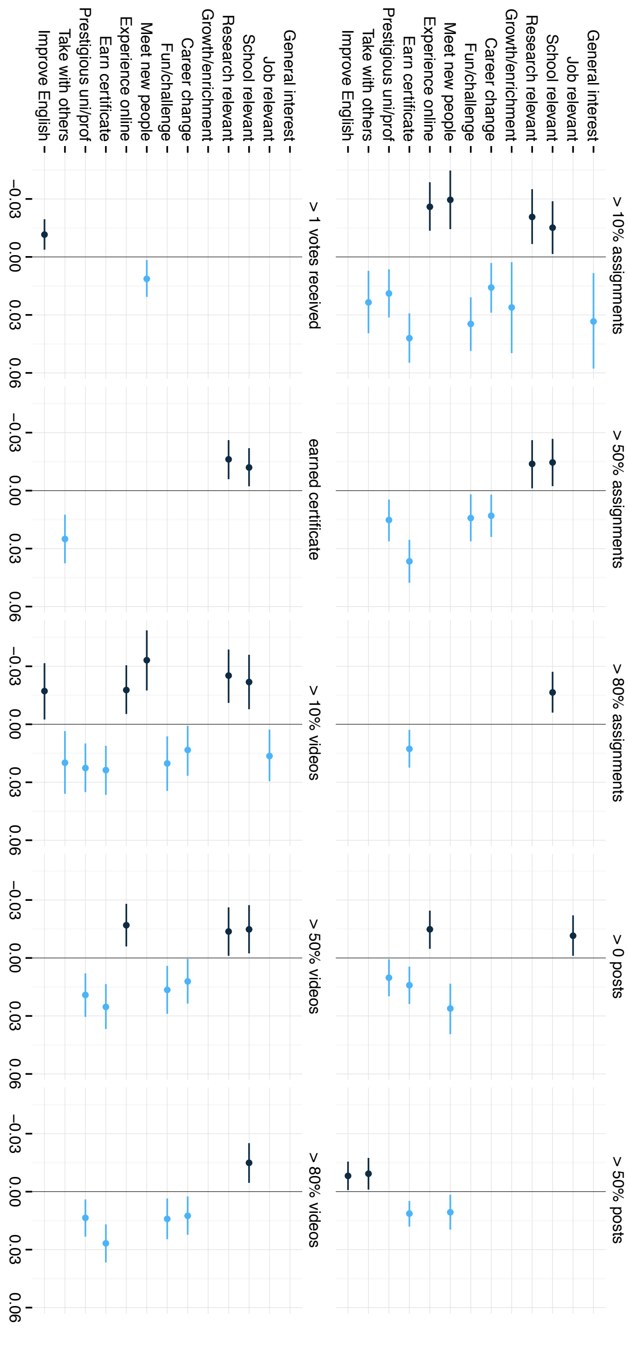
5.2. Results

To address which motivations are predictive of behaviors in MOOCs and how predictive they are (RQ3), we estimated changes in ten behavioral outcome measures predicted by each enrollment intention (Figure 1). Error bars are 99% confidence intervals and estimates that were not significant at this level were omitted for clarity (for the complete list of coefficients and confidence intervals, see Tables VII, VIII, IX in the electronic appendix). For incremental behavioral outcomes, such as attempting 10, 50, and 80% of assessments, it is possible for an association to be strong and stable enough to be selected for lower increments (e.g., $>$ 10% of assessments) but not higher ones (e.g., $>$ 50% of assessments). This may occur because a smaller number of people achieved the higher increment, which reduced the statistical power with which significant associations could be identified.

Each enrollment intention was found to be predictive of at least one of the behavioral outcome measures. The predicted change in the probability of achieving these behavioral outcomes varied between -4.5% and 6% with 99% confidence. Note that predicted changes are additive. For instance, learners who intended to earn a certificate and also took the course with friends or colleagues—intentions that were moderately correlated (r = .2)—were approximately 6% more likely on average to attempt over 10% of assessments.

We hypothesized that the intention to earn a certificate is predictive of taking more assignments, watching more lectures, and an increased likelihood of earning a certifi-

Fig. 1. Change in probability of achieving ten binary behavioral outcomes predicted by enrollment intentions. Estimates computed using a linear model with fixed effects for each of ten courses and bootstrapped 99% confidence intervals. Points represent averages of bootstrapped coefficients and are omitted if not significant at 99% for variable selection. Changes in probabilities are additive across enrollment intentions.



cate (H1). Consistent with this hypothesis, learners who intended to earn a certificate were more likely to watch most video lectures and attempt most assessment in the course. Lectures provide learners with exposure to the topic of instruction whereas assessments provide an opportunity for monitoring their understanding of the topic and receiving feedback. Learners with the intention to earn a certificate were additionally more likely to engage actively on the discussion forum, which was previously found to be associated with behaviors that are on the pathway to earning a certificate [Kizilcec et al. 2013]. Counter to our expectation, the intention to earn a certificate was not predictive of an increased likelihood of earning a certificate. This could be a result of doing poorly on assessments, or not completing assessments by the deadline required to earn the certificate. This partial disconnect between the intention and actual behavior is discussed further in Section 6.2.4.

The intention to meet new people was hypothesized to be predictive of posting more on the discussion forum (H2). Consistent with this hypothesis, this social intention was a strong predictor of active engagement on the discussion board and even of receiving social recognition from peers in the form of votes. However, these learners were also less likely to engage with lectures and assessments, compared to learners who did not express the intention to meet new people. A possible explanation for this observation is that neither lectures nor assessments facilitated social interactions in the courses under observation. Hence, learners with the intention to meet new people appear to seek a social space to engage with like-minded people.

We further hypothesized that the intention to take the course with colleagues or friends is predictive of taking more assignments, and watching more lectures (H3). This intention relates to the social context in which learners engage with the course, rather than a particular outcome within the course. Consistent with our hypothesis, learners who intended to take the course with friends or colleagues were more likely to watch at least 10% of lecture videos, and attempt at least 10% of assessments. Moreover, these learners were even more likely to earn a certificate. Learners with this intention were, however, less likely to engage heavily on the discussion forum, possibly because they communicated via other channels with the people they took the course with. As discussed in Section 6.2.1, this reveals a promising trend for bolstering engagement through local cohorts taking courses together.

Two moderately correlated enrollment intentions reflect academic motivations, namely enrolling due to relevance to school or degree program and relevance to academic research (r = .43). We posed a research question about the extent to which these motivations predict learner behavior (RQ3a). Learners with academic motivations were less likely to watch video lectures and take assessments, and ultimately less likely to earn a certificate. This trend may suggest that learners with these motivations regard the learning environment as a set of resources available for reference, comparable to looking something up in a text book or on Wikipedia. Alternatively, these learners may seek exposure to a different perspective on a familiar topic but are not inclined to follow a course structure, as discussed in Section 6.2.2.

A second set of enrollment intentions describe vocational motivations, namely enrolling due to relevance to one's job or for a career change. We posed a research question about the extent to which vocational motivations predict learner behavior (RQ3b). Learners who enrolled due to job relevance were more likely to watch at least 10% of lecture videos, but less likely to actively engage on the forum. In contrast, learners who enrolled because they aspired a career change were more likely to watch over 80% of video lectures and complete over half of the assessments in the course. While learners who enrolled due to job relevance seemingly sought to learn new skills or better understand a topic by watching a few lectures, those who enrolled for career change appeared to be more committed to learning a new skill or understanding new concepts to serve them on their new career path.

The extent to which the motivation to improve their English predicts learner behavior is another research question of interest (RQ3c). Learners with the intention of improving their English skills were less likely to heavily engage on the discussion forum or receive votes from peers on their posts. They were also less likely to watch more than 10% of lecture videos. This may be a sign of a language barrier hindering learners' progress in the course. Opportunities to improve English come through communicating with others and engaging with course content, so it is concerning that many English language learners had lower activity levels in these areas. As discussed in Section 6.2.3 below, online courses could be designed to specifically support learners who are seeking to improve their English language skills.

A final set of three enrollment intentions describe interest-related motivations, including general interest in the topic, to gain personal growth and for enrichment, and for fun and challenge. We posed a research question about the extent to which these interest-related motivations predict learner behavior (RQ3d). Learners who enrolled out of general interest in the topic and those with the intention to gain personal growth and enrichment were more likely to attempt at least 10% of assessments. These learners appear to be less invested in fully engaging with the topic. In contrast, enrolling for fun and challenge was predictive of attempting more than half of the assignments and watching most lecture videos. Learners who enrolled for fun and challenge probably include lifelong learners who strive to learn new topics and learners who were already immersed in the topic and got pleasure from exploring it further.

6. DISCUSSION

The current work offers a lens for understanding the motivations of learners in open online learning environments. We developed the Online Learning Enrollment Intentions (OLEI) scale and applied it to explore the motivations of learners in 14 different MOOCs, the predominant example of the current generation of open online learning environments. Learners reported diverse motivations for enrolling in MOOCs, including to earn a certificate, to improve English skills, and a variety of social, academic, vocational, and interest-driven motivations. The prevalence of some motivations varied substantially across courses, which we interpreted to reflect differences in the target audience for these courses. The various motivations did not reduce to a smaller number of key motivators, but are each predictive of certain behavioral outcome measures representing milestones in course activities. We interpret the diverse behavioral patterns of learners as expressions of different needs that learners seek to gratify by engaging with the courses.

6.1. Designing for Success

Technology designers for open online learning environments provide generic, flexible tools that abstract from specific teaching and learning practices. The affordances of these tools shape the structure of the course content and activities developed by course designers. The technology and pedagogy of the majority of MOOCs today is strongly influenced by the "grammar of schooling"—the structures and rules which implicitly organize the work of instruction in traditional classrooms [Tyack and Tobin 1994]. Yet the diverse patterns of learner behavior observed in the current work indicate that while many learners are operating within the expectations of the grammar of schooling, many others have gone beyond institutionalized roles and are operating according to the evolving culture and norms of online media. With these two models at play, how might we understand if these designs are successful?

When defining success in MOOCs, do we adhere to a traditional, top-down model of authority, where the institution and instructors establish the standards for success? Or do we prefer a metaphor of consumption? The answer is between the extremes. Thinking about success more broadly than traditional definitions opens the gates to understanding learners' perspectives more deeply and, instead of adhering to the constraints of physical classrooms, designing in terms of the potential of online technology. As demonstrated in the next section, understanding the needs of learners can inform design directions that better meet the varied needs of learners; and in some cases, support them in achieving instructor-defined goals such as completing the course with a certificate.

6.2. Data-Driven Design Choices for Open Online Learning Environments

From the diverse behavior patterns of learners, we can infer different needs that learners seek to gratify by engaging with the course. Four of these needs communicate clear signals for design changes to the current generation of open online learning environments. Two needs stand out across multiple motivations—the need for social engagement and the need for well-organized and accessible course content—and two follow directly from the apparent mismatch of particular intentions expressed by learners and their behavior—learning English and earning a certificate.

6.2.1. Learning with Peers. The OLEI scale contains two items related to social engagement and we observed distinct associated behaviors with each one: learners who enrolled in a course with colleagues or friends were more likely to be engaged with course materials (watching videos and taking assignments) than those who did not enroll with others, and learners who enrolled to meet new people participated on the discussion forum at relatively high rates relative to those without the intention to meet new people. The need for social engagement is seen throughout the literature on education, online learning, and online communities; we draw on much of this prior work to suggest design directions for open online learning environments.

Learners who enroll in groups could range from casual autodidacts to students in a brick-and-mortar institution using the MOOC materials for support in a particular class; distinguishing among these groups is not possible in our dataset. The composition and study habits of groups of learners who enroll together is an area for further research. A group can mutually reinforce engagement by holding one another accountable or simply by playing the role of enthusiastic interlocutors. In one case study, learners who enrolled in a MOOC as part of a traditional course preferred conversations and problem-solving with their local peers over participating on the discussion board [Bruff et al. 2013]. We suggest some simple initial steps in supporting groups who enroll together—particularly those that are not guided by an instructor in a traditional educational setting. First, it should be possible for learners to self-identify as a group taking the course together. This physically co-present group may also benefit from its own shared space or discussion area on the site. Second, if there was a way for each individual to specify goals for the course in terms of progress through materials or other quantifiable behaviors, a group dashboard could allow members to hold each other accountable to their goals. Finally, the course designer could also include guidance for offline activities or discussions among these groups.

Introducing more social features throughout the platform would better meet the needs of distributed learners who are seeking a social space within open online learning environments. One approach could be to create a social environment independent of any particular course. The existence of such a community would allow people to develop and maintain connections with like-minded individuals over time, while sharing resources on related topics or advice for getting the most out of open online learning environments. The design of this environment should look to prior work on supporting ongoing online communities, which emphasizes trust, group norms, and reciprocity among participants [Kraut et al. 2012].

Within a course, distinguishing between types of social interactions will be important for targeting the design of particular technologies. Discussion forums currently serve as a catch-all for personal connections between learners, homework help, and in-depth conversations about course topics. Personal connections could be further enabled by features such as learner profiles and the ability to search for others with similar interests or geographic locale. Homework help, on the other hand, would likely benefit from a well-organized question-and-answer forum, following the model of successful Q&A communities. Research on the effectiveness of Stack Overflow suggests that, in addition to design decisions, community organization and its regular involvement were key contributors to the success of this Q&A platform [Mamykina et al. 2011].[5]

In-depth conversation around course topics could be further supported in various ways, depending on whether these conversations were synchronous or asynchronous. Asynchronous conversations could be more closely linked to content with annotations of particular moments in lecture videos or textbooks. Without the need to switch contexts to the discussion forum, learners may be more likely to ask questions or to help others. Synchronous discussions embedded in lectures—asking students to work through a problem in small groups, then polling the room for potential answers, and finally explaining the right solution [Crouch and Mazur 2001; Smith et al. 2009]—is one approach that has seen positive effects in classrooms and has been adapted successfully to MOOCs [Lim et al. 2014]. More deeply incorporating social features into the presentation of instructional content could be leveraged to better support learning.

Learners who intend to meet new people may be particularly inclined to participate in group projects, which can promote the development of critical thinking and collaboration skills—valuable competencies in educational settings and the workforce [Darling-Hammond et al. 2008; Pellegrino et al. 2013]. While collaboration could occur via existing external tools, we see two key advantages of keeping group communication inside the learning environment. The first is that collaborations can be observed, which would allow researchers and designers to continue to develop a deeper understanding of learner needs and behaviors. The second is that collaborations can be scaffolded according to principles of productive group work derived from in-person environments, including assigning each individual to a role or providing "scripts" for the steps that a project should take [Johnson et al. 2000; Fischer et al. 2013]. The choice of communication medium influences the richness of possible interactions; recent work has tested the potential of group video chats [Cambre et al. 2014]. Open questions around small, distributed groups include strategies for team composition—for example, regarding the relative benefits of teams with more or less geographic and cultural diversity [Kizilcec 2013]—and strategies to take when group members disengage from the course.

6.2.2. Modularization to Promote Accessibility and Organization of Course Content. Learners for whom course content was relevant to their current academic endeavors (school

[5] Such a system would require a mechanism for aggregating responses and assessing their usefulness, as well as a reputation system for incentivizing contributions and providing more information on authors of answers. Experimental deployments of simple reputation systems in MOOC forums have led to increased engagement [Coetzee et al. 2014; Anderson et al. 2014], but more research is needed on the appropriate framing of incentive structures in this setting. While some learners reported that they were more motivated by altruism to help their peers than by earning reputation points [Coetzee et al. 2014], findings from randomized experiments suggest that appeals to collectivism could reduce engagement in MOOC discussion forums [Kizilcec et al. 2014].

or research) watched fewer video lectures and attempted fewer assignments. These behaviors suggest that learners with these academic intentions use the environment as a reference source. This may demonstrate a need for on-demand, well-organized information, which would be at odds with the presentation of MOOC content as a course in two ways. First, courses have a set start and end date, which means that the course content is only available for a limited period of time. On the one hand, courses with a start and end date create a cohort of learners that supports discussions and peer assessment, as well as a schedule which can be useful for staying organized and focused. On the other hand, an always-available archive enables quick access to specific content, which may be particularly beneficial for those who are motivated by the course's relevance to their academic endeavors. A recent comparison of the same MOOC offered as a cohort-based course and as a self-study course suggests that the self-study format is a viable alternative [Mullaney and Reich 2014].

Second, courses are organized with a pedagogical intent of guiding the learner chronologically through materials. However, learners who are coming to the MOOC as a reference source may find it easier to navigate by another organization scheme, such as the concepts that are covered in each video and assignment. Breaking apart MOOC content into separately tagged modules would allow reference-style usage. It also has the added benefit of enabling remixing and sharing of pieces of the MOOC—a function central to the adaptation of course content to local instructional contexts [Bruff et al. 2013].

Conceptualizing the MOOC as an archive positions the content in a larger information ecosystem. Individuals who are seeking granular information within a given MOOC are likely to be seeking related information elsewhere on the Web, or in different MOOCs, and this behavior could be supported by tools for linking and organizing content outside of any particular course. The vision of personal knowledge management has prevailed since Vannevar Bush described "trails" among related content in an integrated information space [Bush 1945]—and while it may be far-fetched to suggest that open online learning environments could play the role of the Memex, it is a vision to keep in mind as we continue to create structured opportunities for sharing information online.

6.2.3. Learning English. Accounting for language skills in courses that reach a global population of learners can be critical. Nearly thirty percent of survey respondents in a typical course reported the desire to improve their English skills, but it is unclear from the available data whether these learners were able to learn more English. We do know, however, that they were significantly less likely to participate in course activities that may have helped them reach this goal, specifically watching videos and participating on the discussion forum. The digital nature of online learning environments allows designers and instructors to meet the multiple goals that individuals have for enrolling in the course by adding a layer of content that is only applicable for some individuals.

Individuals who are actively seeking to improve their English skills can be supported through integrating features specifically to facilitate language learning. The literature on language learning and instruction is vast (for a review, see Bygate et al. 2013), but two simple suggestions directly address the behaviors which these learners were less likely to engage in. First, courses could provide video transcripts and subtitles in multiple languages. While some course providers currently offer transcripts, Enligsh language learners would likely benefit from subtitles that have been specifically designed for vocabulary learning. For example, a subtitling tool that enables learners to pause the video and ask for a translation is a promising direction which could be adapted to open online learning environments [Kovacs 2013].

Second, speaking practice is as important as comprehension, and the global population of participants could be a good source for conversation partners via video chat. A study on assigning learners into small groups to discuss course content via video chat has yielded promising results [Cambre et al. 2014], and this framework could easily be adapted to conversations designed to support language learning. Each of these design directions would also give English language learners greater access to course content and to the other learners in the course. More generally, the finding that many learners enroll in MOOCs to improve their English raises the question if this trend generalizes to learning other languages in, for instance, German or Chinese MOOCs.

6.2.4. Learners Seeking Certificates. The MOOCs in our study offered certificates of completion to learners who completed the majority of assignments with a passing grade. Learners who reported the intention to earn a certificate were not more likely to actually earn a certificate than those who did not intend to earn one—despite the fact that they were more likely to attempt assignments. One interpretation of this surprising result is that some learners who did not initially intend to earn a certificate became more engaged with the course than they had initially expected, and eventually earned a certificate. A second interpretation may be that learners who stated their intention to earn a certificate were more extrinsically motivated than others, and chose to do the bare minimum on assignments in order to earn the certificate, rather than engaging deeply with the materials [Lepper et al. 1973].

Finally, it could be the case that some certificate-seekers did not complete assignments on time or performed too poorly to earn a certificate. This could potentially reflect lower levels of preparedness or overconfidence from the learner's side. If further research found support for this interpretation, the self-directed learning literature describes processes that could be built into the technology in order to provide extra guidance. These processes include "diagnosing learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies and evaluating learning outcomes" [Knowles 1975, p. 18]. For example, learners could be given tools for mapping goals and motivations to concrete actions in the environment [Kicken et al. 2008] or to-do lists could be automatically generated by the system [Cheng et al. 2013]. In terms of feedback for learners, the design of individual dashboards is an important area for future research, as there are open questions about what data is useful to learners and how to best represent data to help learners understand what they already know and what steps they can take to make progress [Bienkowski et al. 2012].

6.3. Limitations

It is worth explicating certain limitations of the approach taken in the current work. First, data on learner motivations was collected for learners who responded to optional course surveys. These learners self-selected into taking the survey which was distributed at the beginning of each online course, and they were probably more engaged with the course in general. Thus, the predictions about behaviors may be an overestimate relative to the full population of learners with the same intentions, due to survey response bias (see [Kizilcec 2014], for preliminary work on non-response bias in MOOC surveys). Moreover, the validity of these self-report data relies on learners' thoughtful and truthful survey responses.

Given that learners did not report their enrollment intentions at exactly the time of enrollment, it cannot be ruled out that learners were affirming or rationalizing the behavior they had already practiced in the course. Surveys were distributed in the first three weeks of the course, which gave many learners time to familiarize themselves with the course materials and develop an informed opinion about the course.

Although this may have influenced learners' responses, we expect their responses to be a good measure of their motivations, because the enrollment intentions question specifies a fixed point in time and the concrete act of enrolling, which leaves less room for interpretation.

Another limitation of the current work is that conclusions are based on data that was exclusively collected in MOOCs. Although courses on a variety of topics in different academic disciplines were included in the analysis, MOOCs represent a specific incarnation of open online learning environments. While this challenges the generalizability of our findings to other open online learning environments, we would expect the OLEI scale to provide a robust measure of learner intentions in future generations of such environments. This prediction is based on the iterative development strategy and the large number of open responses from a set of learners with diverse backgrounds that were employed in determining the final set of intentions for the scale.

7. CONCLUSION AND FUTURE WORK

A central takeaway of this work is that learner motivations are a useful lens for understanding learners. They are predictive of actual behaviors and inform design directions. Prior work has highlighted discrepancies between learners' actual behaviors and those behaviors that would align with instructor-defined goals. In this study, only about half of the learners surveyed in a typical course—an already more engaged sample—reported enrolling with the intention to earn a certificate of completion. The next generation of open online courses—massive or not—should better account for learner motivations in their designs. For instance, a substantial number of learners were found to take online courses for social reasons, despite the fact that the learning experience was designed primarily for individuals. Many online courses separate learning resources and assessments from any of their social features, even though learners can be an invaluable resource to each other. Neither watching lectures, nor taking assignments, and not even peer grading provide affordances for social interactions in the current generation of MOOCs.

Allowing individuals to choose how to engage with courses is another strategy for supporting the diversity of learner needs. A fundamental tradeoff in designing digital learning environments is the extent to which learners are guided through a set of course content and assignments toward specific learning goals, as opposed to being granted autonomy to make their own choices. Providing online learners with information on the costs and benefits of multiple pathways could strike the appropriate balance between autonomy and guidance; for instance, noting that assessments take an extra hour per week but provide feedback which is valuable for learning.

In order to enable comparisons across studies and translation of findings across open online courses, future research should report enrollment intentions using the OLEI scale as a standardized metric—in addition to the commonly reported demographic distributions of learners. Future work should investigate learner motivations for those who do not self-select into optional surveys to identify if these learners have additional motivations currently not included on the OLEI scale.

Combining motivational with behavioral data provided insights into learners' needs, and future work should examine other factors that affect both behavioral choices and motivations. Beyond demographic differences, levels of prior knowledge [Kalyuga et al. 2003] and cognitive ability, preference, and style [Mayer and Massa 2003] are other individual differences between learners that have been identified in the literature as important mediators of the effectiveness of instructional techniques. A third important factor, as discussed above, is the extent to which learners in open learning environments are self-directed and possess the metacognitive skills to know what activities they need to engage in to achieve their individual goals [Garrison 1997]. A final topic for further investigation is the extent to which the suggested design changes for open online courses could actually support learners with specific motivations. Measuring and accounting for individual differences and supporting learners accordingly could go a long way toward improving learning experiences and the accessibility of open online learning environments.

ACKNOWLEDGMENTS

The authors would like to thank Jeremy Bailenson, Candace Thille, and our three anonymous reviewers for their helpful comments. We thank Stanford's office of the Vice Provost for Online Learning for supporting this research. Emily Schneider was supported by a Stanford Graduate Fellowship.

REFERENCES

Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*. International World Wide Web Conferences Steering Committee, 687–698.

Joshua D Angrist and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Tara S Behrend, David J Sharek, Adam W Meade, and Eric N Wiebe. 2011. The viability of crowdsourcing for survey research. *Behavior research methods* 43, 3 (2011), 800–813.

Marie Bienkowski, Mingyu Feng, and Barbara Means. 2012. Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology* (2012), 1–57.

Jay G Blumler and Elihu Katz. 1974. *The uses of mass communications: Current perspectives on gratifications research*. Vol. 3. Sage publications Beverly Hills, CA.

John D Bransford, Ann L Brown, Rodney R Cocking, and others. 2000. *How people learn*. National Academy Press Washington, DC.

Derek O Bruff, Douglas H Fisher, Kathryn E McEwen, and Blaine E Smith. 2013. Wrapping a MOOC: Student Perceptions of an Experiment in Blended Learning. *Journal of Online Learning & Teaching* 9, 2 (2013).

Vannevar Bush. 1945. As We May Think. *The Atlantic Monthly* 176, 1 (1945), 101–108.

Martin Bygate, Merrill Swain, and Peter Skehan. 2013. *Researching pedagogic tasks: Second language learning, teaching, and testing*. Routledge.

Julia Cambre, Chinmay Kulkarni, Michael S Bernstein, and Scott R Klemmer. 2014. Talkabout: Small-group Discussions in Massive Global Classes. In *L@S 2014 First (2014) ACM Conference on Learning @ Scale*. ACM, 161–162.

Justin Cheng, Chinmay Kulkarni, and Scott Klemmer. 2013. Tools for predicting drop-off in large online classes. In *Proceedings of the 2013 conference on Computer supported cooperative work companion*. ACM, 121–124.

Derrick Coetzee, Armando Fox, Marti A Hearst, and Björn Hartmann. 2014. Should your MOOC forum use a reputation system?. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1176–1187.

David Colton and Robert W Covert. 2007. *Designing and constructing instruments for social research and evaluation*. John Wiley & Sons.

Catherine H Crouch and Eric Mazur. 2001. Peer instruction: Ten years of experience and results. *American Journal of Physics* 69, 9 (2001), 970–977.

Linda Darling-Hammond, Brigid Barron, P David Pearson, Alan H Schoenfeld, Elizabeth K Stage, Timothy D Zimmerman, Gina N Cervetti, and Jennifer L Tilson. 2008. *Powerful learning: What we know about teaching for understanding*. Jossey-Bass San Francisco.

J DeBoer, A D Ho, G S Stump, and L Breslow. 2014. Changing "course": Reconceptualizing educational variables for massive open online courses. *Educational Researcher* (2014).

Frank Fischer, Ingo Kollar, Karsten Stegmann, and Christof Wecker. 2013. Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist* 48, 1 (2013), 56–66.

National Science Foundation. 2013. Women, minorities, and persons with disabilities in science and engineering: 2013. Special Report NSF 13-304. (2013).

D Randy Garrison. 1997. Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly* 48, 1 (1997), 18–33.

Sara Goldrick-Rab. 2010. Challenges and opportunities for improving community college student success. *Review of Educational Research* 80, 3 (2010), 437–469.

Philip J Guo and Katharina Reinecke. 2014. Demographic Differences in How Students Navigate Through MOOCs. In *L@S 2014 First (2014) ACM Conference on Learning @ Scale*. ACM, 21–30.

Laura Horn, S Nevill, and J Griffith. 2006. Profile of undergraduates in US postsecondary education institutions: 2003-2004.(NCES 2006-184). US Department of Education. *National Center for Education Statistics. Washington, DC: US Government Printing Office* (2006).

David W Johnson, Roger T Johnson, and Mary Beth Stanne. 2000. Cooperative learning methods: A meta-analysis. http://www.ccsstl.com/sites/default/files/Cooperative%20Learning%20Research%20.pdf. (2000).

Slava Kalyuga, Paul Ayres, Paul Chandler, and John Sweller. 2003. The expertise reversal effect. *Educational psychologist* 38, 1 (2003), 23–31.

Wendy Kicken, Saskia Brand-Gruwel, and Jeroen JG van Merriënboer. 2008. Scaffolding advice on task selection: a safe path toward self-directed learning in on-demand education. *Journal of Vocational Education and Training* 60, 3 (2008), 223–239.

René F Kizilcec. 2013. Collaborative Learning in Geographically Distributed and In-person Groups. In *Proceedings of the Workshops at the 16th International Conference on Artificial Intelligence in Education, AIED 2013*. 67–74.

René F Kizilcec. 2014. Reducing non-response bias with survey reweighting: applications for online learning researchers. In *Proceedings of the first ACM conference on Learning@Scale conference*. ACM, 143–144.

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

René F Kizilcec, Emily Schneider, Geoff Cohen, and Daniel McFarland. 2014. Encouraging Forum Participation in Online Courses with Collectivist, Individualist, and Neutral Motivational Framings. *eLearning Papers* 37 (2014), 13–22.

Malcolm S Knowles. 1975. *Self-directed learning*. Association Press New York.

Geza Kovacs. 2013. Smart subtitles for language learning. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2719–2724.

Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. MIT Press.

Jon A Krosnick. 1999. Survey research. *Annual review of psychology* 50, 1 (1999), 537–567.

Mark R Lepper, David Greene, and Richard E Nisbett. 1973. Undermining children's intrinsic interest with extrinsic reward: A test of the" overjustification" hypothesis. *Journal of Personality and social Psychology* 28, 1 (1973), 129.

Seongtaek Lim, Derrick Coetzee, Bjoern Hartmann, Armando Fox, and Marti A. Hearst. 2014. Initial Experiences with Small Group Discussions in MOOCs. In *Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14)*. ACM, New York, NY, USA, 151–152. DOI:http://dx.doi.org/10.1145/2556325.2567854

Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. 2011. Design Lessons from the Fastest Q&A Site in the West. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2857–2866. DOI:http://dx.doi.org/10.1145/1978942.1979366

Richard E Mayer and Laura J Massa. 2003. Three Facets of Visual and Verbal Learners: Cognitive Ability, Cognitive Style, and Learning Preference. *Journal of educational psychology* 95, 4 (2003), 833.

Sandra D Mitchell. 2009. *Unsimple truths: Science, complexity, and policy*. University of Chicago Press.

Tommy Mullaney and Justin Reich. 2014. Staggered versus All-at-Once Content Release in Massive Open Online Courses: Evaluating a Natural Experiment. (2014). http://harvardx.harvard.edu/files/harvardx/files/ssrn-id2499729.pdf

Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge. 2014. Survey Research in HCI. In *Ways of Knowing in HCI*. Springer, 229–266.

James W Pellegrino, Margaret L Hilton, and others. 2013. *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.

Paul R Pintrich. 2003. A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational Psychology* 95, 4 (2003), 667.

Paul R Pintrich, David AF Smith, Teresa García, and Wilbert J McKeachie. 1993. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement* 53, 3 (1993), 801–813.

Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

Dale H Schunk and Barry J Zimmerman. 2012. *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.

Daniel T Seaton, Yoav Bergner, Isaac Chuang, Piotr Mitros, and David E Pritchard. 2014. Who does what in a massive open online course? *Commun. ACM* 57, 4 (2014), 58–65.

Michelle K Smith, William B Wood, Wendy K Adams, Carl Wieman, Jennifer K Knight, Nancy Guild, and Tin Tin Su. 2009. Why peer discussion improves student performance on in-class concept questions. *Science* 323, 5910 (2009), 122–124.

Leonard Springer, Mary Elizabeth Stanne, and Samuel S Donovan. 1999. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of educational research* 69, 1 (1999), 21–51.

Thomas F Stafford. 2005. Understanding motivations for Internet use in distance education. *Education, IEEE Transactions on* 48, 2 (2005), 301–306.

Vincent Tinto. 1997. Classrooms as communities. *Journal of higher education* 68, 6 (1997), 599–623.

Roger Tourangeau, Kenneth A Rasinski, and Norman Bradburn. 1991. Measuring happiness in surveys: A test of the subtraction hypothesis. *Public Opinion Quarterly* 55, 2 (1991), 255–266.

David Tyack and William Tobin. 1994. The grammar of schooling: Why has it been so hard to change? *American Educational Research Journal* 31, 3 (1994), 453–479.

Lev S Vygotsky. 1934 (1986). *Thought and language*. MIT Press.

Julia Wilkowski, Amit Deutsch, and Daniel M Russell. 2014. Student Skill and Goal Achievement in the Mapping with Google MOOC. In *L@S 2014 First (2014) ACM Conference on Learning @ Scale*. ACM, 3–10.

Received June 2014; revised October 2014; accepted October 2014

Online Appendix to:
Motivation as a Lens to Understand Online Learners: Towards Data-Driven Design with the OLEI Scale

RENE F. KIZILCEC, Stanford University
EMILY SCHNEIDER, Stanford University

A. ADDITIONAL INFORMATION ON ITERATIVE DEVELOPMENT

This section provides details on the iterative development of the OLEI scale.

A.1. Mechanical Turk Setup

We used Mechanical Turk in three batches during the development of the scale to code open-ended responses from online learners about their reasons for enrolling in the online course. For each batch, a different set of reasons was used and an adjusted codebook was provided.

The title, description, instructions, and target population of the Mechanical Turk job was the same in all three batches.

Target population of Mechanical Turk workers: Categorization Masters. *Title:* Categorize Reasons for Enrolling
Description: People were asked why they enroll in an online class. We would like to categorize their responses.
Instructions:

> The response below is in answer to “Why did you enroll in this online course?”
> Select ALL categories of reasons contained in the response below. You can select multiple categories.
> Use the “Other reason” category if the response, or part of it, does not fit into any other categories.
> Make sure you know the categories before you start categorizing reasons. If the survey response text is blank, in a foreign language, or makes no sense, just mark it under the “Spam response” category. Make sure you complete a substantial number of HITs. If you complete less than 20 HITs, our system will flag your work for review. Make sure you categorize each item as would typical American citizens who speak English natively, and who are *paying attention* to the task. If your categorizations are consistently unusual compared to categories that other workers provide, our system will flag your work for review and ALL of your HITs could be disqualified.

The number of coders, pay, and reasons to enroll responses varied across batches.

In the first batch, 4 coders coded 300 reasons for 5 Cents per reason ($9 average hourly rate). In the second batch, 2 coders coded 200 reasons for 6 Cents per reason ($9 average hourly rate). In the third batch, 3 coders coded 300 reasons for 5 Cents per reason ($9.47 average hourly rate).

DOI: http://dx.doi.org/10.1145/0000000.0000000

A.2. Iterative Changes to the Codebook

The initial codebook consisted of the 12 reasons to enroll: Course offered by prestigious university/professor; Earn a certificate/statement of accomplishment; Relevant to academic research; For career change; Experience an online course; Take with colleagues/friends; Relevant to job; Relevant to academic coursework; Explore an unfamiliar topic; Explore a familiar topic; For fun and challenge; Meet new people. The following two options were added to flag invalid reasons and identify reasons that may not be covered in the codebook: Other reason; Spam response.

The 300 reasons coded in the first batch were randomly sampled from open responses on course surveys of a Political Science MOOC and an Economics MOOC—there were a total of 2,825 and 2,376 open responses available from each course survey, respectively. Agreement across coders for each code was computed as the proportion of reasons for which the majority of coders were in agreement with each other (at least 3 out of 4 selected the code). Agreement across coders was highest for the following two codes: Course offered by prestigious university/professor; Earn a certificate/statement of accomplishment. Agreement was lowest for the following two codes: For fun and challenge; Meet new people. The most frequently selected codes were Explore a familiar topic and Explore an unfamiliar topic. The least frequently selected one were Take with colleagues/friends and Earn a certificate/statement of accomplishment.

Two independent researchers manually coded all open responses that were coded as Other reason by at least two Mechanical Turk coders. This revealed a missing code in the codebook: taking the course for personal growth and enrichment. This was added to the codebook in the second batch.

In the second batch 200 open responses were randomly selected from a set of 2,934 responses that learners provided on a course survey of a Computer Science MOOC. In this case, however, learners were presented with the reasons in the codebook to select from and then asked to specify any other reasons for enrolling.

The most frequently selected codes were Explore a familiar topic and Relevant to job. As expected, many the Other reason code was also frequently selected for this batch. Two independent researchers coded reasons that were marked as not fitting into the codebook. This revealed that many learners indicated enrolling to improve their English language skills. We added this to the codebook in the third batch. Moreover, the Exploring a (un)familar topic code appeared not to be a useful distinction in the context of enrollment intentions. This conclusion was based on the high frequency with which both codes were selected and the mismatch in the language used by learners with the words familiar and unfamiliar. Learners tended to state what they were interested in learning and their familarity or unfamiliarity with the topic was only apparent from context. The codes were therefore replaced by a more general code—General interest—in the topic in the codebook for the third batch.

The third batch used the final codebook, which contained the same reasons that are on the OLEI scale. A different set of 300 responses from the Economics and Political Science MOOCs were randomly sampled and coded. The frequency with which 'Other' was selected was lower than in the earlier batches. Open responses that were coded as 'Other' were coded by two researchers and found to fit into the existing codebook. There were no medium or high correlations between codes on the codebook.

In a final validation step to check if the OLEI scale actually included the most common enrollment intentions, learners were given the opportunity to report any other reasons for enrolling in the 14 course surveys. Although 7% of learners responded to the question, responses were either repetitions or alternative phrasings of reasons on the OLEI scale, or comments specific to a course (e.g., "to learn R"). Hence, no additional modifications to the OLEI scale were made.

B. GENDER AND EDUCATION DIFFERENCES IN MOTIVATION

Table V. Gender differences in motivation. Percentage point differences in the proportion of females relative to males who selected each enrollment intention in each course and the median difference.

Enrollment Intentions	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Median
General interest	-0	-2	0	0	-1	-1	6	27	11	1	0	6	-0	0.4
Job relevant	-0	-1	1	-5	-5	-1	0	16	14	-6	1	6	4	0.4
School relevant	0	20	1	0	-6	9	3	14	13	-16	5	5	11	4.9
Research relevant	4	15	-1	-1	-3	15	-2	6	7	-9	6	2	14	3.5
Growth/enrichment	-2	2	3	2	1	-0	7	30	14	2	0	7	-1	1.8
Career change	1	4	7	7	-1	-5	-1	5	4	1	5	3	-2	3.4
Fun/challenge	-3	-6	-1	-2	2	-3	9	35	12	2	-4	7	-5	-1.0
Meet new people	3	-2	-3	-4	7	-3	-1	5	1	3	10	-2	-2	-0.8
Experience online	-2	9	1	-1	-0	3	4	11	9	-2	1	6	4	3.2
Earn certificate	-0	7	3	5	2	-1	2	11	9	-4	5	3	2	2.8
Prestigious uni/prof	-0	3	-1	-0	4	-1	4	14	11	5	3	3	-1	3.0
Take with others	4	2	-0	-1	6	-1	0	5	-1	3	9	-1	4	2.1
Improve English	6	0	-2	2	3	-0	-1	-7	8	-3	8	1	3	1.0

Table VI. Education differences in motivation. Percentage point differences in the proportion of learners holding a college degree or above relative those without one who selected each enrollment intention in each course and the median difference.

Enrollment Intentions	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C12	C13	Median
General interest	-7	3	-1	-1	-0	-2	2	1	1	-1	1	-1	-0.5
Job relevant	-9	11	12	9	-8	28	13	26	8	3	12	19	11.2
School relevant	18	-32	-25	-18	16	-16	-26	-9	-13	21	-12	-20	-14.9
Research relevant	8	-7	-9	-5	11	1	-4	-4	1	9	-6	-0	-2.0
Growth/enrichment	-6	4	-1	-1	1	-7	-1	-1	-2	-0	-2	-0	-1.3
Career change	-2	2	5	-2	-1	4	-2	8	-1	7	2	6	2.2
Fun/challenge	-5	2	-5	-4	2	-7	-1	-6	-8	5	-7	-6	-5.0
Meet new people	3	-0	-5	-2	2	-0	-3	4	-4	6	-5	-0	-0.4
Experience online	11	-12	-9	-11	7	-9	-13	-10	-13	8	-11	-3	-9.5
Earn certificate	18	-10	-8	-11	8	-10	-12	-8	-8	16	-13	-8	-8.2
Prestigious uni/prof	-1	2	-3	3	-0	5	0	2	6	3	2	5	2.0
Take with others	-1	-3	-2	-2	-2	2	-2	5	-3	10	-2	2	-1.6
Improve English	6	-9	-9	-2	4	-10	-14	-6	-12	16	-10	-11	-9.1

C. RESULTS OF PREDICTIVE ANALYSIS

Table VII. Results of the predictive model (Part 1). Coefficients, standard errors, and confidence intervals for each enrollment intention and each predicted outcome.

Enrollment Intention	Binary Outcome	Mean	Std. Error	Lower 95% CI	Upper 95% CI
General interest	$>$ 10% assignments	0.033	0.010	0.008	0.058
General interest	$>$ 50% assignments	0.012	0.008	-0.010	0.033
General interest	$>$ 80% assignments	0.009	0.007	-0.010	0.027
General interest	$>$ 0 posts	0.004	0.007	-0.015	0.022
General interest	$>$ 50% posts	0.012	0.005	-0.001	0.024
General interest	$>$ 1 votes received	0.010	0.005	-0.003	0.023
General interest	earned certificate	0.013	0.007	-0.004	0.031
General interest	$>$ 10% videos	0.009	0.010	-0.017	0.035
General interest	$>$ 50% videos	0.016	0.009	-0.006	0.038
General interest	$>$ 80% videos	0.016	0.007	-0.003	0.033
Earn certificate	$>$ 10% assignments	0.042	0.005	0.029	0.055
Earn certificate	$>$ 50% assignments	0.037	0.004	0.025	0.048
Earn certificate	$>$ 80% assignments	0.013	0.004	0.003	0.022
Earn certificate	$>$ 0 posts	0.014	0.004	0.004	0.024
Earn certificate	$>$ 50% posts	0.011	0.003	0.005	0.018
Earn certificate	$>$ 1 votes received	0.007	0.003	-0.000	0.015
Earn certificate	earned certificate	0.004	0.004	-0.006	0.013
Earn certificate	$>$ 10% videos	0.024	0.005	0.011	0.037
Earn certificate	$>$ 50% videos	0.025	0.004	0.013	0.037
Earn certificate	$>$ 80% videos	0.027	0.004	0.017	0.037
Prestigious uni/prof	$>$ 10% assignments	0.019	0.005	0.006	0.031
Prestigious uni/prof	$>$ 50% assignments	0.015	0.004	0.005	0.026
Prestigious uni/prof	$>$ 80% assignments	0.008	0.004	-0.001	0.018
Prestigious uni/prof	$>$ 0 posts	0.010	0.004	0.001	0.020
Prestigious uni/prof	$>$ 50% posts	0.005	0.003	-0.001	0.012
Prestigious uni/prof	$>$ 1 votes received	0.003	0.003	-0.004	0.010
Prestigious uni/prof	earned certificate	0.005	0.004	-0.004	0.014
Prestigious uni/prof	$>$ 10% videos	0.023	0.005	0.010	0.035
Prestigious uni/prof	$>$ 50% videos	0.019	0.004	0.008	0.030
Prestigious uni/prof	$>$ 80% videos	0.014	0.004	0.004	0.023
Take with others	$>$ 10% assignments	0.024	0.006	0.007	0.039
Take with others	$>$ 50% assignments	0.009	0.006	-0.006	0.023
Take with others	$>$ 80% assignments	0.002	0.005	-0.011	0.014
Take with others	$>$ 0 posts	-0.009	0.005	-0.023	0.004
Take with others	$>$ 50% posts	-0.009	0.003	-0.017	-0.001
Take with others	$>$ 1 votes received	-0.005	0.004	-0.014	0.004
Take with others	earned certificate	0.025	0.005	0.012	0.038
Take with others	$>$ 10% videos	0.020	0.006	0.003	0.036
Take with others	$>$ 50% videos	0.005	0.006	-0.010	0.019
Take with others	$>$ 80% videos	0.003	0.005	-0.009	0.016
Improve English	$>$ 10% assignments	-0.005	0.006	-0.020	0.009
Improve English	$>$ 50% assignments	-0.002	0.005	-0.014	0.011
Improve English	$>$ 80% assignments	0.006	0.004	-0.005	0.017
Improve English	$>$ 0 posts	-0.010	0.004	-0.021	0.001
Improve English	$>$ 50% posts	-0.008	0.003	-0.016	-0.001
Improve English	$>$ 1 votes received	-0.012	0.003	-0.020	-0.004
Improve English	earned certificate	0.000	0.004	-0.010	0.010
Improve English	$>$ 10% videos	-0.017	0.006	-0.032	-0.002
Improve English	$>$ 50% videos	-0.011	0.005	-0.024	0.002
Improve English	$>$ 80% videos	-0.010	0.004	-0.020	0.001

Table VIII. Results of the predictive model (Part 2). Coefficients, standard errors, and confidence intervals for each enrollment intention and each predicted outcome.

Enrollment Intention	Binary Outcome	Mean	Std. Error	Lower 95% CI	Upper 95% CI
Job relevant	> 10% assignments	0.004	0.005	-0.009	0.017
Job relevant	> 50% assignments	-0.001	0.005	-0.013	0.011
Job relevant	> 80% assignments	-0.001	0.004	-0.012	0.009
Job relevant	> 0 posts	-0.012	0.004	-0.022	-0.001
Job relevant	> 50% posts	-0.006	0.003	-0.014	0.001
Job relevant	> 1 votes received	-0.002	0.003	-0.010	0.006
Job relevant	earned certificate	0.008	0.004	-0.002	0.017
Job relevant	> 10% videos	0.016	0.005	0.003	0.029
Job relevant	> 50% videos	0.005	0.005	-0.007	0.017
Job relevant	> 80% videos	0.007	0.004	-0.003	0.017
School relevant	> 10% assignments	-0.015	0.005	-0.029	-0.002
School relevant	> 50% assignments	-0.015	0.005	-0.027	-0.002
School relevant	> 80% assignments	-0.016	0.004	-0.027	-0.006
School relevant	> 0 posts	-0.008	0.004	-0.018	0.003
School relevant	> 50% posts	-0.007	0.003	-0.014	0.000
School relevant	> 1 votes received	-0.004	0.003	-0.012	0.003
School relevant	earned certificate	-0.012	0.004	-0.022	-0.002
School relevant	> 10% videos	-0.022	0.005	-0.036	-0.008
School relevant	> 50% videos	-0.015	0.005	-0.027	-0.002
School relevant	> 80% videos	-0.015	0.004	-0.025	-0.005
Research relevant	> 10% assignments	-0.021	0.005	-0.035	-0.007
Research relevant	> 50% assignments	-0.014	0.005	-0.026	-0.001
Research relevant	> 80% assignments	-0.007	0.004	-0.018	0.004
Research relevant	> 0 posts	-0.009	0.004	-0.020	0.002
Research relevant	> 50% posts	-0.005	0.003	-0.013	0.002
Research relevant	> 1 votes received	-0.008	0.003	-0.016	0.000
Research relevant	earned certificate	-0.016	0.004	-0.026	-0.006
Research relevant	> 10% videos	-0.025	0.006	-0.039	-0.011
Research relevant	> 50% videos	-0.014	0.005	-0.026	-0.001
Research relevant	> 80% videos	-0.010	0.004	-0.020	0.001
Growth/enrichment	> 10% assignments	0.026	0.009	0.003	0.050
Growth/enrichment	> 50% assignments	0.020	0.008	-0.001	0.040
Growth/enrichment	> 80% assignments	0.017	0.007	-0.001	0.034
Growth/enrichment	> 0 posts	0.013	0.007	-0.004	0.029
Growth/enrichment	> 50% posts	0.007	0.004	-0.004	0.017
Growth/enrichment	> 1 votes received	0.002	0.004	-0.010	0.013
Growth/enrichment	earned certificate	0.016	0.007	-0.001	0.033
Growth/enrichment	> 10% videos	0.017	0.010	-0.008	0.043
Growth/enrichment	> 50% videos	0.017	0.008	-0.005	0.039
Growth/enrichment	> 80% videos	0.010	0.007	-0.008	0.027
Career change	> 10% assignments	0.016	0.005	0.003	0.029
Career change	> 50% assignments	0.013	0.004	0.002	0.024
Career change	> 80% assignments	0.009	0.004	-0.001	0.019
Career change	> 0 posts	0.009	0.004	-0.001	0.019
Career change	> 50% posts	0.004	0.003	-0.003	0.011
Career change	> 1 votes received	0.003	0.003	-0.004	0.011
Career change	earned certificate	0.009	0.004	-0.000	0.018
Career change	> 10% videos	0.013	0.005	0.001	0.027
Career change	> 50% videos	0.012	0.004	0.000	0.024
Career change	> 80% videos	0.013	0.004	0.002	0.022

Table IX. Results of the predictive model (Part 3). Coefficients, standard errors, and confidence intervals for each enrollment intention and each predicted outcome.

Enrollment Intention	Binary Outcome	Mean	Std. Error	Lower 95% CI	Upper 95% CI
Fun/challenge	$>$ 10% assignments	0.035	0.005	0.021	0.049
Fun/challenge	$>$ 50% assignments	0.014	0.005	0.002	0.026
Fun/challenge	$>$ 80% assignments	0.009	0.004	-0.002	0.019
Fun/challenge	$>$ 0 posts	0.007	0.004	-0.004	0.018
Fun/challenge	$>$ 50% posts	0.006	0.003	-0.002	0.013
Fun/challenge	$>$ 1 votes received	0.002	0.003	-0.006	0.010
Fun/challenge	earned certificate	-0.000	0.004	-0.010	0.009
Fun/challenge	$>$ 10% videos	0.020	0.005	0.006	0.034
Fun/challenge	$>$ 50% videos	0.017	0.005	0.004	0.029
Fun/challenge	$>$ 80% videos	0.014	0.004	0.004	0.025
Meet new people	$>$ 10% assignments	-0.030	0.006	-0.045	-0.014
Meet new people	$>$ 50% assignments	-0.009	0.005	-0.023	0.004
Meet new people	$>$ 80% assignments	-0.001	0.005	-0.013	0.010
Meet new people	$>$ 0 posts	0.026	0.005	0.013	0.039
Meet new people	$>$ 50% posts	0.011	0.003	0.002	0.020
Meet new people	$>$ 1 votes received	0.011	0.004	0.002	0.021
Meet new people	earned certificate	-0.005	0.004	-0.016	0.005
Meet new people	$>$ 10% videos	-0.033	0.006	-0.049	-0.018
Meet new people	$>$ 50% videos	-0.012	0.005	-0.025	0.002
Meet new people	$>$ 80% videos	-0.009	0.004	-0.020	0.003
Experience online	$>$ 10% assignments	-0.026	0.005	-0.039	-0.014
Experience online	$>$ 50% assignments	-0.010	0.004	-0.021	0.000
Experience online	$>$ 80% assignments	-0.009	0.004	-0.019	0.000
Experience online	$>$ 0 posts	-0.015	0.004	-0.025	-0.005
Experience online	$>$ 50% posts	-0.006	0.003	-0.012	0.001
Experience online	$>$ 1 votes received	-0.006	0.003	-0.013	0.001
Experience online	earned certificate	-0.006	0.003	-0.015	0.004
Experience online	$>$ 10% videos	-0.018	0.005	-0.031	-0.005
Experience online	$>$ 50% videos	-0.017	0.004	-0.028	-0.006
Experience online	$>$ 80% videos	-0.009	0.004	-0.019	0.000