Ihantola, Petri; Fronza, Ilenia; Mikkonen, Tommi; Noponen, Miska; Hellas, Arto

Deadlines and MOOCs: How Do Students Behave in MOOCs with and without Deadlines

Published in:
2020 IEEE Frontiers in Education Conference (FIE)

DOI:
10.1109/FIE44824.2020.9274023

Published: 21/10/2020

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Ihantola, P., Fronza, I., Mikkonen, T., Noponen, M., & Hellas, A. (2020). Deadlines and MOOCs: How Do Students Behave in MOOCs with and without Deadlines. In *2020 IEEE Frontiers in Education Conference (FIE)* Article 9274023 (IEEE Frontiers in Education Conference (FIE)). IEEE. https://doi.org/10.1109/FIE44824.2020.9274023

Deadlines and MOOCs: How Do Students Behave in MOOCs with and without Deadlines

Petri Ihantola
University of Helsinki
Helsinki, Finland
ORCID: 0000-0003-1197-7266

Ilenia Fronza
Free University of Bozen-Bolzano
Bozen-Bolzano, Italy
ORCID: 0000-0003-0224-2452

Tommi Mikkonen
University of Helsinki
Helsinki, Finland
ORCID: 0000-0002-8540-9918

Miska Noponen
University of Helsinki
Helsinki, Finland
ORCID: 0000-0001-5961-5438

Arto Hellas
Aalto University
Helsinki, Finland
arto.hellas@aalto.fi

***Abstract*—**

***Full research paper*—Online education can be delivered in many ways. For example, some MOOCs let students to proceed with their own pace, while others rely on strict schedules. Although the variety of how MOOCs can be organized is generally well understood, less is known about how the different ways of organizing MOOCs affect retention. In this work, we compare self-paced and fixed-schedule MOOCs in terms of retention and work-load. Using data from over 8.000 students participating in two versions of a massive open online course in programming, we observe that drop-out rates at the beginning of the courses are greater than towards the end of the courses, with self-paced MOOC being more extreme in this respect. Mostly because of different starts, the fixed-schedule course has a better overall retention rate (45%) than its self-paced counterpart (13%). We hypothesize that students initial investment of time and effort contributes to their persistence in their course, meaning that they do not want to let their initial investment go to waste. At the same time, in both self-paced and fixed-schedule MOOCs, there are students who receive almost full points from one week but fail to continue to the next week. This suggests that the issue of drop-outs in MOOCs may also be related to participants struggling to take up new tasks or schedule their work over a longer time period. Our results support scheduling student activities in open online courses and opens up new research directions in engaging students in self-paced courses.**

***Index Terms*—Student persistence, open online course, introductory programming, MOOC**

I. Introduction

The term massive open online course (MOOC) became popular in late 2011 with the "Introduction to Artificial Intelligence" MOOC with over 160,000 enrolled students. The perceived success of the course led to many institutions offering their courses for free for anyone. At the peak of the hype, MOOCs were present also in the popular media, including New York Times, which dubbed the year 2012 as "The Year of the MOOC" [1].

While MOOCs were supposed to revolutionize higher education, creating an opportunity where anyone anywhere could attend high quality courses, it soon became evident that this would not happen [2]. High enrollment rates were followed by poor completion rates, which were suggested to be even under 5% [3], [4]. Moreover, those who complete MOOCs often have prior experience in MOOCs and are typically older and more educated, with an intent to complete the course [5]. MOOCs have also been reported to decrease gender balance in already segregated fields such as computing [6].

Participation in MOOCs has been discussed using term "funnel of participation" [7]. This funnel includes four steps: (i) Awareness of a MOOC exiting, (ii) Registration to the MOOC, (iii) Activity on the course, and (iv) Meaningful learning progress. Based on [7], only a small fraction of students continue to the next step. Using such terminology does not, however, account for when a student drops out, assuming that the student is active in the course.

To make matters more complicated, MOOCs come in many shapes, forms and flavors [8]; For example, some MOOCs are meant for self-paced learning, where there are no deadlines or tutoring is available. In contrast, some other MOOCs are associated with a fixed schedule, where the course mimics a traditional university course, advancing in a step-wise fashion in accordance to a predefined timeline, following strict, measurable learning goals.

In this work, we describe a MOOC in programming, and discuss our experiences in students dropping out from two versions of the course, one being self-paced and the other relying on a fixed schedule. In particular, our focus is on at which point of course week students drop out, reflecting on how the dropouts provide information to the course designers.

The rest of the paper is organized as follows. Next, in Section II we provide some background for the paper, and present our research questions. Then, in Section III we outline the course and data that is used to describe our experiences. This is followed by an analysis of the data to support our experiences in students dropping out from the course in Section IV. Section V discusses the results, indicating possible explanations for the observations. Finally, Section VI draws some final conclusions.

II. Programming MOOCs and Dropout

A. Retention in MOOCs and Programming

Over the years, MOOCs have gained popularity and have attracted millions of online users. Despite their obvious advantages, MOOCs have been criticized for the low completion ratio [9], [10]. A systematic mapping study published in 2019 identified 18 key topics in MOOC research, highlighting that the five most researched topics associated with MOOCS were retention/dropout, instructional design, engagement, student behavior, and assessment [11]. A general finding is that the student's ability to find and manage time effectively is one of the reasons why students complete or drop out of MOOCs [12]. Consequently, considering the difference between self-paced and fixed-schedule MOOCs requires more thorough investigation.

Programming courses are, unfortunately, no exception to the rule of low student retention in MOOCs. Khalil and Ebner [13] quote statistics that show popular programming courses, organized on the MOOC platforms Coursera and edX, have a student retention rate ranging from 0.9% to 20% which is aligned with MOOCs across domains. Moreover, in programming, the high dropout rates are not restricted to MOOCs, and even in computing education research, statement that "programming is difficult" is one the most used premises [14]. There are some positive exceptions to the high dropout rates, where programming MOOCs have had a retention rate of over 60% [15]. Previous studies have outlined several factors that can lead to improved student retention rates in programing MOOCs.

The course content itself has been identified is a key factor driving retention [16]–[19]. These studies show that content can, e.g., provide a suitable level of difficulty for students, allow them to learn by doing or provide automatic feedback to keep the student engaged in the course. Feedback in its many forms is often seen as the most powerful influences on learning [20].

Social learning also improves student retention. Several studies have emphasized the engaging effect of interacting with an instructor or other students [16], [18], [19], [21], [22]. Conversely, the lack of interaction can cause students to feel isolated and thus increase the likelihood of dropout [16].

In addition to these attributes of the course, students' personal characteristics affect retention. Factors such as the student's level of motivation and their aptitude in online learning have been found to have an effect [16], [19], [22]. Also individual contextual factors, such as lack of time or conflicts with other areas of life such as family or work are significant [13], [19], [22]. The latter observation supports a demand for the present study, linking course scheduling and student retention.

Student dropout is not a consistent phenomenon throughout a programming MOOC. Hone and El Said [16] found that dropout is more likely at the first half of the course, with similar observations reported by other studies [5], [17]. More specifically, students are most likely to drop out at the very beginning of the course with a decreasing likelihood as the course progresses. Hone and El Said [16] speculate that this may be due to loss aversion when students have already invested time and effort into the course and risk losing the possible gains by dropping out.

B. Self-paced and fixed schedules

There is an ongoing discussion about the benefits of self-paced [17], [23]–[25] – sometimes called asynchronous [17] – MOOCs in comparison to scheduled MOOCs. The latter have also been dubbed as synchronous [17], instructor-paced [17] MOOCs, or as MOOCs with pre-defined [23] or fixed schedule [23], [25]. When access to instructors is included in the process, a term tutoring MOOC [24] is also used.

On one hand, the analysis presented in some articles supports a self-paced setting [3], [13], [25], [26]. In particular, MOOCs should have longer periods for assignments [27] and include mechanisms to cope with unexpected life events rather than motivational messages to reduce the dropout rate [28]. The analysis of indicators related to self-regulation (e.g., amplitude of intervals between logins) can provide insights about students' sense of timeliness [29] and ability to organize one's own learning in a MOOC, which is related to clear objectives [30]. Khalil and Wong found that learning session duration impacts students completion rate in MOOCs [31], which could be taken into account when defining deadlines in fixed schedule MOOCs.

On the other hand, in some other articles, fixed-schedule MOOCs are suggested to obtain greater engagement of students [17], and self-paced progress and assessments are indicated as solutions that make students more lethargic [32]. Sannicandro et al. [24] found no difference in likelihood of dropout between the two types of MOOCs. Vitiello et al. [23] attempted to predict dropouts from early course behavior, comparing self-paced and fixed-schedule MOOCs; however, the effect of schedule could not be isolated, due to the differences of topics, target group and organizing university.

C. Research Questions

Although previous research has shed light on retention in online education as well as on the role of external guidance in web based learning, details of how self-paced and fixed-schedule MOOCs differ in terms of participatory pattern are not well understood at the moment. Therefore, in this research, we will answer the following questions:

1) What are the differences in retention between the students in the self-paced and fixed-schedule MOOCs?
2) How does the effort (e.g., number of exercises or active days) differ between the students in the self-paced and fixed-schedule MOOCs?

III. Methods

A. Course

The study was conducted in two versions of a free and open online introductory programming MOOC offered by the University of Helsinki in 2019. The MOOC teaches basics of programming with Java, ranging from procedural programming to object-oriented programming, covering also principles of algorithms (sorting and searching) and testing of programs. The workload of the course is 5 ECTS, which corresponds to approximately 150 hours of study work. Students participating in the MOOC have access to an online workbook with theory, questions, and programming assignment handouts. The materials are divided into seven units, which correspond to study weeks.

While the programming assignment handouts are a part of the workbook, students work on the programming assignments on their own using an integrated development environment. Each week, there are dozens of programming assignments which combine into larger programs and demonstrate how programs are constructed using a divide-and-conquer approach. Worked examples of similar programs are embedded in the materials. The programming assignments are returned to an automated assessment system, which grades the work and provides feedback.

The two versions of the course, which we study, differ in their schedule and grading. One of the courses set weekly release dates and fixed deadlines (later referred to as **fixed-schedule**), while the other had no deadlines except a course expiration date (later referred to as **self-paced**). Table I outlines the schedule for fixed-schedule version, as well as the assignment count in each part for both courses. The materials for each part of the self-paced MOOC were released after the deadline of that particular part in the fixed-schedule MOOC, and the course could be started and worked on until January 2020.

TABLE I
SCHEDULE OF THE PROGRAMMING ASSIGNMENTS FOR THE FIXED-SCHEDULE MOOC.

Part	Release	Deadline	Weeks	Assignments
I	14.12.2018	21.1.2019	5.5	41
II	28.12.2018	28.1.2019	4.5	33
III	11.1.2019	4.2.2019	3.5	34
IV	25.1.2019	11.2.2019	2.5	28
V	1.2.2019	18.2.2019	2.5	17
VI	8.2.2019	25.2.2019	2.5	14
VII	15.2.2019	4.3.2019	2.5	9

The grading of the fixed-schedule MOOC is based on the number of completed assignments (50% of the overall course points) and an online exam (remaining 50% of the overall course points), which contains similar programming assignments to those given in the course. In order to pass the course, one must receive at least 50% of the overall course points and at least 50% of the points available in the online exam. The highest grade is obtained by gaining at least 90% of the overall available points. In the self-paced MOOC, one can proceed to the next part of the course once they have completed at least 90% of the programming assignments in the previous part, and the grade is formed solely based on the end of course online exam.

B. Participants and data collection

At the beginning of the course, all participants were asked a consent, which allows the use of their data for research purposes. For the fixed-schedule MOOC, 4162 students provided research consent and completed at least one programming assignment, while for the self-paced MOOC the corresponding number is 5309. 616 students had participated on both course versions and were removed from the analysis. After that, we had data from 3546 students participating the fixed schedule course and 4693 students participating the self-paced course.

The course was given for free for anyone, and in addition to those not affiliated with any university or college, it is also attended by participants from various universities and colleges, including the University of Helsinki that offers the course. The number of local students was marginal compared to the course size. There was no mandatory registration for the course, except at the end of the course if the participants wish to receive an official diploma or credits for the course.

The programming environment used in the course collects keystroke data from course participants that is used for analysis of participants' difficulties during the course as well as monitoring excessive collaboration or plagiarism. In this study, we had access to combined logs which include the number of daily typing events, number of times when a program has been tested, and number of submits. Typing events were further divided into insert, edit, and paste events.

C. Measures and approach

Based on the typing data, we calculated daily statistics on how much participants worked on each assignment; how much they inserted text by typing, removed or pasted text, or how many times each assignment was submitted. Inserts, removals and paste events were all combined into a single measure illustrating the typing activity. Unit-tests used for grading were available for the participants and distributed together with the assignments. Thus, submissions were a relatively accurate proxy for getting the programming assignments correct (to some extent). In this study, we had no access to the actual study records and grading data.

We defined that a student was active on a week if they worked on any of the programming assignments related to that week by editing the source code. Based on this, we calculated how many participants were active each week, and corresponding weekly retention rates. In addition to retention, we also calculated how many participants exited each week. A participant was defined as exiting if they were active in the week but not in any following weeks. Chi-square test was used to compare if distributions of exit points between the course versions were similar.

Two proportional exit rate measures were calculated for each week separately for each course. One was relative to the whole population (just like retention), and illustrated the percentage of all the participants exiting each week. Another rate was calculated proportional to the active students of the week (instead of whole population). This measure "forgets" participants who have exited already in earlier weeks and presents the proportion of the same week actives who exited the course.

Our second research question was related to effort – which is a complicated construct [6]. In the present study, when comparing how much effort participants on different course versions were willing to put in, we focus on the number of programming assignments worked on, number of submitted programming assignments, total amount of typing events and for how many calendar days the work was distributed. The effort of students exiting at different stages of the course (e.g., weeks 1 and 5) are likely very different. Thus, we looked at sub-populations exiting at each week separately. Multiple Wilcoxon rank sum tests were conducted to compare the effort estimates between course versions. Comparisons were conducted separately for the the cumulative activity that happened during and before the exit week (i.e., week when the dropout happens) and only the last week only.

IV. Results

A. Retention

For each week, the number of active students and the number of exiting students are provided in Table II. Exiting students are calculated also in the *active* column as they are active in the week when they drop out. Retention rates, together with the exit rate measures are illustrated in Figure 1. Chi-square tests of homogeneity were conducted to examine distributions of exits within the courses. In both cases, the exits were unevenly distributed with $\chi^2(5) = 689.17$, $p < .000$ and $\chi^2(5) = 5393.4$, $p < .000$ respectively for fixed-schedule and self-paced courses. Retention rates at the end of course were significantly different ($\chi^2(1) = 1055.3$, $p < .000$); 45% of the students in the fixed schedule MOOC are active during the last week of the course, while for the self-paced MOOC the number is 13%. In other words, the retention rate of the fixed schedule is nearly 3.5 times higher than the retention rate of the self-paced MOOC.

A Chi-square test of independence was calculated comparing the distribution of the weekly exits between course versions. A significant interaction was found ($\chi^2(6) = 1411.8$, $p < .000$). The related residuals are provided in Table II, indicating that the majority of differences are related to week one, although differences in weeks 2, 3, 4 and 6 are also identifiable (i.e., abs(residual)>2). The residuals of the last week merely confirm the earlier analysis on the differences in the final retention.

B. Effort

Effort comparison of cumulative effort metrics is provided in Table III. Because it is possible to skip a week (or do only few assignments in a week) number of working days is provided also as normalized for the number of assignments. The number of typing events is provided only as normalized. Each row in the table illustrates comparison of students exiting the course at the same week. Sample sizes can be seen from the exiting column of Table II. For each metric, we illustrate the medians, Wilcoxon tests measure (W), p-value corrected corrected for multiple comparison with Bonferroni approach (p.adj), and Cliff's delta as an easy way to interpret the effect size measure. Effect sizes greater than 0.33, indicating at least medium difference between the populations are highlighted.

TABLE II
NUMBERS OF ACTIVE AND EXITING STUDENTS EACH WEEK AND STANDARDIZED RESIDUALS OF CHI SQUARE TEST COMPARING DROPOUT DISTRIBUTIONS OF THE COURSE VERSIONS.

	fixed-schedule (n=3546)		self-paced (n=4696)		standardized residuals (exiting between groups)	
week	active	exiting	active	exiting		
W1	2983	593	4653	2328	-30.89	30.89
W2	2708	550	2340	869	-3.58	3.58
W3	2398	340	1481	374	2.59	-2.59
W4	2056	180	1112	179	2.78	-2.78
W5	1877	105	937	146	-0.39	0.39
W6	1770	155	793	163	2.09	-2.09
W7	1623	1623	634	634	32.51	-32.51

While the cumulative effort combines all the activity from the course, we also investigated the differences in the exit week activities separately – here, exit week refers to the last week on the course on which the student is active. The results are provided in Table IV

Interestingly practically no differences between the exit week submits were observed, although different grading schemes of the courses could have caused such differences. To better understand the exit week behavior, we also investigated the distribution of the number of exit week submits. These are illustrated in Figure 2. Peaks at the high end of the histograms indicate students that got full points from that week but did not continue any further. For example, 29 and 74 of the first week dropout submitted all the assignments, respectively for the fixed-schedule and self-paced courses.

V. Discussion

A. Differences in retention between the MOOCs

To begin with, the main differences in retention between the MOOCs are visible early on in the course. Students in the fixed-schedule version are more likely to persist in the course early on, while students in the self-paced MOOC are more likely to drop out early on. In the fixed-schedule, approximately 17% of the students dropped out in the first week, while in the self-paced MOOC approximately 50% of the students dropped out in the first week. After the first week, exit behavior is almost similar between the course versions.

Previous studies have suggested that dropping out from MOOCs happens early on in the course [5], [17]. Our results support this observation especially in the context of the self-paced MOOC, where a significant dropout is observable early on. At the same time, in the fixed-schedule version, the ratio between the active and exiting students is the same for the first two weeks. Our results, combined with those from previous studies, suggest the existence of a power law in retention, which is partially moderated by the course type.

We also highlight a phenomenon that has not previously been studied in research related to MOOC dropouts. When

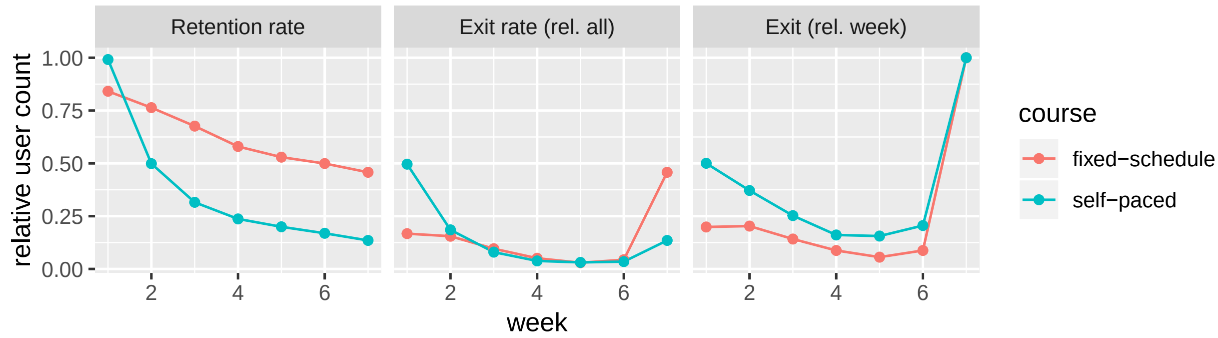

Fig. 1. Comparison of retention rates and exit relative to the whole population and the week active separately (i.e., number of students exiting each week divided by the population size or the number of active students in the same week).

TABLE III
COMPARISON OF CUMULATIVE EFFORT METRICS SEPARATELY FOR EACH EXIT WEEK.

	exit week	self-paced	fixed-schedule	W	p.adj	Cliff's d
assignments	1	11	10	668340.50	1.000	-0.032
	2	58	47	140811.00	**0.000**	**-0.411**
	3	92	74	29188.50	**0.000**	**-0.541**
	4	119	93	7296.50	**0.000**	**-0.547**
	5	150	116	2154.50	**0.000**	**-0.719**
	6	161	124	5654.50	**0.000**	**-0.552**
	7	177	150	219139.50	**0.000**	**-0.574**
submits	1	10	9	650746.00	1.000	-0.057
	2	57	45	136679.50	**0.000**	**-0.428**
	3	90	68	28070.50	**0.000**	**-0.559**
	4	117	85	6967.50	**0.000**	**-0.568**
	5	148	113	2119.00	**0.000**	**-0.724**
	6	160	121	5616.00	**0.000**	**-0.555**
	7	177	146	217692.50	**0.000**	**-0.577**
days	1	2	2	585540.50	**0.000**	-0.152
	2	9	6	139756.50	**0.000**	**-0.415**
	3	14	9	39235.00	**0.000**	**-0.383**
	4	18	12	9788.50	**0.000**	**-0.392**
	5	25	16	4566.50	**0.000**	**-0.404**
	6	25	19	9000.00	**0.000**	-0.288
	7	27	26	493888.00	1.000	-0.040
days/assignments	1	0.25	0.20	607763.50	**0.000**	-0.120
	2	0.17	0.15	227381.50	1.000	-0.049
	3	0.16	0.15	61729.00	1.000	-0.029
	4	0.16	0.15	14805.50	1.000	-0.081
	5	0.17	0.16	6782.50	1.000	-0.115
	6	0.16	0.16	11978.50	1.000	-0.052
	7	0.15	0.18	599711.50	**0.000**	0.166
edits/assignments	1	152.64	149.00	682779.00	1.000	-0.011
	2	264.03	254.52	227369.00	1.000	-0.049
	3	267.09	254.44	61079.00	1.000	-0.039
	4	275.93	248.95	15006.00	1.000	-0.069
	5	307.53	302.33	7094.00	1.000	-0.074
	6	367.06	339.41	11560.00	1.000	-0.085
	7	426.65	455.24	557909.50	**0.000**	0.084

TABLE IV
COMPARISON OF LAST WEEK EFFORT METRICS SEPARATELY FOR EACH EXIT WEEK.

	exit week	self-paced	fixed-schedule	W	p.adj	Cliff's d
assignments	1	11	10	668340.50	1.000	-0.032
	2	16	14	220430.00	1.000	-0.078
	3	17	15	56926.50	1.000	-0.105
	4	10	9	15422.00	1.000	-0.043
	5	14	11	6536.50	1.000	-0.147
	6	8	7	12188.50	1.000	-0.035
	7	9	9	495096.50	1.000	-0.038
submits	1	10	9	650746.00	1.000	-0.057
	2	15	12	214541.00	**0.000**	-0.102
	3	16	14	55380.00	**0.000**	-0.129
	4	10	8	14559.50	1.000	-0.096
	5	12	10	6353.50	1.000	-0.171
	6	7	5	11719.50	1.000	-0.072
	7	9	9	496442.00	1.000	-0.035
days	1	2	2	585540.50	**0.000**	-0.152
	2	4	3	191532.00	**0.000**	-0.199
	3	4	3	44890.00	**0.000**	-0.294
	4	3	2	12080.00	**0.000**	-0.250
	5	4	3	4758.50	**0.000**	**-0.379**
	6	3	3	10231.00	**0.000**	-0.190
	7	5	5	455858.50	**0.000**	-0.114
days/assignments	1	0.25	0.20	607763.50	**0.000**	-0.120
	2	0.30	0.25	207594.50	**0.000**	-0.131
	3	0.27	0.20	48455.00	**0.000**	-0.238
	4	0.33	0.26	13349.50	**0.000**	-0.171
	5	0.46	0.31	6010.50	**0.000**	-0.216
	6	0.50	0.45	10871.50	1.000	-0.139
	7	0.62	0.56	455059.00	**0.000**	-0.116
edits/assignments	1	152.64	149.00	682779.00	1.000	-0.011
	2	316.74	304.99	228317.50	1.000	-0.045
	3	258.54	221.34	56776.00	1.000	-0.107
	4	288.57	262.20	14018.00	1.000	-0.130
	5	524.83	461.75	6161.00	**0.000**	-0.196
	6	831.29	655.08	9728.00	**0.000**	-0.230
	7	2063.79	1782.33	437461.50	**0.000**	-0.150

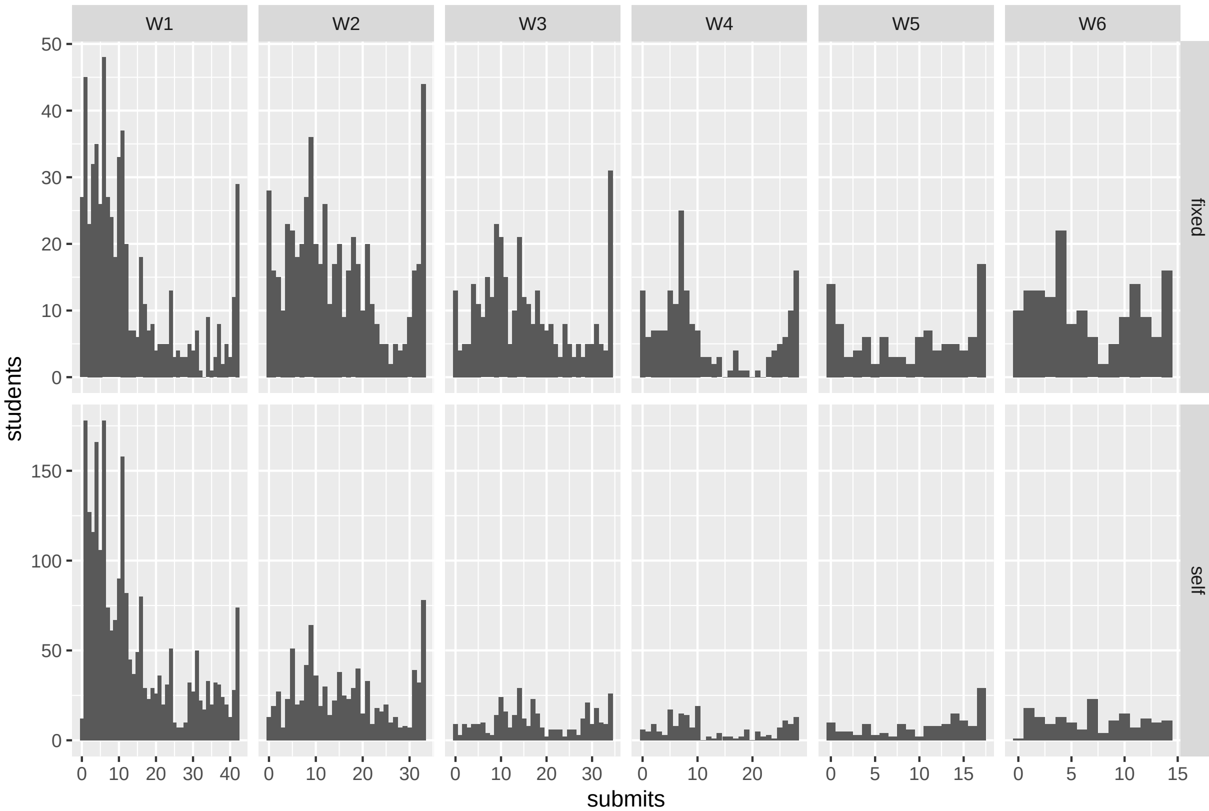

Fig. 2. Histograms of submits in the exit week.

analyzing when the dropouts occur, we observed that both MOOC versions have students who complete all programming assignments in a week, but do not continue further. For example, out of the students who chose to not continue into the second week of the MOOC, approximately 5% completed all the assignments in the first week. It is unclear what contributes to this phenomenon. As the workload of a single week in the course is over 10 hours, it is not clear to us why some would invest such amount of time just to drop out next week. It is possible that these students are such who, when they start to work on something, put their effort to it, but who may struggle with starting their work. Similarly, it is also possible that some notice that the course is not for them, or notice that the workload is too high. Further research is required to investigate this phenomenon.

As mentioned above, it is possible that the dropouts are influenced by other factors in addition to the schedule, one of which is grading. In the self-paced MOOC, students were expected to complete at least 90% of the programming assignments from each week before they could proceed to the subsequent week. This was not the case in the fixed-schedule MOOC, where students could hop in to the course even after the course had started. Such hopping in to the fixed-schedule MOOC is visible especially during the second week of the course, where approximately 421 new students entered the course. At the same time, the grading scheme of the fixed-schedule incentives students to complete more assignments to reach a good grade. Indeed, we observed (Table IV) that there is no significant difference in the number of assignments that students complete during the first week of the course.

Other factors such as students' self-regulation, motivations, and aspirations may influence the outcomes as well. We plan to address these factors as well in the future. Next, we look into the differences in students' effort between the courses to identify tacit behaviors that could be linked to some of these factors.

B. *Differences in effort between the courses*

The differences in effort between the MOOCs was analyzed using data collected by the programming environment. The data was aggregated to create measures that quantified the number of programming assignments that each student worked on, the number of programming assignments that each student submitted, the total number of typing events, and the number of days to which each student distributed their effort per programming assignment. These were aggregated to weekly statistics that took into account only those students who were active on that particular week (note that here weeks refer to sections in study material, not calendar weeks).

When comparing the cumulative effort in the course (Table III), students in the self-paced version complete more assignments and work on the course on more days than the students in the fixed-schedule version. At the same time, when we study the number of days that each assignment took to complete, or the typing events related to each assignment a student worked with, no significant differences exist between the MOOCs for the majority of the weeks. These results on the cumulative effort can partially be explained by the grading scheme differences between the courses: students in the self-paced course were expected to complete more assignments and for them to continue until a particular week, they must have completed over 90% of the assignments in the previous weeks. At the same time, students in the fixed-schedule version can continue in the course even if they do not complete any assignments. But, if students in the fixed-schedule version do not complete the majority of the assignments, the grading scheme of the course leads to a poor or a failed grade, which might further incentive the students in the fixed-schedule course to drop out.

Throughout the course, when comparing the courses and considering only the last week on which the student was active (Table IV), there is no difference in students' effort in terms of assignments worked on. Significant differences between the courses exist in the number of assignments submitted on weeks 2 and 3, but the differences are very small. This speaks against the role of the grading scheme explaining the differences in retention and hints that those differences are related more closely to scheduling. The main difference between the courses is the number of days on which the students work on the assignments. Students in the self-paced version space out their work on more days than students on the fixed-schedule version.

Overall, our research is in line with previous studies that have linked effort with progress [10]. Moreover, while we did not explicitly look into the effort invested in subsequent days or in shorter time spans, evidence on the relationship between high activity and performance within study sessions exist [23]. Such behavior and consequently higher learning outcomes may be encouraged in fixed-schedule courses, which provide students a time frame and deadlines within which they are expected to work.

C. Implications for practitioners

Based on our results and previous research, we highlight a set of guidelines for managing MOOC schedules. First, as we observe that students in the fixed-schedule MOOC are less likely to drop out from the course, practitioners seeking to maximize retention in MOOCs should consider running their course as one or more fixed-schedule courses. Further, simple email reminders can be an effective way of keeping the students aware about the deadlines and schedules [33]; alternatively, mechanisms could be included to cope with unexpected life events [28].

Second, as we observe that the early start (or early investment) in the course is a key factor that determines whether students continue in the course, practitioners should invest effort into making the first investments as easy as possible. This could be realized, e.g., through using smaller assignments as suggested by Denny [34], who have observed that students are more likely to start working early on small assignments than on large assignments, consequently increasing the overall effort that students invest into the course.

Third, as there are participants in the MOOCs who drop out, also likely to reasons out of their control, mechanisms for bringing dropouts back should be considered. Practitioners could, for example, invite dropouts from previous course versions to participate in the next version that will be launched. Furthermore, one could consider creating a buddy system where students would support each others as proposed in [35].

Finally, we also observe that there exists a proportion of students in the self-paced course version who complete the course. It is possible that some of these students have such commitments in their life that they cannot participate in fixed-schedule courses. Thus, practitioners should consider ways to offer self-paced course versions to students, although our results imply that it might be beneficial to direct students into self-paced or fixed-schedule courses based on some tacit background factors, which we are currently unaware of.

D. Implications for researchers

Our work highlights multiple research directions for researchers. First, as we observed significant differences in dropout between the courses, looking deeper into the factors that contribute to students dropping out (or staying) in courses with different schedules could lead to interventions that can increase retention.

Second, as mentioned previously, the main dropout happens early on in the course. Thus, research into how one could engage students to increase early investment and whether such investment could increase retention is called for. Possibilities in this area include, e.g., gamification, course material design, as well as the previously mentioned use of smaller assignments [34].

Finally, ways to balance grading and course format should be studied. In our case, students who perform poorly and complete only few assignments in the fixed-schedule version are penalized, which may lead to unnecessary drop outs. Similarly, for some students in the self-paced course, a boundary lower than the 90% completed assignments could be more effective. Identifying what types of students benefit from fixed schedules and what types of students benefit from self-paced schedules, and studying the interaction of these profiles with the effect of grading on behavior could lead to new and more effective ways of organizing and grading courses.

E. Limitations of work

Our study comes with a range of limitations, which we address next. We acknowledge that the courses under study have a specific topic (programming), which may attract particular type of students. Thus, further research that would look into the generalizability of these results into other areas and topics is needed. Second, as we have noted, the courses that we studied have different grading schemes, which are bound to influence students' behavior. Again, further research on the effect of grading schemes on MOOC completion is called for. Third, we acknowledge that we did not have information of students actual grades or course points, and did not look into whether students learned in the courses or not. At the same time, we quantified students behavior from the data logged by the programming environment, and we consider it unlikely that there would be a large number of students who would continue working on the course assignments, going from one to the next, without being able to actually complete them. Fourth, students were able to choose between self-paced and fixed-schedule MOOCs, which means there may be differences in the students who choose the more strict scheduling option. Further research is needed to account for this, for instance by randomly assigning students to each version and seeing if the differences hold.

VI. Conclusions

In this article, we analyzed students' retention and effort in self-paced and fixed-schedule MOOCs. The topic of the MOOC was programming, and the measures with which retention and effort were quantified were related to students work on the programming assignments.

To summarize the results, our answers to the research questions are as follows.

RQ1 *What are the differences in retention between the students in the self-paced and fixed-schedule MOOCs?* **Answer:** There are significant differences in retention between fixed-schedule and self-paced MOOCs. Based on our data, when measured through the effort and persistence in the first week, students in a fixed-schedule MOOC are 3.5 times more likely to persist in the course after than students in a self-paced MOOC. While the drop out is most noticeable early on in the course, the difference in dropout (students in self-paced MOOC being more likely to drop out than students in the fixed-schedule MOOC) persists almost all the way to the end of the course.

RQ2 *How does the effort (e.g., number of exercises or active days) differ between the students in the self-paced and fixed-schedule MOOCs?* **Answer:** When focusing on the effort on the week on which students drop out, no significant differences in the number of assignments worked and submitted exist between the fixed-schedule and self-paced MOOC. At the same time, students in the self-paced MOOC space their work over more days than students in the fixed-schedule MOOC. If we consider the differences in cumulative effort, students who stay in the self-paced course work on more assignments, submit more assignments, and work on more days on average. We acknowledge that this may be a product of the different grading schemes of the courses, where the grading of the fixed-schedule MOOC effectively encourages students who complete only a handful of assignments to drop out, while the grading of the self-paced MOOC enforces that students must complete at least 90% of the assignments for a particular week to continue to the next week.

Overall, we observed the common observation in MOOCs – both studied MOOCs have high drop-out rates. At the same time, the dropouts between the courses differ, meaning that the course scheduling has an effect on dropout rates, especially early on in the course. We also observed an interesting phenomenon that should be studied further: in both courses, a noticeable share of the participants complete all or almost all assignments in a week, but do not continue to the subsequent week. It is possible that there are some tacit factors involved, such as a perceived future workload, which may cause procrastination and inhibit students from starting the next week.

References

[1] L. Pappano, "The Year of the MOOC," *The New York Times, Nov. 2, 2012*.

[2] M. Rohs and M. Ganz, "MOOCs and the claim of education for all: A disillusion by empirical data," *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 6, 2015.

[3] W. Feng, J. Tang, and T. X. Liu, "Understanding dropouts in MOOCs," *Association for the Advancement of Artificial Intelligence*, 2019.

[4] J. Reich, "MOOC completion and retention in the context of student intent," *EDUCAUSE Review Online*, 2014.

[5] J. A. Greene, C. A. Oswald, and J. Pomerantz, "Predictors of retention and achievement in a massive open online course," *American Educational Research Journal*, vol. 52, no. 5, pp. 925–955, 2015.

[6] J. Leinonen, L. Leppänen, P. Ihantola, and A. Hellas, "Comparison of time metrics in programming," in *Proceedings of the 2017 ACM conference on international computing education research*, 2017, pp. 200–208.

[7] D. Clow, "MOOCs and the funnel of participation," in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 185–189.

[8] N. Wagner, K. Hassanein, and M. Head, "Who is responsible for e-learning success in higher education? a stakeholders' analysis," *Journal of Educational Technology & Society*, vol. 11, no. 3, pp. 26–36, 2008.

[9] K. M. Alraimi, H. Zo, and A. P. Ciganek, "Understanding the MOOCs continuance: The role of openness and reputation," *Computers & Education*, vol. 80, pp. 28–38, 2015.

[10] A. Hellas, P. Ihantola, A. Petersen, V. V. Ajanovski, M. Gutica, T. Hynninen, A. Knutas, J. Leinonen, C. Messom, and S. N. Liao, "Predicting academic performance: a systematic literature review," in *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, 2018, pp. 175–199.

[11] R. A. Rasheed, A. Kamsin, N. A. Abdullah, A. Zakari, and K. Haruna, "A systematic mapping study of the empirical MOOC literature," *IEEE Access*, vol. 7, 2019.

[12] T. Eriksson, T. Adawi, and C. Stöhr, ""time is the bottleneck": a qualitative study exploring why learners drop out of moocs," *Journal of Computing in Higher Education*, vol. 29, no. 1, pp. 133–146, 2017.

[13] H. Khalil and M. Ebner, "MOOCs completion rates and possible methods to improve retention-a literature review," in *EdMedia+ innovate learning*. Association for the Advancement of Computing in Education (AACE), 2014, pp. 1305–1313.

[14] J. Sorva, "Splashing the surface of research: A study of koli abstracts," in *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, 2019, pp. 1–2.

[15] M. Lepp, P. Luik, T. Palts, K. Papli, R. Suviste, M. Säde, and E. Tõnisson, "Mooc in programming: A success story," in *Proceedings of the International Conference on e-Learning*, 2017, pp. 138–147.

[16] K. S. Hone and G. R. El Said, "Exploring the factors affecting MOOC retention: A survey study," *Computers & Education*, vol. 98, pp. 157–168, 2016.

[17] C. Alario-Hoyos, I. M. Estévez Ayres, J. M. Gallego Romero, C. Delgado Kloos, M. C. Fernández Panadero, R. Crespo García, F. Almenares Mendoza, M. B. Ibáñez Espiga, J. Villena Román, J. Ruiz Magaña *et al.*, "A study of learning-by-doing in MOOCs through the integration of third-party external tools: comparison of synchronous and asynchronous running modes," *Journal of Universal Computer Science*, 2018.

[18] A. Vihavainen, M. Luukkainen, and J. Kurhila, "Multi-faceted support for mooc in programming," in *Proceedings of the 13th annual conference on Information technology education*, 2012, pp. 171–176.

[19] P. Luik, M. Lepp, T. Palts, M. Säde, R. Suviste, E. Tõnisson, and M. Gaiduk, "Completion of programming mooc or dropping out: Are there any differences in motivation," in *Proceedings of the 17th European Conference on e-Learning ECEL*, 2018, pp. 329–337.
[20] J. Hattie and H. Timperley, "The power of feedback," *Review of educational research*, vol. 77, no. 1, pp. 81–112, 2007.
[21] J. Warren, S. Rixner, J. Greiner, and S. Wong, "Facilitating human interaction in an online programming course," in *Proceedings of the 45th ACM technical symposium on Computer science education*, 2014, pp. 665–670.
[22] D. F. Onah, J. Sinclair, and R. Boyatt, "Dropout rates of massive open online courses: behavioural patterns," *EDULEARN14 proceedings*, vol. 1, pp. 5825–5834, 2014.
[23] M. Vitiello, S. Walk, V. Chang, R. Hernandez, D. Helic, and C. Guetl, "MOOC dropouts: A multi-system classifier," in *European Conference on Technology Enhanced Learning*. Springer, 2017, pp. 300–314.
[24] K. Sannicandro, A. De Santis, C. Bellini, and T. Minerva, "Analysis of completion and dropout rates in EduOpen MOOCs," *ITALIAN JOURNAL OF EDUCATIONAL RESEARCH*, pp. 27–42, 2019.
[25] W. R. Watson, J. H. Yu, and S. L. Watson, "Perceived attitudinal learning in a self-paced versus fixed-schedule MOOC," *Educational Media International*, vol. 55, no. 2, pp. 170–181, 2018.
[26] P. Tulsi, "Design of mooc on research in technical education," in *2018 Learning With MOOCS (LWMOOCS)*, 2018, pp. 5–8.
[27] Y. Chen and M. Zhang, "MOOC student dropout: Pattern and prevention," in *Proceedings of the ACM Turing 50th Celebration Conference - China*, ser. ACM TUR-C '17. New York, NY, USA: Association for Computing Machinery, 2017.
[28] I. Borrella, S. Caballero-Caballero, and E. Ponce-Cueto, "Predict and intervene: Addressing the dropout problem in a MOOC-based program," in *Proceedings of the Sixth (2019) ACM Conference on Learning @ Scale*, ser. L@S '19. New York, NY, USA: Association for Computing Machinery, 2019.
[29] Z. Niu, W. Li, X. Yan, and N. Wu, "Exploring causes for the dropout on massive open online courses," in *Proceedings of ACM Turing Celebration Conference - China*, ser. TURC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 47–52.
[30] T. Staubitz and C. Meinel, "Graded team assignments in MOOCs: Effects of team composition and further factors on team dropout rates and performance," in *Proceedings of the Sixth (2019) ACM Conference on Learning @ Scale*, ser. L@S '19. New York, NY, USA: Association for Computing Machinery, 2019.
[31] M. Khalil and J. Wong, "The essence of time: Taking a look at learning sessions in MOOCs," in *2018 Learning With MOOCS (LWMOOCS)*, 2018, pp. 131–133.
[32] J. Sinclair, R. Boyatt, C. Rocks, and M. Joy, "Massive open online courses: A review of usage and evaluation," *Int. J. Learn. Technol.*, vol. 10, no. 1, p. 71–93, May 2015. [Online]. Available: https://doi.org/10.1504/IJLT.2015.069450
[33] J. Martin, S. H. Edwards, and C. A. Shaffer, "The effects of procrastination interventions on programming project success," in *Proceedings of the Eleventh Annual International Conference on International Computing Education Research*, ser. ICER '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 3–11. [Online]. Available: https://doi.org/10.1145/2787622.2787730
[34] P. Denny, A. Luxton-Reilly, M. Craig, and A. Petersen, "Improving complex task performance using a sequence of simple practice tasks," in *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, ser. ITiCSE 2018. New York, NY, USA: ACM, 2018, pp. 4–9. [Online]. Available: http://doi.acm.org/10.1145/3197091.3197141
[35] M. Nelimarkka and A. Hellas, "Social help-seeking strategies in a programming mooc," in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 2018, pp. 116–121.