The impact of extra credit incentives on students' work habits when completing online homework assignments

Zachary Felker and Zhongzhou Chen
Department of Physics, University of Central Florida, 4111 Libra Drive, Orlando, Florida, 32816-2385

The current paper examines the effectiveness of offering a small amount of extra credit as an incentive to encourage proper work distribution and reduce procrastination and cramming among college students completing introductory physics homework assignments in the form of online learning modules. Students' distribution of work over time is systematically measured by clustering clickstream log events into study sessions according to a cutoff determined empirically using mixture model analysis. Significantly more study sessions are initiated well before the assignment due date when extra credit is offered compared to data from a previous semester. Using two proxy variables designed to capture the distribution and duration of work, we found that in addition to starting the assignments earlier, students also spent a longer time on the assignments. Finally, the benefit of extra credit in encouraging work distribution is not limited to high-performing students, as shown by a reduction in score gap between early and late starters on a midterm exam administered prior to the release of the homework assignment.

2020 PERC Proceedings edited by Wolf, Bennett, and Frank; Peer-reviewed, doi.org/10.1119/perc.2020.pr.Felker

I. INTRODUCTION

Many college students frequently procrastinate on assignments and cram before due dates and exams. Several studies have shown that those behaviors are frequently associated with lower academic performance or learning outcomes. While earlier researches relied on surveys or self-reports [1–5], more recent studies have utilized the log data from online learning platforms to measure quantitatively students' work distribution [6–10].

In addition to observing and measuring the level of cramming or procrastination, fewer studies have tried to answer the more important question of how to reduce such behavior and incentivize students to distribute the work. For example, Cavanaugh, Lamkin and Hu [11] sent a checklist to students to encourage early submission; Yeomans and Reich [12] sent planning prompts to students taking a massive open online course (MOOC). Ackerman and Gross [3] found that providing incentives for early completion could significantly reduce procrastination, while two much earlier papers measured the effectiveness of such incentives [4,5]. One way of providing this type of incentive is by giving extra credit for early completion of assignments, which rewards early submission and distribution of work but does not penalize those who are unable to do so. Advancement in online learning technology makes it easy to implement flexible forms of extra credit on major learning management systems such as Canvas [13].

This paper examines the effectiveness of offering such extra credit on students' work distribution and study time, as they complete online homework in the form of a sequence of Online Learning Modules (OLMs) [14–17]. **The first research question** (**RQ1**) that we answer is: How can we accurately measure students' distribution of work over time? We adopt an approach similar to that of Miyamoto et al. [10] by grouping learning events into "study sessions" separated by a time gap. Compared to the more common approach of analyzing event frequencies [6,7,9,18], study sessions better capture students' work distribution by providing information on not only the timing of work but also the duration of work. Unlike the Miyamoto study, which used a 30-minute minimum gap duration, we performed a more careful analysis on the distribution of gaps between learning events to estimate the separation between different study sessions. In addition, we also exclude those sessions that were likely generated by students quickly browsing through the learning materials. Based on the identified study sessions, we will answer **the second research question** (**RQ2):** Can a small amount of extra credit incentivize students to start working on assignments earlier?

In addition, we examine two potential side effects of using extra credit incentives to reduce cramming and procrastination. First, research on student self-regulated learning (SRL) suggests that some learners may focus more on achieving extrinsic goals, in this case acquiring the extra credit [19–21]. For those learners, it is possible that they will replace one cramming session with multiple cramming sessions to earn the extra credit without increasing their study time on learning materials. Second, students with higher SRL skills are not only more likely to earn the extra credit through proper planning and distribution of work, but are also likely to perform better in other aspects of the course such as exams [22]. Therefore, it is possible that extra credit will increase the performance gap between students by "making the rich richer". To measure the extent of those side effects, we will explore the answer to the following two research questions:

RQ3: To what extent do extra credit incentives lead to an increase in study time on the assigned learning materials?

RQ4: Are our extra credit opportunities benefiting predominantly high-performing students?

In Section II, we will describe in detail the implementation of extra credit, the collection and selection of log data as well as analysis methods to answer each of the four research questions. In Section III, we present the outcome of the analysis, followed by a discussion on the answers to those four questions in Section IV.

II. METHODS

A. Implementation of online homework assignments and extra credit

Online Homework Assignments in the form of OLM sequences are created on the free and open source Obojobo platform [23], developed by the Center for Distributed Learning at the University of Central Florida, which is integrated with the Canvas Learning Management System via Learning Technology Interoperability standards. Each module covers a single concept or one type of problem using an instructional component containing text and practice problems, and an assessment component containing 1 or 2 multiple choice problems. Several modules form an OLM sequence on a given topic and students must complete each module in order. The sequence in the current study consists of 10 modules on the topic of conservation of mechanical energy. The sequence was assigned as homework to be completed over two weeks in both Fall 2018 and Fall 2019 semesters. Students in each semester enrolled in the same calculus-based college introductory physics course taught in lecture format by the same instructor.

Extra credit for early completion: In Fall 2019 semester, students who completed the first 3, 6 and 9 out of 10 modules of the Energy sequence could each access one of three additional Canvas quizzes named "Treasure Troves". Each contained only a single question asking if students want to claim their extra credit. The three quizzes expire 10, 7 and 2 days before the assignment due date, and are worth 2, 2, and 3 points in extra credit, respectively. All "Treasure Trove" quizzes in the course (across all module sequences) contain a total of 21 points and are worth 5% of course credit.

B. Identifying valid study sessions from event logs (RQ1)

Collection of Online Event Log Data: A subset of all log events are selected for analysis. These events correspond to students navigating between instructional pages and assessments, viewing and answering questions, and entering and leaving the modules.

Identifying study sessions by clustering of log events: Our method for identifying students' study session from the log data is based on the following hypothesis: Longer breaks between study sessions and shorter breaks within a study session arise from different and independent student behavior. The former are closely related to the students' decision to distribute work over time, while the latter stem from the student taking breaks or running errands during a continued period of work. We measure these breaks using the time elapsed between a "viewer-exit" event, which triggers when either the module is closed or the browser window is minimized or remain inactive for more than 10 minutes, and the following "viewer-enter" event, triggered when the student returns. Therefore, the distribution of time elapsed between consecutive exit and enter events should contain two or more separate distributions, which can be separated by fitting the data using multi-component mixture model, similar to the analysis in several earlier papers [14,24–26].

Based on the fitting result (as shown in Fig. 1), we estimate the maximum separation for two consecutive log events to be considered to belong to the same study session. Students' study sessions can then be identified by clustering all events that took place within the maximum separation into one cluster, using the DBSCAN clustering algorithm from Scikit-Learn [27].

Selecting valid study sessions: Some of the study sessions identified via this method arise from students briefly glancing over the homework assignments. To ensure that the identified study sessions corresponds to actual engagement with the learning material, we filter out sessions that are either too short, contain too few events, or have a time density of events that is too low, according to the distribution of study sessions (Fig. 2) and methods explained in the next section.

C. Answering research questions 2-4

To answer RQ2 regarding whether extra credit can motivate students to distribute their work, we compare the distribution of the start dates of all valid study sessions with respect to the assignment due date and examine if the distributions are significantly different before and after the implementation of extra credit incentives (Fig. 3).

To examine the relationship between distribution of work and engagement with learning resources (RQ3), we create two proxy variables for each student. A student's distribution of work is captured by $\bar{d}$, the weighted average start date of all study sessions with respect to the due date:

$$\bar{d} = \frac{\sum_s m_s d_s}{\sum_s m_s}$$

where m_s is the number of modules accessed in study session s, and d_s the session start date relative to due date. The summation is over all valid study sessions s by the student. A student's engagement with learning resources is estimated by $\bar{t}$, the average study time per module:

$$\bar{t} = \frac{\sum_s t_s}{M}$$

where t_s is the duration of study session s and M is the total number of unique modules accessed by the student prior to the assignment due date. We then divide students into four categories based on the combined population median of each variable, as shown in Fig. 4 and listed in Table I, and compare the fraction of students in each category between Fall 2018 and Fall 2019.

To answer RQ4, which addresses who primarily benefits from extra credit, we compare the difference in normalized scores (z-scores) on a midterm exam administered before the homework assignment, between students in each of the four categories. We hypothesize that in the absence of extra credit, students who have lower exam scores are also more likely to "cram" close to the due date and engage less with the learning resources. If those students were incentivized to distribute their work, then the gap in normalized score between different categories will be reduced. Likewise, the gaps would increase if only high-performing students were motivated by the extra credit.

III. RESULTS

A. Identifying valid study sessions (RQ1)

We plot in Fig. 1 the distribution of time elapsed between all exit and enter events from the same student on the assigned modules for both semesters on a log scale. The data set is best fit with a three-component normal mixture model according to the Bayesian Information Criterion [28]. A reasonable interpretation is that the three components correspond to three types of behaviors: short interruptions

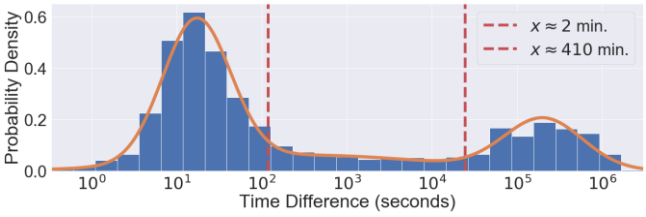

FIG. 1. Distribution of the time elapsed between viewer enter and viewer leave events. Plotted in orange is the three-component log-Gaussian mixture model fit. Dashed lines at $10^{2.08}$ (2 min.) and $10^{4.39}$ (410 min.) show the equiprobability boundaries between the three components.

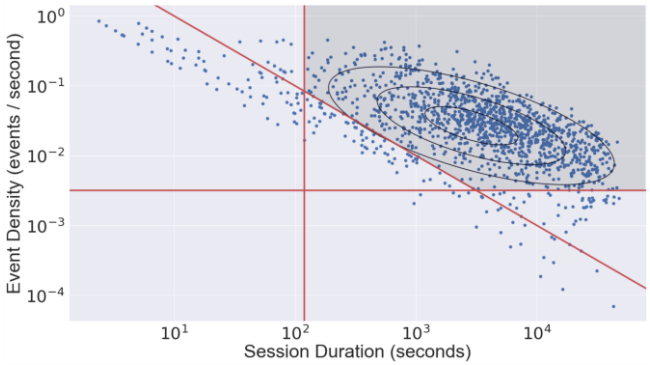

FIG. 2. Event density versus session duration of all sessions in both semesters (Fall 2018 and Fall 2019). Contour lines are based on a 2D log-normal fit and correspond to probability densities 0.15, 0.30, and 0.45. Note that this Gaussian fit is normalized in log-space and not for the axes as labeled. Red lines correspond to criteria for validity: duration ≥ 120 seconds, event density ≥ $10^{-2.5}$ events per second, and event number ≥ 10 events. The irregular shaded pentagon upper right contains "valid" sessions.

less than a minute during study (such as advancing from one module to the next), medium length interruptions (such as answering a phone call or taking a lunch break) that happen less frequently, and the separation between two study sessions longer than 7 hours that likely stem from deliberate distribution of work over multiple days. Therefore, in the current study we adopt $\varepsilon = 410$ minutes as the minimum gap separating two study sessions.

We identify a total of 689 study sessions for Fall 2018 semester and 811 sessions for Fall 2019 semester, which are plotted in Fig. 2 according to their duration and event densities. Of these, 79% reside within the boundaries: duration ≥ 120 seconds, event density ≥ $10^{-2.5}$ events per second, and event number ≥ 10 events. These boundaries are approximately tangential to the 0.15 contour line of the best fit 2D log-Gaussian distribution shown in the figure.

Therefore, we consider those 1191 study sessions within the boundaries valid study sessions. Among which, 542 were from Fall 2018 semester and 649 from Fall 2019 semester.

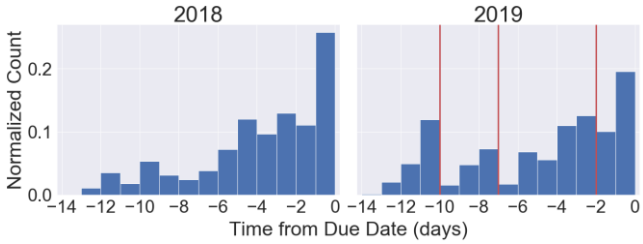

FIG. 3. Distribution of study session start dates with and without extra credit incentives (2018 and 2019, respectively). Sessions without extra credit are closer on average to the due date. Red lines indicate the extra credit deadlines in 2019: 10, 7, and 2 days before the due date. When extra credit is offered, we see clear peaks in graph corresponding to the first two extra credit deadlines: 10 and 7 days before the due date.

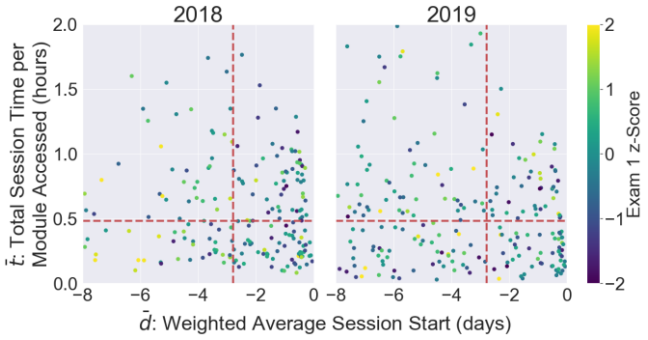

FIG. 4. Total session time per accessed module ($\bar{t}$) versus weighted average session start date ($\bar{d}$), colored by normalized Exam 1 score. A dotted line represents the combined median for each variable (lines have same position for both subplots). These lines divide the field into four behavior categories, starting upper right and moving counterclockwise: long late, long early, short early, and short late.

The difference in proportion of valid sessions between years is not statistically significant.

B. Impact of extra credit on session start dates (RQ2)

In Fig. 3 we show the distribution of start dates measured with respect to the due date for all valid sessions from each semester. The distributions differ significantly (Mann-Whitney U test, $p < 0.01$), with the Fall 2018 start dates grouped closer to the due date and Fall 2019 starting dates having two more peaks. The two additional peaks correspond to the expiration date of the first two "Treasure Trove" extra credit quizzes (10 and 7 days before due date).

TABLE I: Definition of the four behavior categories and number of students in each.

Category Name	Definition	No. of students (2018)	No. of students (2019)
Long Late (LL)	$\bar{d} \geq -2.82$ days $\bar{t} \geq 0.482$ hours	65	48
Long Early (LE)	$\bar{d} < -2.82$ days $\bar{t} \geq 0.482$ hours	44	71
Short Early (SE)	$\bar{d} < -2.82$ days $\bar{t} < 0.482$ hours	42	70
Short Late (SL)	$\bar{d} \geq -2.82$ days $\bar{t} < 0.482$ hours	60	55

C. Impact of extra credit on study time (RQ3)

To examine whether improvement in work distribution also resulted in increased study time (RQ3), we plot in Fig. 4 the distribution of the two proxy variables, $\bar{d}$ and $\bar{t}$ for all students in each semester. As can be clearly seen from the figure, the distribution shifted left, indicating an earlier start for the assignment, and up, indicating increased time spent

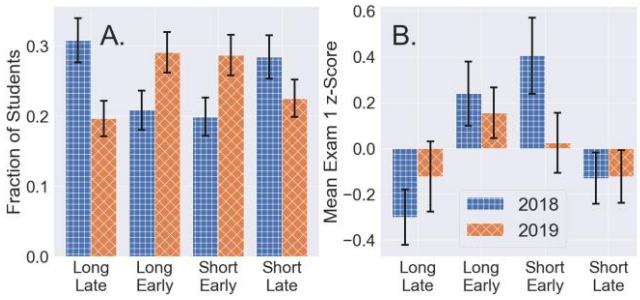

FIG. 5. Fraction of students and Exam 1 z-scores for students in each category and year. Error bars extend one standard error from the mean.

per module. To quantify this shift, we divide the data into four behavior categories based on the combined population medians for each variable, as listed in Table I. In Fig. 5A we plot the fraction of students in each category for both semesters. In the Fall 2019 semester, population in both LE and SE categories increased by almost 10%, whereas the population in LL and SL categories decreased by a similar amount. The differences between the two semesters are statistically significant ($p < 0.01$, extended Fisher's exact test).

D. Correlation between work distribution and exam performance (RQ4)

To answer RQ4, we plot the average exam scores for each of the four behavior categories in Fig. 5B. In the Fall 2018 semester, there is a significant gap between the two early categories (SL and LL) and two late categories (ANOVA, $df = 3, F = 5.8, p < 0.01$), which is in agreement with some earlier studies as well as common belief that low-performing students are also prone to more cramming and procrastination [1,2]. In the 2019 semester, the differences between the four categories are no longer statistically significant (ANOVA, $df = 3, F = 1.1,\ p = 0.34$). Since the exam scores are normalized, the result suggests that the extra credit motivated students who scored lower on the previous midterm exam to move from the two late categories into the two early categories. RQ4 asks whether extra credit benefits primarily high-performing students, "making the rich richer". We see here that while the extra credit is more likely to be earned by high performers (since they are still overrepresented in the early categories) the extra credit benefits low-performing students by motivating them to complete the assignments earlier than they otherwise would have.

IV. DISCUSSION AND FUTURE DIRECTIONS

In this paper, we first analyze the distribution of gaps between study events and determine an appropriate minimum duration for separating different study sessions (410 minutes). Compared to the 30-minute cutoff used in an earlier study [10], this longer gap duration better reflects students' deliberate choice to distribute work on different days, rather than taking a break or temporarily attending to a different task during a period of continued work. Compared to simpler metrics such as work done on each day, our clustering method can correctly identify study sessions that start just before midnight as a single session that spans two calendar days—a common occurrence with college students.

Based on the identified valid study sessions, we observe the following effects of extra credit on students' distribution of work: First, we see a significant reduction in cramming prior to the assignment due date, as more students begin work earlier in order to collect the extra credit. Second, the reduction in cramming is accompanied by an increase in number of students with longer study time. It is likely that some students are able to spend more time on the assignments as a result of starting earlier. Several students commented in the course evaluation survey that the extra credit motivated them to start working early. Finally, the reduce in score gap on an earlier exam between the four categories suggest that the change in work distribution is not limited to high-performing students, as more students who scored below average also started the assignment earlier compared to previous years.

Our results suggest that assigning extra credit for completing parts of the assignment early can be an effective method to encourage better work distribution and longer study time for both high and low performing students. Since extra credit can be easily implemented on most existing learning management systems such as Canvas, it serves as a valuable tool for both instructors and students, especially during the current mandatory distant learning period resulting from the COVID-19 outbreak.

One of the most prominent questions that need to be answered in follow up studies is whether better work distribution leads to better learning gains from the assignments, and how extra credit affects this relationship. In addition, it will also be valuable to examine how the total weight of extra credit assigned impacts both its effects and its side effects. Finally, the analysis methods developed in the current paper will allow us to study how students work habits change over the semester, especially during the Spring 2020 semester to study the impact of COVID-19 outbreak on students' work distribution and level of engagement.

ACKNOWLEDGMENTS

The authors are grateful to the Instructional Systems and Technology team at the University of Central Florida's Center for Distributed Learning for developing the Obojobo platform and providing the log data for analysis. This study is supported by NSF DUE 1845436.

[1] S. H. McIntyre and J. M. Munson, "Exploring cramming: student behaviors, beliefs, and learning retention in the principles of marketing course," J. Mark. Educ. **30**, 226 (2008).

[2] S. A. Nonis and G. I. Hudson, "Performance of college students: impact of study time and study habits," J. Educ. Bus. **85**, 229 (2010).

[3] D. S. Ackerman and B. L. Gross, "My instructor made me do it: task characteristics of procrastination," J. Mark. Educ. **27**, 5 (2005).

[4] L. L. Lamwers and C. H. Jazwinski, "A comparison of three strategies to reduce student procrastination in psi," Teach. Psychol. **16**, 8 (1989).

[5] R. A. Reiser, "Reducing student procrastination in a personalized system of instruction course," Educ. Commun. Technol. **32**, 41 (1984).

[6] J. Park, K. Denaro, F. Rodriguez, P. Smyth, and M. Warschauer, "Detecting changes in student behavior from clickstream data," Proc. Seventh Int. Learn. Anal. Knowl. Conf. - LAK '17 21 (2017).

[7] F. Rodriguez, M. J. Rivas, R. Yu, M. Warschauer, J. Park, and B. K. Sato, "Utilizing learning analytics to map students' self-reported study strategies to click behaviors in stem courses," ACM Int. Conf. Proceeding Ser. 456 (2019).

[8] D. T. Seaton, G. Kortemeyer, Y. Bergner, S. Rayyan, and D. E. Pritchard, "Analyzing the impact of course structure on electronic textbook use in blended introductory physics courses," Am. J. Phys. **82**, 1186 (2014).

[9] G. Kortemeyer, "Work habits of students in traditional and online sections of an introductory physics course: a case study," J. Sci. Educ. Technol. **25**, 697 (2016).

[10] Y. R. Miyamoto, C. A. Coleman, J. J. Williams, J. Whitehill, S. Nesterko, and J. Reich, "Beyond time-on-task: the relationship between spaced study and certification in moocs.," J. Learn. Anal. **2**, 47 (2015).

[11] T. Cavanaugh, M. L. Lamkin, and H. Hu, "Using a generalized checklist to improve student assignment submission times in an online course," J. Asynchronous Learn. Netw. **16**, 39 (2012).

[12] M. Yeomans and J. Reich, "Planning prompts increase and forecast course completion in massive open online courses," in *ACM Int. Conf. Proceeding Ser.* (Association for Computing Machinery, 2017), pp. 464–473.

[13] Instructure Inc., "Canvas learning management system," .

[14] Z. Chen, G. Garrido, Z. Berry, I. Turgeon, and F. Yonekura, "Designing online learning modules to conduct pre- and post-testing at high frequency," in *2017 Phys. Educ. Res. Conf. Proc.* (American Association of Physics Teachers, Cincinnati, OH, 2018), pp. 84–87.

[15] Z. Chen, S. Lee, and G. Garrido, "Re-designing the structure of online courses to empower educational data mining," in *Proc. 11th Int. Educ. Data Min. Conf.*, edited by K. Elizabeth Boyer and M. Yudelson (Buffalo, NY, 2018), pp. 390–396.

[16] K. M. Whitcomb, Z. Chen, and C. Singh, "Measuring the effectiveness of online problem-solving tutorials by multi-level knowledge transfer," in *Proc. 2018 Phys. Educ. Res. Conf.* (2018).

[17] Z. Chen, K. M. Whitcomb, and C. Singh, "Measuring the effectiveness of online problem-solving tutorials by multi-level knowledge transfer," in *2018 Phys. Educ. Res. Conf. Proc.* (American Association of Physics Teachers, 2018), pp. 1–4.

[18] D. T. Seaton, Y. Bergner, G. Kortemeyer, S. Rayyan, I. Chuang, and D. E. Pritchard, "The impact of course structure on etext use in large-lecture introductory-physics courses," 2013 Phys. Educ. Res. Conf. Proc. 333 (2014).

[19] P. H. Winne, "Self-regulated learning," *Self-Regulated Learning*, Second Edi (Elsevier, 2015).

[20] P. R. Pintrich, "A conceptual framework for assessing motivation and self-regulated learning in college students," Educ. Psychol. Rev. **16**, 385 (2004).

[21] M. Boekaerts, "Self-regulated learning: where we are today," Int. J. Educ. Res. **31**, 445 (1999).

[22] P. R. Pintrich and E. V. De Groot, "Motivational and self-regulated learning components of classroom academic performance," J. Educ. Psycology **82**, 33 (1990).

[23] Center for Distributed Learning, "Obojobo," .

[24] Z. Chen, M. Xu, G. Garrido, and M. W. Guthrie, "Relationship between students' online learning behavior and course performance: what contextual information matters?," Phys. Rev. Phys. Educ. Res. **16**, (2020).

[25] D. L. Schnipke and D. J. Scrams, "Modeling item response times with a two-state mixture model- a new approach to measuring speededness," J. Educ. Meas. **34**, 213 (1999).

[26] C. L. Barry, S. J. Horst, S. J. Finney, A. R. Brown, and J. P. Kopp, "Do examinees have similar test-taking effort? a high-stakes question for low-stakes testing," Int. J. Test. **10**, 342 (2010).

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: machine learning in python," J. Mach. Learn. Res. **12**, 2825 (2011).

[28] G. Schwarz, "Estimating the dimension of a model," Ann. Stat. **6**, 461 (1978).