Experiences With and Lessons Learned on Deadlines and Submission Behavior

Francisco Enrique Vicente Castro
francisco.castro@nyu.edu
New York University
New York, United States

Juho Leinonen
juho.2.leinonen@aalto.fi
Aalto University
Espoo, Finland

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

ABSTRACT

Course exercises are typically given so that the time it takes to finish them fits in the time constraints of the academic system. Exercises come with deadlines that are considered to help students plan their schedules and consequently help get the exercises done. Without deadlines, exercises that need to be done may easily slide away to make room for other tasks that are seemingly more important. Even with deadlines, however, some students procrastinate and leave their tasks without attention until the very last moment. In this article, we study computer science course exercise deadlines by analyzing data from a course that had different deadline placements over the years. The deadline placements of the course were varied to identify a deadline that would be suitable for the majority—if not all—of students. Our analyses from six different deadlines demonstrate that some deadlines seem to reduce last-minute work on exercises. Our findings highlight that not all deadlines are the same and serves as a call for more research into deadline placement and their potential impacts on student time management and performance.

CCS CONCEPTS

• **Social and professional topics** → *Computing education.*

KEYWORDS

deadline placement, deadlines, procrastination, time management, submission behavior

ACM Reference Format:
Francisco Enrique Vicente Castro, Juho Leinonen, and Arto Hellas. 2022. Experiences With and Lessons Learned on Deadlines and Submission Behavior. In *Koli Calling '22: 22nd Koli Calling International Conference on Computing Education Research (Koli 2022), November 17–20, 2022, Koli, Finland.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3564721.3564728

1 INTRODUCTION

Identifying deadlines that work for all students can be challenging. Deadlines given in courses may not always be favored by course participants, even if the deadlines would have justifications behind them. While deadlines can be helpful, they may also create anxiety and stress [44] among students. Some benefits to deadlines include supporting students' study strategies and, when designing deadlines in collaboration with other courses, helping avoid situations where multiple deadlines are clumped together.

Research into deadlines and earliness of study work (*i.e.*, when students start working on exercises) has provided some insight into work behavior and how that relates to study performance. Edwards et al. [11] and Parson and Seidel [38] observed that starting early led to better outcomes when compared to starting late, and Leinonen et al. [28] observed that students who started early tended to have better grades than those who started late.

Instructors' beliefs and notions may also influence which days instructors pick to schedule deadlines. Some may, for example, prefer to place deadlines on Fridays, with the hope that this would allow students to relax during the weekend. On the other hand, some may prefer to place deadlines on Sundays, with the hope that this would give students enough time to work on the exercises. There are also preferences and beliefs related to the time of the deadline. For example, some may place deadlines in the mornings, which would allow grading during the day, while others may prefer middle-of-the-night deadlines to also allow students to work for a whole day on the day of the deadline. While the notions above were identified in our own informal discussions with the course faculty responsible for the data that we have at our disposal, these beliefs and notions are also commonly discussed in informal contexts such as in mailing lists and at academic conferences. However, while such beliefs exist, quantitative evidence in favor or against them is lacking in research.

Our work provides a starting point for filling this gap in research. We quantitatively analyze data from course iterations with different deadlines. We first study to what extent course exercise deadlines relate to when students submit their work, after which we examine to what extent when students are working relates to the correctness of their work. To quantify *when*, we look into two metrics: *time of day* (*i.e.*, if they submit during the day or night, which hours of the day do students submit their work) and *distance to deadline* (*i.e.*, the number of hours and days from the submission to the deadline). Formally, our research questions are as follows:

- **RQ1** How are course exercise deadlines related to when exercises are submitted?
- **RQ2** How are the *time of day* and the *distance to deadline* related to the correctness of submissions?

We also look into evidence corroborating prior studies; prior work focusing on course deadlines and students' work has suggested that starting early in general tends to lead to more working days as well as better outcomes [8, 11, 28], but beyond noting that

Koli 2022, November 17–20, 2022, Koli, Finland

ACM ISBN 978-1-4503-9616-5/22/11.
https://doi.org/10.1145/3564721.3564728

students often tend to work close to the deadline [9], the effects of specific deadlines have received little attention. The closest matches to the present work are works studying: (1) the relationship of earliness and outcomes (*e.g.* [11, 15, 30, 38]), (2) interventions designed to nudge students to start their work earlier (*e.g.* [21, 22, 36]), (3) earliness and activity (*e.g.* [28]), (4) work patterns through exercise submissions (*e.g.* [45]), and (5) quality of work and when the work was conducted (*e.g.* [13]).

2 BACKGROUND

2.1 Deadlines and Time Management

When students start their work in relation to course deadlines (*e.g.*, for weekly exercises or course projects) has been studied using data from learning management systems, course materials, and instrumented programming environments (*e.g.*, [11, 24, 25, 28–31, 45]). Such data has also sometimes been combined with survey data to study the relationship between students' metacognitive strategies (*e.g.*, time management) and course outcomes (*e.g.*, [3, 31, 33, 43, 52]).

In general, the lack of time management skills among students has been found to be related to lower academic performance [34], which may manifest through a variety of ways. For example, poor time management can manifest as *procrastination*, where tasks are delayed until they can no longer be completed at an expected level (or at all) [14]. Similarly, poor time management skills can lead to poor study strategies including plagiarism [7], and can cause stress and anxiety [37]. The way individuals handle deadlines can have an effect on teamwork, as poor time management practices of individuals can influence the time management of team members [2]. On the other hand, good time management is linked with higher academic performance [34]. For example, students who start their work earlier are likely to perform better (*e.g.*, [9, 11, 36, 38]) and distribute their work over multiple days (*e.g.*, [8, 28]); this *spacing* of work over multiple days can already have a significant effect on learning [10, 17].

Students' time management is inevitably linked with course deadlines. Some studies within MOOCs have shown that having deadlines can increase retention in contrast to not having deadlines [20]. At the same time, motivations and attitudes related to joining courses can differ depending on whether courses have deadlines or not [53]. Naturally, in cases where courses do not have explicit deadlines, attendees can self-impose deadlines for themselves, although such deadlines are not as efficient for time management or controlling procrastination as compared to evenly spaced deadlines imposed by others [1]. How students behave with deadlines may be related to their feelings or preferences towards the course and to their other commitments; if given two events associated with different values and the possibility to decide whether the event associated with a higher value comes first or not, there are differences in the choices between individuals [32].

2.2 Deadlines and Procrastination

In principle, as some students have the tendency to procrastinate, poorly placed deadlines could potentially lead to students working during sub-optimal hours such as in the middle of the night, which may lead to more mistakes. Developers, for example, are more likely to introduce bugs in code during the night [13], and there is anecdotal evidence that some students may be more susceptible to poor decisions (such as committing plagiarism) during the night [16]. Simply requiring everyone to work during the day may also not be effective due to individual differences in the natural inclination to sleep or be active at certain times—*i.e.*, *chronotypes* [23, 40]—as certain chronotypes are more inclined to perform better early in the day [51]. Giving students flexibility to choose when to work may help students work during hours that fit them [54].

2.2.1 *Temporal motivation theory.*

In studies of students' time management, one theory that has been proposed to explain why deadlines can be helpful to curb procrastination is temporal motivation theory [48]. Temporal motivation theory is an extension of expectancy-value theory. In expectancy-value theory, it is posited that an individual's motivation to perform a task (*i.e.*, the task's *utility*) is reliant on both their prediction of their performance in the task, as well as the subjective benefit or worth of the task to the individual. Temporal motivation theory extends this by arguing that in addition to expectancies and values, the utility of a task is also dependent on time, that is, while motivation may increase when people are confident of achieving a desired outcome, motivation may decrease when there is a large amount of time before the outcome is achieved [49]. In essence, temporal motivation theory posits that the utility of a task, or the motivation to perform a task, increases as the deadline of that task becomes closer [47].

Steel et al. [49] have used the lens of temporal motivation theory to explain findings from prior studies on procrastination. For example, in prior classroom studies, researchers found that students' exercise submission patterns tended to form a hyperbolic curve, with procrastinators' submissions clustering around deadlines (*i.e.*, steeper curves) [18, 41]. Steel et al. explains that this is due to motivation being dependent on a goal's temporal distance, with motivation hyperbolically increasing as the deadline gets closer (in the studies above, submissions served as proxy for motivation). Temporal motivation theory has been used by other work to drive the design and embedding of anti-procrastination strategies within systems, such as the implementation of goal- and reward-setting to incentivize users to complete tasks on time [39], or the use of calls to action (CTAs) within massive online open course (MOOC) platforms, such as the use of *deadline reminders* to call students' attention to the proximity of deadlines and *descriptive norms* that communicate task completion by peers to influence students' perceived self-efficacy and increase motivation to complete tasks [19].

2.3 Good Deadlines

Ideally, *good* deadlines would account for multiple factors, such as avoiding deadline conflicts with other courses, providing ample time for students to complete their exercises, considering students who also have other critical priorities (*e.g.*, students who are parents or who also have professional jobs), or providing space for student well-being (*e.g.*, not having students work late into the night so they can get ample rest and sleep). As discussed, deadlines do matter, but setting *good* deadlines is not easy. There is prior work on the effect of deadlines on student behavior and outcomes [6, 9, 20, 42, 53], for and against using deadlines [1, 4, 5], and for longer deadlines leading to poorer outcomes when no reminders are used [50]. However, outside these examples and a few online resources such as [46], little empirical evidence or discussion exists on *how* deadlines should be placed, and whether some specific times or days are better for deadlines than others. The findings of our work contributes to this space.

3 METHODOLOGY

3.1 Course Contexts

To address our research questions, we analyzed submission timestamp data collected from three courses offered by a research-intensive university in Northern Europe with approximately 30,000 enrolled students. The courses from which we collected submission data include:

- An introductory programming (*INTROCS*) course (6 iterations),
- A web software development (*WEBDEV*) course (6 iterations), and
- An introductory statistics (*INTROSTATS*) course (4 iterations).

Each course is seven weeks long and worth 5 ECTS[1] credits (approximately 125 to 150 hours of study time). A student typically takes two to three courses in an academic quarter. All of the courses have exercises that are submitted to a learning management system for assessment. We used data from multiple courses in our analyses in order to create a form of baseline submission pattern to which we can compare different deadlines. Only the submission data that was collected by the learning management system was available, which did not include student demographic information.

All three courses focused heavily on the completion of course exercises and reading online course materials. The introductory statistics (*INTROSTATS*) course was in an online MOOC format (no face-to-face/synchronous lectures), the web software development (*WEBDEV*) course typically had a single lecture in the beginning of the course with the rest of the course delivered online, and the introductory programming (*INTROCS*) course typically had one face-to-face lecture each week. All courses had online support available through course chat rooms, and the *INTROCS* course had walk-in labs where students could receive support from teaching assistants.

3.1.1 *The web development course.*

Of the three courses we collected submission data from, the *WEBDEV* course had varying deadlines over the iterations, where the course instructor (the third author) had explicitly sought to find a deadline that would reduce procrastination and last minute submissions as much as possible. Within that course, the instructor and the teaching approach remained the same over the years when the data was collected. The *WEBDEV* course also has fewer confounding variables since the content across iterations was only marginally modified and it was offered in the same semester of every year, thus the students taking the course are likely in the same stage of the degree program (second-year students). Hence, our analyses focus on the *WEBDEV* course in more detail.

[1]European Credit Transfer and Accumulation System. One ECTS accounts for approximately 25 to 30 hours of study work.

3.1.2 *Exercises for all the courses.*

The *INTROCS* and *INTROSTATS* courses are typically taken during the first year of studies, while the *WEBDEV* course is taken during the second year of studies. Each course uses a *many small exercises* model and hands out weekly exercise sets with five to twenty exercises per set, and the students have approximately ten to eleven days to work on an exercise set. For all the courses, students have one exercise set deadline per week. All exercises in the courses were assigned to be done individually. Similar to the *WEBDEV* course, across different course instances, the *INTROCS* and *INTROSTATS* course instructors were the same and course materials only had minor modifications. The same instructor taught the *INTROCS* and *WEBDEV* courses, while the *INTROSTATS* course was taught by a different instructor. The instructor for *INTROCS* and *WEBDEV* was the third author of this paper.

The exercises are automatically assessed and the courses have no upper limit to the number of submissions per exercise. This means that a student can submit an exercise as many times as they wish. If a submitted exercise does not pass the automated tests, the student is given feedback that includes, for example, suggestions on the types of inputs to try and the types of outputs to expect, which then helps students to pinpoint issues in their exercises. Late submissions are not allowed.

3.2 Analysis Approach and Data

In the analyses, we excluded exercises that were submitted more than 12 days before the deadline to limit issues caused by differences in the release schedules of exercises. In addition, we only included the submissions (both correct and incorrect) for an exercise up to the first correct submission of that exercise for each student. In case a student had no correct submissions for a particular exercise, all submissions from the student were included in the data. We filtered out duplicate correct submissions to remove noise from the data (*e.g.*, there were a handful of students who submit exercises multiple times just before the deadline to confirm that they have submitted their work).

This led to a data set with 229,589 submissions from 3,514 students[2]. For the *WEBDEV* course instances, the weekly deadlines, number of submissions, number of students, and number of exercises are outlined in Table 1. For "midnight" deadlines (*e.g.*, "Wednesday Midnight"), we mean that students had time to submit exercises until the end of the day (*e.g.*, the midnight between Wednesday and Thursday). In practice, deadlines were explicitly reported to the students using the 24-hour time format (*e.g.* "Wednesday, 23:59") to reduce confusion among students about the time and day of submission; the 24-hour time format is commonly used in Finland where the courses were run.

For *RQ1*, we used all the data outlined in Table 2 for calculating baseline submission behavior, and submissions from the different iterations of the *WEBDEV* course were compared to that baseline. For *RQ2*, we utilized all the submission data from all courses. Note that as some students took the *WEBDEV* course multiple times

[2]Note that a student might be counted multiple times: for example, if they took all three courses (*INTROCS*, *INTROSTATS*, and *WEBDEV*), they will be counted once for each course. The number of students is reported to contextualize the data; all analyses focus on aggregate submission behavior and thus in practice does not consider individual students.

Table 1: *WEBDEV* course instances with varying deadlines, including average and median distances (in hours) to deadline for submissions.

Instance	Weekly Deadline	Submissions	Students	Exercises	Average (hours)	Median (hours)
Fall 2012	Monday 6 AM	5,981	135	37	71.5	62.6
Fall 2013	Wednesday Midnight	5,012	118	41	89.5	77.5
Fall 2014	Thursday 6 PM	7,098	155	47	86.3	74.5
Fall 2015	Monday 4 PM	4,120	121	25	65.7	48.3
Fall 2016	Tuesday Midnight	4,662	110	56	72.5	61.9
Fall 2017	Friday Midnight	3,833	123	30	96.8	83.7

(*e.g.*, due to failing the first time), the numbers of students for the *WEBDEV* course in Table 1 summed for all years does not equal the number of students in Table 2 for that course.

4 RESULTS

4.1 Submission Behaviors Over the Week Before the Deadlines

Using the submission data collected from the six *WEBDEV* courses (deadlines and submission counts outlined in Table 1), we analyzed how the deadlines relate to when students submit their exercises (*RQ1*). Submission behavior during the week (*i.e.*, the last seven days) before the deadline, calculated over all course weeks, is shown in Figure 1 (a-f). The lines display a probability density function that shows the likelihood of a submission being at a specific time in the data (averaged to the closest hour; the line has been smoothed for visualization purposes). The blue solid line represents the specific *WEBDEV* course and the dashed orange line represents the average over all our data (all *WEBDEV*, *INTROCS*, and *INTROSTATS* courses). The red vertical line marks the specific time of the deadline. In practice, when looking at, for example, Figure 1, we could interpret the highest peaks on Wednesday and Thursday so that it is about twice as likely that an exercise is submitted during the peak hour on Thursday than during the peak hour on Wednesday.

First, when visually analyzing students' submission behavior, we observe that for the *WEBDEV* courses where the deadline was varied, students were more likely to work close to the deadline in some iterations. Noticeable peaks are visible for the Monday 6 AM (Figure 1.a), Thursday 6 PM (Figure 1.c), and Monday 4 PM (Figure 1.d) deadlines, while for the other deadlines, the peaks are more subtle. For the deadlines placed at midnight, there are no considerable spikes, even during the day. Additionally, it seems that for deadlines placed on Mondays, students are more likely to work on the exercises in the preceding weekend, as opposed to deadlines in the latter part of the week (*e.g.*, Thursday or Friday), where the students could also have started their work early during the previous weekend.

Acknowledging that the figures only show activity during the last week before the deadline, which excludes those who submitted prior to the last 168 hours before the deadline (7x24 hours), we calculated the average and median distance of submissions to deadline (in hours) for the *WEBDEV* course instances, shown in Table 1. The data shows that the Monday 4 PM deadline yields the smallest average and median distance to deadline, while the Friday midnight deadline yields the largest average and median distance to deadline. In terms of differences in median time to deadline, the difference between the Monday 4 PM deadline and the Friday midnight deadline is approximately 35.4 hours.

In addition to the visual analysis, we conducted a Kruskal-Wallis H test to examine whether there are statistically significant differences in the distribution of the "*distance to deadline in hours*" between the different *WEBDEV* course instances—a significant difference would suggest that submission behavior was different between the courses. The Kruskal-Wallis H test resulted in a very low p-value ($p < 0.0001$, Bonferroni corrected), indicating that the distributions were different across course iterations. We then further conducted pairwise Mann-Whitney U tests between the course iterations to determine which pairs are significantly different. We found that for all of the courses, the distributions are different ($p < 0.01$, Bonferroni corrected) with the exception of the Monday 6 AM deadline (Fall 2012) and the Tuesday midnight deadline (Fall 2016). The effect sizes, calculated using Epsilon squared (ϵ^2) [12], are mostly negligible ($\epsilon^2 < 0.01$, 7 observations) to weak ($\epsilon^2 < 0.04$, 5 observations), with the exception of moderate effect size ($\epsilon^2 < 0.16$, 2 observations) between the Wednesday midnight deadline (2013) and the Monday 4 PM deadline (2015), as well as between the Monday 4 PM deadline (2015) and the Friday midnight deadline (Fall 2017). These results are summarized in Table 3.

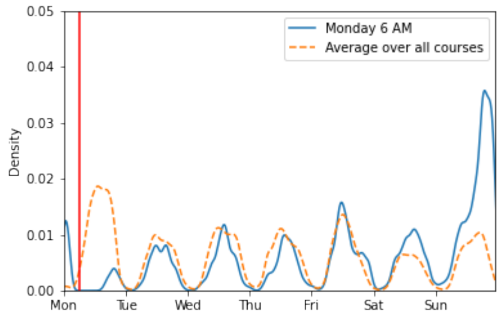
4.2 Submission Behaviors Closer to the Deadlines

To explore the extent of submissions done closer to the deadline, we quantitatively analyzed the proportion of submissions within a set of *windows* before the deadline, calculated over all course weeks (Table 4). For these windows, we chose 72 hours, 24 hours, 12 hours, 6 hours, and 2 hours before the deadline. From Table 4, we can see that the Monday 4 PM deadline has the most submissions done in the last 72 hours, while the Friday midnight deadline is the opposite and has the least submissions done. Similarly, but not surprisingly, the Monday 6 AM deadline has the least submissions during the last 2 hours before the deadline, while the Tuesday midnight deadline has the most submissions in the last 2 hours before the deadline. We also observed differences in the full data set when compared to the *WEBDEV* course; we elaborate on this further in the limitations (Section 5.4).

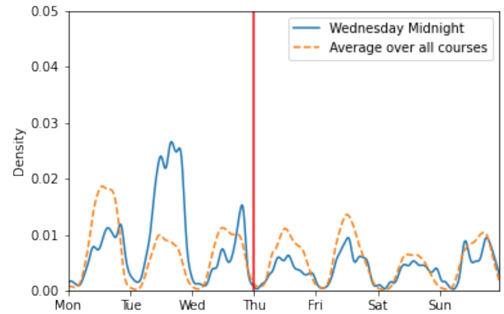
We then analyzed the submission behavior during the last 24 hours before the deadline. This is shown in Figure 2 (a-f), where the lines show the probability density function of the specific *WEBDEV*

Table 2: Years offered, total student counts, and total submission counts for each of the three courses.

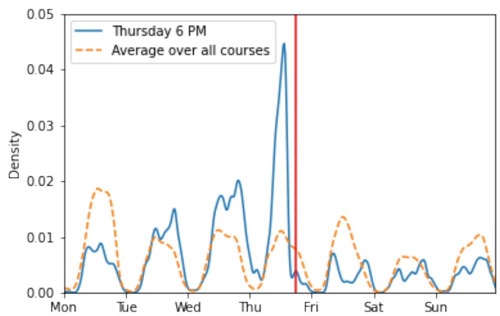
Course	Years	Total students	Total submissions
Introduction to Programming (*INTROCS*)	2012-2018	1,775	175,686
Introduction to Statistics and R (*INTROSTATS*)	2019-2020	1,333	23,197
Web Software Development (*WEBDEV*)	2012-2017	701	30,706



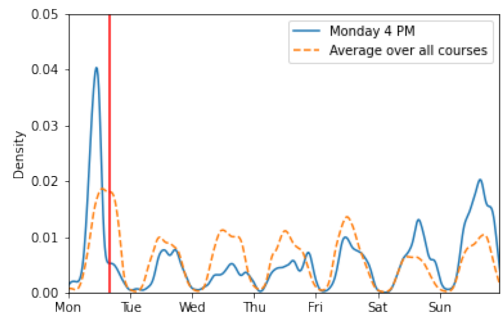
(a) 2012: Monday 6 AM deadline



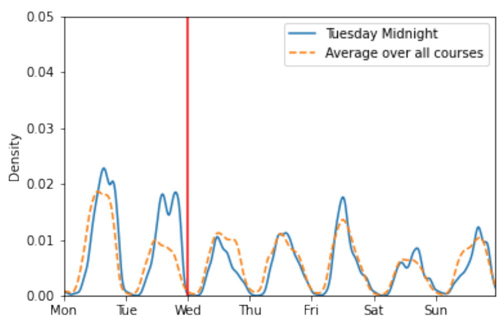
(b) 2013: Wednesday midnight deadline



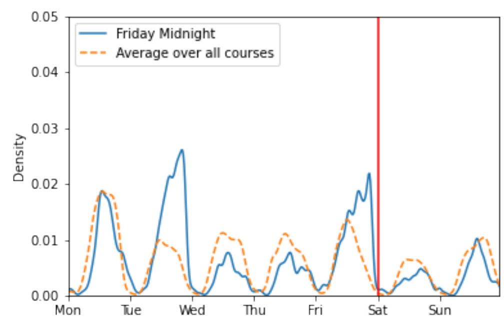
(c) 2014: Thursday 6 PM deadline



(d) 2015: Monday 4 PM deadline



(e) 2016: Tuesday midnight deadline



(f) 2017: Friday midnight deadline

Figure 1: Distribution of submissions over the last week before the deadline for each of the *WEBDEV* course instances. Blue solid line represents the specific *WEBDEV* course and the dashed orange line represents the average over all the data.

course instances and the full dataset (all *WEBDEV*, *INTROCS*, and *INTROSTATS* courses) similar to Figure 1 (a-f). A visual analysis

Table 3: Pairwise Mann-Whitney U tests between the different *WEBDEV* deadlines on whether the distance to deadline differs. The upper triangle shows the p values for the pairwise tests ($n.s.$ indicates non significant after Bonferroni correction; threshold 0.01). The lower triangle shows the effect sizes calculated using Epsilon squared and their interpretations (N=negligible, W=weak, M=moderate, R=relatively strong).

	2012	2013	2014	2015	2016	2017
2012: Monday 6 AM	-	$\frac{1.3}{10^{53}}$	$\frac{1.2}{10^{16}}$	$\frac{5.4}{10^{5}}$	$n.s.$	$\frac{1.1}{10^{56}}$
2013: Wednesday Midnight	0.022, W	-	$\frac{1.4}{10^{17}}$	$\frac{2.2}{10^{91}}$	$\frac{2.7}{10^{46}}$	$\frac{3.2}{10^{5}}$
2014: Thursday 6 PM	0.005, N	0.006, N	-	$\frac{8.0}{10^{36}}$	$\frac{2.6}{10^{11}}$	$\frac{1.1}{10^{21}}$
2015: Monday 4 PM	0.002, N	0.045, M	0.139, W	-	$\frac{5.3}{10^{8}}$	$\frac{2.3}{10^{97}}$
2016: Tuesday Midnight	$n.s.$	0.021, W	0.004, N	0.003, N	-	$\frac{6.4}{10^{50}}$
2017: Friday Midnight	0.026, W	0.002, N	0.008, N	0.055, M	0.026, W	-

Table 4: Percentage of submissions in the last 72, 24, 12, 6, and 2 hours before the deadline for the *WEBDEV* course instances, the *WEBDEV* course instances combined (All *WEBDEV*), and the full dataset including all courses (All Data).

Instance	Weekly Deadline	% Last 72 hrs	% Last 24 hrs	% Last 12 hrs	% Last 6 hrs	% Last 2 hrs
Fall 2012	Monday 6 AM	58.4%	33.5%	23.9%	8.3%	1.0%
Fall 2013	Wednesday midnight	46.7%	11.2%	9.6%	6.7%	2.3%
Fall 2014	Thursday 6 PM	49.7%	29.2%	18.4%	14.6%	3.6%
Fall 2015	Monday 4 PM	62.6%	34.0%	19.6%	16.9%	5.5%
Fall 2016	Tuesday midnight	51.4%	27.9%	25.7%	15.5%	6.3%
Fall 2017	Friday midnight	32.4%	19.6%	16.4%	10.1%	4.1%
All *WEBDEV*	-	50.7%	26.5%	19.0%	12.0%	3.6%
All Data	-	36.8%	18.0%	12.6%	7.2%	2.0%

shows that the submission behavior (blue line in the subfigures) differs between the courses. We conducted pairwise Mann-Whitney U tests to determine whether the submission behavior during the last 24 hours is different between the different *WEBDEV* course iterations. The results of the pairwise tests are shown in Table 5. From the table, we can observe that in most cases, there are significant differences (p < 0.01 after Bonferroni correction). However, in most of the cases, the effect sizes are weak to negligible with a few exceptions. Notably, the effect size between the Monday 6 AM deadline and the Tuesday midnight deadline is relatively strong, and the effect sizes between the Monday 6 AM and Wednesday midnight and Friday midnight deadlines are moderate.

4.3 Deadlines, Times of Work, and Correctness of Submissions

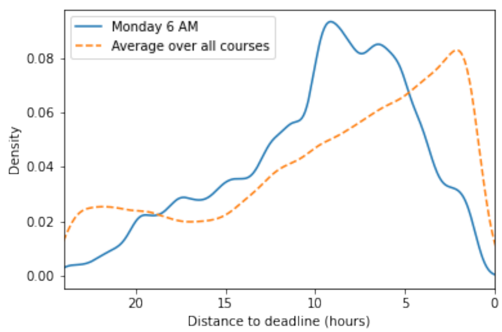
When analyzing correctness of submissions, we looked into (a) the average correctness over the week and (b) the average correctness over the hour of data. We performed the analysis on (1) submissions until, and including, the first correct submission, and on (2) all submissions (*i.e.*, including those after the first correct submission). The analysis here uses data from all the courses at our disposal—*i.e.*, the *INTROCS*, the *INTROSTATS*, and the *WEBDEV* courses.

4.3.1 Submissions until (and including) the first correct submission. On average, when looking at submissions in the data until, and including, the first correct submission (n=229,589), 74.0% of the submissions passed the automated tests. There are differences between the courses, however. The percentage of passing submissions was 83.0% for *INTROCS*, 64.0% for *WEBDEV* and 19.1% for *INTROSTATS*. A chi-squared test between the different course iterations suggests that there are differences in the number of passing and failing submissions ($p < 0.0001$, Bonferroni corrected). This is true also when only testing for differences between the *WEBDEV* course iterations ($p < 0.0001$, Bonferroni corrected).

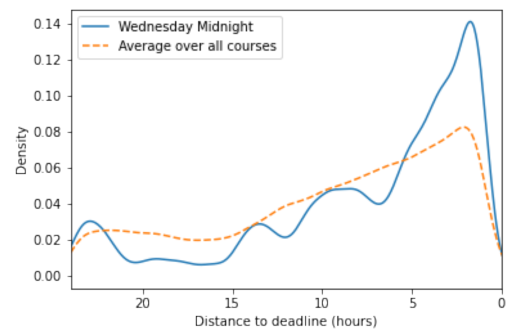
We then analyzed the correctness of submissions made within the seven days preceding the deadline by calculating the average correctness of the submissions based on distance to deadline and then observing trends in the data. Overall, as shown in Figure 3a, there seems to be a trend that submissions made closer to the deadline are less likely to be correct. A Mann-Kendall statistical test for trend [26, 35] confirms the observation; there is a statistically significant decreasing trend in average correctness when approaching the deadline (z=-3.74, τ=-0.89, $slope$=-0.022, $p \approx 0.0002$).

Additionally, to exclude the possibility of this being due to weaker students working closer to the deadline, we separately analyzed the trend for students who worked both close (<12h) and further away (>72h) from the deadline. The Mann-Kendall result for these students also reveals a decreasing trend in submission correctness closer to the deadline (z=-2.75, τ=-0.59, $slope$=-0.01, $p \approx 0.006$).

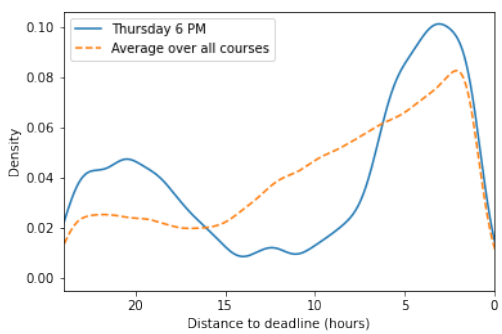
We further looked into if *when* the work was conducted influenced correctness of submissions. Overall, as shown in Figure 3b, the average correctness of submissions seems to be somewhat related to the time of day. Using definitions for *night* and *day* based



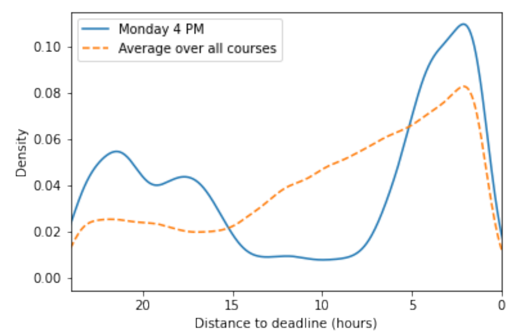
(a) 2012: Monday 6 AM deadline



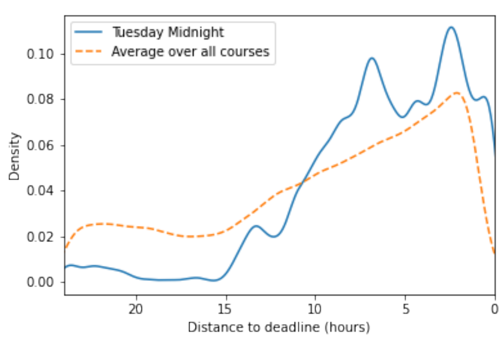
(b) 2013: Wednesday midnight deadline



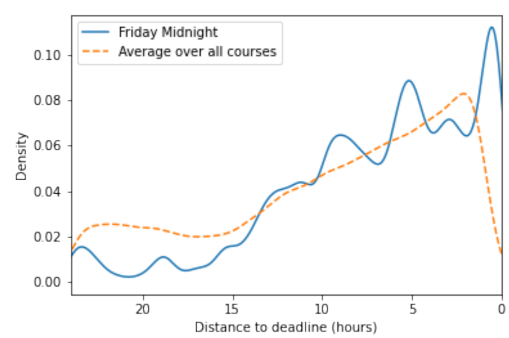
(c) 2014: Thursday 6 PM deadline



(d) 2015: Monday 4 PM deadline



(e) 2016: Tuesday midnight deadline



(f) 2017: Friday midnight deadline

Figure 2: Distribution of submissions over the last 24 hours before the deadline for each of the *WEBDEV* course instances. Blue solid line represents the specific *WEBDEV* course and the dashed orange line represents the average over all the data.

on Finnish law regarding night work[3], we divided the data into *day* (6 AM to 11 PM, n=219,839 submissions) and *night* (11 PM to 6 AM, n=9,750 submissions). During the day, approximately 74.7% of the submissions were correct, while during the night, approximately 66.0% of the submissions were correct. A chi-squared test indicated that the groups differ ($\chi^2 \approx 339.7, df = 1, p \approx 0.0$); Cramer's V ($V \approx 0.038$) indicates that the effect is negligible.

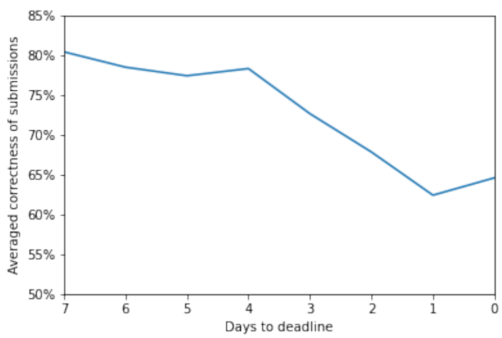
[3]Definition of night work from Finnish law (translated from Finnish): "Work that is conducted between 23 (11 PM) and 6 (6 AM) is night time work", https://www.finlex.fi/fi/laki/ajantasa/2019/20190872

Additionally, again to exclude the possibility of the time of day results being due to weaker students working during the night, we separately analyzed only students (n=1493) who worked both during the day and during the night. A chi-squared test shows ($\chi^2 \approx 197.2, df = 1, p \approx 0.0$) that for these students too, submissions during the night have lower average correctness (66.7%) compared to submissions during the day (73.3%).

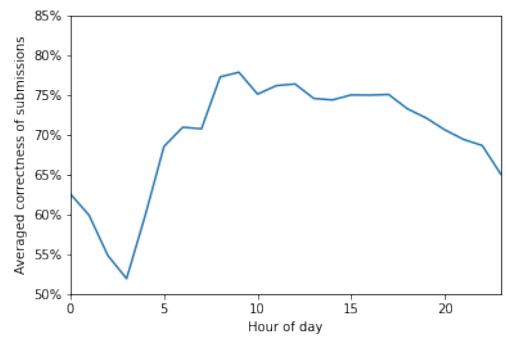
4.3.2 All submissions. For completeness, we also considered all submissions in addition to the submissions until, and including,

Table 5: Pairwise Mann-Whitney U tests for the submission behavior between the different *WEBDEV* deadlines for the last 24 hours before the deadline. The upper triangle shows the p values for the pairwise tests ($n.s.$ indicates non significant after Bonferroni correction; threshold 0.01). The lower triangle shows the effect sizes calculated using Epsilon squared and their interpretations (N=negligible, W=weak, M=moderate, R=relatively strong).

	2012	2013	2014	2015	2016	2017
2012: Monday 6 AM	-	$\frac{2.7}{10^{51}}$	$\frac{3.9}{10^{32}}$	$\frac{1.7}{10^{15}}$	$\frac{6.5.}{10^{120}}$	$\frac{3.7}{10^{49}}$
2013: Wednesday Midnight	0.089, M	-	$\frac{3.2}{10^{8}}$	$\frac{3.2}{10^{8}}$	$n.s.$	$n.s.$
2014: Thursday 6 PM	0.036, W	0.013, W	-	$n.s.$	$\frac{6.7}{10^{20}}$	$\frac{3.1}{10^{7}}$
2015: Monday 4 PM	0.019, W	0.017, W	$n.s.$	-	$\frac{5.3}{10^{8}}$	$\frac{2.3}{10^{97}}$
2016: Tuesday Midnight	0.166, R	$n.s.$	0.026, W	0.031, W	-	$\frac{1.8}{10^{5}}$
2017: Friday Midnight	0.080, M	$n.s.$	0.010, W	0.016, W	0.007, N	-



(a) Average submission correctness calculated over (rounded) days to deadline.



(b) Average submission correctness over hours of the day.

Figure 3: Relationships between (a) submission correctness and days to deadline using submissions until, and including, the first correct submission and (b) submission correctness and time of day.

the first correct submission. When considering all submissions (n=234,336), the observations are similar to the observations focused on submissions until, and including, the first correct submission. In all submissions, 74.3% of the submissions are correct. First, when considering the correctness of the submissions over the seven days preceding the deadline, Mann-Kendall statistical trend test shows a statistically significant decreasing trend in average correctness (z=-2.85, τ=-0.86 , slope=-0.024, $p \approx 0.004$).

Second, when considering the correctness of the submissions during day and night, 74.6% of the submissions (n=224,031) made during the day are correct, while 67.7% of the submissions (n=10,305) made during the night are correct. Chi-Squared test indicated that the groups differ ($\chi^2 \approx 250.8, df = 1, p \approx 0.0$); Cramer's V ($V \approx 0.033$) indicates that the effect is negligible.

5 DISCUSSION

While there is some research into deadlines such as exploring the impact of "optional" early deadlines for feedback [9] and comparing courses with and without deadlines [20], the placement of deadlines with regards to time of day and day of the week has not been studied extensively. Most research into deadlines thus seems to implicitly assume that all deadline placements are equal. The results of our study contest this assumption as we found that students' submission behavior was different when the placement of the deadline—both regarding time of day and day of week—was different.

5.1 Deadline Placement and Students' Work

We found that *WEBDEV* students from different course iterations, each with different deadlines, exhibited different submission behaviors. In general, across the different deadlines, students tended to work closer to, than further away, from the deadline, but perhaps not to the extent we suspected a priori. Out of the *WEBDEV* deadline placements that we analyzed (Table 1), the Monday 4 PM deadline had the lowest average (65.7 hours) and median (48.3 hours) time to deadline for submissions, and the Friday midnight deadline had the highest average (96.8 hours) and median (83.7 hours) time to deadline. One likely explanation for this is that the Monday deadline is directly preceded by the weekend while the Friday deadline is further away from the preceding weekend. For example, if there are students who only have time to work on course exercises on the weekend, their work (submissions) will be closer to the deadline with deadlines that occur in the beginning of the week, *e.g.*, Monday, compared to deadlines that occur further away from the preceding weekend, *e.g.*, Friday. We also observed that deadlines placed early in the week seem to have more students submitting during the preceding weekend, while deadlines towards the end of the week (*e.g.*, Thursday and Friday) show fewer students submitting exercises during the weekend.

Interestingly, we found that in our context, the placement of deadlines affected whether there were noticeable peaks with regards to day of the week in students' submission activity (see Figure 1). Specifically, when the deadline was on Monday 6 AM, Thursday 6 PM, and Monday 4 PM, students were very active in the last 24 hours before the deadline; while for Wednesday midnight, Tuesday midnight, and Friday midnight, the activity peaks were not as noticeable. Our results thus seem to suggest that a midnight deadline could discourage working close to the deadline. However, even in the courses with high activity peaks near the deadline, there were many submissions further away from the deadline too; considering all the data (the *WEBDEV*, *INTROCS*, and *INTROSTATS* courses), approximately 36.8% of the submissions (50.7% considering only the *WEBDEV* course) were made during the last 72 hours before the deadline.

A peak in activity near the deadline could indicate procrastination, and it is possible that in these cases, students would have preferred to continue to work on exercises after the deadline, suggesting that a later deadline would have been preferable to them. On the other hand, we do not have information on why the peaks happened. For example, the deadlines of other courses students are taking concurrently likely affects students' submission behavior.

Regardless of the chosen deadline, we observed that the day of the deadline had either the highest (Monday 6 AM, Thursday 6 PM, Monday 4 PM) or the second highest (Wednesday midnight, Tuesday midnight, Friday midnight) peak in submission activity (shown in Figure 1). This result is in line with temporal motivation theory [47, 48], supporting the notion that the utility value of the task, or the motivation to do a task (*e.g.*, to complete exercises, receive points from exercises, or learn about the topic from the exercises), increases towards the deadline. Interestingly, it seems that for the midnight deadlines, the peaks were on days other than the day of the deadline (Figure 1). Future work should examine whether this is due to students perceiving the midnight deadline to be closer than, for example, a 6 PM deadline (affecting the "time" aspect of temporal motivation theory), or due to other factors affecting the utility of the task.

5.2 Deadlines, Time of Work, and Correctness of Submissions

Overall, our analysis showed that submissions (across all courses and course iterations) made during the night (between 11 PM and 6 AM) were more likely to be incorrect as compared to the submissions made during the day (between 6 AM and 11 PM). Similar results have been previously observed in the context of software development, where code committed to version control systems at night was more likely to contain bugs [13]. Our analysis extends and provides further value on these prior results by observing this phenomenon in the educational context. This information could be taken into account when designing deadlines and could potentially be used in the design of automated assessment systems, for example, by disallowing or limiting submissions during the night.

Quantity-wise, only approximately 4% of the submissions (across all courses and course iterations) were done during the night, which in general indicates that the majority of the students tend to work during the day (*i.e.*, 6 AM to 11 PM); although we did find that 1509 students had at least some submissions during the night (*i.e.*, 11 PM to 6 AM) and 16 students worked exclusively during the night. There was also evidence of submissions made closer to the deadline being more likely faulty—this could partially be explained by increased time pressure related to the closeness of the deadline, which is known to decrease software quality in software development [27]. In particular, these results also held when analyzing only students who had submissions both close and further away from the deadline, and students who had submissions both during the day and during the night, indicating that the effects are not solely due to more poorly performing students working during the night and closer to the deadline.

We note that the submission behavior differs between students—some need multiple submissions to reach a correct solution, while others need only a few (or only one). Some students never reach a correct solution for all the exercises. Overall, submission behavior could be used, for example, for tailored and data-driven interventions, as students who work close to the deadline might need different support than students who work during the night. One could also prioritize such interventions by, for example, giving higher priority to students with clearer time management challenges.

We observed that students' submissions were somewhat more likely to be incorrect closer to the deadline (see Figure 3a). There are multiple possible explanations for this. Firstly, it is possible that this is due to struggling students wanting to submit at least something before the deadline and thus more likely having errors in their programs, whereas better performing students might have already submitted their exercises earlier in the week. Another explanation based on the temporal motivation theory [48] is that students whose expectancies of performance on the task are low (*e.g.*, due to having less programming experience or having struggled in previous weeks) only start work on the exercises closer to the deadline when the utility of the task increases as posited by the theory. Indeed, based on temporal motivation theory, those with higher expectancies of performance—for example, students with more prior programming experience—will start work earlier as their motivation will be higher due to the higher expectancy. Another factor that may play a part is survivorship bias: students who complete all exercises are removed from the pool of submitters. However, we found that there is a trend of submissions tending to be of lower quality closer to the deadline also for students who worked both close and further away from the deadline, suggesting that survivorship bias is not the sole reason for the observed trend.

Our observation that submissions further away from the deadline were more likely correct supports much of the prior work (*e.g.*, [9, 11, 28, 36]) that has found that students who start their work early tend to perform better. As there were differences in submission patterns with different deadlines, with some deadlines having median submissions over a day earlier than others, one concrete way instructors could nudge students to start their work earlier would be adjusting the placement (day in the week and hour of day) of course deadlines, which potentially might then affect when students start work on course exercises and help students reach better performance and in the end, better learning outcomes.

5.3 Suggestions on Deadline Policies

Overall, based on our data, if it is desired that submissions happen well before the deadline, one could consider placing the deadline towards the end of the week as we found that the average and median distance to deadline was largest for deadlines near the end of the week (Thursday and Friday). Regarding the time of day of the deadline, we found that in our context, the submissions in the courses with the midnight deadlines seemed to be more spread out over the week as evidenced by more equal peaks in Figure 1. Thus, if instructors want students to spread out their work over the week, our results suggest that a midnight deadline could help with that. We acknowledge that the issue is more complex than this, however. We have no information on deadlines from the other courses that the students were taking, and if all courses would place the deadlines at similar days and times (*e.g.*, towards the end of the week and at midnight), it would be only up to the students to prioritize their work—having courses with different deadlines can help students manage their time. We also do not have information on students' specific circumstances (*e.g.*, day jobs outside of school, coursework schedule, extracurriculars) that could have influenced students' priorities.

The optimal placement of the deadline also differs depending on what we desire to "optimize". For example, the Monday 6 AM deadline had the fewest submissions in the few hours directly preceding the deadline, likely because most students were asleep then. On the other hand, if we want to purely optimize consistent work over the week, the Monday 6 AM deadline would not work as well as the Tuesday midnight deadline, which had the most consistent submission patterns across different days of the week (see Figure 1). If, instead, we want to optimize the deadline regarding early start of work, the Friday midnight deadline had the largest average and median distance to deadline. We recommend that instructors who may want use our findings should compare it with data from their own contexts and potentially craft deadline scheduling plans collaboratively with their colleagues if and when possible.

5.3.1 Opinions of the WEBDEV instructor. In discussions about the *WEBDEV* course instructor's (*i.e.* the third author's) motives for trying out different deadlines, the driving observations were that there were always some students for whom the deadlines were not suitable. Some asked for time to finish exercises during the weekends as they were working during the weekdays, while some asked for time to finish work during the weekdays as they had jobs in the weekends. The instructor also noted that the only deadline that they would not try again was the 6 AM deadline, highlighting that they were often awake late until the night to support the few students who were working last minute and to check that everything was working, despite having office hours at 9 AM. Additionally, the instructor noted that courses may receive a stigma—although the course was well-liked among students (and highly ranked in departmental statistics), it was still referred to as "the course with the wacky deadlines" by some students in 2017. As a final note, the *WEBDEV* course instructor also remarked that "*after years of trying out these different deadlines, it seemed that no matter what the deadline was, there were always some students who complained about them*". While we agree that there likely is no perfect deadline, our findings point to course deadline setups that could inform course designs and deadline policies to better support students. Interesting future work could examine the extent to which our deadline setups in conjunction with *calls to action* [19] could help students start work earlier, manage time, or perform better on exercises.

5.4 Limitations and Threats to Validity

5.4.1 Generalizability towards students. We acknowledge that there are multiple student-related factors that can affect submission behavior that we have not taken into account. For example, students are likely working on multiple courses at the same time and we do not have information on the deadline setups or policies for those courses. Similarly, students can have different circumstances—some students may be studying full time while others may have jobs or family responsibilities (*e.g.* childcare) that could naturally affect when they can work on their studies. An additional factor that likely affects students' behavior regarding submission patterns is their prior programming experience and competence in programming in general. As we had not collected data on these student-related factors, we are unable to study their effects on submission patterns post hoc.

Our main analysis, however, compares the same course over multiple years, and we have no reason to believe that the general distribution of different student factors would significantly differ between the years. Essentially, even though, for example, having a job versus studying full time likely affects submission behavior, the proportion of students who are studying full time and who are working a job is likely similar across the years and thus any effects this has on the results should be similar across the years. Future work should examine how the circumstances of individual students affect submission behavior: the focus of our work is studying the effects of deadlines in the aggregate.

5.4.2 Generalizability towards courses. There are many aspects of the *WEBDEV* course studied here that should be taken into account when considering whether the results we report might generalize beyond our study context. Firstly, in all the courses included in this study (*WEBDEV*, *INTROCS*, *INTROSTATS*), students could submit their exercises as many times as they wished to the automatic assessment system. Had the number of submissions been limited, we might have observed different submission patterns (*e.g.*, students might have tried to "save" their submissions).

Additionally, the *WEBDEV* course where we varied the deadlines is a second year optional course. As the course is typically taken in the second year of studies, students are likely already somewhat accustomed to the environment of university studies, which could affect submission behavior. For example, many studies in computing education research focus on introductory programming courses (CS1), which are typically the very first courses that computer science students take at university. As can be observed in Table 4, comparing the full data (including introductory courses) to only the *WEBDEV* course, there are more submissions close to the deadline for the *WEBDEV* course.

It is possible that replicating this study in CS1 would yield different results because, among others, students might not have yet adjusted or become accustomed to university studies and thus may, for example, be less experienced or adept at estimating the time it takes to complete exercises or at time management (or the other way around). Additionally, introductory first-year courses are likely to have a more varied student population, for example, in terms of proficiency or competence, or in terms of student interest (*e.g.*, there are more students in introductory courses who are only taking a CS minor); by the second year, the worst-performing students are unfortunately likely to have dropped out or changed programs. The fact that the course was optional (*i.e.*, not required to graduate) can also affect the results as students may, for example, prioritize mandatory courses they might have running parallel to the *WEBDEV* course.

5.4.3 Internal validity. In the *WEBDEV* course, students typically had a little over a week to work on an exercise set from the release of the exercise set to the deadline of that set. However, the exact time window they had for each exercise set varied slightly: for example, in some weeks, the exercises might have been released ten days before the deadline, while they might have been released eleven days before the deadline on another week.

Part of the analysis combined data from *WEBDEV*, *INTROCS*, and *INTROSTATS*; but *INTROSTATS* had considerable differences in the number of passing submissions compared to the other two courses. While we believe that this should not affect submission behavior significantly, it could affect the results related to correctness. In future work, we will analyze each course separately to explore potential differences in submission behavior in relation to average correctness.

When looking at the submission data, we did not consider the difficulty of the exercises. We acknowledge that some exercises in the latter weeks of the course are more complex, which could affect the results. For example, if students work on exercises in the order they are presented in the course, the latter exercises are more likely to be worked on close to the deadline, and might also require more submissions (*e.g.*, students might make submissions for partial points) if students struggle more with them.

The choice of what is considered *day* and *night* time (see Section 4.3.1) may affect the results related to the time of day analysis. However, studying different options for this choice is out of scope of this paper, but could be interesting future work. Lastly, for *RQ2*, we used data from multiple iterations of three different courses (*INTROCS*, *INTROSTATS*, and *WEBDEV* courses), which differed in pedagogy and content. Looking at data over three different types of courses can miss some course-specific factors that can affect the results, but on the other hand makes the findings more robust.

6 CONCLUSION AND FUTURE WORK

In this work, we analyzed submission timestamp data from a learning management system to explore differences between a variety of deadlines. We summarize our findings as follows:

RQ1: *How are course exercise deadlines related to when exercises are submitted?* Our findings suggest that the placement of deadlines does correlate with when students work, and in many cases there is an observable peak in submissions close to the deadline. Out of the deadlines we analyzed, Friday midnight led to the largest average and median distance of submissions (in hours) from the deadline; the difference in median distance of submissions between the largest median and the smallest median was over 30 hours. Additionally, the Monday 6 AM deadline was the deadline where students were least likely to work very close to the deadline (see Table 4 and Figure 2).

RQ2: *How are the time of day and the distance to deadline related to the correctness of submissions?* We observed that submissions made during the night were more likely to be incorrect than those made during the day. Submissions made closer to the deadline were more likely to be incorrect compared to those made further away from the deadline. However, it is possible that the latter result is in part due to students potentially completing easier exercises first.

Our results have implications for both research and teaching. Firstly, our work provides quantitative evidence on deadline placement correlating with student submission behavior. This opens new research avenues, such as exploring why students are more likely to submit their work on the day of the deadline when the deadline was something else besides midnight, understanding the effect of student prioritization (*e.g.*, prioritization between deadlines across different courses, or student-specific contexts such as day jobs), and exploring the placement of deadlines quantitatively in general. We hope to see replications of these analyses in other contexts to increase our understanding of how contextual factors affect the results and the extent to which findings generalize to other contexts. Regarding teaching, our results shed light on the effect of deadlines on the correctness of students' work, and more broadly provide evidence and guiding points for how deadlines might affect student submission behavior, which can be used to inform course designs and deadline policies and how to best schedule deadlines in ways that support students meaningfully.

Future work could explore the effects of different deadline placements in other contexts. Additionally, it is possible that there are exercise and student-specific factors that could affect which deadlines work best, for example, whether there are potential differences between setting deadlines for larger projects compared to smaller exercises. Other avenues for research include exploring differences in submission behaviors of novice and more experienced students, whether submission or time management behaviors differ for varying difficulty of exercises or policies regarding late work, and how different deadlines might work in conjunction with strategies for combating procrastination (*e.g., calls to action* [19]). As part of our future work, we are interested in continuing this work on data-driven deadline placement by analyzing dynamic deadlines where students would have personalized deadlines, for example, by being assigned (or suggested) optimal deadlines based on their previous activity in the course (and the ethical implications of such an approach). In addition, we are considering comparing multiple submission setups such as (1) optionally submitting multiple times, (2) being required to submit multiple times (*e.g.*, as milestones/check-ins), and (3) submitting only once—these could help us tease out more specific factors that influence student submission behavior.

ACKNOWLEDGMENTS

We acknowledge the SIGCSE members mailing list discussion related to "Cinderella deadlines", which reminded us of our exploration of deadlines that had happened years before the discussion and inspired us to publish our findings.

REFERENCES

[1] Dan Ariely and Klaus Wertenbroch. 2002. Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science* 13, 3 (2002), 219–224.

[2] Tapio Auvinen, Nickolas Falkner, Arto Hellas, Petri Ihantola, Ville Karavirta, and Otto Seppälä. 2020. Relation of Individual Time Management Practices and Time Management of Teams. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[3] Susan Bergin, Ronan Reilly, and Desmond Traynor. 2005. Examining the role of self-regulated learning on introductory programming performance. In *Proceedings of the first international workshop on Computing education research*. 81–86.

[4] Alberto Bisin and Kyle Hyndman. 2020. Present-bias, procrastination and deadlines in a field experiment. *Games and economic behavior* 119 (2020), 339–357.

[5] Nicholas Burger, Gary Charness, and John Lynham. 2011. Field and online experiments on self-control. *Journal of Economic Behavior & Organization* 77, 3 (2011), 393–404.

[6] Jennifer Campbell, Andrew Petersen, and Jacqueline Smith. 2019. Self-paced Mastery Learning CS1. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 955–961.

[7] Rubén Comas-Forgas and Jaume Sureda-Negre. 2010. Academic plagiarism: Explanatory factors from students' perspective. *Journal of Academic Ethics* 8, 3 (2010), 217–232.

[8] Paul Denny, Andrew Luxton-Reilly, Michelle Craig, and Andrew Petersen. 2018. 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* (Larnaca, Cyprus) *(ITiCSE 2018)*. Association for Computing Machinery, New York, NY, USA, 4–9. https://doi.org/10.1145/3197091.3197141

[9] Paul Denny, Jacqueline Whalley, and Juho Leinonen. 2021. Promoting Early Engagement with Programming Assignments Using Scheduled Automated Feedback. In *Australasian Computing Education Conference*. 88–95.

[10] Hermann Ebbinghaus. 1885. *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.

[11] Stephen H Edwards, Jason Snyder, Manuel A Pérez-Quiñones, Anthony Allevato, Dongkwan Kim, and Betsy Tretola. 2009. Comparing effective and ineffective behaviors of student programmers. In *Proceedings of the fifth international workshop on Computing education research workshop*. 3–14.

[12] Hans-Magne Eikeland. 1975. Epsilon-squared should be preferred to Eta-squared. (1975).

[13] Jon Eyolfson, Lin Tan, and Patrick Lam. 2011. Do time of day and developer experience affect commit bugginess?. In *Proceedings of the 8th Working Conference on Mining Software Repositories*. 153–162.

[14] Joseph R Ferrari and Catherine A Roster. 2018. Delaying disposing: examining the relationship between procrastination and clutter across generations. *Current Psychology* 37, 2 (2018), 426–431.

[15] Joonas Häkkinen, Petri Ihantola, Matti Luukkainen, Antti Leinonen, and Juho Leinonen. 2021. Persistence of Time Management Behavior of Students and Its Relationship with Performance in Software Projects. In *Proceedings of the 17th ACM Conference on International Computing Education Research*. 92–100.

[16] Arto Hellas, Juho Leinonen, and Petri Ihantola. 2017. Plagiarism in take-home exams: help-seeking, collaboration, and systematic cheating. In *Proceedings of the 2017 ACM conference on innovation and technology in computer science education*. 238–243.

[17] Douglas L Hintzman. 1974. Theoretical implications of the spacing effect. (1974).

[18] Andrew J. Howell, David C. Watson, Russell A. Powell, and Karen Buro. 2006. Academic procrastination: The pattern and correlates of behavioural postponement. *Personality and Individual Differences* 40, 8 (June 2006), 1519–1530. https://doi.org/10.1016/j.paid.2005.11.023

[19] Ni Huang, Jiayin Zhang, Gordon Burtch, Xitong Li, and Peiyu Chen. 2021. Combating Procrastination on Massive Online Open Courses via Optimal Calls to Action. *Information Systems Research* 32, 2 (June 2021), 301–317. https://doi.org/10.1287/isre.2020.0974

[20] Petri Ihantola, Ilenia Fronza, Tommi Mikkonen, Miska Noponen, and Arto Hellas. 2020. Deadlines and MOOCs: How Do Students Behave in MOOCs with and without Deadlines. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[21] Kalle Ilves, Juho Leinonen, and Arto Hellas. 2018. Supporting self-regulated learning with visualizations in online learning environments. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. 257–262.

[22] Michael S Irwin and Stephen H Edwards. 2019. Can Mobile Gaming Psychology Be Used to Improve Time Management on Programming Assignments?. In *Proceedings of the ACM Conference on Global Computing Education*. 208–214.

[23] David A. Kalmbach, Logan D. Schneider, Joseph Cheung, Sarah J. Bertrand, Thiruchelvam Kariharan, Allan I. Pack, and Philip R. Gehrman. 2016. Genetic Basis of Chronotype in Humans: Insights From Three Landmark GWAS. *Sleep* 40, 2 (Dec. 2016), zsw048. https://doi.org/10.1093/sleep/zsw048

[24] Ayaan M Kazerouni, Stephen H Edwards, T Simin Hall, and Clifford A Shaffer. 2017. DevEventTracker: Tracking development events to assess incremental development and procrastination. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education*. 104–109.

[25] Ayaan M Kazerouni, Stephen H Edwards, and Clifford A Shaffer. 2017. Quantifying incremental development practices and their relationship to procrastination. In *Proceedings of the 2017 ACM Conference on International Computing Education Research*. 191–199.

[26] MG Kendall. 1975. Rank correlation methods. 2nd impression. *Charles Griffin and Company Ltd. London and High Wycombe* (1975).

[27] Miikka Kuutila, Mika Mäntylä, Umar Farooq, and Maelick Claes. 2020. Time pressure in software engineering: A systematic review. *Information and Software Technology* 121 (2020), 106257.

[28] Juho Leinonen, Francisco Enrique Vicente Castro, and Arto Hellas. 2021. Does the Early Bird Catch the Worm? Earliness of Students' Work and its Relationship with Course Outcomes. In *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1*. 373–379.

[29] Juho Leinonen, Paul Denny, and Jacqueline Whalley. 2022. A Comparison of Immediate and Scheduled Feedback in Introductory Programming Projects. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education*. 885–891.

[30] Juho Leinonen, Leo Leppänen, Petri Ihantola, and Arto Hellas. 2017. Comparison of time metrics in programming. In *Proceedings of the 2017 acm conference on international computing education research*. 200–208.

[31] Leo Leppänen, Juho Leinonen, and Arto Hellas. 2016. Pauses and spacing in learning to program. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research*. 41–50.

[32] George F Loewenstein and Dražen Prelec. 1993. Preferences for sequences of outcomes. *Psychological review* 100, 1 (1993), 91.

[33] Krista Longi. 2016. *Exploring factors that affect performance on introductory programming courses*. Master's thesis.

[34] Therese H Macan, Comila Shahani, Robert L Dipboye, and Amanda P Phillips. 1990. College students' time management: Correlations with academic performance and stress. *Journal of educational psychology* 82, 4 (1990), 760.

[35] Henry B Mann. 1945. Nonparametric tests against trend. *Econometrica: Journal of the econometric society* (1945), 245–259.

[36] Joshua Martin, Stephen H Edwards, and Clfford A Shaffer. 2015. The effects of procrastination interventions on programming project success. In *Proceedings of the eleventh annual International Conference on International Computing Education Research*. 3–11.

[37] Ranjita Misra and Michelle McKean. 2000. College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction. *American journal of Health studies* 16, 1 (2000), 41.

[38] D Parson and Allison Seidel. 2014. Mining Student Time Management Patterns in Programming Projects. In *Proceedings of FECS'14: 2014 Intl. Conf. on Frontiers in CS & CE Education*. 21–24.

[39] Shih-Yu Pu, Nan-Ching Tai, and Kah-Hoe Ng. 2019. Development of Innovative Online Shopping Experience to Overcome Procrastination through Self-Rewarding. In *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*. 1–2. https://doi.org/10.1109/ICCE-TW46550.2019.8991946

[40] Arcady A. Putilov, Dmitry S. Sveshnikov, Alexandra N. Puchkova, Vladimir B. Dorokhov, Zarina B. Bakaeva, Elena B. Yakunina, Yuri P. Starshinov, Vladimir I. Torshin, Nikolay N. Alipov, Olga V. Sergeeva, Elena A. Trutneva, Michael M. Lapkin, Zhanna N. Lopatskaya, Roman O. Budkevich, Elena V. Budkevich, Marina P. Dyakovich, Olga G. Donskaya, Juri M. Plusnin, Bérénice Delwiche, Clara Colomb, Daniel Neu, and Olivier Mairesse. 2021. Single-Item Chronotyping (SIC), a method to self-assess diurnal types by using 6 simple charts. *Personality and Individual Differences* 168 (Jan. 2021), 110353. https://doi.org/10.1016/j.paid.2020.110353

[41] Henri C Schouwenburg and JanTjeerd Groenewoud. 2001. Study motivation under social temptation; effects of trait procrastination. *Personality and Individual Differences* 30, 2 (Jan. 2001), 229–240. https://doi.org/10.1016/S0191-8869(00)00034-9

[42] Clifford A Shaffer and Ayaan M Kazerouni. 2021. The Impact of Programming Project Milestones on Procrastination, Project Outcomes, and Course Outcomes: A Quasi-Experimental Study in a Third-Year Data Structures Course. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 907–913.

[43] Jaemarie Solyst, Trisha Thakur, Madhurima Dutta, Yuya Asano, Andrew Petersen, and Joseph Jay Williams. 2021. Procrastination and Gaming in an Online Homework System of an Inverted CS1. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 789–795.

[44] William G Sommer. 1990. Procrastination and cramming: How adept students ace the system. *Journal of American College Health* 39, 1 (1990), 5–10.

[45] Jaime Spacco, Paul Denny, Brad Richards, David Babcock, David Hovemeyer, James Moscola, and Robert Duvall. 2015. Analyzing student work patterns using programming exercise data. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. 18–23.

[46] Susan Spangler. 2020. Cinderella Deadlines: Reconsidering Timelines for Student Work. https://www.facultyfocus.com/articles/course-design-ideas/cinderella-deadlines-reconsidering-timelines-for-student-work/. Accessed: 2021-04-01.

[47] Piers Steel. 2007. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin* 133, 1 (2007), 65.

[48] Piers Steel and Cornelius J König. 2006. Integrating theories of motivation. *Academy of management review* 31, 4 (2006), 889–913.

[49] Piers Steel, Frode Svartdal, Tomas Thundiyil, and Thomas Brothen. 2018. Examining Procrastination Across Multiple Goal Stages: A Longitudinal Study of Temporal Motivation Theory. *Frontiers in Psychology* 9 (2018). https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00327

[50] Dmitry Taubinsky. 2014. *Essays in Behavioral and Experimental Economics*. Ph.D. Dissertation.

[51] Vincent van der Vinne, Giulia Zerbini, Anne Siersema, Amy Pieper, Martha Merrow, Roelof A Hut, Till Roenneberg, and Thomas Kantermann. 2015. Timing of examinations affects school performance differently in early and late chronotypes. *Journal of biological rhythms* 30, 1 (2015), 53–60.

[52] Christopher Watson, Frederick WB Li, and Jamie L Godwin. 2014. No tests required: comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th ACM technical symposium on Computer science education*. 469–474.

[53] William R Watson, Ji Hyun Yu, and Sunnie L Watson. 2018. Perceived attitudinal learning in a self-paced versus fixed-schedule MOOC. *Educational Media International* 55, 2 (2018), 170–181.

[54] Albina Zavgorodniaia, Raj Shrestha, Juho Leinonen, Arto Hellas, and John Edwards. 2021. Morning or Evening? An Examination of Circadian Rhythms of CS1 Students. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*. 261–272. https://doi.org/10.1109/ICSE-SEET52601.2021.00036