Patterns in assignment submission times: Procrastination, gender, grades, and grade components

Megan Nieberding* and Andrew F. Heckler

Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, Ohio 43210, USA

(Received 12 April 2021; accepted 7 June 2021; published 30 June 2021)

In this study we characterize student procrastination habits and investigate associations between these habits and student performance on graded course components, student beliefs about their own procrastination behavior, and gender. The procrastination habits of calculus-based introductory physics students are measured via the amount of time before the assignment deadline or "completion time" that students submit their work on relatively short (>30 min) weekly online assignments. With the aid of latent profile analysis, we find that one can meaningfully categorize students into 4 completion time classes that clearly distinguish students between their mean completion time, their week-to-week completion time patterns, assignment completion rates, mean course grades, and proportion of women. Consistent with many studies in a variety of contexts, we find that procrastinating students tend to have lower course grades. Closer examination of exam and nonexam grade components reveals that completion time is directly associated with the nonexam component, but its association with the exam component is weaker and completely mediated by the nonexam component grade. This is in contrast to student ACT score, which is strongly associated with exam component but only weakly associated with the nonexam component, and the direct association of ACT with the exam scores is only weakly mediated by nonexam scores. Further, we find that ACT score is at best very weakly correlated with completion time. Taken together this supports the idea that exam and nonexam components are separately predicted by the two somewhat "orthogonal" measures of ACT score and completion time, and we propose that these are measuring so-called cognitive and noncognitive factors, respectively. Regarding gender differences, we found that on average women tended to procrastinate less than men, submitting the assignments on average 8 h earlier than men. Considering previous studies documenting that women tend to score higher than men on nonexam components, we found that completion time completely mediates the gender differences in nonexam components, providing support for the hypothesis that procrastination mediates the gender differences in performance on nonexam components. Finally, we found that the overwhelming majority (90%) of students did not strategically ("actively") intend to delay completion of the assignment, and that students who did indicate actively delaying were 2–3 times more likely to receive a D or E in the course.

DOI: 10.1103/PhysRevPhysEducRes.17.013106

I. INTRODUCTION

Procrastination may be a familiar experience in our lives, and it can be especially prevalent among some students in school. Often procrastination has been associated with poor performance and negative behaviors among students. For example, in physics classes, Kostas *et al.* found that the later students turned in their homework assignments, the lower their final grade [1]. Palazzo *et al.* found that physics students who delay starting online homework assignments until the last day are more likely to copy their online homework assignments, and Kortemeyer found that accessing the online physics homework assignments one per week (presumably just before the deadline) is negatively correlated with exam performance while daily access is positively correlated with exam performance [2,3]. In a meta-analysis of procrastination across many subjects and institutions, Kim and Seo found that the association between procrastination and academic performance depended on whether student procrastination was measured using self-reported data or behaviorally measured data [4]. Specifically, their analysis suggested that behavioral measures predicted academic performance better than self-reported data because students tended to overestimate how much they procrastinated and how well they performed in the class [4]. However, while numerous studies have found that procrastination predicts academic performance, some

*nieberding.17@osu.edu

caution should be taken to generalize this to all cases, because there are also studies that showed very little dependence, highlighting the point that the relation is likely very context dependent [5]. Nonetheless, given the mentioned studies in the context of introductory physics assignments, it seems that procrastination may be an important predictor of performance that could be used to identify students at risk in physics courses.

Procrastination is often framed as a failure of self-regulation [6], and numerous studies have mathematically modeled procrastination using *temporal motivation theory* to include the influence of a variety motivational factors [7,8]. The dependence of procrastination on motivational factors prompted us to think more deeply about the dependence of grades on procrastination. Specifically, grades are often considered as made up of two components [9]: a cognitive component (or academic knowledge component), represented by exam (and quiz) scores, and a "noncognitive" component, represented by tasks such as completion of homework assignments and attendance. A study by our group [10] as well as other recent studies [11,12] have further supported this dual-component nature of grades by showing that ACT scores were a moderate predictor of exams (i.e., cognitive components) but only a weak predictor at best of the "nonexam" (noncognitive) course components. Further, Simmons and Heckler also pointed out that even though ACT score was not correlated with nonexam components, the nonexam components were still moderately correlated with exam components [10]. Therefore, we hypothesize that since procrastination is associated with motivational (noncognitive) factors, it will be a relatively strong predictor of noncognitive course components, unlike the ACT score. Further, we predict that procrastination will likely be a relatively weak predictor of exam components, and this relation is strongly mediated by the nonexam components. In other words, we hypothesize that procrastination will predict components such as homework completion and other nonexam assignments, and this in turn predicts exam grades. This is in contrast to ACT scores, which only appreciably predict exam grades.

Procrastination may also help to explain differences in physics grades between men and women. Studies showed mixed results when comparing grades between men and women [13], especially when considering other factors such as prior physics knowledge or ACT score [10,14]. However, there was a systematic pattern when considering grade components, with women scoring higher than men on nonexam components and lower on exam components [10,11,15]. Given that women tend to have higher academic conscientiousness and overperform on the noncognitive components compared to men [12], one might naturally expect that women would procrastinate less than men, and this may lead, at least in part, to higher nonexam scores. Put another way, we hypothesize that the gender difference in performance on nonexam components is mediated by procrastination.

Finally, since we will show a clear relationship between procrastination and grades, we also administered a procrastination survey to students to get a little more insight into the phenomenon, and perhaps to help learn how to address it. To that end, we used an "active procrastination" instrument by Choi and Moran [16]. In contrast to "passive procrastinators" who delay submissions for reasons associated with failure of self-regulation and have negative performance outcomes, some students who delay submission are "active procrastinators" displaying positive attitudes and self-regulation practices and leading to desirable outcomes. For example, they argue that some procrastinators intentionally decide to procrastinate but are still able to successfully complete tasks on time because the pressure of the approaching deadline motivates them. Therefore, it may be productive to determine, among the procrastinators, which of them are active and which are not. Note that while Choi and Moran refer to this as active procrastination; there are some who cogently argue from theoretical models and empirical data that the term *active delaying* should be used instead [17–19]. Active delaying may play an important role in success for some students. For example, one study suggested that students who report high levels of actively delaying received higher grades [18]. Choi and Moran constructed an active procrastination instrument comprised of four scales [16]. We used all four scales in this study to determine if they are valid in this context, to determine if the self-reported data matched with the measured behavior (submission time from the deadline), and to gain more insight into procrastination in our context.

In sum, this study aimed to investigate the following research questions:

(RQ1) To what extent can one productively categorize and characterize student assignment completion time (i.e., procrastination) behaviors for online homework assignments, and how are these categories related to course performance?
(RQ2) What are the associations between observed student procrastination and various graded course components? Specifically, is there a difference in how student procrastination is related to cognitive versus noncognitive course components?
(RQ3) To what extent do women procrastinate compared to men? Do these differences help to explain differences in grade component performance?
(RQ4) To what extent do student homework completion time habits align with student self-reported procrastination principles?

II. METHODOLOGY

The study was conducted at Ohio State University, a large public research university, over one semester with 1374 students enrolled in the first semester of a two-semester calculus-based introductory physics course, consisting of a majority of engineering majors. The course structure included a lecture section (typically comprised of traditional lectures), a recitation section, and a lab section. The graded course components included a set of nonexam components and exam components. The nonexam course components included a lab grade (11% of the grade), the prelab assignments (3%), online homework (12%), hand-in-homework (3%), and 12 online essential skills assignments (1%). The exam components included weekly in-class quizzes in recitation (15%), 2 midterms (15% each), and a final exam (25%). Therefore, the exam components comprised of 70% of the course grade and the nonexam components comprised of 30% of the grade.

The data in this study were focused on the completion time data of the online essential skills (ES) assignments. The ES assignments were designed to help build accuracy and fluency with basic math and physics skill necessary for the course [20]. The assignments consisted of 3–5 categories each week and were mastery based such that a student must get 4 questions correct in a row to "master" the category. The questions were designed to be simple and straightforward and on average typically took between 30 sec and 2 min to answer each one. A student got full credit for mastering each category or zero credit for any given category not mastered. The full weekly assignments typically took 10 to 30 min to complete, with the assignment window opening on Tuesdays at noon and closing on Sundays at 11:59 pm. While 1374 students enrolled in the physics course, only 1218 of those students completed at least 1 ES assignment, so we report only on the 1218 students' data. Of the 1218 students, 320 (26%) students were female, 894 were male, and 4 did not disclose their gender identity. Considering the students by their URM status, 122 (10%) students identified as a URM while 1034 students did not. 62 students did not indicate their URM status. Based on first-generational college student status, 265 (22%) students were first-generational college students and 953 students were not.

A total of 14 ES units were assigned during the semester. The first and last were pre and post-tests, and the remining were mastery assignments. Here we will only analyze data from assignments 2–13. The 12 units were completed online, we collected timestamps of when students first attempted the assignments, and when students completed the assignments. For the ES assignments, we found that students tended to complete the entire assignment in one sitting, and the analysis of first-attempt times vs time yielded very similar results. Therefore, in this paper we will only investigate what we will label as the *completion time*, which we will define as the time span between the time the assignment was completed and the assignment deadline time. The smaller the completion time, the more the student had procrastinated. For this analysis, we considered a student's submission completed when they earned a 100% completion on the assignment. Any students that earned less than 100% of the points were removed from the dataset because we could not distinguish if the student was still working on the assignment up to the deadline, or if they had stopped working on the assignment, accepting a score less than 100%. We looked specifically at the completion data, not at the attempted-only data, because the procrastination surveys we asked directly addressed *finishing* assignments. Removing the attempted-only data resulted in the removal of 4% of the total number of student assignment attempts or completions.

Student timing data presented here were collected during the autumn semester of 2019. Note that, we ran a pilot study during a previous semester (autumn 2018 semester) and saw results very similar to those reported here in terms of student completion time patterns and their relation to grades.

Latent profile analysis (LPA) was used to classify the observed completion time data in this study. LPA allowed for the reduction of one or many continuous variables into a few discrete subgroups by utilizing fit statistics and probabilities to model the optimal number of completion classes [21]. Because there was variation among students in the number of ES assignments completed, we used the average completion time for each student by summing over the total number of hours from the deadline that a student completed an assignment, and then dividing by the total number of assignments each student completed. The MCLUST package in R (studio version 3.63) ran the LPA analysis and identified the optimal number of profiles [22]. We evaluated each model using the Bayesian information criterion (BIC), integrated completed likelihood (ICL), and the bootstrapped likelihood ratio test (BLRT). The best fitting model (or optimal number of completion classes) will have the highest BIC, highest ICL, and a significant BLRT p value. Based on these criteria, the LPA analysis suggested that 4 completion classes best describe our data.

To compare the observed completion time behaviors of the 4 completion classes to self-reported data on student behavior, we administered Choi and Moran's procrastination survey [16]. The procrastination survey was administered during the 8th online assignment. It contained 4 subscales that defined the characteristics of an intentional delayer: outcome satisfaction ("If I put things off until the last moment in this physics class, I'm not satisfied with their outcomes."), preference for pressure ("I feel tense and cannot concentrate when there's too much time pressure on me in this physics class."), ability to meet deadlines ("I often start things at the last minute and find it difficult to complete them on time in this physics class."), and intentional decision ('To use my time more efficiently, I deliberately postpone some tasks in this physics class") [16]. For this study the subscales we focused on were the ability to meet deadlines and intentional decision subscales. The ability to meet deadlines subscale (henceforth called

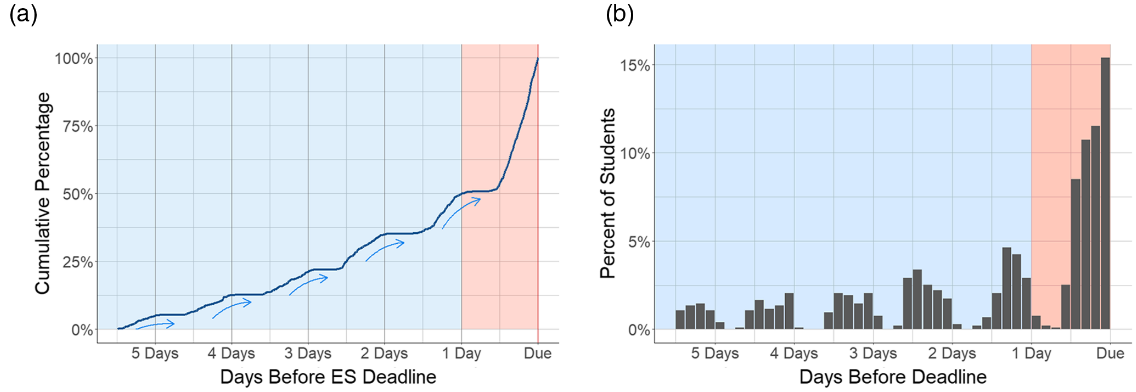

FIG. 1. (a) Cumulative percent of calculus-based physics students completing essential skills assignment #3 before the deadline at 11:59 pm. The blue arrows show the diurnal cycle of student activity. The light blue region highlights the students that completed the essential skills assignment earlier than one day before the deadline. The light red region highlights the students that completed the essential skills assignment within the final 24 h before the deadline. (b) Percent of calculus students completing ES assignment #3 before the deadline binned in 3 h increments. These graphs represent $N = 1034$ students that completed ES assignment #3.

the deadline subscale) measured how well a student believed they could estimate the time needed to complete a task before a given deadline. The intentional decision subscale measured how much a student believed they deliberately postponed completing tasks for the physics course. The 4 subscales were given on a 7-point response scale. For our dataset, each subscale had a level of reliability (Cronbach's α) ranging between 0.84 to 0.92, indicating the subscales were internally consistent.

III. RESULTS

A. Characterizing and classifying completion time

To provide insight into how long before the deadline students completed an ES assignment, we plotted the cumulative percent of students who completed one particular ES assignment before the deadline, and a "completion rate" graph showing the number of students completing the assignment binned in 3 h windows (Fig. 1). The first thing to notice was that about 50% of the students completing this assignment waited until the final 24 hours to complete the assignment. Additionally, these plots had three other interesting features. First, there appeared to be a turning point ("knee") at about 12 h before the deadline. The turning point marks the end of the shallow sloped feature of the cumulative percent curve that begins when the assignment is first opened (5 days before the deadline). The turning point also marks the beginning of the steeply sloped portion of the cumulative percent graph that continues until the assignment is due. The turning point in this graph suggests that there could be at least 2 different groups of students: those that completed the ES assignments "early" and "late." Finally, we point out the periodic curves in the shallow sloped part of the graph, marked by the light blue arrows, and is perhaps more evident in the completion rate graph. These periodic curves are representative of the diurnal pattern: most students completed the assignments between noon and midnight. While the data presented in Fig. 1 were only for ES unit 3, these four features were repeated in all 12 essential skills units throughout the course of the semester.

While looking at individual assignments gives a cross-sectional snapshot of student behavior, we gained a more global and longitudinal sense of when students were completing the assignments by looking at a student's completion times for all completed assignments. For example, this could help determine if students were consistently completing an assignment early or late throughout the semester. One might expect that students created study habits for completing online assignments.

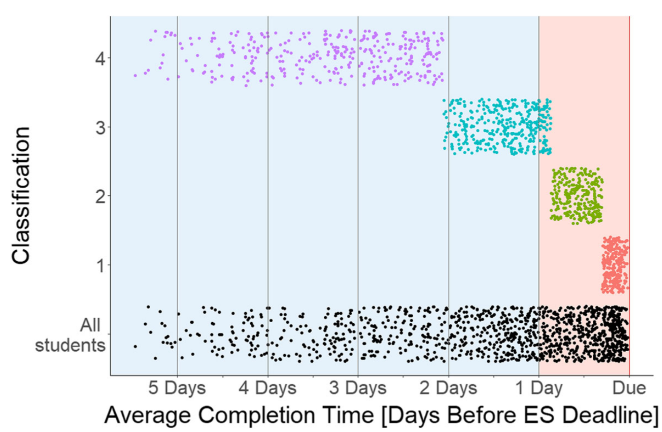

FIG. 2. Mean completion times across all ES assignments (jittered), split by the latent profile analysis classification groups 1, 2, 3, and 4 in red, green, blue, and purple, respectively. Each dot represents the mean completion time for one student. The black dataset represents the entire sample of students. The light blue region highlights the students that on average completed the essential skills assignments earlier than one day before the deadline. The light red region highlights the students that on average completed the essential skills assignment within the final 24 h before the deadline, which is at 11:59 pm.

To begin, we first considered the average completion time of each student. To more formally determine the extent to which students could be placed into categories, we turned to the LPA analysis of the student's average completion time. The best-fitting model grouped students into 4 "completion classes" based on their average completion time (Fig. 2). From this figure, we noticed that two groups (class 1 and 2) on average completed the ES units within the last 24 hours before the deadline. We labeled these two completion classes as the *deadline-driven* students. In contrast, completion classes 3 and 4 on average completed the ES units more than 21 hours before the deadline. We labeled these two classes as the *early-completing* students.

The students we labeled as early-completing either completed their online assignment in the middle of the week on average (i.e., class 3; 1 to 2 days before the deadline), or they completed their assignment immediately after the assignment had opened (i.e., class 4; 3 or more days before the deadline). When looking at the 2 groups of deadline driven students, the LPA analysis distinguishes between students that completed their assignment on average within the last 7 h and students that completed the assignment between 21 and 7 h before the deadline.

While the classification of students into the four classes provided ranges of the average completion times, it did not specify whether students *consistently* completed assignments within a given range throughout the semester. To get a better sense of the within-student consistency of completion times over the course of a semester we considered individual completion times for each assignment. For students in a given completion class, we pooled all student completion times for all assignments and binned them into the four class completion time ranges (Fig. 3). This way we could determine whether students in, say, class 2 always completed the assignments in the class 2 range, or if in fact sometimes they completed an assignment in the class 1 range and sometimes in the class 3 range, keeping in mind that the average time must fall within the class 2 range.

Figure 3 indicates that the within-student completion time consistency depended on the class. Students in classes 1 and 4 were the most consistent: They overwhelmingly completed assignments within their respective completion time windows (80% and 78%, respectively), with the adjacent class being the next-most populated. In fact, class 1 students never completed an assignment in the class 4 time range.

Students in classes 2 and 3 were much less consistent and were not even most likely to complete their assignments

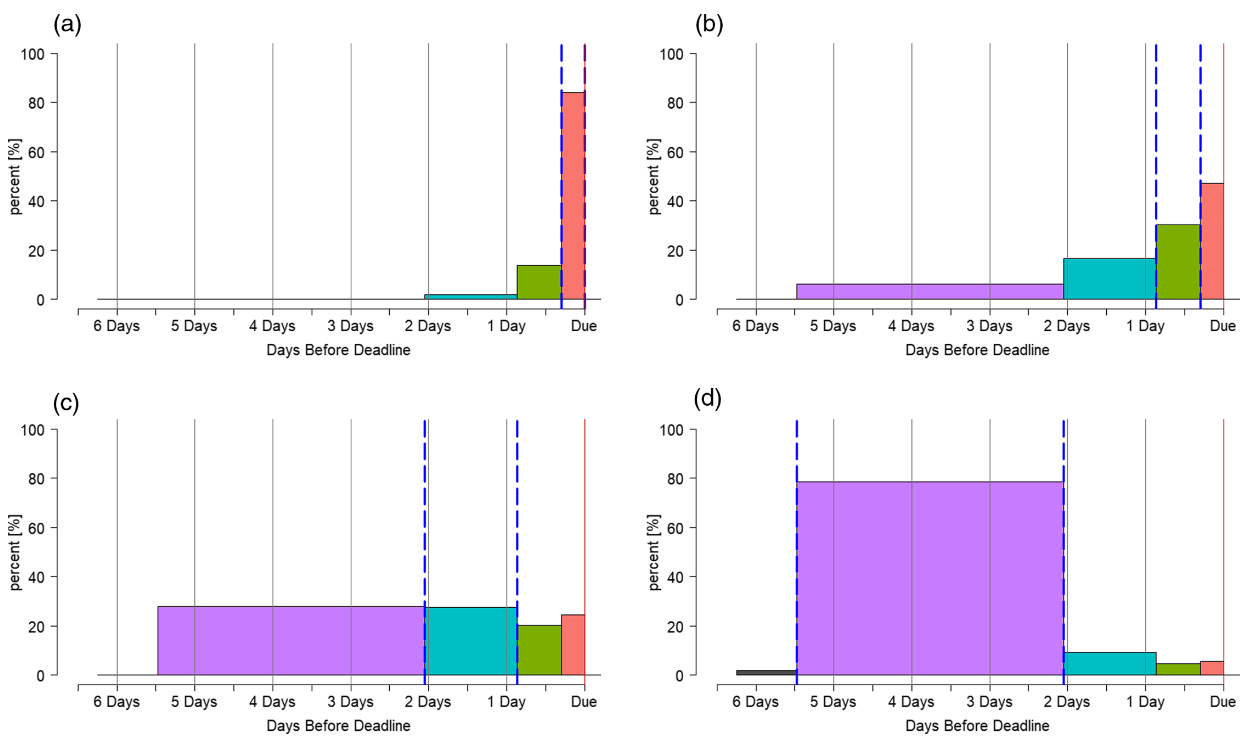

FIG. 3. For each class, we plotted a histograms of student completion times for all ES assignment completed. Figure 3(a) is for completion class 1 students, (b) class 2 students, (c) class 3 students, and (d) class 4 students. The histogram bin widths are based on ranges of completion times for each completion class. The bins were from 0 to 7.1 h, 7.1 to 20.7 h, 20.7 to 49.2 h, and 49.2 to 131.2 h, representing the range of average times of students in class 1,2,3 and 4, respectively. The blue dashed lines are included to emphasize the window that corresponds to the completion class that is being plotted. For example, in Fig. 3(a), 80% of class 1 completion times are between 0 and 7 hours (i.e., between the two blue dashed lines).

TABLE I. Descriptive statistics about the 4 completion classes, such as the number of students in each classification, the average completion time of each classification, the average course GPA for each classification, etc. Bolded numbers indicate significant at $p < 0.001$ level after running an ANOVA between classes.

	Class 1	Class 2	Class 3	Class 4
Number of students	254	268	357	339
Mean completion time [hours before deadline]	**3.92**	**13.6**	**33.0**	**80.0**
Mean course GPA	**2.03**	**2.57**	**2.69**	**2.97**
Mean ACT score	30.0	30.5	30.3	30.7
Mean assignments completed	**7.37**	**9.58**	**10.2**	**10.8**

within their completion class time windows. For example, only 30% of class 2 completion times were within their completion time window, while 47% of class 2 completion times were within the class 1 completion time window. This analysis makes it clearer to see that class 2 students procrastinate differently than class 1 students. For example, 21% of the time class 2 students completed the assignment two or more days before the deadline, while only 1% of class 1 students completed assignments more than 24 h before the deadline.

Interestingly, class 3 students were roughly equally likely to complete assignments in any class time range throughout the semester. Most of the time, class 3 students completed the assignments more than 24 h before the deadline but not necessarily 2 to 3 days before the deadline, unlike the class 3 average completion time of 33 h (Table I) might have suggested. Therefore, the average completion time maybe be a useful descriptor of the completion time behavior of students in classes 1 and 4, but less so in class 2, and could be considered misleading for students in class 3.

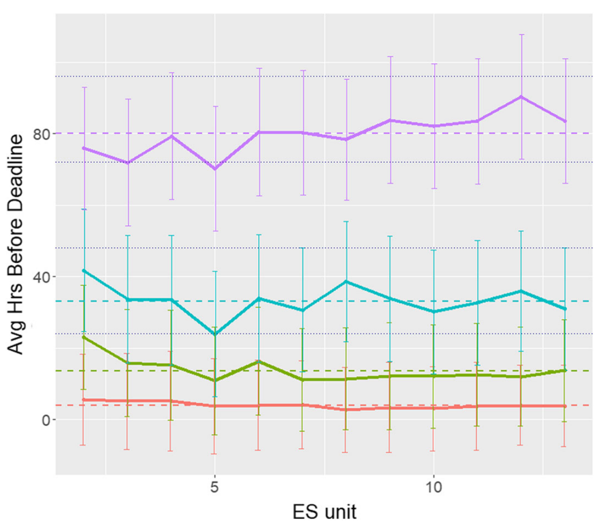

FIG. 4. For each completion class, we plotted the mean completion time for each essential skills unit (measured in units of hours before the deadline). Error bars represent 1 standard error. The dashed colored line represents the mean completion time for each class averaged over all essential skills units.

One might expect an evolution of the completion times throughout the semester, but we did not find any significant results suggesting any time evolution (Fig. 4). We noticed there was a slightly earlier average completion time in the first unit for class 2 and 3, but these times were not significant as compared to the semester's average.

B. Associations between completion time classes, grades, and demographic factors

Table I also reveals significant differences in the course grade of students between each of the completion classes $[F(3, 1214) = 34.79, \ p < 0.001]$. For example, class 1 students had the lowest average course grades at 2.03, and class 4 students were on average one grade point higher than class 1 students. A more in-depth breakdown of the course grades is presented in Fig. 5. It is clear to see from Fig. 5 that the early completers in class 4 dominated the higher grades, with 42% of students earning an A (4.0) belonging to class 4. In contrast, the procrastinating class 1 students were overrepresented in the lower grades, with over 40% of students receiving a D, E (failing) grade, withdrawing, or receiving an incomplete. Furthermore, while both class 1 and class 2 students waited until the final day to complete the assignment, class 1 students earned significantly lower course grades than class 2 students $[F(3, 1214) = 34.79, \ p < 0.001]$. Note that this is one key distinction between class 1 and class 2 students, which we both labeled as deadline-driven students.

As shown in Table I and also represented in more detail in Fig. 6, there were significant differences in the number of assignments that students completed in each of the completion times $[F(3, 1214) = 70.25, \ p < 0.001]$. For example, 50% of the students that completed fewer than half of the online assignments belong to class 1. In contrast, only 7% of the students who completed all 12 of the online assignments belong to class 1, and 42% belong to class 4. Assignment completion reveals another noteworthy distinction between class 1 students and class 2 students. Class 1 students complete significantly fewer online ES

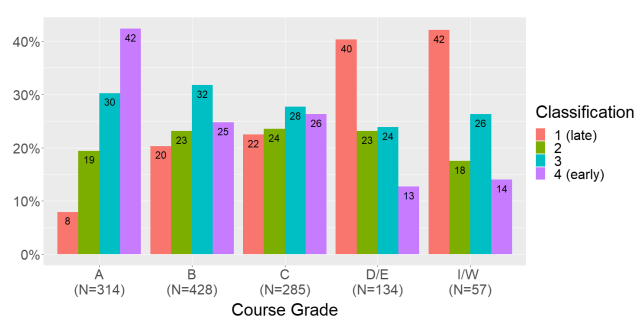

FIG. 5. Proportion of students that received the indicated course grade for each completion classification.

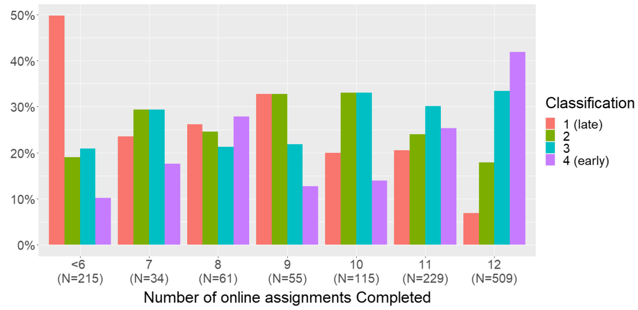

FIG. 6. Proportion of students in each classification group by the number of online assignments the student completed.

assignments than class 2 student [7.37, 9.58, $t(484) = -7.3$, $p < 0.001$, effect size 0.64 (Cohen's d)].

Table I also helped to answer RQ3, showing that the proportion of women differed significantly by completion time class [$X^2(3, N = 1214) = 20.2$, $p < 0.001$)]. More specifically, women tended to procrastinate less than men as shown by the fact that women were overrepresented in classes 3 and 4 and subsequently underrepresented in classes 1 and 2. In fact, women were 1.2 times as likely to be in classes 3 or 4 than males [risk ratio 1.24, 95 C.I. (1.12, 1.37)]. Further analysis reveals that the completion time for women is significantly larger (mean = 40.2 h), than for men (mean = 32.2 h) (Mann-Whitney test, $p < 0.001$).

Another salient way to represent the procrastination differences between men and women is shown in Fig. 7. Males were evenly distributed among the four classes [$X^2(3, N = 894) = 5.3$, $p = 0.151$)]. while women were not [$X^2(3, N = 320) = 40.55$, $p < 0.001$)]. Put another way, while men were roughly equally likely to be in a given completion time class, women were twice as likely to be in classes 3 or 4 than in classes 1 or 2. For the whole population, there were about 33% more students in class 3 and 4 than in classes 1 and 2, so this highlights the fact that women dramatically overpopulated classes 3 and 4.

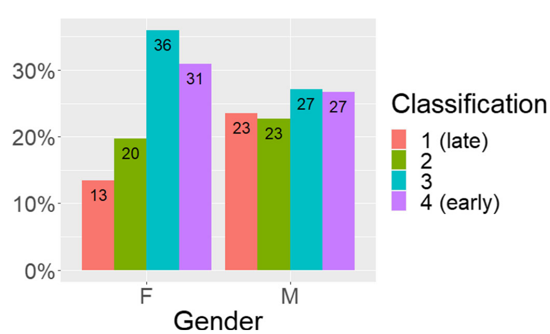

FIG. 7. The percent of females (F) and males (M) in each completion time classification. For example, 13% of females are in class 1.

C. Completion time, ACT score, and grade components

To continue our investigation beyond discrete completion time categories, we considered completion time as a continuous variable. More specifically, we considered the logarithm of each student's mean completion time because the completion time distribution was asymmetric with a long tail, and we found the logarithm of completion time to be a slightly better fit in the linear regressions.

To begin with the big picture, Table II presents several important findings regarding the correlations between (log) completion time, grade components, and ACT scores. First, we replicated past findings that ACT score was moderately strongly correlated with exam components (0.47), but only weakly correlated with nonexam components (0.15) [10]. Second, perhaps as to be expected from the completion class vs course grade results of the previous section, we found a weak-to-moderate correlation of +0.31 between course grade and the completion time, indicating the earlier a student finished the assignments the higher course grade. Third, Table II supports our hypothesis that, in contrast to ACT score, the completion time was moderately correlated with nonexam components (0.38), but only weakly correlated with exam components (0.22), and only very weakly

TABLE II. Correlations between grades, grade component scores, ACT score, and the log of the student's mean completion time. Note: All values are significant at 0.001 level, except for one cell marked **=significant at the 0.01 level.

	Course GPA	MC ACT score	Exam completion	Nonexam completion
ACT score	0.46			
Exam component	0.92	0.47		
Nonexam component	0.66	0.15	0.53	
Log (Completion Time)	0.31	0.08**	0.22	0.38

TABLE III. Results of model 1 predicting the nonexam component of a student's grade, ranging from 0–30 (percent), with the student's (mean-centered) ACT score and the log of their average completion time as the independent variables. The estimates of the coefficients are presented with 95% confidence intervals. The ΔR^2 column corresponds to the increase in R^2 as the model was forward stepped.

Variables	Estimate [95% C.I.]	Standard coefficients	ΔR^2
Intercept	27.11 [26.89, 27.33]	0	⋯
Log (Compl. Time)	1.07 [0.89, 1.26]	0.36	0.143
MC ACT score	0.13 [0.07, 0.20]	0.13	0.006

TABLE IV. Results of model 2 predicting the exam component of a student's grade, ranging from 0-70 (percent), with the student's (mean-centered) ACT score and the log of their average completion time as the independent variables. The estimates of the coefficients are presented with 95% confidence intervals. The ΔR^2 column corresponds to the increase in R^2 as the model was forward stepped.

Variables	Estimate [95% C.I.]	Standard coefficients	ΔR^2
Intercept	48.29 [47.66, 48.92]	0	⋯
MC ACT score	1.52 [1.33, 1.71]	0.46	0.22
Log (Compl. Time)	1.67 [1.14, 2.19]	0.18	0.033

TABLE V. Results of a model predicting the course grade, ranging from 0–4, with the student's (mean-centered) ACT score and the log of their average completion time as the independent variables. The estimates of the coefficients are presented with 95% confidence intervals. The ΔR^2 column corresponds to the increase in R^2 as the model was forward stepped.

Variables	Estimate [95% C.I.]	Standard coefficients	ΔR^2
Intercept	2.58 [2.52, 2.64]	0	
MC ACT score	0.15 [0.13, 0.17]	0.44	0.22
Log (Compl. Time)	0.25 [0.20, 0.30]	0.26	0.06

correlated with ACT score. This suggested that while ACT score and completion time both predicted grade, they were somewhat orthogonal to each other.

Linear regression models predicting the grade components provided a more precise picture of the predictiveness of ACT scores and completion time. The linear regression results of Table III indicated that when predicting nonexam component score, the completion time explained 14% of the variance while the student's ACT score explained a negligible amount, less than 1%, of the variance. Further, Table IV shows that for predicting the exam component score, ACT score explained 22% of the variance in the student's exam components while only 3% of the variance was explained by the average completion time. These results further indicated that ACT score was a good predictor of exam components, but had virtually no predictive power for the nonexam components, while completion time was a good predictor of the nonexam components, but only a very weak predictor of exam components.

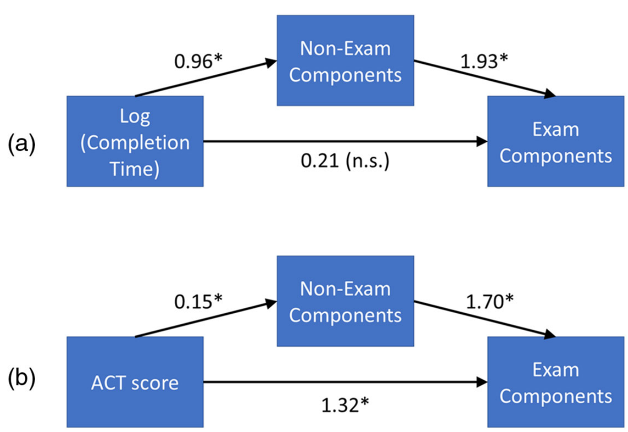

FIG. 8. Mediation model path diagrams depicting the total, direct, and indirect effects of (a) log completion time on exam components and (b) ACT score on exam components with the indirect effect operating through a single mediator, nonexam components. [$*p < 0.001$. (n.s) = not significant.]

To test our hypothesis that nonexam components strongly mediated the relation between completion time and exam components, we ran the mediation model as show in Fig. 8(a) using the process macro in SPSS [23]. The results confirmed our hypothesis. In fact, the nonexam component may completely mediate the effect of completion time on exam component, with a direct effect consistent with zero (0.21, LLCI: -0.28, ULCI: 0.71), and a significant indirect effect (1.85, LLCI: 1.52, ULCI: 2.23). For example, a student completing 1 h earlier than the average completion time was estimated to have a gain of 0.21 (out of 70 points) on their exam component (or 0.3% increase in their exam component score). Additionally, the indirect effect of the log completion time on the exam components through the nonexam components had an increase in 1.85 (out of 70 points) exam components (or a 2.6% increase in their exam component score). For comparison, Fig. 8(b) shows that the nonexam component only weakly mediates the effect of ACT on the exam component, but as expected the direct effect (1.32, LLCI: 1.15, ULCI: 1.48) is much larger than the indirect effect (0.25, LLCI: 0.12, ULCI: 0.37) CI.

Finally, consider how ACT score and completion time together predict the final grade. This is worthwhile because it is well established that ACT score predicts, indeed was designed to predict, course grade, but since completion time is probing a different component of grade, it might add value to efforts when attempting to identify at-risk students. Table V show the regression results for final grade as the outcome variable and ACT score and log completion time are the predictor variables.

Perhaps as expected, since exam components dominated the grade weight, the ACT score accounted for the largest amount of variance (22%), while the completion time only added about 6% to the variance when included in the model.

D. Completion time as a mediator of gender differences

As mentioned earlier, previous research has shown that women tend to score higher than men on nonexam components. We also established that women submitted

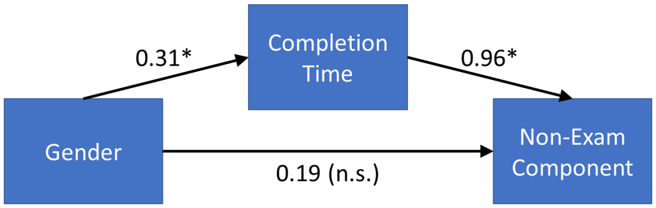

FIG. 9. Mediation model path diagrams depicting the total, direct, and indirect effects of Gender on nonexam components with the indirect effect operating through a single mediator, completion time. [$*p < 0.001$. (n.s) = not significant.]

assignments earlier on average, and that completion time was related to nonexam component scores. Given these results and the links between procrastination, gender and noncognitive factors describe in the introduction, we hypothesize that completion time at least partially mediates the gender difference in nonexam scores.

To test this hypothesis, we ran the mediation model as show in Fig. 9 using the process macro in SPSS [23]. The results confirmed our hypothesis. In fact, completion time may completely mediate the effect of gender on the nonexam component, with a direct effect consistent with zero (0.19, LLCI: −0.16, ULCI: 0.58), and a significant indirect effect (0.30, LLCI: 0.16, ULCI: 0.46). In other words, women perform better on nonexam components, and this is mediated by the fact that they complete ES assignments earlier. This effect is significant even though the effect is small, that is women only score 1.7% better than men on the nonexam components [$t(1157) = -2.45$, $p = 0.01$, effect size −0.163 (Cohen's d)].

E. Procrastination and intentional delay

To address RQ4, we examined student self-reporting of their beliefs and strategies related to the timing of their assignment submissions. While we have shown that delaying completion of the assignment was associated with lower grades and a failure to complete assignments, to what extent did students *believe* they could purposefully delay working on their online homework but still complete it on time? In other words, can we connect student beliefs about intentionally delaying online coursework to their observed behavior? Given that class 1 students on average tended to perform poorly in the course, completed fewer essential skills assignments, and delayed submission to the final few hours, one might have reasonably inferred that these students were passive procrastinators, failing to self-regulate and had mismatches between their beliefs and behavior. This led us to hypothesize that class 1 students believed they could intentionally delay completing assignments in the last hours of the last day but nonetheless failed to successfully complete the essential skills assignments on time. On the other hand, we hypothesized that class 2 students were more in alignment and may be considered more like productive active delayers: they believed they could intentionally delay completing assignments on the last day (though usually not in the last few hours), and they successfully completed the assignments on time. Finally, we hypothesized that class 3 and 4 students were also aligned: they believed they did not intentionally delay working on the ES assignments, and they tended to complete the assignments relatively early.

To test these hypotheses, we compared the scores on the four procrastination subscales defined by Choi and Moran [16] between the four completion classes (Table VI). We drew three conclusions from our analysis. First, two of the subscales, preference for pressure and outcome satisfaction, did not appear to be informative in this context. Students in all completion classes responded similarly on each of these two subscales, and we found the correlations of preference for pressure and outcome satisfaction with mean completion time to be weak (0.13 and 0.07, respectively). These results indicated that at least in this context, no matter what their completion time, all students tended to answer neutrally (score ∼4) in their preference for the challenge (pressure) that a deadline creates, thus this was not a differentiating motivating factor for delaying. Further, all students, regardless of completion time felt the same relatively small lack of outcome satisfaction (score ∼3) in completing the assignments.

Second, scores on the intentional decision scale clearly distinguished between the more deadline-driven student in classes 1 and 2 versus the early-completing students in classes 3 and 4. A Kruskal-Wallis and Dunn test *post-hoc* analysis revealed that students in class 1 and 2 were statistically similar to one another [$H(3) = 66.2$, $p < 0.001$; Dunn test $p = 0.16$], and different from classes 3 and 4 [Dunn Test $p < 0.001$]. Additionally, we found that intentional decision was weakly to moderately negatively correlated (−0.28) with mean completion time. It is

TABLE VI. Mean scores of the procrastination subscales for each of the LPA completion classes. Standard deviations in parenthesis. Level of reliability for each subscale in final column. Scales range from 1 to 7 with 4 being neutral.

	Classa 1	Class 2	Class 3	Class 4	Cronbach α
Satisfaction subscale	2.8 (1.37)	3.0 (1.36)	3.0 (1.27)	3.1 (1.34)	0.84
Pressure subscale	3.5 (1.73)	3.5 (1.59)	3.7 (1.54)	4.1 (1.52)	0.92
Deadline subscale	4.0 (1.42)	4.5 (1.46)	4.8 (1.18)	5.5 (1.35)	0.89
Intentional decision	3.8 (1.42)	3.6 (1.46)	3.2 (1.18)	2.8 (1.35)	0.87

interesting to note that the mean intentional decision scores for all completion classes were below 4.0, the neutral score, indicating that the majority of students in the course do *not* intentionally delay completing assignments. Nonetheless, the higher mean intentional decision scores of the class 1 and 2 students indicated that at least some of the students reported purposefully and strategically delaying work on assignments. In contrast, classes 3 and 4 had significantly lower scores indicating that the overwhelming majority of these students did not purposefully delay working on assignments. Thus, the relative scores of the completion classes indicated some overall consistency with the observed relative student completion time behavior and their reported intent to delay.

Third, the ability to meet deadlines scale was able to distinguish between all four completion classes [$H(3) = 97.9$, $p < 0.001$; Dunn test for all class comparisons $p \leq 0.006$]. Specifically, class 1 students, who completed the fewest number of assignments, tended to answer neutrally when asked about their ability to meet deadlines, where all other classes tended to affirm their ability to meet deadlines. This provides some indication that class 1 students were aware of their inability to meet deadlines, yet they still reported the highest level of intention to delay. In contrast, students that reported they were able to meet deadlines tended to have higher mean completion times, as evidenced by the fact that the ability to meet deadlines was moderately correlated (0.33) with mean completion time. These results suggest that instructors might help class 1 students address their procrastination issues by explicitly mentioning the student's inconsistent completion habits and incentivizing early completion times with extra credit, like what was done at University of Central Florida [24].

Finally, we produced Table VII to gain a better understanding of the numbers and percentages of students in each completion class who reported an intention to delay their submissions in addition to the extent to which they were successful in the course. Here we operationally defined intent to delay as having a mean intentional delay score of 5 or greater (4 is the "neutral" score), and we also conservatively defined success in the course as obtaining a grade of C- (1.7 grade points) or higher. There is one caveat to this table—it represents only 77% of students in this study, namely, those that completed the intentional decision survey.

TABLE VII. Numbers and percentages of students according to intent to delay and grade listed by completion class. Intent to delay is defined here as having a mean intentional decision score of 5 or greater (on scale of 7), and a passing grade is defined here as C- (1.7 GP) or higher. Percentages are expressed in terms of completion class (i.e., by row).

	No intent to delay, Passing grade	No intent to delay, Nonpassing grade	Intent to delay, Passing grade	Intent to delay, Nonpassing grade
Class 1	105 (69%)	19 (12%)	18 (12%)	11 (7%)
Class 2	169 (79%)	18 (8%)	21 (10%)	6 (3%)
Class 3	250 (87%)	23 (8%)	9 (3%)	5 (2%)
Class 4	267 (93%)	11 (4%)	9 (3%)	1 (0.3%)

Perhaps the most notable trend in Table VII is that 90% of students do not (strategically) intend to delay the completion of assignments. Perhaps as to be expected, most (70%) of the students who did report an intent to delay were in completion classes 1 and 2. One might reasonably label these students as active delayers and they comprise of 19% and 13% of class 1 and class 2 students, respectively. In contrast, a very small number (<5%) of class 3 and 4 students reported an intent to delay and can be labeled as active delayers. Another important result from Table VII is that students who did not intend on delaying were more successful: 92% of students who did not intend on delaying received C- or better in the course compared to only 71% of students who did intend on delaying [$\chi^2(2) = 34.3$, $p < 0.001$]. Put another way, students who reported intentionally delaying were 3.5 times more likely to receive a D or E in the course. What is interesting is that this trend still holds true for class 1 students: Even among students who completed the assignment within the last 7 hours, those who reported intentionally delaying were 2.5 times more likely to receive a D or E compared to those who did not intentionally delay.

IV. DISCUSSION AND CONCLUSIONS

Let us summarize the findings of this study by addressing our research questions. For RQ1, we have found that we can meaningfully categorize students into 4 completion time classes that not only clearly distinguish between their mean completion time but also their week-to-week completion time patterns, assignment completion rates, mean course grades, and proportion of women. Overall, the higher the number of the completion class the higher the assignment completion and average course grade. These results are consistent with prior studies documenting that in many contexts students who procrastinate less tend to have higher course performance. We labeled class 1 and class 2 students as deadline-driven students because both classes completed their assignments on average in the last 21 h. Some main distinctions between these two classes included: Class 1 students were more than twice as likely to complete only 50% or less of the assignments, and class 1 students earned 40% of the D/E in our population compared to 23% in class 2. We labeled class 3 and 4 students as early completing students because they typically completed assignments at least one day before the assignment was due. Class 4 students earned 42% of the A's in our population, and 42% of the students that completed all 12 of the assignments are class 4 students, which as a comparison represent 28% of all students in the course. Another interesting distinction between the completion classes was that classes 1 and 4 tended to be on the extremes: 80% of the time students completed the assignment very late, hours before the window closed (class 1), or very early, sometimes hours after the window opened (class 4). In contrast, students in classes 2 and 3 had more distributed completion times. For example, class 3 students completed the assignments equally frequently throughout the week.

Our findings regarding the relationships between completion time, grades, and ACT scores (RQ2) were also very interesting and potentially very important. For example, completion time was found to be moderately correlated with grade, but it was only very weakly correlated with ACT score. Digging deeper, we found that while completion time is moderately correlated with nonexam components of the grade, the association of completion time with the exam component score is completely moderated by the nonexam component score. Further, ACT scores are moderately correlated with exam components and this association is only weakly moderated by the nonexam component scores. Our current interpretation of these results is that there are two factors in grades, the noncognitive and cognitive factors that can be at least partially measured by the orthogonal factors of completion time and ACT scores. Specifically, completion time may be (only) measuring noncognitive attributes of the student that are associated with homework, lab, and participation scores, whereas ACT scores may be (only) measuring the cognitive attributes that are associated with exam and quiz scores. This latter point was also found by Simmons and Heckler [10], and here we propose that completion time is measuring a separate and orthogonal dimension of grades associated with noncognitive components.

Our investigation into gender differences (RQ3) revealed that women on average tend to procrastinate less than men, which may be expected given that prior research documents higher conscientiousness in women on average and they tend to score higher on non-cognitive components. In concrete terms, women submitted the assignments on average 8 h earlier than men, and women were twice as likely to be in class 3 or 4 than in classes 1 or 2, whereas mean were equally likely to be in any completion class. Investigating further, we have uncovered at least a potential partial explanation for the replicated finding that women tend to score higher on nonexam components than exam components. Namely, we have provided evidence for our hypothesis and found that the gender difference in nonexam components is completely mediated by completion time. Of course, this statistical mediation does not establish causality, but this finding does consistently tie together the general ideas that nonexam components are associated with noncognitive components, and that women tend to score higher than men on nonexam components and noncognitive factors such as conscientiousness. Clearly more research is needed to better understand the relationships between non-cognitive factors such as conscientiousness and procrastination and how these may be related to gender. Further, more research is needed to establish potential causal mechanisms between these factors and better performance on noncognitive exam components. Nonetheless, our findings do suggest that the root of the gender differences in nonexam components may arise at least in part from gender differences in non-cognitive factors, and this warrants further investigation.

Our findings on the relationship between measured completion times and student-reported beliefs about factors ostensibly related to procrastination behavior present a complicated picture (RQ4). On one hand, we find that students who procrastinate more are more likely to report an intentional (strategic) decision to delay completion of assignments. But on the other hand, students who procrastinate more also tend report they are less able to meet deadlines and tend to have lower grades. While the idea of intentional or "active" delay is an active area of research, these results call into question how many students are reporting strategic delaying of work on course assignments and are also successful in their courses. We find only a small fraction, about 6%, of students doing both. Overall, we found that only about 10% of students report that they are strategically delaying completion of the ES assignment, but we also found that these students are at an especially high risk of failing, even among students who complete the assignments late.

There are a few limitations to keep in mind when interpreting this work. First, the completion time studies here were for brief (15–30 min) online "essential skills" assignments that are designed to be a practice session with a low level of difficulty rather than a challenging one that is typically found in other important assignments like more traditional homework. The relatively low level of difficultly and short duration of the assignments impacts our ability to generalize to other assignments and our interpretation of what it means to procrastinate. Future work should investigate the submission behavior of more traditional, more difficult, and longer duration homework assignments, such as found on commonly used commercial online homework applications, to see if similar results are found.

In addition, our population sample is skewed because we collected data from only 89% of the students enrolled in the course: only 1218 of 1374 total enrolled students completed at least 1 essential skills assignment. It is likely that the missing students did not do well in the course, since the average GPA for all 1374 students is slightly lower than the 1218 students in our sample (2.55 vs 2.60). Thus, our analysis is likely disproportionately missing low performing students but does not appear to be a large effect. Finally, our population sample is not representative of the broad range of students enrolled in physics nationally, limiting our ability to generalize the results beyond this context. For example, the students were enrolled in the calculus based physics course, and about 75% white males.

While we did observe an expected negative correlation between procrastination and performance in this context, of course we cannot infer from the results of this study that procrastination causes poor performance. However, the findings here, for example that procrastination connects (only) with nonexam grade components that tend to be correlated with other motivational factors, do raise some important questions worthy of further investigation and may indicate a greater importance of assignment completion time than originally assumed.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under DUE IUSE Grant No. 1914709.

[1] P. M. Kotas, J. Finck, and M. Horoi, Homework habits of college physics students, J. College Teach. Learn. **1**, 53 (2004).

[2] D. J. Palazzo, Y. Lee, R. Warnakulasooriya, and D. Pritchard, Patterns, correlates, and reduction of homework copying, Phys. Rev. ST Phys. Educ. Res. **6**, 010104 (2010).

[3] 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).

[4] K. R. Kim and E. Seo, The relationship between procrastination and academic performance: A meta-analysis, Personality Indiv. Diff. **82**, 26 (2015).

[5] S. Cormack, L. Eagle, and M. Davies, A large-scale test of the relationship between procrastination and performance using learning analytics, Assess. Eval. Higher Educ. **45**, 1046 (2020).

[6] P. Steel, The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure, Psychol. Bull. **133**, 65 (2007).

[7] P. Steel, F. Svartdal, T. Thundiyil, and T. Brothen, Examining procrastination across multiple goal stages: a longitudinal study of temporal motivation theory, Front. Psychol. **9**, 327 (2018).

[8] S. Zhang and T. Feng, Modeling procrastination: Asymmetric decisions to act between the present and the future, J. Exp. Psychol. Gen. **149**, 311 (2020).

[9] S. Brookhart, T. Guskey, A. Bowers, J. McMillan, J. Smith, L. Smith, M. Stevens, and M. Welsh, A century of grading research: meaning and value in the most common educational measure, Rev. Educ. Res. **86**, 803 (2016).

[10] A. Simmons and A. Hecker, Grades, grade component weighting, and demographic disparities in introductory physics, Phys. Rev. Phys. Educ. Res. **16**, 020125 (2020).

[11] S. Salehi *et al.*, Gender performance gaps across different assessment methods and the underlying mechanisms: the case of incoming preparation and test anxiety, Front. Educ. **4**, 107 (2019).

[12] H. Keiser, P. Sackett, N. Kuncel, and T. Brothen, Why women perform better in college than admission scores would predict: exploring the roles of conscientiousness and course-taking patterns, J. Appl. Psych. **101**, 569 (2016).

[13] R. L. Matz, B. P. Koester, S. Fiorini, G. Grom, L. Shepard, C. G. Stangor, and T. A. McKay, Patterns of gendered performance differences in large introductory courses at five research universities, AERA Open **3**, 2332858417743754 (2017).

[14] S. Salehi, E. Burkholder, G. P. Lepage, S. Pollock, and C. Wieman, Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics, Phys. Rev. Phys. Educ. Res. **15**, 020114 (2019).

[15] L. E. Kost, S. J. Pollock, and N. D. Finkelstein, Characterizing the gender gap in introductory physics, Phys. Rev. ST Phys. Educ. Res. **5**, 010101 (2009).

[16] J. N. Choi and S. Moran, Why not procrastinate? Development and validation of a new active procrastination scale, J. Soc. Psyc. **149**, 195 (2009).

[17] S. F. Chowdhury and T. Pychyl, A critique of the construct validity of active procrastination, Pers. Individ. Differ. **120**, 7 (2018).

[18] D. Corkin, S. Yu, and S. Lindt, Comparing active delay and procrastination from a self-regulated learning perspective, Learn. Individ. Differ. **21**, 602 (2011).

[19] D. Corkin, S. Yu, C. Wolters, and M. Wiesner, The role of the college classroom climate on academic procrastination, Learn. Individ. Differ. **32**, 294 (2014).

[20] B. Mikula and A. Heckler, Framework and implementation for improving physics essential skills via computer-based practice: Vector math, Phys. Rev. Phys. Educ. Res. **13**, 010122 (2017).

[21] D. Oberski, *Modern Statistical Methods for HCI* (Springer, Cham., 2016), Mixture models: Latent profile and latent class analysis, pp. 274–287.

[22] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria 2020).

[23] A. Hayes, *Introduction to Mediation, Moderation, and Conditional Process Analysis* (Guilford Press, New York, 2017), Vol. 2.

[24] Z. Felker and Z. Chen, The impact of extra credit incentives on students' work habits when completing online homework assignments, in *Proceedings of the 2020 Physics Education Research Conference, virtual conference* (AIP, New York, 2020), pp. 143–148.