FULL PAPER - Implementing Exam Wrappers in a First-Year Engineering Course

Dr. Andrew Charles Bartolini, University of Notre Dame

Director, First-Year Engineering Program, University of Notre Dame Associate Teaching Professor, University of Notre Dame

Implementing Exam Wrappers in a First-Year Engineering Course

Introduction

This full paper examines the implementation of exam wrappers into a first-year engineering computing course. Exam wrappers allow students to reflect on how prepared they were for course exams and how effective their preparation was. Previous studies have shown that exam wrappers generally assisted students, particularly underprepared first-semester STEM students, in increasing their metacognitive skill sets [1, 2].

Methods

Course Structure

This study was conducted at the University of Notre Dame, a medium-sized, private, Midwestern, residential university, in the Spring 2024 semester. Exam wrappers were utilized in a multi-section, team-taught, first-year engineering computing course. There were 11 sections of the course taught by 10 different instructors. In this study, only certain sections elected to use the exam wrapper, which allowed a comparison group of students who did not complete exam wrappers.

The course covered foundational programming concepts in both MATLAB and Python. A total of 14 lectures focused on MATLAB (covering variable assignment statements, inputs, displays, conditionals, loops, arrays, plotting, character arrays, strings, user-defined functions, data structures, cell arrays, app designer, and matrix-based implementation) while six lectures focused on Python (covering variable assignment statements, inputs, displays, conditionals, loops, lists, plotting, dictionaries, data types, and user-defined functions). The remaining lectures were used for project work, project presentations, and exam reviews.

All first-year engineering students took the course (regardless of their major and prior programming experience). The exception was students intending to major in Computer Science and Engineering who demonstrated advanced programming experience through an assessment. These students were able to move to the first programming course in the Computer Science and Engineering curriculum and replace the course identified in this study with an additional technical elective. As discussed in this paper, while this study focused on the exam wrappers, the impact of student's prior programming experience was also investigated to provide context to some of the results found in the study.

Exam Structure

It is important first to understand the structure of the exams. There were three exams: (i) Exam 1: Reading Code, (ii) Exam 2: Writing Code, and (iii) Final Exam: Reading Code. The structure of Exam 1 and the Final Exam were similar other than the duration. Both exams were between 70% and 80% multiple choice questions (where the students primarily read and interpreted code). The remaining portion was a free-response style where students identified what a longer code

segment displayed. Exam 1 covered MATLAB (through the topic of strings in the list previously provided), while the Final Exam was cumulative over MATLAB and Python. Exam 2 required students to write code based on a series of prompts (similar to a homework assignment). It covered all MATLAB topics other than app designer.

Exam Wrapper Structure

The exam wrappers were developed by colleagues in the College of Science and have been utilized for multiple semesters in various College of Science courses. The template was adapted for the engineering computing course that is the focus of this study. The exam wrapper consisted of a series of questions, and the relevant questions to this study are outlined below:

- Approximately how much time do you spend each week studying course material using non-required activities?
- How much total time did you spend specifically preparing for the exam associated with this wrapper?
- How prepared did you feel coming into the exam associated with this wrapper?

The wrapper was used in certain course sections, depending on instructor preference. The wrapper was completed the lecture after the exam scores were released. Of the approximately 500 students in the course, 181 and 147 completed the wrapper after the first and second exams, respectively (and consented to their responses being included in the study).

After the first exam, the exam wrappers were briefly introduced in class (with a single slide) as a reflection exercise, and time was allocated to complete the wrapper. Time was again allocated to complete the second exam wrapper, but there was no further discussion.

Research Questions

The focus of this study was on the results from the exam wrappers; however, because of the hypothesized effect that prior programming experience had on the required study methods and time, along with student's perceived level of preparedness, the impact of prior programming experience was also investigated:

- RQ1: How did the duration of preparing for an exam relate to exam performance?
- RQ2: How did a student's perception of preparedness relate to exam performance?
- RQ3: Was there a difference in performance on subsequent exams for those students who completed an exam wrapper?
- RQ4: Was there a difference in perception of preparedness related to a student's prior programming experience?

Results and Discussion

Each of the following sections will address one of the research questions.

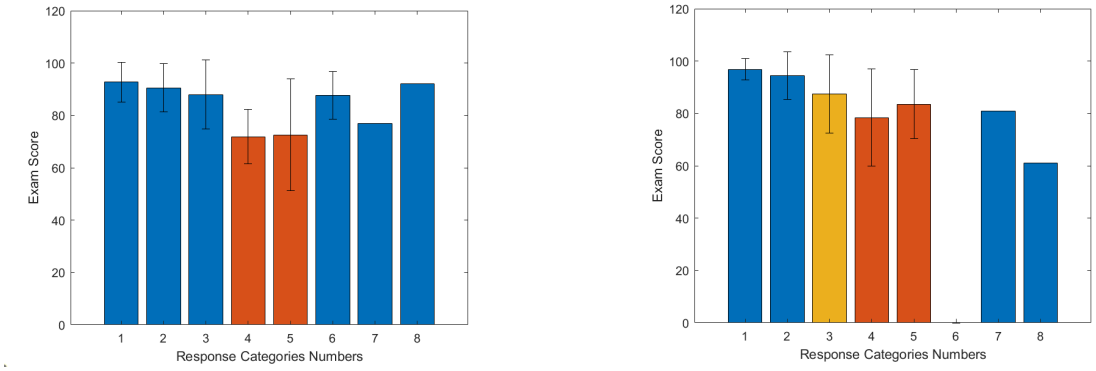
RQ1: Duration of Preparation Related to Student Performance

Figure 1 illustrates the average exam score (along with error bars corresponding to one standard deviation) based on the time a student spent out of class on non-required activities. Figure 2 illustrates the same exam score data based on the time a student spent studying for the particular exam. For Figures 1 and 2, Table 1 outlines the response category numbers and how many students selected each response (i.e., N values). For all the figures, yellow and orange bars represent that the given group differs from the first group (Bin 1) at a statistical significance of $p < 0.01$ and $p < 0.001$, respectively. All statistical tests used an unpaired two-tailed t-test (note: a bar with only one response was not evaluated for statistical significance).

Table 1. Student Responses Based on Amount of Time Studying/Preparing

Bin Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Amount of Time	None	< 1 hr.	1-2 hrs.	3-4 hrs.	5-6 hrs.	7-8 hrs.	9-10 hrs.	10+ hrs.
N (Figure 1a)	39	58	48	8	6	3	1	1
N (Figure 1b)	29	41	37)	23	4	0	1	1
N (Figure 2a)	15	27	47	36	26	6	4	2
N (Figure 2b)	11	18	36	34	25	7	3	1

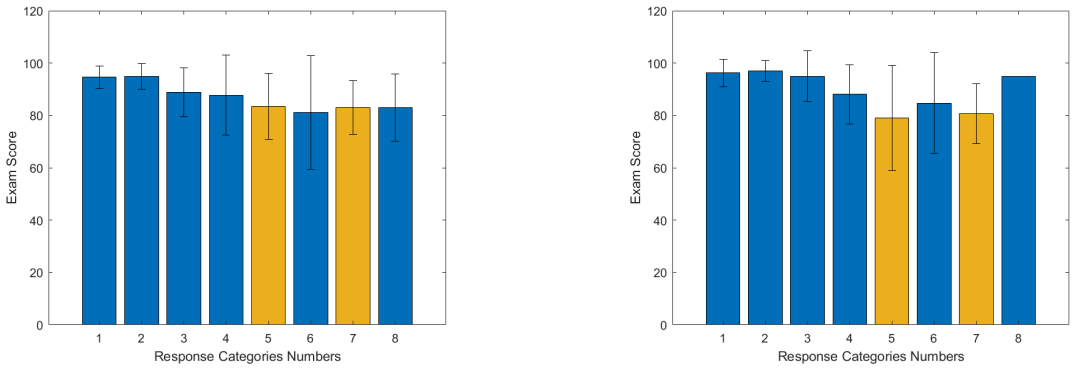
Figure 1a shows no consistent trend related to time spent on non-required activities and student exam performance. Figure 1b demonstrates that students who spent more time on non-required activities scored lower on Exam 2 than those who spent less time.



a) Exam 1 b) Exam 2

Figure 1. Exam Score Related to Amount of Time Spent on Non-Required Activities

The results from Figure 2a indicate that, on average, students who studied longer for the first exam scored lower (although there is a significant spread in the data for students studying between 3 and 8 hours). For the second exam, while students who studied between 5-6 hours earned lower scores than those who studied less, those who studied more than 7 hours outperformed those who studied 5-6 hours on average. For both exams, for each response category from 1 to 6 (i.e., from none to 7-8 hours), at least one student for each response category earned a perfect score. However, for students studying 9 or more hours, the maximum score was less than 100 (although still between 90 and 100).



a) Exam 1 b) Exam 2

Figure 2. Exam Score Related to the Amount of Time Studying for a Particular Exam

One explanation for these results, which at first appear counterintuitive, is that some students enter the course with prior programming experience and therefore may not need as much (or any) studying to feel prepared (about 40% of students in the course reported some prior programming experience – although their proficiency varies from not proficient to very proficient). Another explanation could be that the active learning methods of the course resulted in many students feeling prepared for exams without additional studying or non-required activities. The active learning activities in the course included pre-lecture videos and quizzes on the lecture's topics (flipped classroom), two or three instructor examples during class, a shorter problem started by the students in class on that lecture's topic, and weekly longer homework assignments. Finally, these results potentially highlight that the students who do study more to prepare for exams may not be utilizing techniques that improve their exam performance.

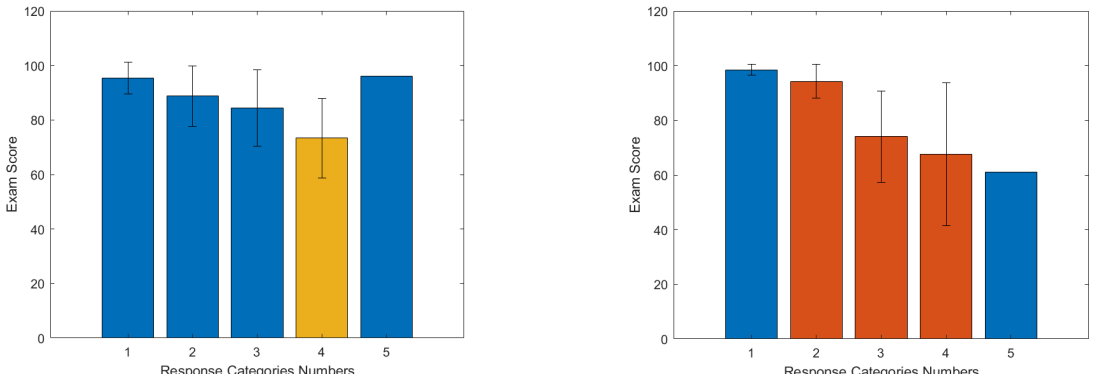
RQ2: Students' Perception of Preparedness Related to Student Performance

Figure 3 illustrates the relationship between students' level of preparedness and their exam scores. Figure 3 is formatted similarly to Figures 1 and 2. Table 2 contains response information.

Table 2. Student Responses Based on Numerical Values.

Bin Number	**1**	**2**	**3**	**4**	**5**
Response	Extremely prepared	Very prepared	Somewhat prepared	A little prepared	Completely unprepared
N (Figure 3a)	6	87	35	5	1
N (Figure 3b)	35	64	33	2	1

With the exception of one response in Figure 3a, Figure 3 demonstrates that students' level of perceived preparedness correlated with their exam scores. The more prepared a student felt for Exam 1 and 2, the higher the student scored on average. It should be noted that these surveys were completed after the exam results had been released; therefore, there is the potential that a student's score biased their responses.



c) Exam 1 d) Exam 2

Figure 3. Exam Score Related to Perceived Preparedness

RQ3: Difference in Performance on Subsequent Exams Based on Completion of Wrapper

The third research question investigated whether students who completed the wrapper performed better on subsequent exams. Table 3 contains data on the median and mean grade changes for both the population of students who completed and did not complete the wrapper. Table 4 contains the same data but only for students without programming experience. It should be noted that the course-wide averages were lower for Exams 2 and the Final Exam compared to Exam 1.

Table 3. Difference Between Exam Scores Based on Exam Wrapper Completion

Complete Exam#1 Wrapper	Difference Between Exams 1 and 2		Difference Between Exam 1 and Final		Complete Exam#2 Wrapper	Difference Between Exam 2 and Final	
	Mean Diff.	Median Diff.	Mean Diff.	Median Diff.		Mean Diff.	Median Diff.
Yes (164)	0.21	1.5	-2.05	-2.75	Yes (136)	-2.92	-3.375
No (281)	-1.18	0	-2.05	-2.25	No (309)	-0.71	-2.5

Table 4. Difference Between Exam Scores Based on Exam Wrapper Completion (Students without Prior Programming Experience)

Complete Exam#1 Wrapper	Difference Between Exams 1 and 2		Difference Between Exam 1 and Final		Complete Exam#2 Wrapper	Difference Between Exam 2 and Final	
	Mean Diff.	Median Diff.	Mean Diff.	Median Diff.		Mean Diff.	Median Diff.
Yes (89)	-0.89	1.0	-1.80	-2.75	Yes (74)	-1.625	-2.25
No (168)	-1.67	0.5	-1.90	-2.50	No (183)	0	-2.25

For both the overall population (Table 3) and just those without prior programming experience (Table 4), students who completed the Exam 1 Wrapper performed better on Exam 2 relative to their Exam 1 score compared to those who did not complete the wrapper. However, there was a negligible difference between Final Exam and Exam 1 scores based on whether or not those students completed the Exam 1 Wrapper. Those students who completed the Exam 2 Wrapper

performed worse on the Final Exam (relative to Exam 2) than those who did not complete the wrapper (however, the median difference between those who did and did not complete the wrapper for students without any programming experience was identical). This finding suggests the initial inclusion of the exam wrapper may have provided a temporary moment for reflection and assessment by the students who helped on Exam 2. However, unless additional time and attention are provided, the positive impact does not persist. Furthermore, it should be noted that none of the aforementioned differences were statistically significant.

RQ4: Students' Perception of Preparedness Related to Prior Programming Experience

The last research question investigated the relationship between prior programming experience and student perceptions of preparedness. Figure 5 illustrates this relationship.

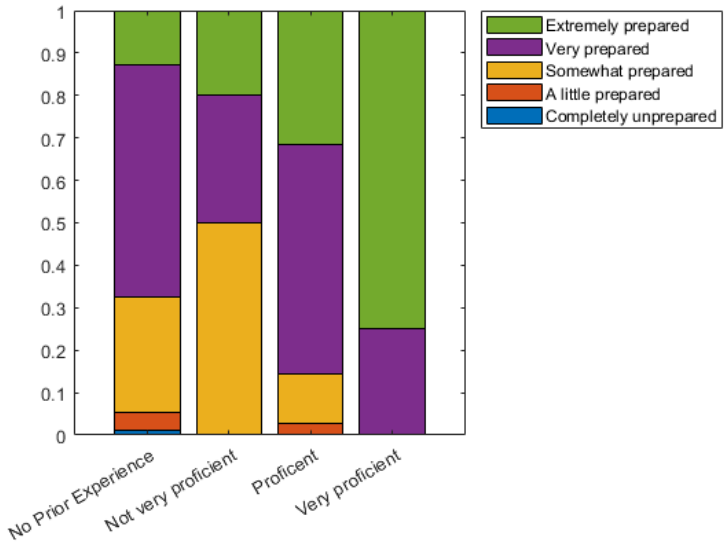


Figure 5. Preparedness Before Exam 1 Related to Prior Programming Proficiency

As expected, a higher percentage of students who self-rated themselves as proficient or very proficient felt very or extremely prepared before Exam 1 compared to those who had no prior programming experience or were not very proficient. However, it was encouraging to see that over 60% of students without any prior programming experience were very prepared or extremely prepared before Exam 1 (although this number was a bit lower for students who had prior experience but were not very proficient).

Future Work

While this study represents using exam wrappers in a minimally intrusive way (i.e., discussing exam wrappers with one slide and allowing a few minutes to complete the survey), future iterations of this work will investigate the impact of using a more thorough and complete overview of exam wrappers and their benefits when introducing exam wrappers to the course.

Additionally, other exam wrapper questions that were not the focus of this study (exam study strategies and exam testing strategies) will be a focal point of future studies.

References

[1] A.T. Stephan, L. Whisler, E.A. Stephan, and B. Trogden. "Using exam wrappers in a self-directed first-year learning strategies course." In 2019 ASEE Annual Conference and Exposition 2019, Tampa, Florida, USA, June 2019.

[2] L.C. Hodges, L.C. Beall, E.C. Anderson, T.S. Carpenter, L. Cui, E. Feeser, T. Gierasch, K.M. Nanes, H.M. Perks, and C. Wagner. “Effect of exam wrappers on student achievement in multiple, large STEM courses.” In Journal of College Science Teaching, 50(1), pp.69-79. 2020