Mastery Grading Approach in a Calculus Course

Dr. Hui Ma, University of Virginia

Dr. Hui Ma is an assistant professor of applied math at the University of Virginia. She holds a Ph.D. in applied mathematics from the University of Alabama at Birmingham. Her current research interest focuses on mathematics education and STEM education, in particular student-focused instruction, and learning-based grading.

Mastery Grading Approach in a Calculus Course

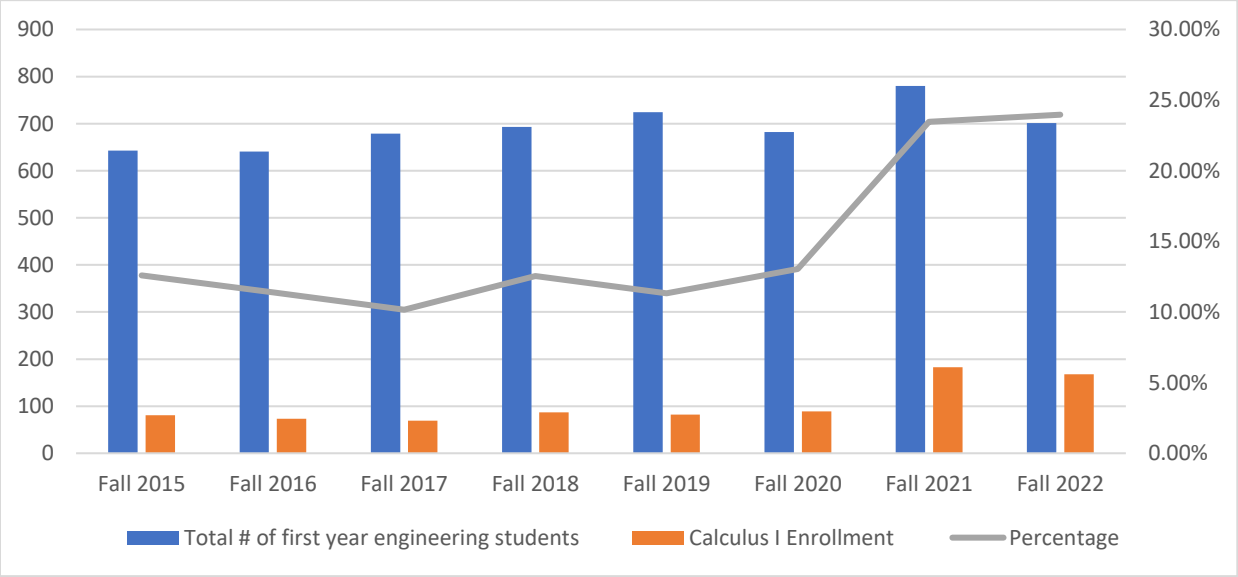
Abstract: In a Calculus I course for approximately 90 first-year engineering students, we adopted the mastery grading approach to reduce stress, foster learning over grades, and enhance growth. This paper reports the key aspects of the mastery grading design and presents initial findings on its impact. Results show that students from the mastery grading sections experienced a significant decrease in test anxiety and a significant improvement in self-efficacy across three categories. No significant changes in performance were observed.

Keywords: Calculus, mastery grading, anxiety, growth mindset, self-efficacy

Introduction

Our university's School of Engineering assumes that most incoming students begin with Calculus II. In general, we have less than 15% of students who begin with Calculus I. Since the pandemic, the number of students who begin with Calculus I has slowly grown to around 24%. Please see Figure 1 for the trend of Calculus I enrollment over the last eight years.

Figure 1. Enrollment in Calculus I over the past eight years



Calculus I students at our engineering school are at a wide range of preparedness levels. Many struggle with this course and the pandemic in the past two years only exacerbated the situation. In fall 2021, 33 out of 183 students (18%) in Calculus I at our engineering school received D/fail/withdraws. Calculus I is considered a gateway course for many engineering students and has been shown to be a critical factor in predicting their success in engineering school. Students in this course – their first engineering and college course – need much more support than those in other courses. They feel anxious about their grades and lack confidence, and a sense of belonging in the classroom [1]. In some cases, students were not prepared because they did not have access to academic support and educational resources before coming to college. Their performance towards the beginning and throughout the semester will not be accurate and will affect their overall course grade. It is also important to remember that not all students learn the

material at the same pace, and some will need a longer time to grasp the concepts. The traditional grading approach reinforces the inequities among students. Consequently, it becomes crucial to rethink how we teach these students and how we align our grading practices to their learning and well-being [3].

In recent years, there has been a significant effort to reconsider how we grade in order to encourage students to prioritize learning over grades, lower students' anxiety, and address some of the inequities of traditional grading schemes [2], [3]. There is a wide range of practices in alternative grading systems, such as "Mastery-based Testing (MBT)" [4], "Standards-Based Grading (SBG)" [5], or "Specifications Grading (Spec)" [6]. These alternative grading approaches are collectively known as "mastery grading".

Key features of mastery grading include: 1) the course material is broken into a list of learning targets. 2) Rather than traditional points or partial credit system, students' work will be evaluated based on their level of understanding of the concept. They will only be given credit for a particular learning target when they display a "mastery" level of understanding. 3)
If they do not master the learning target on the first try, they will have multiple opportunities to demonstrate their learning by reattempting them. They will receive full credit for the learning outcomes once they master it, regardless of when in the semester that occurred, without any penalties.

With the mastery grading approach, instructors and researchers found that students feel less stress or anxiety during timed assessments [7], and they appreciate the opportunities to reattempt the concepts, without being penalized for early mistakes. Instructors also feel that their grades are a better reflection of students' actual learning [1].

Purpose and research questions

In light of the importance of helping students succeed in this class, which sets the foundation for future courses, and the benefit that alternative grading systems can help students reduce their stress levels and focus on learning, the author has implemented the mastery grading approach in her Calculus I class, described below. The following questions guided this pilot study:

1. How, if at all, do student performance and grade distributions differ between mastery and traditionally graded courses?
2. How does mastery grading affect student anxiety, mindset, or beliefs about their ability?

Description of the mastery grading design

Learning Targets:
One of the key components of the mastery grading scheme is to identify the learning objectives for the course. The rest of the course design and assessments will revolve around those key objectives. According to Bloom's Taxonomy [8], there are six levels of cognitive learning: remember, understand, apply, analyze, evaluate, and create. The goal of this course was to build a solid foundation for students and equip them with the tools for their future engineering courses. In the end, the course material is broken into 28 learning targets under four different units: precalculus, limits, derivatives, and integrals, with most of the learning targets falling under the

level of remembering, understanding, and applying. Among the 28 learning targets, 10 of those are designated as core learning targets, which are the most essential topics in Calculus I.

Class Structure:
Students' work in the class follows a pattern that involves them before, during, and after class meetings. Before each class, students read parts of the text, lecture notes, or watch videos that will get them familiar with the basic ideas of new material. Then class time is reserved for doing math, together and individually. The typical class includes warm-up questions, mini-lecture, and group work. The warm-up questions were asked using PollEverywhere, a web-based student response system. This helps ease students into the class in an engaging way and helps me assess their understanding of basic concepts from their reading. Following the mini-lecture, students work in groups on worksheets. This enables students to discuss and explore problems by themselves, thus providing them with opportunities to share ideas and build a learning community. After each class, students practice and apply the ideas we discuss during class through online homework assignments. Online homework assignments are administered via WebAssign to help deepen students' understanding of the material. WebAssign problems can be resubmitted as many times as needed until the deadline.

Assessments and Reassessments:
Most students experience anxiety and stress while taking math exams. The key feature of the mastery grading approach is to provide opportunities for students to reassess until they master the concept. Their final grade will reflect how they eventually understand each topic. There are no exams in this class. Instead, learning targets are assessed on weekly checkpoints. Every checkpoint includes several targets, one problem per learning target. Students will earn a mark for each target, either "mastered" or "progressing". "Mastered" means that students have demonstrated an understanding of the target as well as that the work shown is correct and complete. "Progressing" means that students have demonstrated partial understanding, but with a fundamental error, and the work needs to be reviewed and revised. Each learning target will be offered again in the next three checkpoints so that students can attempt again if they didn't master it. The target is completed if students earn the "mastered" mark twice. Using a 2-time mastery method avoids the situation where students manage to get a mastery mark once but cannot apply it in a different context. Another hope is that students may be able to retain the material longer when they need to review the material multiple times. In addition to weekly checkpoints in class, students can also demonstrate their mastery of a learning target during office hours. A cumulative points-based final exam was also given at the end of the semester. This gives students an opportunity to demonstrate everything they have learned in the course.

Students' Final Grades:
Students' final grades are determined by a bundle of three things: the number of learning targets (including core ones) they mastered, the number of worksheets they have completed, and their overall WebAssign average. See Table 1 for details. To earn a grade, students will need to meet all the requirements in the row for that grade. Students would, for instance, receive an A- if they mastered 25 out of 28 learning targets (including all 10 core ones), completed 32 out of 36 worksheets, and achieved a 90% or above on WebAssign. Students could choose to take the final exam if they wanted to raise their overall grade. If they achieve a final exam grade of 85%, they will advance by half a letter grade. A whole letter grade will be advanced if they receive a 90%.

Table 1: Final Grades Table

Letter Grade	Learning Targets Completed (out of 28)	All 10 Core Learning Targets Completed?	Worksheets Completed (out of 36)	WebAssign assignments with at least
A-	25	Yes	32	90%
B+	24	Yes	31	87%
B	23	Yes	30	83%
B-	22	Yes	29	80%
C+	21	No	28	77%
C	20	No	27	73%
C-	19	No	26	70%
D+	18	No	24	65%
D	17	No	22	60%
D-	15	No	20	50%
F	Have not **fully** completed **any** row			

- If you score 85% on the final exam, you will move up half a letter grade. For example, A- will become an A.
- If you score 90% on the final exam, you will move up a whole letter grade. For example, B will become an A.

In the traditionally graded sections, worksheets and web-assigned homework made up 15% and 20% of the grade, respectively. The final exam and three midterms count for 65% of the final grade. The final exam was the same for both classes. The main distinction is that students in the mastery grading sections used the final exam as a grade modifier to raise their final base grades.

Methods

This study, which received approval from our Institutional Review Board, was carried out at the engineering school of a four-year, public university with roughly 22,000 students. Four Calculus sections taught by two separate instructors were examined to compare the mastery grading system with the traditional grading system. The author taught two sections using the mastery grading scheme, and a different instructor taught the other two using the traditional grading scheme.

Participants

In fall 2022, 168 students enrolled in Calculus I. All 168 students were invited and 143 consented to participate in the study. Out of 143 students surveyed, only 97 completed both pre- and post-surveys. The demographic breakdown for the Calculus I students is included in Table 2.

Both qualitative and quantitative methods were used to assess the grading approach. Data sources included students' placement test scores, pre-course and post-course surveys, and final exam grades. Pre-course and post-course surveys included Likert questions on their beliefs about their math ability using a math self-efficacy scale [9], and math mindsets (modified from [10]). Math self-efficacy refers to an individual's belief in their own ability to succeed in mathematical tasks and activities. The self-efficacy questions were divided into four categories: mastery experience (i.e., previous success informs belief about math ability), vicarious experience (i.e., the success of others informs belief about math ability), social persuasions (i.e., encouragement

from others informs beliefs about math ability), and physiological state (i.e. anxiety, stress, etc. can inform beliefs about math ability). The math mindset questions focused on students' beliefs about the fixed or growth nature of their abilities in math. The survey also included five test anxiety questions that ask students to report how frequently they experience symptoms of anxiety before, during, and after tests, using a 5-item test anxiety inventory [11].

Table 2. Participant Demographics

Groups	Number	Percentage
Mastery Grading	69	71%
Traditional Grading	28	29%
Gender		
Female	38	39%
Male	59	61%
Ethnicity		
Non-Hispanic, White or Euro-America	43	44%
Black, Afro-Caribbean, or African American	16	16%
Latino or Hispanic American	13	13%
South Asian or Indian American	12	12%
East Asian or Asian American	7	7%
Middle Easter or Arab American	4	4%
Other	1	1%
No response	1	1%

Data analysis

Welch's t-tests at a significance level of 0.05 were employed to compare placement test scores of students in sections with different grading methods. This was done to establish the students' grade baseline and measure their course preparedness. Welch's t-test was also used to determine if there was any difference in the students' performance on the common final exam.

For the test anxiety Likert questions, a score ranging from 5 to 20 was obtained by summing the scores for all five questions, with 1=almost never and 4=almost always. Paired t-tests were performed for both grading methods to identify any changes over the semester after taking the course. In the case of the self-efficacy questions, scores for each category (e.g. mastery experience, vicarious experience, social persuasion, and physiological state) were averaged after being set on a scale of 1=definitely false to 6=definitely true. Paired t-tests were performed for both grading methods to evaluate changes in students' self-efficacy over the semester. For the math mindsets questions, scores ranging from 7 (fixed) to 20 (growth) were obtained by summing individual question scores, with 1=strongly disagree and 6=strongly agree. A paired t-test was conducted to compare changes in students' mindsets across the semester.

Qualitative data was collected through open-ended survey response questions and students' comments from the course evaluation.

Results

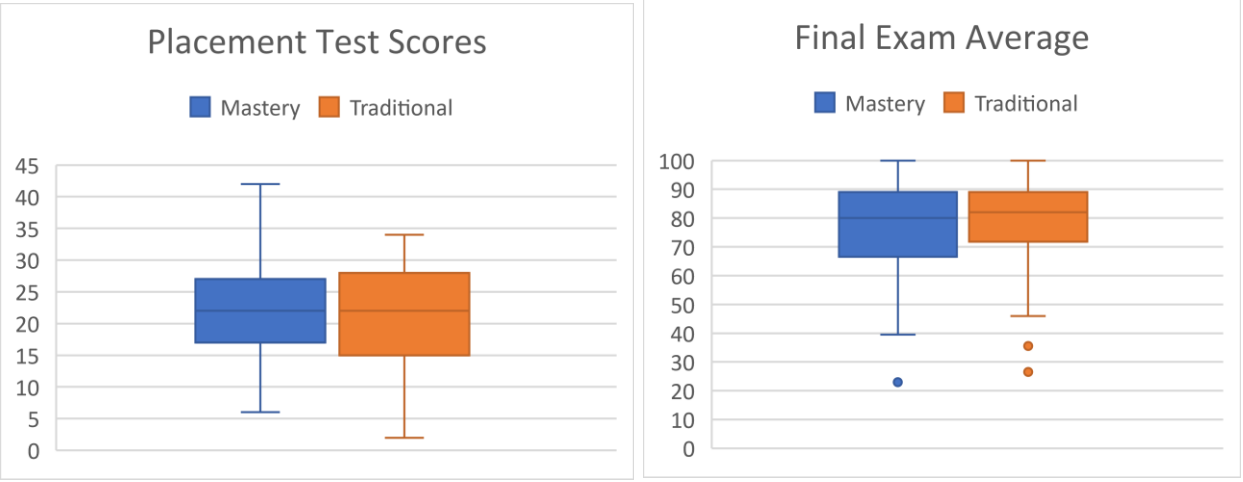
Below we report on the results for the two research questions. Then we discuss the qualitative data collected and students' feedback on the mastery grading approach.

Question 1: How, if at all, do student performance and grade distributions differ between mastery and traditionally graded courses?

Before we compare students' performance at the end of the semester, the author compares students' placement scores and see if there are any differences in students' preparedness between the two groups. The placement test includes 45 questions with 30 pre-calculus and 15 calculus I questions, and each question is graded on a 1-point scale. There was no significant difference between students' performance prior to the start of the course (p-value=0.49). Students from both groups took the common final exam at the end of the semester. The author also didn't find any significant difference between students' performance at the end of the semester (p-value=0.86). Please see Figure 2 for the comparison of students' performance.

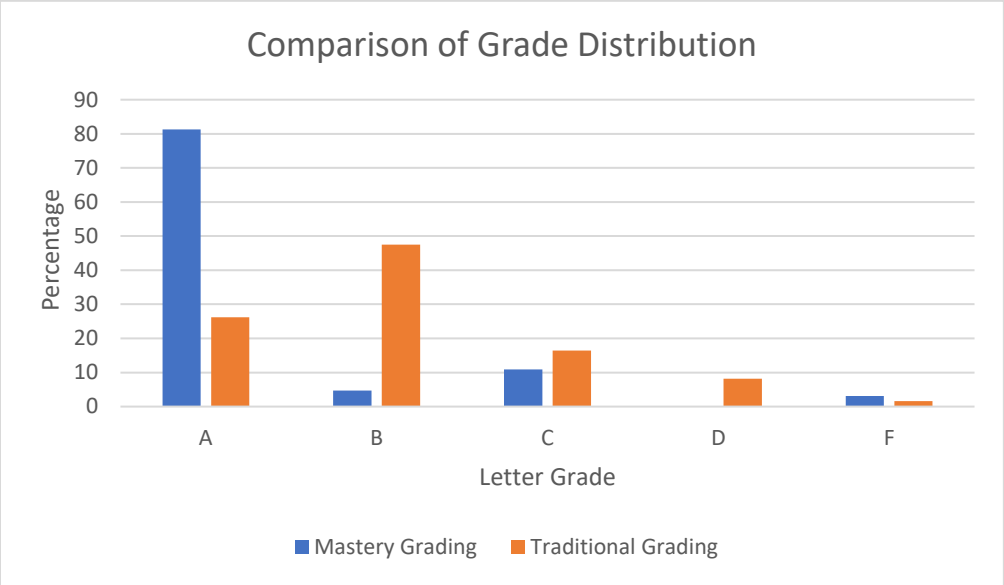
Despite the author's hope that constant revisiting of topics would improve retention and performance, the students in mastery graded sections did not outperform those in traditionally graded sections. It's worth noting that the final exam in mastery graded sections was optional, allowing students to improve their grades if they did well. However, the absence of a negative consequence for not taking the final exam or scoring low may have reduced students' motivation to put in additional effort. Some students submitted their final exams early when they realized a score above 85 was unlikely. Therefore, the final exam score for some students in the mastery group may not reflect their best performance.

Figure 2. Students' Performance



Before the final exam, students in the mastery graded sections can earn up to an A-. Most students did. The final grade distributions between the two groups, as shown in Figure 3 below, indicate a considerable difference, with the majority of students in the mastery group obtaining an A.

Figure 3. Grade Distribution



Question 2: How does mastery grading affect student anxiety, mindset, or beliefs about their ability?

An independent t-test suggested that there was no initial difference in test anxiety between students in the traditional group and the mastery group. However, by the end of the semester, the mastery group showed a significantly lower test anxiety level of 9.78 (p-value < 0.001), compared to 12.64 for the traditional group.

Furthermore, within the mastery group, student's test anxiety significantly dropped from an initial mean of 11.97 to 9.78 over the course of the semester (p-value<0.001). In addition, Students' self-efficacy significantly improved for three categories, mastery experience, vicarious experience, and social persuasion (p-value=0.005, 0.012, 0.018, respectively).

In comparison, test anxiety significantly increased for the traditional group (p-value=0.008) and there is no significant difference in students' self-efficacy in the above three categories.

In both groups, students psychological state (anxiety) score significantly increased (p-value=0.031 for mastery and 0.013 for traditional). It is expected that, regardless of the grading approach, students' anxiety level rises as the semester progresses. In the mastery graded sections, students who haven't mastered many learning targets get stressed and more anxious as it gets near the end of the semester, even when they have opportunities to reattempt.

Comparing students' mindsets, the average post-course score in the mastery group remains higher than the traditional group. However, a paired t-test showed a significant decrease (p-value = 0.007) in students' math mindset pre- and post-course in the mastery group, while there was no difference (p-value = 0.301) in the traditionally graded group. See Table 3 for additional details.

Table 3: Changes in Test Anxiety, Mindset and Self-efficacy Across the Semester

		Pre-Survey Mean (SD)	Post-Survey Mean (SD)	p-value, one tailed
Test Anxiety	Mastery	11.97 (4.06)	9.78 (3.78)	<0.001
	Traditional	11.07 (3.68)	12.64 (3.62)	0.008
Mindset	Mastery	35.61 (4.38)	34.16 (5.03)	0.007
	Traditional	33.71 (5.82)	33.36 (5.68)	0.301
Self-efficacy				
Mastery experience (prior success)	Mastery	4.10 (0.56)	4.25 (0.66)	0.005
	Traditional	4.19 (0.52)	4.12 (0.53)	0.248
Vicarious experience (peer success)	Mastery	4.63 (0.71)	4.78 (0.76)	0.012
	Traditional	4.46 (0.53)	4.52 (0.63)	0.289
Social Persuasions (support for success)	Mastery	4.38 (1.05)	4.54 (0.91)	0.018
	Traditional	4.41 (0.92)	4.45 (0.87)	0.388
Physiological State (anxiety)	Mastery	2.21 (0.96)	2.37 (1.02)	0.031
	Traditional	2.27 (0.99)	2.68 (1.02)	0.013

Note: test anxiety sum from 5 (low) to 20 (high), mindset sum from 7 (fixed) to 42 (growth), all self-efficacy Likert scale from 1 (low) to 6 (high). Significant changes are highlighted, $p<0.05$.

When students were asked to describe the primary factors that they believe affect their success in math classes, an open-ended question in the math mindset instrument, the factors mentioned the most in their responses were:

1. Hard work/diligence and time spent on the material
    - Practice problems/studying outside of class
    - Completing homework/assignments
2. Good attitude/mindset
3. Attending class/lectures
4. Asking for help/taking advantage of resources
5. Ability to learn from failure/mistakes.
6. Good teacher/instructor's teaching style.

Student feedback:

The author also collected feedback from students in the mastery graded sections. Students are generally positive towards the mastery grading approach. When asked about their experience with the class, students commented that 1) grading approach allowed them to revisit ideas and encouraged more practice and better retention of knowledge. 2) the opportunities to reattempt took a lot of stress away, especially for their first semester of college, and 3) the feedback received from the assessments and office hours for reassessment really helped them learn from their mistakes and improved their understanding of the material.

Here are some selected comments:

- The mastery learning system was highly effective. I found myself practicing more than I would have for the same class but with large exams, and I find I can recall information and complete problems from even the first few weeks of the semester with very little effort.

- Because of the checkpoints, I was able to not forget content due to the fact that I was constantly seeing content from the past on the checkpoints which motivated me more to keep reviewing so I could master them.

- The way that the class is designed allows for your grade to only go up and not down. Although the learnings targets may take a while to get used to, I was able to complete them with reassessments during office hours. I knew what I did wrong and improved throughout the year.

- I enjoyed the grading system for this course. It relieved a lot of the stress that often comes with learning math, and I felt like I actually had a chance to make mistakes without failing where in a "normal" grading situation, that wouldn't have been the case.

When asked about the most challenging or confusing parts of the mastery grading system, students mentioned that the grading system can take some getting used to and the learning targets can pile up if they don't make progress, which put stress on them towards the end of the semester. And some students still prefer traditional grading system, one student commented, "I did not like the grading system for this semester at all. I believe that if every category were put on a percentage basis instead of a "mastery" basis it would be much easier to succeed with the same amount of work. The "mastery" system made this course unnecessarily hard."

Discussion

There was no statistically significant difference in students' performance when comparing the mastery and traditional grading approaches. The final grade distributions suggest a significant difference with the mastery grading group getting higher letter grades. This is due to the fact that students have multiple opportunities to reattempt and demonstrate their understanding of the learning targets without any penalty before the end of the semester. Most student who are willing to put the effort ended up achieving A- by the end of the semester. Just as one student commented, "I am considering how much work students have to put in if they continue getting things wrong. Some may give up, but those who are willing can strive through it." However, the author noted that a small portion of students may not be A-grade students. Due to the limited number of versions of each learning target problem, these students were able to memorize the solution patterns by taking the assessments repeatedly and ultimately received a mastery grade without fully comprehending the material. This highlights the need for the instructor to create more diverse and varied assessments that cannot be memorized and to write problems in a way that discourages this approach. Additionally, having more versions of the problems would further prevent this issue.

The mastery grading approach had a significant impact on students' anxiety and beliefs about their ability. The mastery graded group showed a significant decrease in test anxiety and significant improvement in self-efficacy. In contrast, the traditionally graded group showed a significant increase in test anxiety and no significant difference in self-efficacy. It is interesting that the mastery graded group showed a lower mindset score at the end of the semester compared to the beginning of the semester. This could be due to the growth mindset video shown during the first day of class and instructor's buy-in speech, which had helped to raise their initial mindset score.

When implementing the mastery grading approach, the author faced many obstacles and challenges, which required extra thought and planning for the future semesters. For instance, many students procrastinated and only came for reassessments at the end of the semester, causing long lines outside the office during office hours. This was a heavy burden for the instructor and left insufficient time for providing individualized feedback, which was the purpose of the office hour reassessments. The unlimited number of attempts did not sufficiently motivate students to perform better on their first tries. It also generated an excessive amount of grading. The author needs to reevaluate the number of reassessments allowed and encourage students to reflect and review before reattempting.

Additionally, there have been practical challenges such as grading and syncing grades with Gradescope, tracking students' progress, recording reassessments, and communicating transparently with students about their learning progress. Since the pandemic, Gradescope, an online grading tool that allows paper-based assignments to be scanned, submitted, and graded, has been used for grading our students' work. As Gradecope was designed primarily for the traditional grading, it had several limitations with regards to the binary grading approach used in mastery grading.

Limitations and future work

This pilot study was limited to a single course, taught by two different instructors, and for one semester only. Many factors could impact students' performance and their experience with the course in each group. To gain a better understanding of student performance, it would be useful to collect additional data in subsequent semesters by introducing a non-optional common, final assessment. Additionally, the author is interested in examining which students would derive the greatest benefit from the mastery grading approach. Future work might also involve expanding the investigation to other courses and seeing how students' mastery grading experience affect their performance and experience in later calculus and engineering courses.

References

[1] Hendry, G. D., White, P., & Herbert, C. (2016). Providing exemplar-based 'feedforward' before an assessment: The role of teacher explanation. *Active Learning in Higher Education*, 17(2), 99-109. https://doi.org/10.1177/1469787416637479

[2] Robert Campbell, David Clark & Jessica OShaughnessy (2020) Introduction to the Special Issue on Implementing Mastery Grading in the Undergraduate Mathematics Classroom, PRIMUS, 30:8-10, 837-848, DOI: 10.1080/10511970.2020.1778824

[3] Adriana Streifer, Michael Palmer (2020). Alternative Grading: practices to support both equity and learning. https://cte.virginia.edu/blog/2020/12/04/alternative-grading-practices-support-both-equity-and-learning

[4] Amanda Harsy (2020) Variations in Mastery-Based Testing, PRIMUS, 30:8-10, 849-868, DOI: 10.1080/10511970.2019.1709588

[5] Megan E. Selbach-Allen, Sarah J. Greenwald, Amy E. Ksir & Jill E. Thomley (2020) Raising the Bar with Standards-Based Grading, PRIMUS, 30:8-10, 1110-1126, DOI: 10.1080/10511970.2019.1695237

[6] Carlisle, S. 2020. Simple specifications grading. *PRIMUS*. 30(8–10): 926–951. https://doi.org/10.1080/10511970.2019.1695238.

[7] Lewis, D. (2022). Impacts of Standards-Based Grading on Students' Mindset and Test Anxiety. Journal of the Scholarship of Teaching and Learning, 22(2), 67-77.

[8] Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into practice, 41(4), 212-218.

[9] Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. Contemporary educational psychology, 34(1), 89-101.

[10] Dweck, C. S. (2014). Mindsets and math/science achievement.

[11] Taylor, J., & Deane, F. P. (2002). Development of a short form of the Test Anxiety Inventory (TAI). The Journal of general psychology, 129(2), 127-136.