Comparison of Partial Credit and Mastery Assessment on Student Learning and Retention

Shraddha Sangelkar, William Lasher, Russell L. Warley, and Adam S. Hollinger
Mechanical and Industrial Engineering Department
Penn State Behrend
Erie PA, USA

Abstract— **For several years, the faculty in the Mechanical Engineering program at Behrend have been using a modified mastery grading system in three sophomore-level courses: Statics, Strength of Materials, and Thermodynamics. In this system students are given up to three opportunities to solve each problem correctly. A solution must be essentially correct (minor errors only) in order for points to be awarded, otherwise a retake is required on a similar problem with the points awarded for a correct solution reduced with each retake. Although the effectiveness of this system has been evaluated qualitatively through student and faculty opinion, there has been no performance-based evaluation of the difference in student learning between the mastery system and the traditional partial credit system. This paper describes the results of a comparative study conducted during the spring 2015 semester on four sections of Strength of Materials (SOM). Students in each section (117 total students) of the SOM class were divided into a "mastery" group or a "partial credit" group such that there was no statistical difference between the GPA's of the two groups. Both groups took the same five initial exams during the semester, with the mastery group exams graded as described above and the partial credit group exams graded using a partial credit system. The partial credit group was not able to retake any of the exams. All students took a common final exam, which was graded on a partial credit basis, and the results of the two groups were compared. The final exam consisted of ten concept questions, five fundamentals problems, and three typical exam problems with all problems graded using partial credit. The results of this study showed that there was no significant difference in the total exam score between the two groups, nor was there a difference on the concept questions, fundamentals problems, and the first two typical exam problems. The students in the mastery group did perform better on the third typical exam problem (p-value 0.03). Based on these results it appears that mastery grading does not improve student comprehension of fundamental concepts, but that repeated practice does make them better able to solve specific problems. Sixty-two of these students were enrolled in a course for the fall 2015 semester which requires SOM as a prerequisite. The students were fortuitously evenly distributed as to GPA and the exam method from the previous course. During the first week of the fall semester, a prerequisite quiz consisting of the concept questions and fundamentals problems from the spring SOM final was administered, and the results from the two groups were compared. Based on this study, there is not enough evidence to indicate that student recall of material is impacted by the grading system.**

Keywords—mastery grading, assessment methods

I. INTRODUCTION

For several years the faculty in the Mechanical Engineering program at Penn State Behrend have been using a modified mastery grading system for exams in three sophomore-level courses: Statics, Strength of Materials, and Thermodynamics. The basic mastery learning approach has been used in education for decades. The main idea behind mastery learning is that the course material is broken down into modules and the students are required to demonstrate their mastery of each module before moving to the new module. In the modified mastery system, implemented at Penn State Behrend, students are given up to three opportunities to complete each problem. A student must get a problem perfectly correct or almost correct (e.g., all equations were correct but there is a minor calculation error) to get credit; if not, then they have to retake that particular problem. The problems are relatively straight-forward and focus on basic concepts. With each retake, the problems become slightly simpler and the maximum points available decreases. Details of the mastery grading system can be found in [1-3].

Results of the previous studies indicated that students do not appreciate the benefits of mastery learning for statics after they take future engineering courses but not while they are enrolled in the course [1]. Basically, mastery learning helps to cultivate a habit of learning from mistakes and gives students additional opportunities to learn the material if they missed it the first time compared to partial credit grading. With mastery grading, the students more often go back and look at the problems on which they did not do well. Thus, mastery learning helps to cultivate a culture of trying over and over till the students get it right.

From the faculty perspective, a mastery approach helps to gather rich data on proportions of students who pass each concept since scores are necessarily tracked on a problem basis rather than summary exam scores. The passing rate of each concept could be a valuable indicator for assessment of ABET student outcomes [1]. For instance, the fact that 80% of students can solve 2D rigid-body equilibrium problems helps the instructor to evaluate a specific course outcome, which is on-going work leading to a future publication. Also, based on the passing rate from previous semesters we can predict the concepts that are difficult for most of the students and better plan for instruction in future offerings of the course.

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

The authors also conducted faculty interviews in a prior study to understand the benefits and pitfalls of a mastery approach [2]. The instructors agreed that the mastery learning provided them with a deeper insight into students' achievements and needs, made the grading easier, and improved the learning experience. On the other hand, instructors identified some challenges to the approach. The mastery learning approach reduces the flexibility in writing exam questions, and instructors might get negative ratings and comments from students who are frustrated by the approach. Some of the advantages of the modified mastery approach over the partial credit approach are summarized below:

- Students have to demonstrate that they can solve all problems correctly. In a partial credit system, students can fail one problem but pass the exam by getting a good score on the rest of the problems.
- Students who are weak or have a "bad day" have additional opportunities to demonstrate their knowledge through repeated attempts.
- Grading of the problems is simplified and is quicker because the grader does not have to assign points to partially correct work since there are only three possible scores correct, almost correct, or repeat - with pre-defined point values.
- The simplified grading allows faculty to conduct more frequent exams on fewer topics.
- The approach directly motivates students to solve more problems if they have not shown that they are able to.

There are also some disadvantages to the mastery system such as faculty concerns over whether the problems are too easy and focus on problem-solving as opposed to understanding of the fundamentals. The other disadvantage is the time required to set up and administer multiple exams.

The effectiveness of the mastery grading system has been evaluated qualitatively through student and faculty opinion. However, there has been no quantitative evaluation of the difference in student learning between the mastery system and the traditional partial credit system [1, 2]. This reality led to a desire to undertake a comparative study so that the effect of the grading system on student learning could be evaluated directly. This paper describes the results of such a study conducted during the spring 2015 semester on four sections of Strength of Materials. The objective of the study was to compare the performance of the students in two similar groups (one group graded using the mastery system, the other group graded using a partial credit system) on a common final exam graded using partial credit. The two groups were also evaluated on a prerequisite quiz in a follow-up Intermediate Mechanics of Materials (IMM) course taken during the Fall 2015 semester.

II. Research Method

This paper addresses two research questions, namely,

1) Do the students subjected to mastery assessment perform differently than the students subjected to partial credit assessment on a comprehensive final exam in the course?

2) Do the students subjected to mastery assessment retain information differently than the students subjected to partial credit assessment on the prerequisite quiz in a follow-up course?

To test these two hypotheses, all the students enrolled in the Strength of Materials (SOM) course in the spring 2015 semester were divided into a "mastery" (M) group or a "partial credit" (P) group. The sample size for both groups was the same when they were divided into two groups, but some students dropped the course, hence the students who dropped the course are not included in the analysis. There was no statistical difference between the grade point averages (GPA) of the two groups. All students took the same five exams during the semester, with the mastery group exams graded as described earlier and the partial credit group exams graded using a traditional partial credit system. The partial credit group was not permitted to retake any of the exams.

To test the first hypothesis, all students enrolled in the SOM course were given a common final exam, which was graded completely on a partial credit basis. To test the second hypothesis, the students were given a portion of the same final exam used in the SOM course in a follow-up course on Intermediate Mechanics of Materials (IMM). The next section describes the population investigated in this study.

A. Study Population

The study population consisted of students enrolled in the sophomore-level classes in the School of Engineering at Penn State Behrend. Students in the course were planning on majoring in one of a variety of different programs offered within the Penn State system. Table 1 shows the number of students enrolled in each section of the SOM course in the spring 2015 semester. It also shows the number of students subjected to mastery and partial credit assessment. One student from the mastery group scored an overall grade of less than 30% and preferred to not drop the course; that student was excluded from the analysis.

Table 1. Sample size from each section of Strength of Materials course in spring 2015 semester

Condition	Section 1	Section 2	Section 3	Section 4	Total
Mastery	19	18	17	6	60
Partial Credit	15	18	18	5	56
Total	**34**	**36**	**35**	**11**	**116**

The students from this cohort of spring 2015 were tracked in the junior-level IMM course in fall 2015. We were able to track a total of 62 students between the two semesters and compare their scores. Interestingly, the number of students (thirty-one) in the IMM course who had been in the mastery group was identical to the number who were in the partial credit group. The next section explains the results of the study.

III. Results and Discussion

The final exam consisted of ten concept questions, five fundamentals problems, and three typical problems. The exams were graded on a partial credit basis using increments of half a point, with the point total for each type of question explained as follows. Each of the concept questions was worth one point and focused on one concept such as the compatibility condition for a statically-indeterminate problem or identifying the boundary conditions for a beam deflection problem. Each of the fundamentals problems were worth two points and covered simple problems such as the application of the parallel-axis theorem or the free body diagram of a beam showing the development of the equation for the internal moment. The typical problems were each worth ten points and were similar to those taken by the students during the regular exams. The sum of scores on the ten concept questions was added for each student and the mean of the sum of scores was compared between the mastery and partial credit groups as shown in part (A) of Fig. 1. Similarly, part (B) of Fig. 1 compares the means of the sum of the scores on the five fundamentals problems between the two groups. Finally, the comparison of the total final exam scores is shown in part (C) of Fig. 1. A two-tailed t-test was performed in each case to compare the mean values and the p-values are reported along with the figures. Fig. 1 shows that there is no significant difference in the mean scores for the concept questions scores, fundamental problems score, or total exam score between the two groups.

Fig. 2 shows the comparison of the means between the two groups on the three typical problems asked on the final exam. The students in the mastery group performed significantly better on the first typical problem (p-value < 0.05). The first problem was a combined loading problem, which was covered on the last two exams, so including the retakes, the mastery group had up to six opportunities to practice this type of problem. On the other two typical problems, there was no significant difference in the performance of the two groups. Based on these results it appears that mastery grading does not improve student comprehension of fundamental concepts, but that repeated practice does make them better able to solve specific types of problems.

Table 2 shows the comparison between mastery and partial credit groups across each section. Similar to the aggregate data, results of the section-wise comparison indicate that there is no statistically significant difference between the mean scores of the two groups within each section. This observation suggests that the section's instructor does not affect the overall performance between the mastery and partial credit group.

Sixty-two of these students also took the follow-up IMM course during the fall 2015 semester. During the first week of the semester, a prerequisite quiz, which consisted of the ten concept questions and five fundamentals problems from the spring SOM final, was given. Fig. 4 shows a comparison of the means using a two-tailed t-test for the sum of scores on the concept questions, sum of scores on the fundamental problems, and the total scores between the two conditions. There was no statistical difference between the two groups, which indicates that the mastery grading system did not significantly impact the students' recall of the material over the summer.

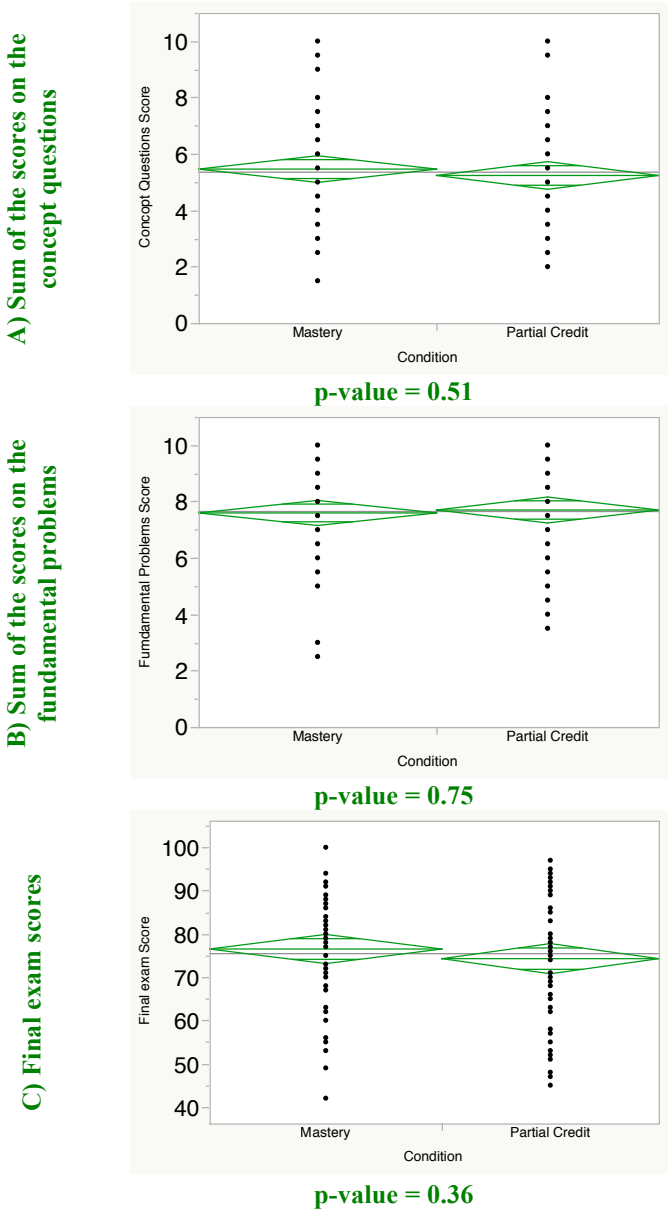

Fig. 1. Results of the two-tailed t-test to compare the mastery and partial credit scores for the (A) Concept questions, (B) Fundamental problems, (C) Total final exam score in the Strength of Materials course.

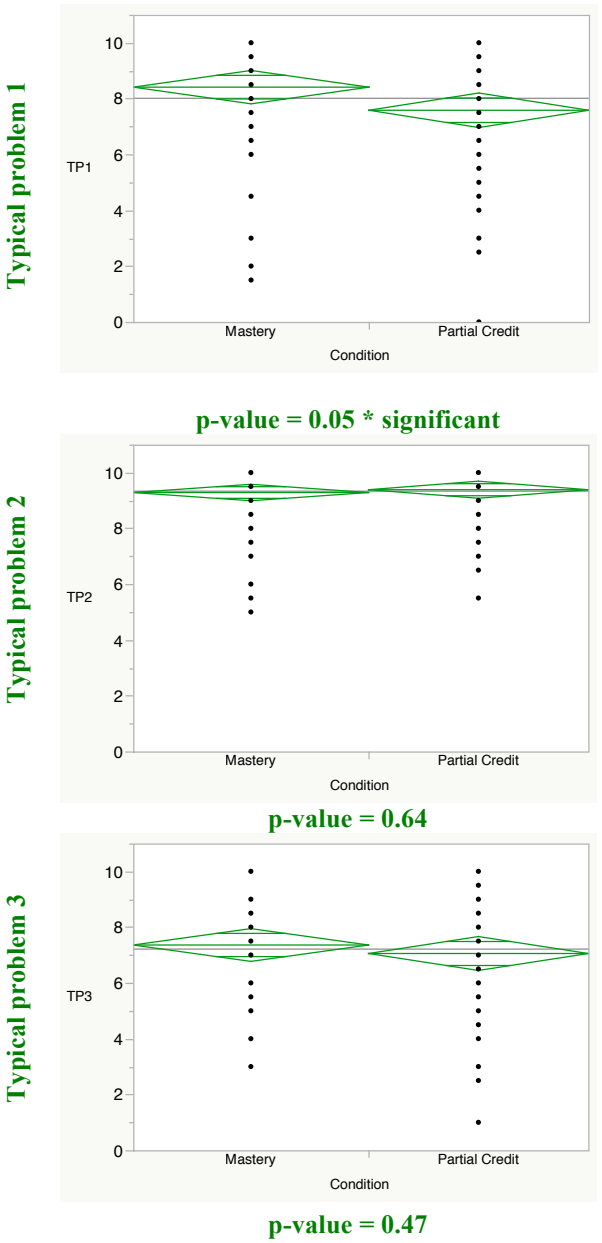

Fig. 2. Results of two-tailed t-test to compare the mastery and partial credit scores on the three typical problems in the Strength of Materials course. There is a significant difference on the scores for the first problem.

Since we were tracking the same set of students between the two semesters it would be interesting to see how the performance of each individual student impacted the comparison between the two groups. The students' score on the final exam in the SOM course was subtracted from their corresponding score on the prerequisite quiz of the IMM course. A positive change for a student would indicate that the student retained the material better. A negative change would indicate that the student forgot concepts over the summer. Fig. 3 shows the comparison of the mean differences using a two-tailed t-test between the mastery and partial credit groups. The results indicated that there was no significant difference in the material retained between the mastery and partial credit groups. Similar results were observed for comparison between the sum of scores on the concept questions, and the sum of scores on the fundamental problems.

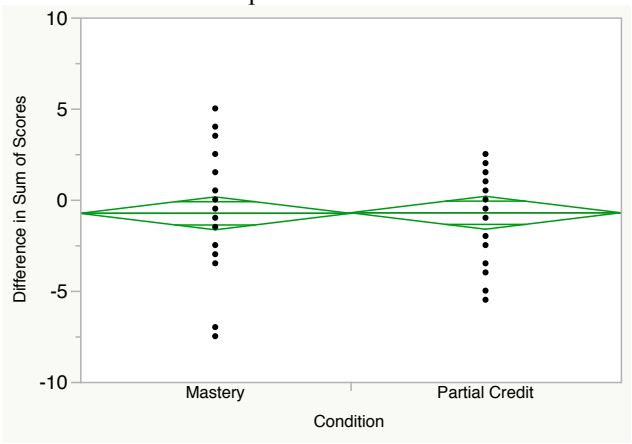

Fig. 3. Results of two-tailed t-test to compare the mean of the change in students' score between the two semesters.

The authors were somewhat surprised that the mastery group did not outperform the partial credit group, except on one typical problem, as the intended purpose of the mastery approach is to ensure that the students have "mastered" each of the concepts taught in the course. There are several possible explanations for this result, each of which requires further investigation:

Table 2. Section-wise comparison of the mastery and partial credit group (no significant difference) for the final exam scores, concept question scores, fundamental problem scores, and typical problem scores for the Strength of Materials course.

			Final Exam Score			Sum of Scores on the Concept Questions			Sum of Scores on the Fundamental Problems			Sum of Scores on the Typical Problems		
Section	Condition	Sample Size	Mean	Std Error	p-value	Mean	Std Error	p-value	Mean	Std Error	p-value	Mean	Std Error	p-value
1	M	19	82.5	2.18	0.84	6.7	0.42	0.12	7.7	0.40	0.34	26.4	0.62	0.30
1	P	17	83.1	2.45		5.7	0.47		8.2	0.45		27.4	0.70	
2	M	15	75.0	2.82	0.77	4.6	0.43	0.23	8.0	0.30	0.44	24.9	0.97	0.20
2	P	18	73.8	2.82		5.4	0.43		8.3	0.30		23.1	0.97	
3	M	18	71.9	3.55	0.38	4.8	0.37	0.93	7.1	0.47	0.73	24.1	1.22	0.25
3	P	6	67.6	3.45		4.8	0.36		6.9	0.46		22.1	1.19	
4	M	18	74.8	5.96	0.85	5.9	0.85	0.48	7.4	0.78	0.50	23.9	2.21	0.98
4	P	5	73.2	6.52		5.0	0.93		6.6	0.86		24.0	2.42	

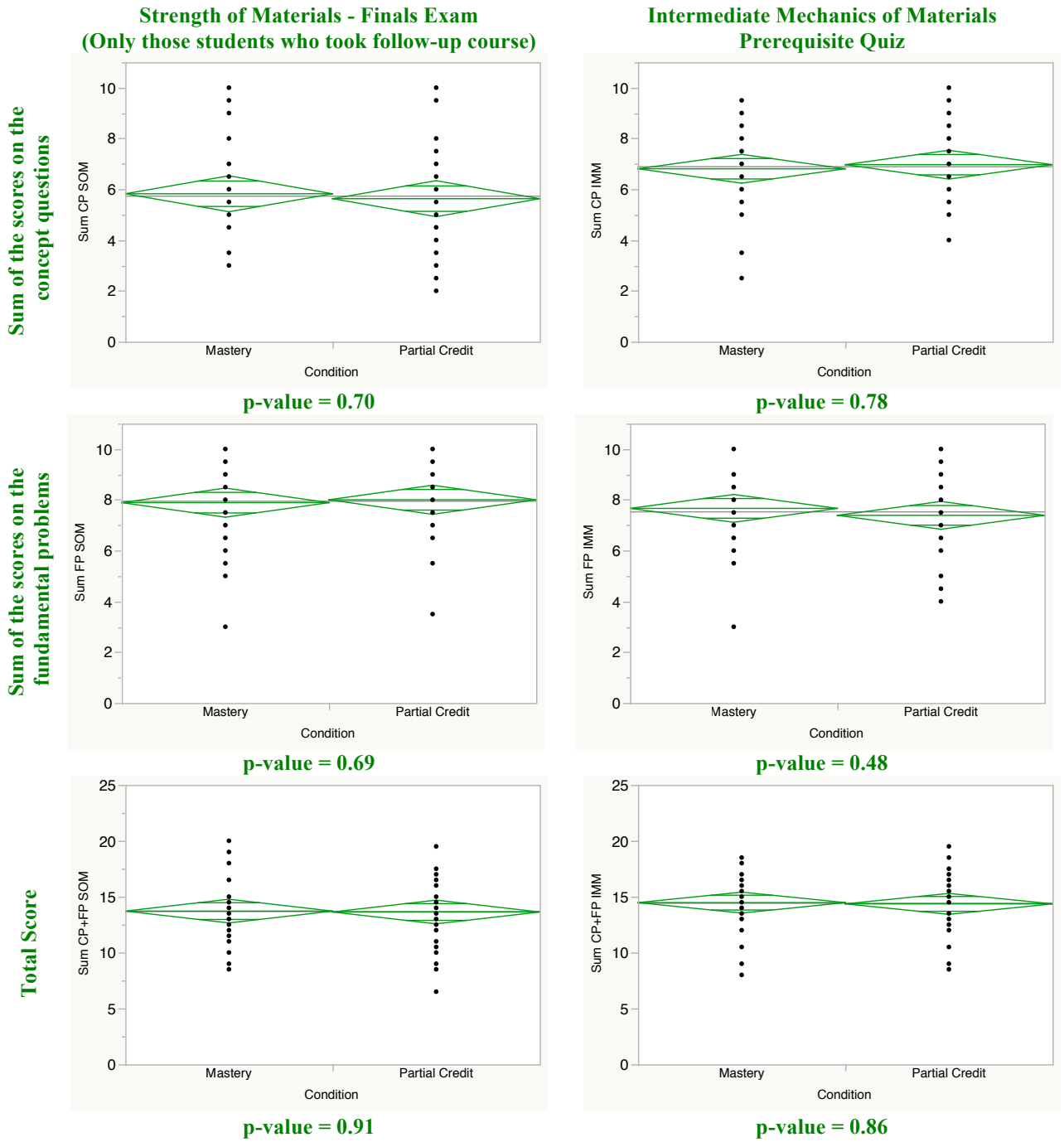

Fig. 4. Results of two-tailed t-test to compare the mastery and partial credit scores in the two classes based on the concept questions and the fundamental problems (no significant difference in the means between the two groups).

1. The five exams taken during the semester were all typical problems, so the students did not have practice in answering concept or fundamental questions beyond those asked in class using clickers. This lack of practice may have contributed to the similar scores.
2. Typical problems evaluate the ability of a student to solve a problem, whereas concept and fundamental questions are directed at the student's understanding of the concept. It may be possible that students are solving problems heuristically based on their study of similar problems, rather than understanding basic concepts and problem solving techniques
3. Students' approaches to preparing for an exam may be different between the two groups. For example, students in the mastery group may see the first exam as an opportunity to find out what type of problems they will have to solve, since they know they will have a second or third opportunity to re-take the problem, whereas students in the partial credit group

know that they only have once chance at solving the problem.

4. Student approaches to taking an exam probably differ depending on how the exam is graded. A student in the mastery group may spend more time working on a particular problem knowing that they will not get credit unless it is correct or almost correct, whereas students in the partial credit group will put anything down that they think might earn them points.

As noted above, the better performance of the mastery group on the first typical problem suggests that practice does indeed make perfect. If the objective is to be able to get students to solve a particular type of problem, the repeated exposure of a mastery approach appears to offer an advantage.

IV. Summary

Based on the results of this study, a mastery grading approach does not appear to improve students understanding of fundamental concepts but, if the students are given enough opportunities to practice a particular type of problem, it does improve their ability to solve that type of problem. Furthermore, the mastery grading approach does not improve students' recall of concepts over a summer, which is consistent with the insignificant difference between the two groups on concepts for the Strength of Materials final exam.

It should be noted that both students and faculty tend to have a preference of one grading system over another. Historically, all sections of a course during a given semester have been graded using either the mastery approach or the partial credit approach for consistency and to avoid a sense of unfairness amongst the students. The results of this study suggest that both approaches work equally well in terms of student learning, at least in terms of understanding of concepts, and that allowing flexibility for instructors of different sections should not impact learning (although it could impact student perception of fairness). One suggestion that has been made is to have courses listed as either "mastery grading" or "partial credit grading" and allow students choose whichever approach they prefer. It would be interesting to try this approach and see the results of a similar study.

Another option would be to modify the mastery approach to include at least some concept questions, which could be graded either mastery or partial credit. This modification would allow students to practice both types of questions during the semester and may improve their performance on the final exam.

Finally, the approach used here is a modified version of the mastery system originally proposed by Bloom [4]. The original mastery system requires mastery of a topic before the students progress; in the modified mastery system, students can continue even if they "fail" all three retakes of an exam. Inclusion of such elements of Bloom's approach may improve students' performance [4].

References

[1] S. Sangelkar, O. M. Ashour, R. L. Warley, and O. J. Oladipo, "Mastery Learning in Engineering: A Case Study in Statics," in *ASEE Annual Conference and Exposition*, Indianapolis, IN, 2014.

[2] O. M. Ashour, S. Sangelkar, R. L. Warley, and O. Onipede, "Redesign the engineering teaching and assessment methods to provide more information to improve students' learning," in *Frontiers in Education Conference (FIE), 2014 IEEE*, 2014, Madrid, Spain.

[3] O. Onipede and R. Warley, "Rethinking Engineering Exams To Motivate Students," presented at the 26th Annual Lilly Conference on College Teaching, Miami University, Oxford, OH, 2007.

[4] B. S. Bloom, "Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1," *Evaluation comment,* vol. 1, p. n2, 1968.