Paper ID #17615

Adapted Mastery Grading for Statics

Cmdr. Linda E. Craugh P.E., United States Naval Academy

CDR Craugh is a Permanent Military Professor in the Mechanical Engineering Department of the United States Naval Academy, a Surface Warfare Officer, and a registered Professional Engineer.

Adapted Mastery Grading for Statics

Abstract

This paper presents a study of an adapted mastery grading scheme applied to an introductory Statics course. Previous studies have identified student characteristics that predict performance in this course at this school. This study will compare those predictors for the entire enrollment (14 sections) of the course against those of the three sections of students subjected to the alternative grading scheme. The alternative scheme was applied to five in-class exams unique to the three sections; the final exam was common to all sections and was graded using standard partial-credit methods. Discussion of the particular implementation of the mastery grading scheme as well as student feedback will also be included.

Introduction

Initially discussed by Bloom[1] in 1968, mastery learning is an instructional strategy designed to minimize achievement gaps and have all students learn well. In Bloom's formulation, course material is broken into discrete units, and students must demonstrate their mastery of current material before moving on to subsequent units.

Sangelkar[2] and her colleagues implemented a variation of this philosophy course-wide for an introductory Statics course at Penn State, Behrend. In that implementation, lectures continued at an instructor-defined pace, but exams were administered in a separate setting (an evening testing window) and graded for mastery. That is, students were given up to three opportunities to demonstrate their understanding of a concept through testing. That demonstration of mastery was pre-requisite to their continuing on to demonstrate their understanding of subsequent concepts.

This paper presents an adaptation of Sangelkar's implementation for select sections of introductory Statics at the United States Naval Academy. Previous research by Burkhardt[3,4] has identified strong predictors of success in this course. In that work, data collected over eleven years was analyzed and the factors that most strongly correlated to Statics performance (final course grade) were previous semester quality point rating (QPR, akin to grade point average), Chemistry I grade, and Calculus I grade. Lesser correlations were found for SAT scores (both math and verbal), and gender and ethnicity did not correlate to course performance. The quantitative portion of this paper will use these previously identified predictors to compare the expected performance of students to their actual performance on a common final exam as a measure of the efficacy of mastery grading, if not learning.

Implementation

In fall 2015, Statics (EM211) was offered to fourteen sections of students taught by six instructors. The students were predominately sophomore engineering majors (Mechanical, Systems, and Ocean Engineering) taking their first courses in their respective majors. The author taught three sections that each met for seventy-five minutes (a period-and-a-half) twice a week on Tuesdays and Thursdays; they will be referred to as the "mastery" sections. The other eleven sections each met for fifty minutes three times a week on Mondays, Wednesdays and Fridays; they will be referred to as the "traditional" sections. All sections administered five instructor-generated exams on specified content throughout the semester. The team of instructors collaborated to write a comprehensive common final exam, which was team graded using traditional partial credit methods. The grading scheme discussed in this paper was applied to the five in-class exams administered to the mastery students throughout the semester.

The class day at the Naval Academy consists of six, fixed, fifty-minute periods with ten minutes allotted for travel between classes. Students in the period-and-a-half sections, therefore, cannot have another class scheduled for the thirty-five minutes between the end of their regular Statics class and the beginning of the next period. This particular schedule offered a built-in window of opportunity that could be used for the subsequent mastery demonstration attempts without significant impact to the students' schedules. It also offered the opportunity to use the full double-period for the administration of exams, minimizing the effect of time constraints on exam performance.

Borrowing from Sangelkar's implementation, the grading scheme shown in Table 1 was used for the five in-class exams administered to the mastery sections. The exam first administered to all mastery students for each set of concepts was assessed problem-by-problem as either "A" level work (correct or almost correct) or "try again" (retake or "R"). Students only re-took problems (concepts) for which they had not demonstrated the requisite level of competency. Subsequent problems were drawn from the same concept area but of a lower degree-of-difficulty, and were worth fewer points.

The ideal timeline for an exam cycle was to return the graded first attempts during the next meeting of the class and begin their subsequent attempts the class meeting after that. This allowed students to re-engage the material and seek extra instruction (EI) if necessary before testing again. On at least one occasion, external schedule imperatives dictated that the first re-takes for an exam were to be administered the very next class day after the initial exam. When this happened, students were notified via e-mail of their requirement to re-do a problem (or more) and which concept area was to be re-addressed. This enabled them to seek extra instruction to correct their deficiencies in the concept area(s) despite the shortened timeline. Second attempts (or first re-takes) were administered in the immediate post-class window mentioned above. Frequently, students were able to leave those sessions knowing if their work that day met the standard or if they would need a third attempt. Between a first and second attempt, students were left to their own devices to determine how to remediate their understanding. They were welcome to seek extra instruction, but not required to do so. Between the second and third attempt, however, they were required to discuss their particular issues with the instructor before continuing the exam cycle. Institutional administrative constraints prevented dis-enrolling students from the course for not demonstrating required command of the material.

Requiring extra instruction between the second and third attempts at a concept was one of two significant changes made during the semester this scheme was used. The other was a broadening of the retake criteria. After the first exam cycle, the "Almost Correct" score window was widened to include the high B, making its floor 88% – still higher than possible on the second attempt, but alleviating some of the grading burden (at a school with no graduate teaching assistants).

Table 1: Exam scoring scheme (initial)

Assessment		Attempt at Problem		
		1st	2nd	Final
Correct	Correctly solved problem with no mistakes	100%	85%	70%
Almost Correct	Satisfactorily solved problem with a minor error	90%	75%	60%
Retake	Student did not show ability to solve the problem correctly	0	0	0
Not Attempted	Either no attempt or essentially no attempt at a solution	-15%	-10%	-5%

Results

The enrollment in EM211 in the fall of 2015 was 241 students; the data below includes 227 students. Those excluded from the data were any who were required to take an alternate final exam (because of scheduling conflicts) and those who were not due-course students. Due-course students are sophomore engineering majors; the non-due-course students were upperclassmen who took statics that semester as an elective and a few international exchange students. Of the 227 students included in the final population, 47 were in the mastery sections and 180 were in the traditional sections.

Quantitative

The three main predictors of Midshipman academic performance in Statics have historically been their second semester freshman year quality point rating (QPR), their grade in Foundations of Chemistry I, and their grades in Calculus I.[3,4] In addition to QPR having a stronger statistical correlation to Statics grades, their grades in the Chemistry and Calculus are one of four discrete values (A, B, C, or D), making the previous semester QPR a much richer data set with which to work. Figure 1 shows the cumulative distribution function of previous semester QPR for the two populations under consideration: mastery and traditional. The closeness of the two plots illustrates that the two populations were comparable at the beginning of the semester; neither had a marked advantage or disadvantage. Figure 2 shows the cumulative distribution function of the final exam scores for the two populations. Again, the two groups are comparable.

The correlation coefficients between QPR and exam scores are 0.530 for the mastery population and 0.612 for the traditional population. Thus, there is insufficient evidence to say the data are not correlated.[5] Figure 3 shows the raw data points for both populations as well as a linear approximation of Statics final exam score versus previous semester QPR for both populations.

The data show that students who entered the mastery sections with lower QPRs scored better on the final exam than their classmates in the traditional sections with similar entering conditions. The break even point appears to be at an entering QPR of 3.6. That is, students who had stronger previous academic performance benefited less from the mastery grading scheme; their grades did not suffer because of it either.

Qualitative

The initial implementation of this mastery grading scheme did require more time and effort: generating three problems of appropriate difficulty for each concept area was the largest factor, but the logistics of additional testing and grading cannot be discounted. On the flip side, grading of the first and second attempts at problems was dramatically streamlined once the "partial credit" habits were broken by the grader. That is, development of an "A" versus "R" attitude significantly sped up the process. A secondary benefit of that attitude manifested in the development of future exam problems: "automatic R" criteria were identified for problems in the creation process.

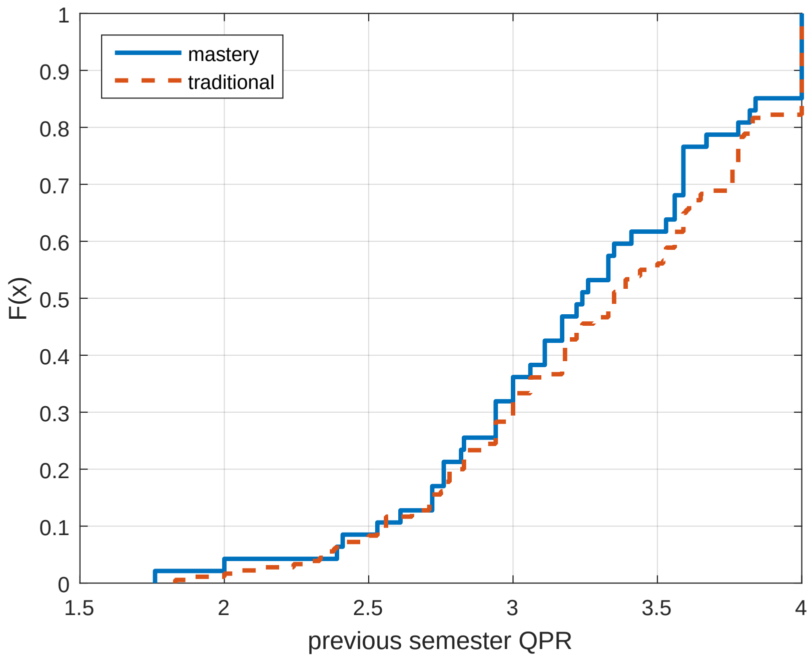

Figure 1: Cumulative Distribution Function of previous semester QPR

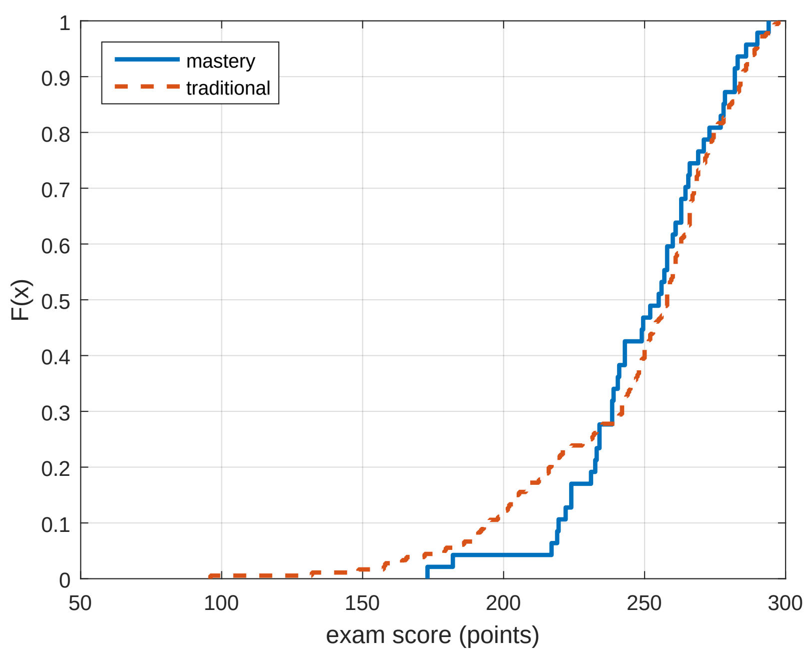

Figure 2: Cumulative Distribution Function of final exam score

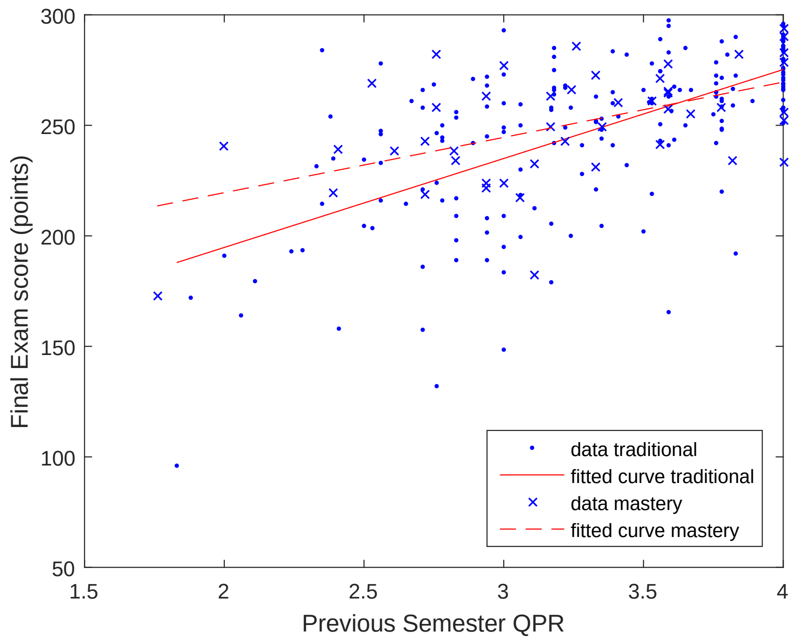

Figure 3: Final exam score (out of 300) vs. Previous semester's QPR (max 4.0) raw data plus linear approximation

As part of the standard end-of-semester student opinion forms, mastery students were asked to specifically address the mastery grading scheme. These surveys were administered anonymously in class during the last week of the semester, more than a week before the final exam. In general, they intellectually appreciated the concept, but did not necessarily care for the logistics in the throes of the semester. Select comments follow:

> "The retake tests are great if you don't know how to solve the problem, but they add a lot of stress if you know how to work the problem but just mess up while doing so."

> "I loved the testing style, it really helps you understand the material because you can go back and re-learn the material multiple times and although at times frustrating retaking exams, it helps learning the content completely!"

> "At first I didn't like the test taking set up, but then it really grew on me. I think it is a great way to evaluate students learning and it really helped having to retake certain problems. It made me focus and learn where I otherwise wouldn't have gone over it if I was just given partial [credit]."

> "The mastery concept of learning the material is stressful, but it does assist in helping to learn the material. I personally like to take a test and be done with it, but the retakes cause the student to continually worry about the grade. However, in the concept I had to retake, I found myself learning the material better."

> "I really liked the test system because it took the focus off the stress of having a test,

and instead put the focus on thoroughly understanding the material."

'The master learning test method forced me to understand the material."

'While I like how the test retakes makes sure that you learn the material, it draws out a test so much. Instead of taking it just for one day, I felt like I was taking it for a whole week. A week long test is really stressful. Maybe give an option that you don't have to retake if you don't want to, so you would have the choice of either retaking the problem or accepting the grade that you received on the problem if you didn't correct it."

"I believe the mastery grading is helpful. I liked the retakes as they would force you to understand the material before moving on."

"Mastery grading really helped as the material in this course is dependent upon prior chapters."

"The mastery grading system was good because it made sure that the students knew the material. Even though it took extra time for re-takes, it was well worth the time."

"For test [regrades], if possible I believe that students should be given the option whether or not to retake. For instance, if a retake conflicts with studying for another test or creates an otherwise greater burden for the student, the student should have the option to decline if they are fine with an 85 the first time and want to focus on their next test."

The suggestion made in the final comment shown (give the student the option to retake or accept a partial credit grade) merits consideration. Such an option is quite tempting, but it would return the focus to the numerical grade as opposed to some demonstration of understanding. This is a highly competitive institution, and academic grades are the single most significant factor in our graduates eventual choices of military specialty. That combination tends to distract Midshipmen from their education.

While no quantitative analysis of student performance on the mastery graded exams has been conducted, casual observation indicates that eventual exam grades ended up in the same general zone as they would have had the initial exams been graded using partial credit.

Conclusion

This paper described an adapted mastery grading scheme implemented for three of fourteen sections of Statics (EM211) at the United States Naval Academy in the fall semester of 2015. The logistics of an extended testing regime were cumbersome, the data appears to show that students who entered the semester with lower predictors of success performed better than expected on the common final exam after being forced to demonstrate competency throughout the semester instead of taking their (grading) lumps and moving on.

Note: a previous version of this paper was presented at a 2016 ASEE Section Conference.

References

[1] Benjamin S Bloom. Learning for mastery. *UCLA CSEIP Evaluation Comment*, 1(2):1–12, 1968.

[2] S. Sangelkar, O. M. Ashour, R. L. Warley, and O. Onipede. Mastery learning in engineering: A case study in statics. *121st ASEE Annual Conference & Exposition*, (2010):1–19, 2014.

[3] John Burkhardt. The effectiveness of additional class contact time on student performance in statics. *International Journal of Mechanical Engineering Education*, 41(2):169–177, 2013. doi: 10.7227/IJMEE.41.2.8.

[4] John Burkhardt. The effect of additional statics class time on at-risk student performance. In *2015 ASEE Annual Conference and Exposition*, 2015. doi: 10.18260/p.24859.

[5] Anthony J. Wheeler and Ahmad R. Ganji. *Introduction to Engineering Experimentation*. Prentice Hall, 3rd edition, 2010.