The Effects of Second-Chance Testing on Learning Outcomes in a First-Year STEM Course in Engineering

Prof. Christopher D. Schmitz, University of Illinois at Urbana-Champaign

Christopher D. Schmitz is a Teaching Associate Professor, Chief Undergraduate Adviser and an Education Innovation Fellow (EIF) in the Grainger College of Engineering at the University of Illinois Urbana-Champaign. An Electrical Engineer, his research interests include algorithmic fault-tolerant adaptive systems, software defined radio, multi-user cellular communication, electrically-small devices, and pedagogies of teaching and learning. An amateur beekeeper, he mentors a project for applying engineering solutions to the plight of honeybees.

Dr. Geoffrey L. Herman, University of Illinois at Urbana-Champaign

Dr. Geoffrey L. Herman is a teaching associate professor with the Department of Computer Science at the University of Illinois at Urbana-Champaign. He also has a courtesy appointment as a research assistant professor with the Department of Curriculum & Instruction. He earned his Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign as a Mavis Future Faculty Fellow and conducted postdoctoral research with Ruth Streveler in the School of Engineering Education at Purdue University. His research interests include creating systems for sustainable improvement in engineering education, conceptual change and development in engineering students, and change in faculty beliefs about teaching and learning.

Prof. Timothy Bretl, University of Illinois at Urbana-Champaign

Timothy Bretl is an Associate Professor of Aerospace Engineering at the University of Illinois at Urbana-Champaign. He received his B.S. in Engineering and B.A. in Mathematics from Swarthmore College in 1999, and his M.S. in 2000 and Ph.D. in 2005 both in Aeronautics and Astronautics from Stanford University. Subsequently, he was a Postdoctoral Fellow in the Department of Computer Science, also at Stanford University. He has been with the Department of Aerospace Engineering at Illinois since 2006, where he now serves as Associate Head for Undergraduate Programs. He holds an affiliate appointment in the Coordinated Science Laboratory, where he leads a research group that works on a diverse set of projects (http://bretl.csl.illinois.edu/). Dr. Bretl received the National Science Foundation Early Career Development Award in 2010. He has also received numerous awards for undergraduate teaching in the area of dynamics and control, including all three teaching awards given by the College of Engineering at Illinois (the Rose Award for Teaching Excellence, the Everitt Award for Teaching Excellence, and the Collins Award for Innovative Teaching).

The Effects of Second-Chance Testing on Learning Outcomes in a First-Year STEM Course in Engineering

Introduction

Students entering college in an engineering discipline often struggle in their first semesters. While the reasons for their difficulties may vary, it is often the summative course assessments (exams) that serve as the first "wake-up call." Many under-performing students are suddenly convinced to reassess their approach to college-level studies. Increased frequency of assessment has been found to improve retention of the learning objectives [1, 2]. In our first-year engineering course, ECE 110 *Introduction to Electronics* at the University of Illinois, we have decided to quickly leverage this high-level of motivation following each assessment.

In our course, each of three midterm assessments is immediately followed by an opportunity for mastery improvement. Successful completion of that exercise allows for a second-chance exam to be taken. Of course, after applying a weighted average, there is a tightening of the grade distribution for that exam. What is interesting, however, is its effect on long-term retention of the material. Specifically, we have observed that although the students who pursue second-chance testing perform, on average, two letter grades worse than the rest of the student population at the time of the midterm, they ultimately perform within a fraction of a letter grade by the assessment of the cumulative final examination.

In our study, we will look at the mode of mastery improvement exercises and methods for attracting students to these activities. We will discuss our methodology in second-chance testing while providing the rationale for our choices. We hope that this paper will serve as a practical guide to implementation of second-chance testing, speeding implementation while addressing typical concerns and avoiding potential flaws. A statistical analysis of the assessment scores demonstrate the afore-mentioned improvement in grades by struggling students and supports the use of second-chance testing within a couple weeks of the first-chance exam.

The first-chance exam

In our first-year electrical engineering course, three 50-minute midterm examinations are administered in the semester, roughly one each month. These exams are offered online and taken in the Computer-Based Testing Facility (CBTF) [3], a computer lab staffed with proctors. These exams use the same software tool, PrairieLearn, as the homework assignments so that no additional training is required. The students have access to a standardized help sheet (available on the computer), calculator, online software, pencil and paper, but no materials are allowed to enter or leave the examination room. In this way, students cannot merely imitate the process used in solutions to similar problems they have seen and details of the exam content are more difficult to extract from the room. In fact, exam problems not only have randomized numerical variables, but they are also likely to have differences in structure or be replaced with an entirely different question altogether that serves to assess the same learning objective. Although the exam period runs over multiple days, evidence that leaked information might lead to improved performance of future exam takers is low [4, 5].

The course is not "curved," that is, it does not distribute grades on exams or in the course according to a predetermined allotment of As, Bs, Cs, Ds, and Fs. Rather, it adheres to the belief that a pedagogy that truly supports mastery learning for all students can lead to a high level of mastery for most students [6]. It is desired that our summative assessments are constructed in a manner that the grades themselves will reflect the students' mastery of the learning objectives and that higher levels of collaboration during study can be encouraged.

The second-chance exam

The second-chance exam varies in two aspects from the first-try exam. Of the nine questions, typically two of them will pivot upon one learning objective or shift to an entirely different learning objective within the material to be assessed. Also, in the second-chance exam, the help sheet is removed as a crutch. Students are asked to depend upon memory through practice for the recollection of basic circuit formulae. The remaining seven questions continue the same randomization procedures as in the first-try exam. In the fall of 2019, each midterm exam was followed by a second-chance exam roughly one week separated and spread across the 15-week semester.

The formula combining the first-chance and second-chance exam does not penalize students who do worse on the second-chance exam by setting the first-chance exam score as a minimum adjusted score. We desire to encourage students who have performed below their potential. At the same time, the formula allows for the largest incremental increase to students who most-desperately desire the bump in grade. Students who perform well on the first-chance exam do not feel as compelled to increase their effort through a mastery-building assignment for a more-modest increase in their exam score.

The midterm exam adjustment formula (presented to the students in tabular form) is an important tool in recruiting the target population to the second-chance exam. The adjustment formula provides a weighted average that progresses from heavy weight on the second-chance exam for students who did most poorly on the first-chance exam to a weighted average that weighs the first-chance exam more heavily for students who did well on the first-chance exam. The formula also utilizes a "max" function to prevent penalty for a second-chance attempt.

$$E_{adj} = MAX(E_{FC}, (1-\alpha)\mathrm{E_{FC}} + \alpha E_{SC}), \text{where } \alpha = \frac{1-E_{FC}}{1-72}$$

Here, E_{FC} is the first-chance exam score, E_{SC} is the second-chance exam score, and E_{adj} is the adjusted exam score based on the two offerings. Students with a first-chance exam score below 72 will receive an adjusted exam score that weights the second-chance score more heavily. The value of 72 was selected to motivate retry attempts by the target group and is not altered during the semester. The mapping of this formula is demonstrated in Table 1 which also highlights that there is no penalty for a poor grade on the second-chance exam and that students who do well on the first-chance exam still retain the potential to obtain the highest adjusted score.

The use of a weighting formula might be interpreted as a form of "curving" the class. However, that interpretation can be refuted. The availability of a retry exam for each of three midterm exams offered through a 15-week course equates to an assessment every 2.5 weeks for the

students so-inclined to take them. The improvement is *earned by the student* only after accepting extra work while utilizing feedback gathered from review of the first-chance exam, the mastery exercise, office hours, and/or the peer-led team learning. Our allowance of more-frequent assessment with this form of weighting function blurs the line between lower-grade-cost formative assessment and higher-grade-cost summative assessment without an a priori expectation of grade distributions.

A typical argument for allowing the second-chance score to reduce the adjusted exam grade is that students will take the exam just *hoping* for an easier version. However, we believe the mastery exercise already helps to avoid this issue. Many students not serious about *improving* their understanding of the material will also be turned away by the requirement of an additional exercise. We have also heard concern that students may neglect preparing for the first-chance exam knowing that a second-chance exam is present. However, the weighted grade used provides the opportunity of the best adjusted score to the students who do the best on the first-chance exam, thus there remains substantial motivation for preparing well for the first-chance exam. In addition, randomized exams that uses a small pool of questions have been shown to typically provide a fair assessment despite some information leak [5].

E_{adj}		E_{SC}																				
		0	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**	**100**
E_{FC}	**0**	0	3.91	7.81	11.7	15.6	19.5	23.4	27.3	31.3	35.2	39.1	43	46.9	50.8	54.7	58.6	62.5	66.4	70.3	74.2	78.1
	5	5	5	8.86	12.7	16.6	20.4	24.3	28.2	32	35.9	39.8	43.6	47.5	51.3	55.2	59.1	62.9	66.8	70.7	74.5	78.4
	10	10	10	10	13.8	17.6	21.4	25.3	29.1	32.9	36.7	40.5	44.3	48.1	51.9	55.8	59.6	63.4	67.2	71	74.8	78.6
	15	15	15	15	15	18.8	22.5	26.3	30	33.8	37.6	41.3	45.1	48.8	52.6	56.4	60.1	63.9	67.7	71.4	75.2	78.9
	20	20	20	20	20	20	23.7	27.4	31.1	34.8	38.5	42.2	45.9	49.6	53.3	57	60.7	64.4	68.1	71.9	75.6	79.3
	25	25	25	25	25	25	25	28.6	32.3	35.9	39.6	43.2	46.8	50.5	54.1	57.8	61.4	65	68.7	72.3	76	79.6
	30	30	30	30	30	30	30	30	33.6	37.1	40.7	44.3	47.9	51.4	55	58.6	62.1	65.7	69.3	72.9	76.4	80
	35	35	35	35	35	35	35	35	35	38.5	42	45.5	49	52.5	56	59.5	63	66.5	69.9	73.4	76.9	80.4
	40	40	40	40	40	40	40	40	40	40	43.4	46.8	50.2	53.6	57	60.5	63.9	67.3	70.7	74.1	77.5	80.9
	45	45	45	45	45	45	45	45	45	45	45	48.3	51.6	54.9	58.3	61.6	64.9	68.2	71.5	74.8	78.1	81.4
	50	50	50	50	50	50	50	50	50	50	50	50	53.2	56.4	59.6	62.8	66	69.2	72.4	75.6	78.8	82.1
	55	55	55	55	55	55	55	55	55	55	55	55	55	58.1	61.2	64.2	67.3	70.4	73.5	76.6	79.7	82.7
	60	60	60	60	60	60	60	60	60	60	60	60	60	60	62.9	65.9	68.8	71.8	74.7	77.6	80.6	83.5
	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	67.8	70.6	73.3	76.1	78.9	81.7	84.4
	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	72.6	75.2	77.8	80.3	82.9	85.5
	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	77.4	79.7	82.1	84.4	86.8
	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	82.1	84.2	86.3	88.3
	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	86.7	88.5	90.2
	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	91.3	92.6
	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95.8
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table 1. E_{adj} as calculated using the midterm exam adjustment formula for various values of E_{FC} and E_{SC}.

The mastery exercise

The ultimate goal for second-chance testing is to improve student mastery of the course learning objectives in a way that carries through the entire semester. We provide an optional formative

assessment in the form of an additional assignment covering some of those same learning objectives. This assignment is offered through PrairieLearn, the same online tool as the homework and exams, and contains fewer overall questions than a typical homework (the average length of an assignment is 15 questions while the mastery exercise contains 10 questions).

The details of the mastery-improvement exercise has evolved across several years of offering second-chance testing. Initially, no mastery exercise was required and, indeed, many students did appear to approach the exam with little-to-no additional training on the material. The adoption of an additional (optional) assignment was productive in this area, but, to encourage second-chance testing, a modest grade of at least 60% was required. Instead of appreciation for a low bar to take the second-chance opportunity, several students actually ignored the requirement and took the second-chance exam only to have it refused by the gradebook. The students complained that such a "light" requirement should just be waived by the instructor. This approach was quickly modified. Currently, the assignment, while still only 10 questions in all, is fairly demanding and requires a grade of at least 80% before the second-chance exam will be accepted. The complaints by students have significantly dropped and the time spent on the exercises increased.

As a further incentive to complete the mastery exercises, the points accumulated are allowed to be applied to lost homework points lost in the past, present, or future assignments. That is, the homework points are calculated by adding the points earned from the mastery exercises to the points earned from the homework assignments, but capped at 100%. Since the homework assignments already have unlimited attempts until the closing ("due") date each week, it is often the students who are struggling with the material that may see the most benefit in completing the mastery exercise and also earning the right to take the second-chance exam. In fall of 2019, the three mastery exercises could replace up to 30 points lost from the semester total of 200 homework points. This was also greatly appreciated by students who joined the course late, missing one or two assignments.

More sophisticated mastery exercises might be provided through "levels" of problems, each level requiring some measure of mastery to be shown prior to advancement and video solutions served as correctives [7], but the method used here was both convenient to implement and robust to issues of student frustration.

Other support for mastery improvement is available to our students. The non-competitive nature of the course has led to students regularly seeking and finding strong help from peers in addition to feedback in the large number of office hours available. Following each of the first two exams (first-chance on exam 1 and exam 2), an email is sent targeting those students with strong attendance and homework grades who performed below an 80% on the recent exam. The email recognizes that the student is working diligently, but likely not satisfied with the exam grade. The email goes on to invite the student to a study group with the instructor where students can gain corrective feedback and then participate in peer-led team learning activities [8] to enhance comprehension and improve self-efficacy. After allowing these students an opportunity to join, the rest of the class is also invited to these sessions. Additional sessions may be added until we feel all interested students have been served.

The final exam

The final exam in the course is a 3-hour cumulative examination. In the fall of 2019, it consisted of 36 questions, most numeric short-answer but several multiple-choice "concept" questions as well. The final exam carries as much as 35% of the course grade (nominally 25%, but replacing the lowest midterm grade, which carries 10% of the course grade, when appropriate) and is designed to thoroughly cover key learning objectives. We will leverage the final exam as our best available assessment of student mastery of the course learning objectives. There is no second-chance assessment for the final examination.

Questions to address

Our course staff was introduced to second-chance testing in 2015 as we were just beginning to adopt a new software, PrairieLearn, for online homework assignments. PrairieLearn plus the CBTF provided, what seemed to us, a low-cost method for offering a second-chance opportunity on our written exams. Like many of our readers, we had several questions as we approached the use of second-chance testing. It would be several semesters before we could both formulate what mattered most and also fairly evaluate what we were witnessing. Three key questions arose...

Which students would take second-chance exams? While we would make them available to all students, we wish to attract the student pool that most needed extra effort to master the learning objectives. Students who have already mastered the material could find second-chance testing to be a significant distraction from other courses or worthwhile activities. Yet, weaker students should not delay in their efforts to learn even while the course continues to move onto new learning objectives.

Does the second-chance exercise actually improve mastery of the learning objectives? Granted a true assessment of mastery is difficult to attain, can we show that the improvements attributed to offering a second-chance exam were significant with our assessment? If there is no penalty for second-chance testing (your exam grade cannot worsen), then isn't it just a self-fulfilling prophecy?

And, most importantly, does increased mastery demonstrated at the time of a second-chance exam also map to retention of that material at the semester's end?

Many variations of assessment can be done for this course based on the ***three*** first-chance midterms and their corresponding second-chance exams. This is further complicated by the fact that students could have elected to take any combination of these second-chance exams. After some initial studies of the data, we choose to focus specifically on the performance of students who elected to take the second-chance opportunity of midterm ***exam 2*** and, by extension, those who specifically did not.

Focus on exam 2 (the second midterm) makes a good study for several reasons. Material from the second midterm's learning objectives were regularly cited by students as the most-challenging of all learning objectives during the entire semester. Also, learning objectives associated with exam 2 are highly central to the course and, due to the cumulative nature of the

final examination, are assessed at some level by nearly 50% of the final exam. Statistically, second-chance exam 2 was the most-selected second-chance opportunity with 208 of 404 students (51%) electing to take it. Of the 139 students who took second-chance exam 1, 96 (69%) also took second-chance exam 2. Of the 118 students who took second-chance exam 3, 78 (66%) also took second-chance exam 2. The large population of students who took second-chance exam 2 also captures many students who took at least one of the other second-chance exams. Therefore exam 2 and second-chance exam 2 would be highly representative of the benefits of second-chance testing to the students of this course.

Which students would take second-chance exams?

We desired to attract the students who needed additional mastery exercises and not the students who were already performing at a higher level who might have better academic outcomes placing their efforts into other courses or extracurricular activities. To see if we were successful, let's examine the histogram of scores from the first-chance exam 2 shown in Figure 1. Shown are the number of students who earned grades in each 5-percent bin from 0 to 100. The histogram is typical of many of the midterm exams we have seen in this course, showing a distribution that is tell-tale of the broad diversity in student backgrounds, approximately half of which come in with some significant background in the topic and half of which finds the material to be new and sometimes very unfamiliar. It also carries effects of many of the issues that first-year college students encounter such as poor time management. Many of these students are horrified to discover that the approach they took to succeeding in their high school courses does not translate to success on college exams.

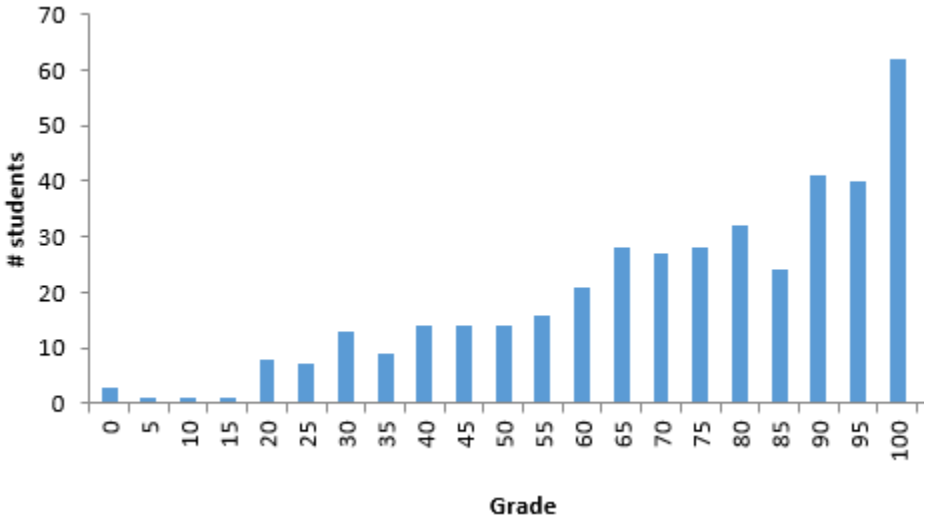

Figure 1: Number of students at each score for first-chance exam 2.

Figure 2 emphasizes the difference between the students who, soon after, elected to take the second-chance exam 2 (orange) and those who elected not to take it (blue). It is not surprising that students who scored in the 90's were not interested in the mastery exercise or the second-chance exam. The majority of students who chose to take the second-chance exam consists of those who scored at and below 80% (C, D, and F grades) and, especially, those in the long tail on this first-chance assessment. While the mean grade on the first-chance exam 2 for all 404

students was 70.9%, the mean grade (standard deviation = 23.9) of those who later *elected to take second-chance exam 2 (N=208)* was only 62.2% (standard deviation = 18.9) which is nearly a full letter grade lower than the class average. Even more significantly, this mean grade of those who elected later to take second-chance exam 2 was nearly *two letter grades lower* than the 79.0% mean grade of those who elected *not* to take the second-chance (N = 196, standard deviation = 25.8). Using an unequal-variance two-tailed t-test, we found $p < .00001$ (Cohen's d = 0.743). This statistic will be valuable when we later look at the final exam statistics.

It is surprising that there is a significant, almost uniformly-distributed, set of students with grades ranging from 20 to 90 that also chose not to pursue second-chance testing, but this might be partially attributed to disengaged students who remained unaware or uninterested in the second-chance opportunity. Otherwise, the students attracted to the second-chance exam are, indeed, representative of the target pool. Hitting the target pool can be important in both resource allocation for the instructor as well as encouraging high-achieving students to tackle more important pursuits in this course, other courses, or valuable extra-curricular activities.

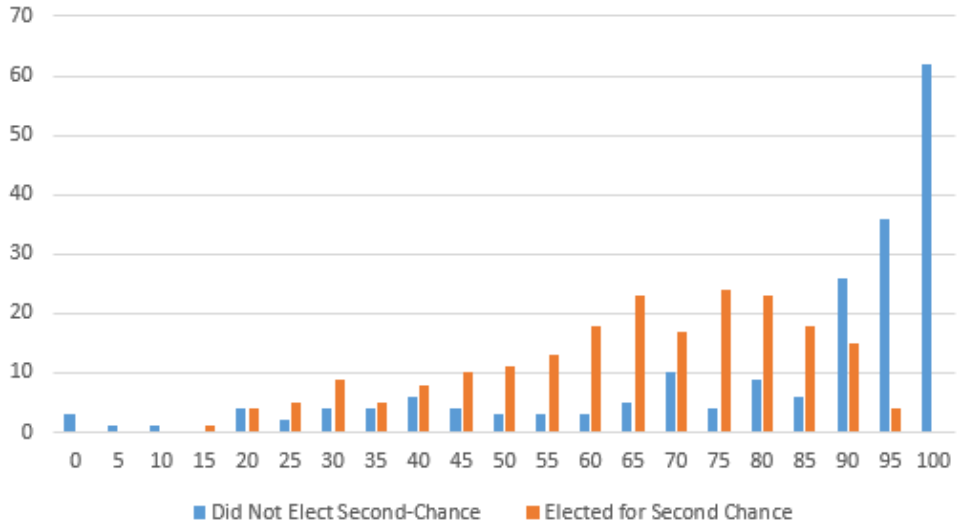

Figure 2: Demonstrating the discrepancy of the number of students at each score for first-chance exam 2: those who did not later elect to take second-chance exam 2 and those who did.

Second-chance testing's effect on students' mastery on an individual exam

We wanted to know that improved mastery was significant. Figure 3 shows the histogram of grades for the 208 students on both their first-chance and on their second-chance exams. Note the significant shift in the distribution that includes a "shorter" tail on the grade distribution (second-chance in blue). The mean grade on the second-chance exam was 73.8% (up from 62.2% on the first-chance exam and higher than the 70.9% class average from the first-chance exam).

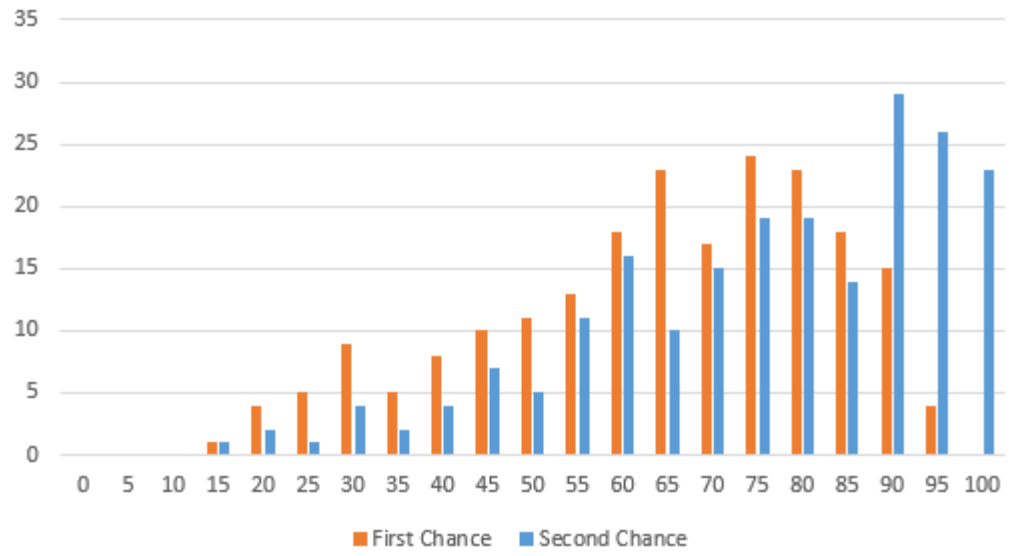

Figure 3: The improvement of the grade distribution between the first-chance (orange) and the second-chance (blue) exams for 208 students.

Once the weighting formula was implemented (distribution of adjusted scores is not shown in this paper), the mean grade on the adjusted midterm 2 for students who elected the retry exam became 71%. This is the same as the overall class average (based on the entire 404 students) on the first-try exam, although still close to one letter grade below the 79% mean of those 196 students who elected not to take the second-chance exam. But, we will not provide any additional focus on the weighted grades as any instructor might choose to weight the first-chance and second-chance in any number of ways, thus altering the distributions in significant ways. It is enough to note that *at this point in the semester*, the mastery and second-chance opportunity *appear* to have closed the gap on mastery of the learning objectives between the two groups of students and that the choice in our midterm exam adjustment formula still provides an incentive for doing well on the first-chance exam.

An examination of the improvement in raw scores of the second-chance exam 2 over the first-chance exam 2 is shown in the histogram of Figure 4. The individual students overwhelmingly did better on the second-chance exam with only 23% doing the same or worse.

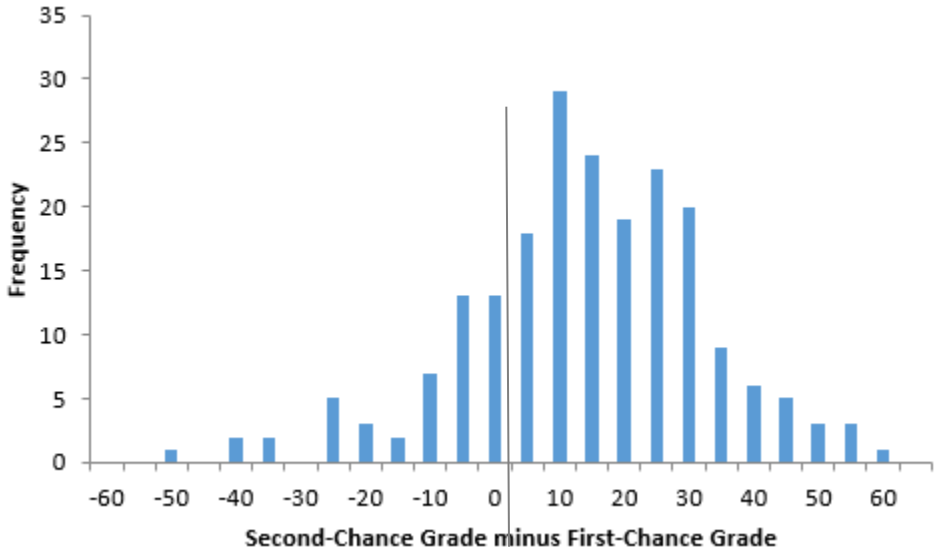

Figure 4: The difference between the raw score of the second-chance exam 2 and the original raw score of the first-chance exam 2 for the students who took both exams.

Second-chance testing's effect on students' end-of-term mastery

Finally, we address the critical question: Does the apparent increased mastery demonstrated at the time of the second-chance exam also map to *retention* of the course's learning objectives *at the semester's end*? Since the final exam is cumulative, we can compare student performance on the final exam for students who elected to take the second-chance exam 2 to the performance of the entire class and, more specifically, to those who did not take the second-chance exam 2. Although, on the final exam, some questions spanned learning objectives from multiple midterm topics, approximately 50% of the final exam questions have content that maps well to exam 2 learning objectives. Recall that not only did students generally feel that the most-challenging learning objectives belonged to exam 2, about 2/3 of the students who elected to take second-chance exam 1 and 2/3 of the students who elected to take second-chance exam 3 also took second-chance exam 2. Therefore, we also believe that statistical observations concerning the second-chance exam 2 and the final exam will be indicative of the statistical observations of second-chance testing for our course in general.

In Figure 5, we can compare the histogram of final exam grades for the entire course of 404 students to the histogram of the 208 students who elected to take second-chance exam 2. The mean grade on the final exam for the entire class was 82.4%. The mean grade on the final exam for the students who (earlier in the semester) took second-chance exam 2 was a numerically-similar 79.6%.

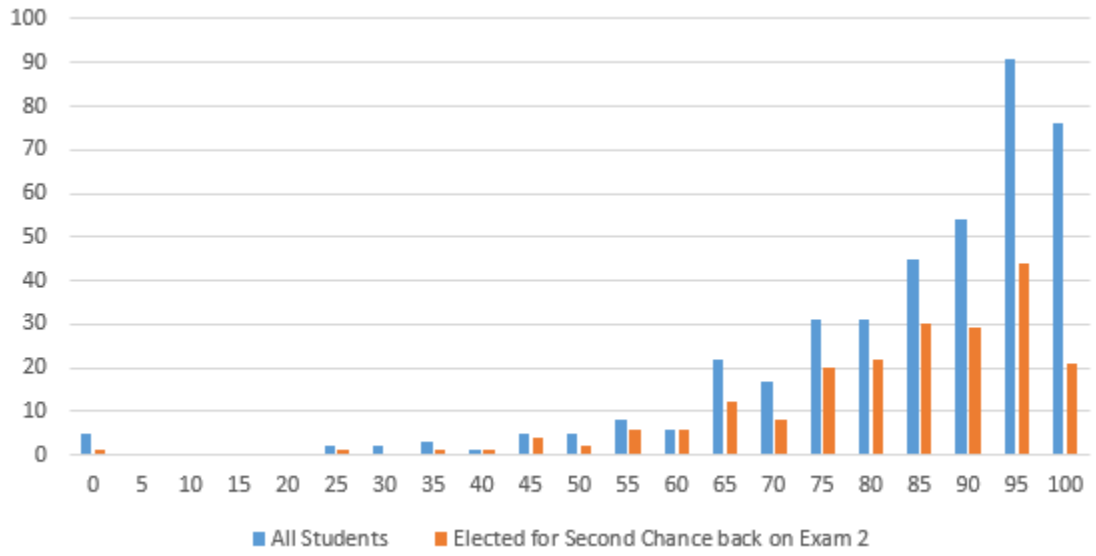

Figure 5: The histogram of the final exam scores for the entire course plus the histogram of final exam scores for students who earlier in the semester took second-chance exam 2.

For clarity, in Figure 6, we also plot the histogram for the final exam grades for the 208 students who did take second-chance exam 2 with the histogram for the final exam grades for the 196 students who did not. The students who elected to take second-chance exam 2 still account for the majority of grades in the tail of the histogram, but have largely normalized their performance with the remainder of the course by performing at an 79.6% average (standard deviation = 15.5) as compared to the 83.3% average (standard deviation = 19.4) for those who did not take the second-chance exam 2. To validate a null hypothesis that the two groups performed nearly equally, we performed an unequal-variance two-tailed t-test and found p=0.037 (Cohen's d = 0.211). The final exam performance is still significantly different, but the size of the effect was dramatically reduced from almost two full letter grades to less than half a letter grade. These findings suggest that while the groups are still different, the performance gap was closed.

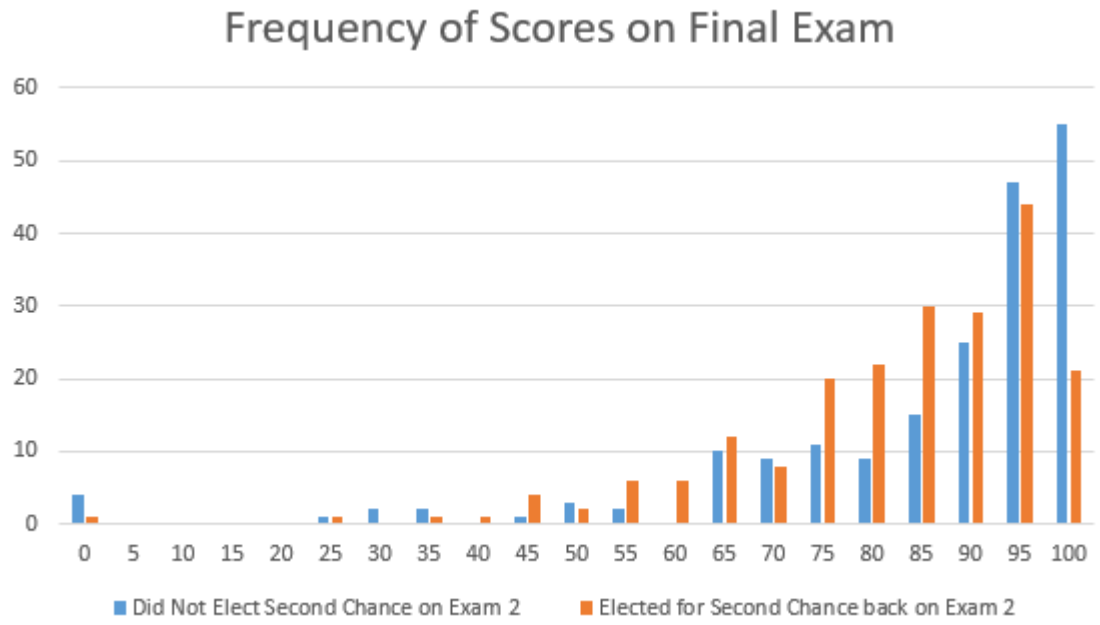

Figure 6: The histogram of the final exam scores for those who did and did not take the second-chance exam 2.

Conclusions

In this study, we examined the early performance of students in a first-year course in engineering on the exam assessments performed throughout the semester while providing suggestions and rationale on our methodology. We limited some of the detailed analysis of our study to the most-crucial midterm of the semester, the second of three midterm examinations. We found that a student population accounting for half of the class who initially performed with a mean grade on a first-chance exam of 62.2%, which is 17.8% or nearly *two letter grades* below the mean of the rest of the population of students later assessed on the final exam with a mean grade of 79.6%, only 3.7 percent or *less than a numerical half-letter grade* below the 83.3% mean of the rest of the students. We attribute this performance improvement to the motivation provided by a weak summative assessment followed by an opportunity for mastery improvement exercises rewarded by exam adjustments following a second-chance examination.

References

1. R. L. Bangert-Drowns, J. A. Kulik, and C.-L. C. Kulik, “Effects of frequent classroom testing,” Journal of Educational Research, vol. 85, 1991.
2. Herman, G. L., Varghese, K., and Zilles, C., Second-chance testing course policies and student behavior, In Proceedings of the 49th ASEE/IEEE Frontiers in Education Conference, Cincinnati, OH, Oct. 16-19, 2019.
3. C. Zilles, M. West, D. Mussulman, and T. Bretl, “Making Testing Less Trying: Lessons Learned from Operating a Computer-Based Testing Facility,” in Frontiers in Engineering (FIE), 2018.
4. B. Chen, M. West, and C. Zilles, “Do performance trends suggest wide-spread collaborative cheating on asynchronous exams?,” in Proceedings of the Fourth ACM Conference on Learning at Scale, 2017.
5. B. Chen, M. West, and C. Zilles, “How much randomization is needed to deter collaborative cheating on asynchronous exams?”, in Proceedings of the Fourth ACM Conference on Learning at Scale, 2018.

6. B. Bloom, “Learning for mastery,” Evaluation Comment, vol. 1, no. 2, pp. 1–12, 1968.
7. B. Gutmann, G. Gladding, M. Lundsgaard, and T. Stelzer, “Mastery-style homework exercises in introductory physics courses: Implementation matters,” Phys. Rev. Phys. Educ. Res., vol. 14, p. 010128, May 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.14.010128
8. Gosser, D. K., Cracolice, M. S., Kampmeier, J. A., Roth, V., Strozak, V. S., and Varma-Nelson, P., Peer-Led Team Learning: A Guidebook, Upper Saddle River, N.J.: Prentice Hall, 2001.