Want to Reduce Guessing and Cheating While Making Students Happier? Give More Exams

James T. Laverty, Wolfgang Bauer, Gerd Kortemeyer, and Gary Westfall, Michigan State University, East Lansing, MI

It is almost universally agreed that more frequent formative assessment (homework, clicker questions, practice tests, etc.) leads to better student performance and generally better course evaluations.[1] There is, however, only anecdotal evidence that the same would be true for more frequent summative assessment (exams). There may be many arguments against giving more exams, including the general "pain" associated with examinations, as well as reduced teaching time, since classroom sessions are dedicated to exams rather than lecturing. We present evidence that increasing the number of exams in fact does lead to better learning success, less cheating and guessing on homework, and better student course evaluations.

Introducing additional midterms

Our first evidence of more exams being more highly favored by students was inadvertently gathered in an introductory physics course for mostly premedical students: as the instructor changed between two semesters, the number of exams was reduced from six midterms to only three midterms.Written course evaluation from a large percentage of the students indicated that they liked the old exam system better, and that in fact the new system with fewer exams was "unfair." The data presented in this study come from three other classes, in two different calculus-based courses, taught over two semesters: two E&M courses, one with the old system (two midterms, 294 students) and one with the new system (13 weekly midterms, 442 students), and one mechanics course with the new system (13 weekly midterms, 464 students). The additional exams replaced the exam-retake mechanism, which had been used for several years;[2] in the old system, after each of the two midterms, students were allowed to retake the exam again online for partial credit, while in the new system with weekly exams no such opportunity was provided. Both in the old and the new system, midterm exams were taking up one previously scheduled and announced class period (50 minutes) each. In character, these midterms were like homework questions, but required transfer to new contexts. Some of the final exam problems in turn were similar to problems appearing on the midterms. All exams were administered using multiple-choice bubble sheets in connection with a learning content management system (LON-CAPA[3]) to manage test banks. The system's shared content pool includes over 100,000 homework problems on the introductory physics level, including 1500 problems specifically coded for our textbook.[4] With some experience, a complete exam can be produced in our system with only about two hours of work from the instructor. Exams are automatically randomizing, so every student has a different version of the problems: different numbers, options, graphs, etc. Thus, students cannot copy answers from each other. The exams are identified by a cover sheet, which has the student's name and photo. Figure 1 shows an example of a midterm in the new E&M course with identifying information removed.

The same instructor taught the old-system mechanics and half of the new-system E&M course. In both systems, the classes did not include recitation sessions, as it was found that those sessions did not have any significant learning benefit.[5] The final grades were based on multiple categories: homework (30% old E&M, 23% new mechanics, and 25% new E&M), midterms (40% total old E&M, 52% each new mechanics, and 50% new E&M), and final exam (30% old E&M, 25% new mechanics, and new E&M). The classes had varying bonus points for pre-class reading quizzes (up to 5% old system E&M) and clicker participation (up to 5% in all courses). At the end of each of the semesters, the students were given a questionnaire, primarily intended to obtain feedback on the class and the textbook.

Cheating on homework

Formative assessment is only useful if students are truly taking advantage of it to evaluate their learning progress rather than perceiving it as a chore that should be circumvented in the most efficient way. Copying homework answers from other students, be it in person or through online "cheat websites," is detrimental to learning success.[6] Cheating is different from collaborating, where students truly work together to solve problems and teach each other, which in turn can be a signature of desirable learning strategies.[7] In order to encourage collaboration, LON-CAPA[3] has different versions of the same problem for every student, as well as monitored discussion boards attached to the problems. It has been shown that use of these monitored discussion boards (sanctioned collaboration) is positively correlated with success in the course, while the use of external cheat websites is negatively correlated with success in the course.[8]

For our studies reported on here we collected student responses through anonymous questionnaires, which were not traceable to individual students. The questionnaire was filled out by 193 students in the old system E&M course, 186 students in the new system E&M course, and 269 students in the new system mechanics course. Our most notable finding

1 pt

A copper coil has a resistance of 0.561 Ω at room temperature. What is its resistance when it is cooled to -53.9°C?

(*in Ohm*)

1. A ❍ 0.0542 B ❍ 0.0721 C ❍ 0.0960
D ❍ 0.1276 E ❍ 0.1697 F ❍ 0.2257
G ❍ 0.3002 H ❍ 0.3993

1 pt

A rectangular wafer of pure silicon, with resistivity p = 2300. Ω m, measures 7.00 cm by 8.00 cm by 0.019 cm. Find the maximum resistance of this rectangular wafer between any two faces.

(*in Ohm*)

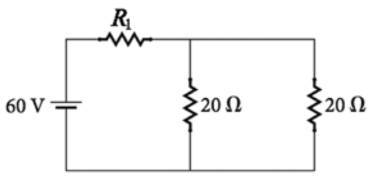
2. A ❍ 5.393×10^6 B ❍ 6.310×10^6
C ❍ 7.383×10^6 D ❍ 8.638×10^6
E ❍ 1.011×10^7 F ❍ 1.182×10^7
G ❍ 1.383×10^7 H ❍ 1.619×107

1 pt

A resistor of unknown resistance and a 33.0-Ω resistor are connected across a 135.0-V emf device in such a way that an 11.0-A current flows. What is the value of the unknown resistance?

(*in Ohm*)

3. A ❍ 4.420 B ❍ 6.409 C ❍ 9.293
D ❍ 13.475 E ❍ 19.539 F ❍ 28.332
G ❍ 41.082 H ❍ 59.569



1 pt

What is the current in the R_1 = 15-Ω resistor in the circuit in the figure?

(*in A*)

4. A ❍ 1.36 B ❍ 1.80 C ❍ 2.40 D ❍ 3.19
E ❍ 4.25 F ❍ 5.65 G ❍ 7.51 H ❍ 9.99

1 pt

Two cylindrical wires of identical length are made of copper and lead. If they carry the same current and have the same potential difference across their length, what is the ratio of their radii (copper/lead)? (See Table 25.1.)

5. A ❍ 1.188×10^{-1} B ❍ 1.581×10^{-1}
C ❍ 2.102×10^{-1} D ❍ 2.796×10^{-1}
E ❍ 3.719×10^{-1} F ❍ 4.946×10^{-1}
G ❍ 6.578×10^{-1} H ❍ 8.749×10^{-1}

1 pt

In an emergency, you need to run a radio that uses 24.9 W of power when attached to a 12.0 V power supply. The only power supply you have access to provides 25.0 kV, but you do have a large number of 30.0 Ω resistors. If you want the power to the radio to be as close as possible to 24.9 W, how many series resistors should you use?

6. A ❍ 43 B ❍ 63 C ❍ 91 D ❍ 132
E ❍ 191 F ❍ 277 G ❍ 401 H ❍ 582

0.5 pt

How much money will a homeowner owe an electric company if he turns on a 100.-W incandescent light bulb and leaves it on 7 months? (Assume that the cost of electricity is $0.120/kW h and that the light bulb lasts that long. Give your answer in dollars.)

7. A ❍ 26.08 B ❍ 34.69 C ❍ 46.14
D ❍ 61.36 E ❍ 81.61 F ❍ 108.54
G ❍ 144.36 H ❍ 192.00

0.5 pt

The same amount of light can be provided by a 26.0-W compact fluorescent light bulb. What would it cost the homeowner to leave one of those on for a year?

8. A ❍ 8.97 B ❍ 13.01 C ❍ 18.86
D ❍ 27.35 E ❍ 39.66 F ❍ 57.50
G ❍ 83.38 H ❍ 120.90

Fig. 1. Example of a midterm in the new E&M course.

is that student self-reported cheating using websites dropped off substantially. Figure 2 shows a histogram of the percent of homework problems students claim they cheated on in each semester. The number of students who cheated on more than 20% of their homework problems dropped to almost zero in the semesters that had 13 exams. The presumed reason behind this is that it doesn't make any sense for students to cheat on the homework due Sunday only to try and learn everything for the test on Monday.

The second surprising finding was that students both used and contributed to the sanctioned discussion boards more frequently. Figure 3 shows that a large majority of the students in the semester with only two midterms used the discussion boards for only 10-20% of their homework. By contrast, students in the classes with 13 exams used the discussion boards much more. In addition to their use of the discussion boards, students also contributed significantly more frequently to the sanctioned course discussions: the E&M class under the old system had an average of 0.015±0.014 posts per homework problem per student, whereas the same class under the new system had an average of 0.026±0.011 posts per homework problem per student; using Welch's t-test, the difference is significant at the $p < 0.0001$ level.

Several other components of the class did not change

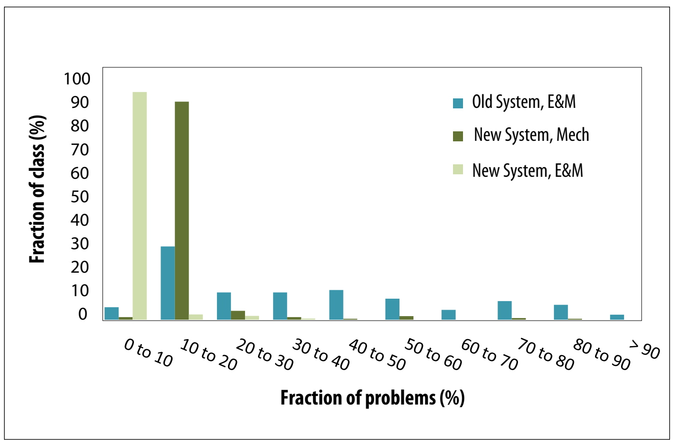


Fig. 2. Self-reported use of "cheating sites."

between the two different exam schedules. For example, the self-reported use of the help room (a collaborative learning center staffed by graduate and undergraduate teaching assistants) and the percentage of students actually doing the reading assignments were very similar between all three classes.

Guessing on homework

Most online homework systems allow for multiple submissions, so students have more than one opportunity to get a

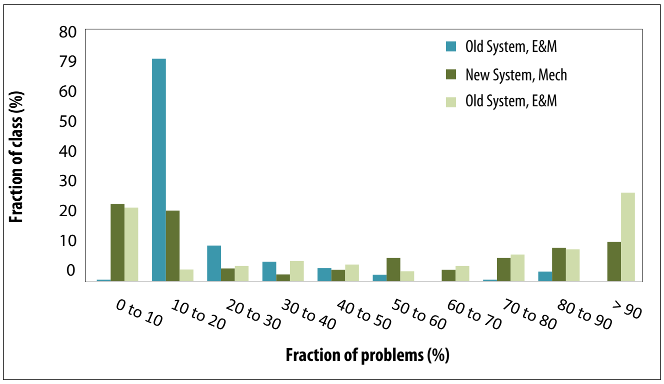


Fig. 3. Self-reported use of sanctioned internal discussion board.

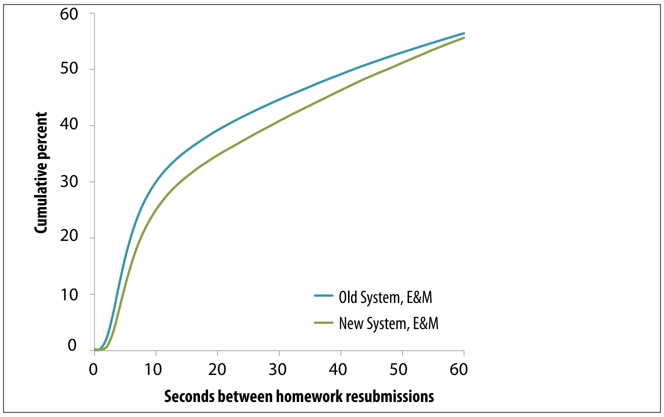


Fig. 4. Cumulative percentage of resubmissions versus seconds between resubmissions.

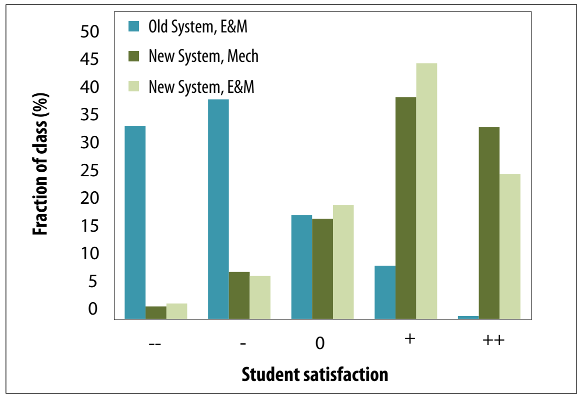


Fig. 5. Student satisfaction with exam schedule

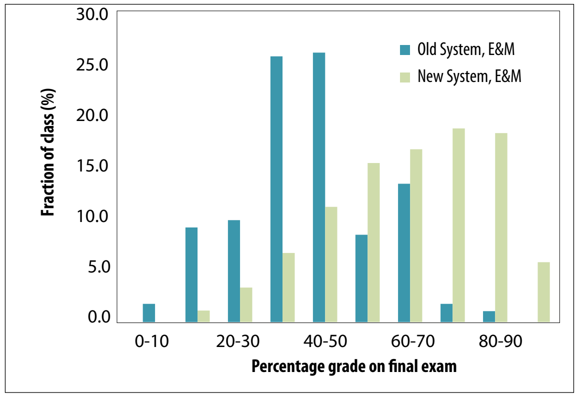


Fig. 6. Final exam grade distribution

problem correct and master the associated concept. Unfortunately, students frequently take advantage of multiple tries by guessing on different answers, randomly changing signs, factors of two, orders of magnitude, etc. A signature of such guessing behavior is a short time interval between two subsequent submissions[5,9] after getting a problem wrong, students submit another attempt within just a few seconds.

Both E&M classes used essentially the same online homework problems, which were mostly numerical in nature. Figure 4 shows the percentage of the total number of resubmissions versus the time between resubmission on the same problem. It turns out that in the old system, students more quickly resubmit after a failed attempt than in the new system. For example, in the old system half of the resubmissions occur within 43 seconds, compared to 48 seconds in the new system. Resubmissions that occur in less than 10 seconds are clearly guesswork; in the old system, 30% of the resubmissions occurred within less than 10 seconds, while in the new system, 25% did so.

The distributions of resubmission times are significantly different up to a 145-second cutoff (approx. 2.5 minutes) on a two-sample Kolmogorov-Smirnov test ($p < 0.05$), while for longer cutoffs, the difference is not significant anymore—the resubmission time distributions differ for guessing answers, but not when it comes to actually reworking the problems. In fact, both distributions have very long tails, and it turns out that the average time between resubmissions for the old course is about 31minutes, while for the new course it is 17 minutes—students in the new course were faster in reworking the problems.

Outcomes at the end of the semester

An obvious concern when increasing the number of exams given in a class is the reception by the students.While they seemed unhappy with the system when they first heard about it, they ultimately were much happier with it than having only two midterm exams. Figure 5 shows a significant increase in student satisfaction with the new exam schedule over the old one, with the number of likes and dislikes basically swapping between the two exam schedules.

Of course, the most important question is whether or not the students learned more with the new system. While we have no direct measure of this, one obvious comparison is the results on the comprehensive final exams. Figure 6 shows the grade distributions on the final exam in the two E&M classes. The average grade on the final exam for the class using the new system was 67%, whereas the average for the class using the old system was 42%. While this is not conclusive evidence that the students learned more, it is at least another indicator that this new exam schedule is beneficial for students. The loss of lecture time did not appear to negatively affect the learning outcomes, which is hardly surprising.[10]

Conclusion

Alongside with more frequent exams, student self-reporting of cheating significantly decreased, while desirable discussion and peer-teaching behavior increased. Also, timing signatures of guessing on homework significantly decreased. As an exam is "always around the corner," homework is likely perceived more as a time to study than as a chore, and there is some urgency to actually understanding the material rather than just getting the points and postponing learning until the night before each midterm.

Student performance on solving standard numerical problems increased: giving students the task to sit down each week and really solve problems all by themselves accomplishes much more than recitation sessions (where often TAs solve problems and the students just wait for others to do the work). Like online exam retakes[2,11] the high frequency of exams also partially turned them into formative venues. However, students appear to prefer more exams to the online retakes, likely because each exam is less high stakes. As students get more immediate feedback on their learning progress, any shortcomings can be remedied before the student falls behind hopelessly. While increasing the number of exams given in a semester can be a bit taxing on the instructors, it ultimately seems to improve student performance and attitude. The effort to generate this large number of exams can be reduced by using a learning content management, such as LON-CAPA,[7] to manage exam question test banks.

References

1. *How People Learn,* edited by J. R. Bransford, A. L. Brown and R. R. Cocking (National Academy Press, Washington, DC, 2000).
2. Gerd Kortemeyer, Wolfgang Bauer, Walter Benenson, and Edwin Kashy, "Retaking a test online," *Phys. Teach.* **44**, 235–239 April 2006).
3. Gerd Kortemeyer, Edwin Kashy, Walter Benenson, and Wolfgang Bauer, "Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses," *Am. J. Phys.* **76**, 438–444 (April 2008).
4. Wolfgang Bauer and Gary Westfall, *University Physics* (McGraw-Hill, 2011).
5. Deborah A. Kashy, Guy Albertelli, Edwin Kashy, and Michael Thoennessen, "Teaching with ALN Technology: Benefits and costs," *J. Eng. Educ.* **89**, 499 (2001).
6. David J. Palazzo, Young-Jin Lee, Rasil Warnakulasooriya, and David E. Pritchard, "Patterns, correlates, and reduction of homework copying," *Phys. Rev. ST Phys. Educ. Res.* **6,** 010104 (2010).
7. Gerd Kortemeyer, "An analysis of asynchronous online homework discussions in introductory physics courses," *Am. J. Phys.* **74**, 526–536 (June 2006).
8. Deborah A. Kashy, Guy Albertelli, Wolfgang Bauer, Edwin Kashy, and Michael Thoennessen, "Influence of non-moderated and moderated discussion sites on student success," *J. Asynch. Learn. Net.* **7**(1), 31–36 (2003).
9. Gerd Kortemeyer, "Gender differences in the use of an online homework system in an introductory physics course," P*hys. Rev. ST Phys. Educ. Res.* **5**, 010107 (2009).
10. Carl Wieman and Katherine Perkins, "Transforming physics education," *Phys. Today* **58**, 36–41 (2005).
11. Charles Henderson and Kathleen A. Harper, "Quiz corrections: Improving learning by encouraging students to reflect on their mistakes," *Phys. Teach.* **47**, 581–586 (Dec. 2009).