Re-Taking a Test Online

Gerd Kortemeyer, Wolfgang Bauer, Walter Benenson, and Edwin Kashy
Michigan State University

PACS:
01.40Gb Teaching Methods and Strategies
01.50H Computers as Educational Aids
01.50K Testing Theory and Techniques

Tests and midterms given during the running semester are in the mindset of most educators located somewhere between formative and summative assessment: more serious than homework, but still – as opposed to the final exam – mostly a learning opportunity. In the mindset of most learners, however, these venues are purely summative – they “flunked” or “did well on” a test, but mostly, they got it over with. Few students come to office hour to understand what they did wrong. If they had a bad day, they have no second chance to correct their mistakes, but more importantly, they do not receive any immediate incentive or reward to address detected deficiencies, and for deeper learning, reviewing and understanding of the material after the test is over.

The paper discusses a mechanism used by a number of physics educators at Michigan State University to give students a second chance on succeeding on a test, and to encourage them to learn the expected material, even after the test is done. Students get to earn partial credit by solving a differently randomized online version of the test they just took. For a quantitative discussion and illustration of the mechanism, the paper uses the example of the first three tests in a second semester 200-student introductory calculus-based physics course. The same techniques can also be used in classes or sections with smaller enrollments, but they work best for introductory courses, where we typically ask the students for mastery of simpler concepts.

Randomized Exams

In large-enrollment introductory courses without teaching assistants, unfortunately bubble sheets are the only sustainable means to give frequent tests over the course of the semester. The online system used in this paper, LON-CAPA [1], can generate randomized versions of the same questions, e.g., different graphs, images, numbers, options, etc, from student to student. It also enables instructors to turn any question into a bubble sheet format suitable for printing, where for example numerical answer fields are rendered as answer options with wrong answers either generated by the system or algorithmically pre-specified by the question author. Using both functions together results in 200 different versions of the same exam for our course, see for example Fig. 1. Randomized bubble sheet exams have been used for several years in our department, and were found to be an effective way to reduce cheating [2,3].

The LON-CAPA system is by no means the only course management system in which computerized exams can be given. Other systems, such as for example WebAssign [4], Angel [5], Blackboard [6], or WebCT [7] have similar features that can be utilized. For a comparison see reference [8]. However, LON-CAPA may be the only system that allows the very same randomizing question to be used in an automatically graded bubble sheet and online mode.

Taking the Same Test – Again

The students in our course are using to the same online system, LON-CAPA, for their reading and homework assignments. Since both the test questions and the students are already in the system, it is a matter of a few minutes to make the test questions available online to the students. Immediately after the written test is over, an online copy of the test opens up. This new version of the test can have exactly the same questions that the students encountered during the written exam, or problems with the same words and different numbers for the variables as the written exam, or a different selection of multiple-choice concept choices in different random ordering. Which of these choices is implemented depends on the preferences of the instructor and the capabilities of the online system used. The approach is the reverse of the frequent use of computer-based assessment as practice test *before* an exam (e.g. [9]) – in our courses the online homework serves this function.

Fig.2 shows an online version of the question from Fig. 1.In the first question part, the system rendered a free-form answer box and left the input of the correct physical unit to the student.

In our course, students are given 30 hours to complete the online version of the test, with 2 or 3 attempts to arrive at the correct answer. Since we have no way of monitoring student interactions during this time period, we explicitly allow students to collaborate on solving the retake exam, and we do not block any of the online threaded discussions that the system allows around individual questions. We feel comfortable doing this, since the randomization of the questions inhibits mindless copying of answers – instead, since we also do not curve the course grades, we have the added benefit of the constructive collaborations as students explain the physics to each other. Using the “block discussion” option inside of LON-CAPA would likely just shift the discussion to other online forums [10]. Surprisingly, in spite of allowing collaborations, in an end-of-semester survey, the statement “I did the retakes by myself” received a rating of 0.6±1 on a scale from “-2” (“strongly disagree”) to “2” (“strongly agree”), suggesting that about half of the students decided to work alone.

It is, of course, technically possible that a student has a friend take the entire online correction for him/her. At present, none of the authors has taken steps to prevent this, and in fact, we are not aware of any technology in a completely online setting to make this impossible.

Participation in the online retake, often referred to as 'partial credit by corrections', is completely voluntary, and helps defuse the ubiquitous student issue in computer scored examinations, “Why can't I get partial credit?“ The final score is then calculated according to the following formula:

n : written (bubbled) score

N : online (retake) score

P : final score

$$P = \begin{cases} n \text{ for } N \leq n \\ n + 0.3(N - n) \text{ for } N > n \end{cases}$$

The students can thus earn "30 cents on the dollar" for the difference between their written and online score. As an example, a student might have a written score of $n=8$, and an online score of $N=12$, which results in a final score of $P=9.2$ for the test.

One may ask what the right level of partial credit for the retake is. Clearly, partial credit significantly above 50% makes the written exams much less meaningful. Also, giving partial credit of less than 10% diminishes the attraction of the retake exercise in the eyes of the students. We find empirically, that 30% partial credit for the retake strikes the right balance. At this level more than 95% participate in the retake opportunity.

In our course, online retakes are not offered for the final exam, but only for the six topical tests that replace the two midterm exams traditionally given in this course.

Experiences

Experiences with this mechanism have been excellent for both students and instructors. Students, of course, will first and foremost appreciate the "bonus" points they can earn online, but also the confidence they can gain, even after a test that did not go so well for them. The rate of correct online solutions is consistently over 90 percent.

Instructors appreciate a chance to catch a still highly motivated audience in a teachable moment with only a small additional effort on their part. Once a test is in place, it takes approximately 10 minutes to also provide an online version. In spite of the fact that students are basically spending twice the time on task and doing twice the work, no protest is heard. In fact, even complaints about a test being "too hard" or "unfair" have been greatly reduced by the new mechanism. Finally, if in the bubble sheet version some correct answers were selected by luck, students must now solve these problems in the re-take.

In terms of actual grade changes, Table 1 shows the respective average scores for n, N, and P, for the first three tests in the aforementioned course. For tests with a good written score (e.g., Test 2), the online retake does not make much of a difference – for Tests 1 and 2, the average grade went up by half a number grade from 2.0 to 2.5 and from 2.5 to 3.0, respectively. For tests that did not go so well (e.g., Test 3), the online retake makes a bigger difference; here, the average grade went up by a complete number grade from 1.5 to 2.5.

Student participation turned out slightly lower for the test that went better: for Tests 1 and 3, the student participation in the retake was 97.8%, while for Test 2, "only" 95% of the students participated in the retake opportunity.

In all cases, though, the online retake did not lead to grade inflation in the sense that every student automatically obtains a perfect grade. In the implementation presented here, retaking exams also does not lead to a loss of discrimination between weaker and stronger students. If instructors are worried that the overall class average is raised, they are still free to employ a curve. One may also argue that students deserve a better grade if they learn better, and that if the class learns better as a whole, a higher average grade may be justified. This is also illustrated in the histograms Fig. 3. The online retake by its very design makes the largest contribution in the extreme lower end of the score distribution where the most remediation is needed.

We realize that the reliance on multiple-choice exams is not without problems, see for example [12]. If the in-class version of the test was of the simple "plug-and-chug" variety, adding a retake component will not improve the quality of the test. But as the Force Concept Inventory [13] has shown, it is possible to extract information on concept mastery with these simple tools, if applied properly.

We find that using the right concept questions can lead to deeper reflections on the part of the students during the retake process. In the survey, the statement that the retakes were worthwhile apart from getting more points was rated 1.1±0.9, and the statement that retakes helped learning was rated 1±0.9.

Retakes can even be used as an effective means of peer instruction: the statement that the online discussion around the exam retakes helped learning was rated 1.3±0.7. At present, we are working to incorporate peer instruction questions as developed by Mazur and his group [14] into our question pool.

In summary, the retakes increase student time-on-task and appear to help students master the difficult concepts they encounter in introductory physics.

.

Outlook

In LON-CAPA, online resources, including problems, can be shared across semesters, courses, and institutions [11]. As questions are being deployed, both written and online, the system automatically collects information on their degrees of difficulty and discrimination, and associates those with the problem resources themselves. As more and more of this usage-based data gets collected, we will be able to reliably increase the randomization of exams to a level where we can substitute or re-arrange whole question parts without generating unfairly easy or hard exam versions. This mechanism may enable fully mastery-based exams, where students (with decreasing credit towards scores) can take retakes of retakes until they have demonstrated sufficient mastery of the respective topics and concepts.

Conclusions

Online retakes of tests are providing an additional teaching venue within large enrollment physics courses, and are a means to stress the formative aspects of tests. They motivate students to revisit material not mastered on the tests, and to spend additional time on task learning it.

A capacitor is completely charged with 650 nC by a voltage source that had 350 V.

$\boxed{1\ pt}$ What is its capacitance? *(in* F*)*

7.A○ 1.49×10^{-9} **B**○ 1.86×10^{-9} **C**○ 2.32×10^{-9}
D○ 2.90×10^{-9} **E**○ 3.63×10^{-9} **F**○ 4.53×10^{-9}
G○ 5.67×10^{-9} **H**○ 7.08×10^{-9}

$\boxed{1\ pt}$ Now the plates of the charged capacitor are pushed together with the voltage source already disconnected.

8. **A**○ The charge on the plates increases.
B○ The energy stored in the capacitor remains the same.
C○ The capacitance increases.
D○ The voltage drop between the plates increases.
E○ The energy stored in the capacitor increases.

$\boxed{1\ pt}$ The initial air gap was 8 mm. What is the stored energy if the air gap is now 6 mm? *(in* J*)*

9.A○ 0.00 **B**○ 8.53×10^{-5} **C**○ 1.14×10^{-4}
D○ 1.30×10^{-4} **E**○ 1.52×10^{-4} **F**○ 3.41×10^{-4}
G○ 3.44×10^{-4} **H**○ 4.87×10^{-4}

A capacitor is completely charged with 670 nC by a voltage source that had 350 V.

$\boxed{1\ pt}$ What is its capacitance? *(in* F*)*

7.A○ 1.91×10^{-9} **B**○ 2.39×10^{-9} **C**○ 2.99×10^{-9}
D○ 3.74×10^{-9} **E**○ 4.67×10^{-9} **F**○ 5.84×10^{-9}
G○ 7.30×10^{-9} **H**○ 9.13×10^{-9}

$\boxed{1\ pt}$ Now the plates of the charged capacitor are pulled apart with the voltage source already disconnected.

8. **A**○ The voltage drop between the plates increases.
B○ The energy stored in the capacitor remains the same.
C○ The charge on the plates increases.
D○ The capacitance increases.
E○ None of the above.

$\boxed{1\ pt}$ The initial air gap was 6 mm. What is the stored energy if the air gap is now 11 mm? *(in* J*)*

9.A○ 0.00 **B**○ 6.40×10^{-5} **C**○ 1.17×10^{-4}
D○ 2.15×10^{-4} **E**○ 2.91×10^{-4} **F**○ 3.63×10^{-4}
G○ 4.39×10^{-4} **H**○ 5.42×10^{-4}

Fig 1: Written version of a question from Test 2 in two randomizations

Problem 6

Due on Tuesday, Feb 22 at 10:00 am

A capacitor is completely charged with 640 nC by a voltage source that has 375 V.

What is its capacitance?

Submit Answer Tries 0/3

Now the plates of the charged capacitor are pulled apart with the voltage source still connected.

- The capacitance increases.
- The voltage drop between the plates increases.
- The energy stored in the capacitor increases.
- The energy stored in the capacitor remains the same.
- None of the above.

Submit Answer Tries 0/2

The initial air gap was 5 mm. What is the stored energy if the air gap is now 10 mm?

Submit Answer Tries 0/3

Fig 2: Online version of the question in Fig 1 in another randomization

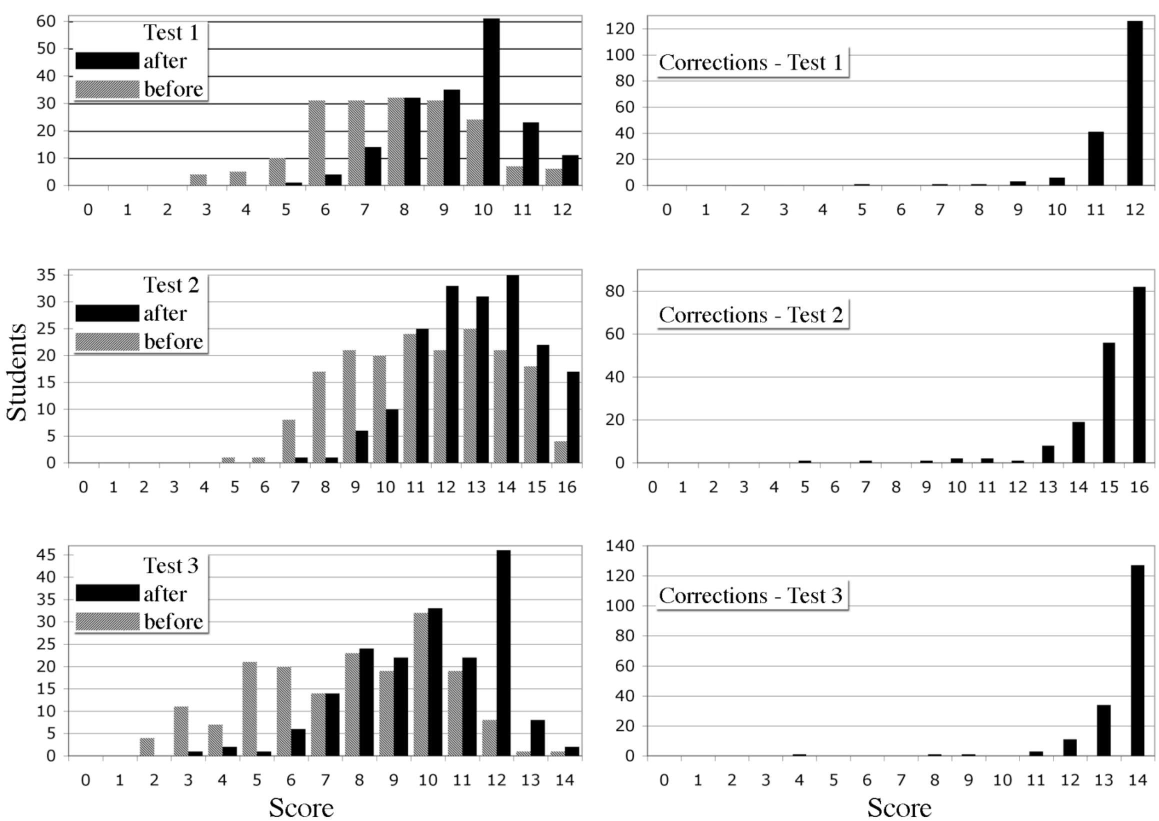

Fig 3: Histograms for Tests 1, 2, and 3 (columns, left to right) of written score *n*, online score *N*, and final score *P* (rows, top to bottom)

Table 1: Maximum, average written, average online, and average final score for Tests 1, 2, and 3

	Max Pts	n_{ave}	N_{ave}	P_{ave}
Test 1	12	7.8 (65%; 2.0)	11.6 (97%)	8.8 (73%; 2.5)
Test 2	16	11.4 (71%; 2.5)	15 (94%)	12.3 (77%; 3.0)
Test 3	14	7.8 (56%; 1.5)	13.5 (96%)	9.3 (66%; 2.5)

References

1. The Learning*Online* Network with Computer-Assisted Personalized Approach software system, Michigan State University (1992-2005), available as open-source freeware under the GNU General Public License from http://www.lon-capa.org/
2. Edwin Kashy and David Morrissey, in Sheila Tobias and Jacqueline Raphael, “The Hidden Curriculum” (Plenum Press, New York, 1997), pp. 165-167
3. Deborah Kashy, Guy Albertelli, Edwin Kashy, and Michael Thoennessen, “Teaching with ALN Technology: Benefits and Cost,” Journal of Engineering Education 89, 499 (October 2001)
4. http://www.webassign.net/
5. http://cyberlearninglabs.com/products/
6. http://www.blackboard.com/
7. http://www.webct.com/
8. http://www.edutools.info/course/
9. Leith Sly and L. J. Rennie, “Computer managed learning as an aid to formative assessment in higher education,” in S. Brown, J. Bull, and P. Race (Eds), “Computer Assisted Assessment of Students” (Kogan Page, London, 1994), pp. 113-120
10. Deborah Kashy, Guy Albertelli, Wolfgang Bauer, Edwin Kashy, and Michael Thoennessen, “Influence Of Non-Moderated And Moderated Discussion Sites On Student Success,” Journal of Asynchronous Learning Networks 7(1) (January 2003)
11. Cheryl Speier and Gerd Kortemeyer, “Open-Source Objects for Teaching and Learning,” Syllabus Magazine (November 2001)
12. Robert J. Dufresne, William J. Leonard, and William J. Gerace, “Making Sense of Students’ Answers to Multiple-Choice Questions” The Physics Teacher **40**, 174-180 (2002).
13. D. Hestenes, M. Wells, and Swackhamer, “Force Concept Inventory”, The Physics Teacher 30, 141-158 (1992).
14. E. Mazur, “Peer Instruction”, Prentice Hall, New Jersey (1997).

Gerd Kortemeyer is a professor of Physics Education in the Lyman Briggs School of Science and the Division of Science and Mathematics Education at Michigan State University, and the principal investigator of the LON-CAPA project.
Contact: Lyman-Briggs School of Science, Michigan State University, East Lansing, MI 48825, korte@lon-capa.org

Wolfgang Bauer is a professor of Physics and department chairperson at Michigan State University.
Contact: Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, bauer@pa.msu.edu

Walter Benenson is a professor of Physics at Michigan State University.
Contact: National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, benenson@nscl.msu.edu

Edwin Kashy is a professor of Physics at Michigan State University.
Contact: National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, kashy@nscl.msu.edu