

PERSONALIZED EXAMINATIONS IN LARGE ON-CAMPUS CLASSES¹

Guy Albertelli², Gerd Korkemeyer³, Alexander Sakharuk⁴, and Edwin Kashy⁵

Abstract - For the past decade we have used examinations that differed for each student: In algorithmic problems the numerical values vary; for conceptual problems a wide set of choices tests the same concept and the order in which the concepts are presented differs. Grading of such examinations has evolved to become highly automated so that the time and effort required is considerably reduced. The preparation of examination materials in a form suitable for optical scanning in a bubble sheet format has also evolved and is now at the stage where a large number of question types are automatically formatted. These advances have facilitated the use of quizzes to promote attendance and transform examinations to become both formative and summative assessment engines. The current formatting is such that online examinations are also facilitated.

Index Terms – Personalized examinations, Personalized quizzes, Summative assessment, Formative assessment, On-line examinations.

INTRODUCTION

Examinations and quizzes require a significant instructional staff effort, especially in large attendance courses. The reduction of that task using IT tools allows instructors much greater use of these course components to improve education. We have used personalized examinations, where the content differs in some way for each student, to do much more than just discourage or inhibit students from copying or otherwise cheating on exams. Since short computer-scored quizzes are now easily managed, we make considerable use of quizzes as unannounced tasks at the end of lectures. This use has resulted in a large decrease in absenteeism and correlates positively with student performance [1].

Transforming mid-term examinations into formative assessment tools is also greatly facilitated by the technology. While the mid-term exams remain an important summative assessment of progress, students are allowed to “rework” these examinations in a homework assignment format following the test [2], with the incentive that they earn “Partial Credit by Corrections” on the examination by solving and correctly answering the various problems.

The software system used in much of the work described here was CAPA, an acronym for Computer-Assisted Personalized Approach. For the past decade, the

system has been used for preparation and grading and assessment of personalized homework, quizzes and examinations in a large number of courses at MSU and other institutions [3]. The latest work presented here is with LON-CAPA, the Learning On-line Network with CAPA [4], a new, totally web based, learning content and course management system. LON-CAPA is open source and is designed to assist in all aspects of teaching a course. As will be seen below, the preparation and handling of examinations in LON-CAPA represents an important advance as it has built upon the past experience with CAPA.

In the following, we review the evolution over the past decade of the examination format and of the method of grading, and also show examples of questions. We then present the latest implementation of examinations in LON-CAPA. We also discuss the use of quizzes and examinations by various instructors and briefly touch on plans to use basic wireless technology to further reduce the task as well as the time required to provide feedback for students on their performance

PREPARATION AND GRADING WITH CAPA

From 1993 to 1995, the formatting of examinations did not differ from homework. A complete personalized printed copy with each student’s name at the top was distributed. For homework each student entered the answers on-line, initially via telnet and later via the web. Students got instant feedback and had a number of tries to get a correct answer without penalty. Examinations were proctored. Students wrote answers on a sheet of paper where each problem had a pre-formatted area for those answers. The instructor assigned grades by comparing a student’s answers to the answers generated by the system for that particular student. The grades were then entered into the system and student could see results at their next login. Thus any question could be part of either a homework set or of an examination with no formatting change.

With the number and size of classes growing, grading examinations represented a large task. For example, grading a final exam of 20 problems in a class of 500 students and checking each answer sheet after initial grading to minimize errors required several hours for a teaching staff of 4 persons. Optical scanning of bubbles on an answer form appeared to be the solution. The use of the university’s central scanning facility appeared impossible out because of

¹ Work supported by the Andrew W. Mellon and Alfred P. Sloan Foundations and the National Science Foundation (ITR 0085921).

² Guy Albertelli, Michigan State University, Division of Science and Math Education, albertel@msu.edu

³ Gerd Kortemeyer, Michigan State University, Division of Science and Math Education, korte@lite.msu.edu

⁴ Alexander Sakharuk, Michigan State University, Department of Physics and Astronomy, sakharuk@lite.msu.edu

⁵ Edwin Kashy, Michigan State University, Department of Physics and Astronomy and NSCL, kashy@nscl.msu.edu

the different pattern of correct responses for each student, and of the variety of questions to be processed. Since most classes at the time were in physics, a relatively affordable tabletop optical scanner was acquired. In order to accommodate questions requiring students to indicate which of several choices applied to each part (see pulley problem in Figure 4) below), a custom form based upon the content was generated for each examination and formatted to reflect the order and type of question to be answered.

These custom bubble forms represented a significant improvement for students as well as improved reliability. When used for quizzes, a single custom form included both the questions and the bubbles. A code on each form allowed rapid distribution of the quizzes as the system would then recognize to correct answers for that particular version. The data generated by the scanner was fed into the system and students could see their grades usually within a few hours of the examination period.

Operating the tabletop scanner was, however, a significant task, especially when scanning custom answer forms which had been severely mauled by students anxiously taking the exams. This was especially true on wet or snowy days. We also had the paper jam in the device and this required delicate extrication and time.

Current CAPA Implementation

Beginning Fall 1998, the scanning task was turned over to the central scanning office with its fast and accurate scanners. This required rewriting certain questions specifically for examinations in a format that made them answerable on the scoring form by one or several bubbles a line. As usual, essay questions were hand graded and the assigned score entered on the form by the grader. For numerical problems, transformation of a problem from homework to exam style was a rather straightforward task and carried out by the system. An if-statement added in the problem code would identify an examination and the system would generate the correct answer, select a factor from ~ 1.13 to ~ 1.5 , and multiply the correct answer by that factor raised to successive integer powers from -7 to $+7$. Answers are then presented in order of increasing values, with the correct answer randomly located among the 8 options.

Two versions of a typical numerical question are now shown. Figure 1 shows the homework version of the problem. Students must calculate a numerical value and also enter units for their answer. Note that bold printing in the question is used to identify vector quantities.

1. [5pt] Two forces \mathbf{F}_1 and \mathbf{F}_2 are acting on a mass $M=5\text{ kg}$. The force $\mathbf{F}_1 = 7\mathbf{i} + 7\mathbf{j}$, in N . The acceleration of the mass is given by $\mathbf{a} = 1\mathbf{i} + 7\mathbf{j}$, in m/s^2 . Calculate the magnitude of \mathbf{F}_2 .

FIGURE 1
NUMERICAL PROBLEM FORMATTED FOR HOMEWORK

1. [5pt] Two forces \mathbf{F}_1 and \mathbf{F}_2 are acting on a mass $M=5\text{ kg}$. The force $\mathbf{F}_1 = 5\mathbf{i} + 4\mathbf{j}$, in N . The acceleration of the mass is given by $\mathbf{a} = 3\mathbf{i} + 7\mathbf{j}$, in m/s^2 . Calculate the magnitude of \mathbf{F}_2 . (in N).

- A) 1.77×10^1 B) 2.00×10^1 C) 2.26×10^1 D) 2.55×10^1
 E) 2.88×10^1 F) 3.26×10^1 G) 3.68×10^1 H) 4.16×10^1

FIGURE 2

NUMERICAL PROBLEM OF FIGURE 1 IN EXAMINATION FORMAT

The same problem as it would appear on an examination is reproduced in Figure 2. The students select the correct answer and enter it on their scoring form (in this instance bubble F on line 1). Numerical values in the problems of Figure 1 and Figure 2 reflect a different seed for the random number generator which controls these values. The amount of space between the question and answer choices (small in Figure 2) is normally large enough to allow students to work out the question. Note that some guidance is given to the student if none of the choices listed match the answer calculated, they have almost certainly done something incorrectly. There is also the option to include likely errors among the answers instead of using the computer-generated choices.

Figure 3 shows a conceptual problem where one selects all correct answers from a list. No change is required in such a homework problem in order to use it in an examination. All appropriate bubbles on a single line of the scoring form are marked, as in the example shown below. As in the numerical example, the line number is the same as the problem number.

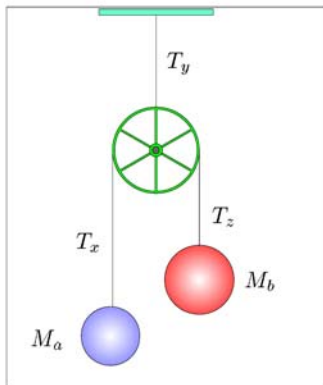
2. [2pt] John is listening to a horn. He knows the frequency of the horn is 300 Hz when both he and the horn are at rest. If he hears a pitch of 330 Hz, there are clearly several possibilities. (Give ALL correct answers, i.e., B, AC, BCD...)

A) Both can be moving and have the same speed.
 B) Both can be moving and have different speeds.
 C) John is moving towards the horn at rest.
 D) The distance between John and the horn is increasing with time.
 E) Both cannot be moving in the same direction.
 F) Both can be moving, in opposite directions.

FIGURE 3
CONCEPTUAL QUESTION WITH SEVERAL CORRECT PARTS.

A correct answer to the example in Figure 3 would require that the student select and mark bubbles A, B, C, and F on line 2 of the scoring form. The grading basis is up to the instructor who can choose from 'strict', 'medium', or 'easy' grading. Credit in 'strict' grading is only given for a completely correct answer. In the 'easy' option, each statement is given proportional credit. 'Medium' grading is similar to 'easy', except that the first error counts as two incorrect, two errors as four, etc. An error consisted of a missed correct choice or of an incorrect one selected.

8. [6pt] A frictionless, massless pulley is attached to the ceiling, in a gravity field $g = 9.81 \text{ m/s}^2$.



Mass M_b is greater than mass M_a . The tensions T_x , T_y , T_z , and the constant g are magnitudes. (For each statement select T-True, F-False, G-Greater than, L-Less than, or E-Equal to, For example, GFLLEG)

- A) T_z is T_x
- B) $(M_b)g$ is ... T_z
- C) The center-of-mass accelerates.
- D) $(M_a)g + (M_b)g$ is T_y
- E) $T_z + T_x$ is T_y
- F) The magnitude of the acceleration of M_b is ... the magnitude of the acceleration of M_a .

FIGURE 4

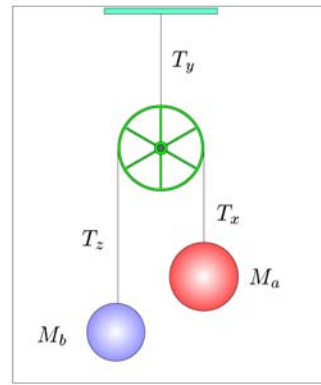
CONCEPTUAL QUESTION, HOMEWORK FORMAT, WITH MULTIPLE ANSWER CHOICES

Figures 4 show a conceptual problem as it appears on a homework assignment. There are three answer responses for some parts, Greater than, Less than, and Equal to and two, True or False for one part. Such problems with the random labeling of elements in the figure work especially well for on-line response as they stimulate considerable discussions among students. To use that complete problem on a test it could be re-written, as two questions. One question could include three of the original choices, with each part reproduced three times to include the three possible responses. The problem is also reasonably well represented in the style of Figure 3. That version, where student select all correct statements from a list, required recoding the original problem and is shown in Figure 5.

For essay questions, or for questions where a formula would be entered in the on-line version, or for numerical questions where the grade depends in part on the work shown by the student, a grade is marked on the form by the teaching staff so that the bubble scoring form has all the necessary information before it is sent for scanning and the resulting file processed by the system.

The transformation of examinations to use standard scoring forms instead of the system generated custom forms represented a significant improvement in the ease of administering examinations and quizzes as responsibility for optical scanning was shifted to the central office. Much greater use was made of quizzes; and for example,

8. [6pt] A frictionless, massless pulley is attached to the ceiling, in a gravity field $g = 9.81 \text{ m/s}^2$.



Mass M_a is greater than mass M_b . The tensions T_x , T_y , T_z , and the constant g are magnitudes. (Select ALL of the correct statements, i.e., A, CD, AEF, BCDEF, ...)

- A) The center-of-mass of M_b and M_a accelerates.
- B) $(M_b)g + (M_a)g$ is equal to T_y
- C) The magnitude of the acceleration of M_a equals that of M_b
- D) T_x is equal to T_z
- E) T_y is greater than $T_x + T_z$
- F) T_z is equal to $(M_b)g$

FIGURE 5

CONCEPTUAL QUESTION OF FIGURE 4 AS RECODED FOR CAPA EXAMINATION FORMAT

approximately 16,000 forms were used in one semester in a 500-student introductory physics course. Many other classes currently use this method, and it works remarkably well.

Since the correct answers and bubble pattern for each student differs, the scanner generates a file on which the student is identified by name, student number, and also by a 6-digit code printed on each individualized exam and entered on the form by the student. One can also choose to have only the code as identifier on an examination or a quiz, allowing a so-called 'anonymous' mode. In that mode, a particular exam paper may be given to any student, and the correct answers for that student can be generated on the basis of that code. It means also that it takes much less time to distribute an examination as a student can use any paper. This mode is especially useful for quizzes near the end of a lecture, as little time is wasted in distributing papers.

For most exams however, instructors have opted to print both name and the student's ID picture (optionally), of the particular student. Typically papers are spread out in alphabetical order with both name and picture showing. Student find their paper and go (another instructor option) to an assigned seat marked on the paper. This takes a few minutes, but students do not open their papers and see the content until advised. It has a great advantage in that there is a lot of redundancy for identifying the student, and eliminates the need to check student IDs. At the close of the exam, and for a number of reasons, the student may elect to tear off and keep the front page with the ID picture.

While the system works well, there is the additional work in having to convert many of the problems used on assignments to make them suitable for machine grading. There is also a technical problem in scanning forms where more than one bubble is allowed on a given line of the scoring form. When single bubbles are expected per line, any form with multiple bubbles on a line can be rejected and set aside for review by the operator. Optical scanners often pick out the mark left by a student who is just 'resting' his pencil on a bubble while deciding on the answer, or the pencil residue on a bubble not completely erased. It is not unusual to have a student show you a black smudge on a line with another bubble on that line and say that it is clear which he has erased. True, but the scanner reads both. This is not a problem when only one bubble is expected, and the operator can take appropriate steps. With multiple bubbles allowed, operators do not reject forms. It is then the instructor's task to look at the forms and in case of bad erasures, insure, at least for the examinations, that the student entries are properly recorded. We also note for that certain forms, especially those using dark purple ink, some scanners are more difficult to adjust to reject erased bubbles that still have some residue of the initial mark.

LON-CAPA EXAMINATIONS

In evolving the examinations method to LON-CAPA, the goal was to eliminate both the re-writing task and the scanning difficulties associated with multiple bubbles on a single line of the scoring form. A single version of a problem now works well for on-line response in a quiz or examination and is suitable for optical scanning response. It should also work well with basic wireless technology as it becomes available since all that is needed is a set of line numbers and letters. This approximates the versatility of on-line responses: pull down selections, strings, numerical values, radio buttons etc. Examples of a few questions types for the printed examinations or quizzes are presented. Note that the formatting for a printed examination is also suitable

12 pt Four blocks are on a horizontal surface.

The blocks are connected by thin strings with tensions T_x , T_y , T_z . The masses of the blocks are $A=25.0$ kg, $B=29.0$ kg, $C=17.0$ kg, $D=17.0$ kg. Two forces, $F_1=48.0$ N and $F_2=77.0$ N act on the masses as shown. Assume that the friction between the masses and the surface is negligible and calculate the tension T_y . (in N)

6. A 3.03×10^1 B 3.43×10^1 C 3.88×10^1
 D 4.38×10^1 E 4.95×10^1 F 5.59×10^1
 G 6.32×10^1 H 7.14×10^1

FIGURE 6
 NUMERICAL PROBLEM IN LON-CAPA FORMAT.

7 pt A fisherman and his young niece are in a boat on a small pond. Both are wearing life jackets. The niece is holding a large floating helium filled balloon by a string. Consider each action below independently and indicate whether the level of the water in the pond, Rises, Falls, or is Unchanged.

- ▷ The niece finds a cup and bails some water out of the bottom of the boat
 15. A Rises B Falls C Unchanged
- ▷ The niece pops the balloon.
 16. A Rises B Falls C Unchanged
- ▷ The fisherman lowers the anchor and it hangs one foot above the bottom of the pond.
 17. A Rises B Falls C Unchanged
- ▷ The fisherman knocks the tackle box overboard and it sinks to the bottom.
 18. A Rises B Falls C Unchanged
- ▷ The fisherman lowers himself in the water and floats on his back.
 19. A Rises B Falls C Unchanged
- ▷ The niece gets in the water and pops the floating helium balloon
 20. A Rises B Falls C Unchanged

FIGURE 7
 LON-CAPA EXAMINATION FORMAT, WITH THREE ANSWER CHOICES

for an on-line examination using radio buttons. For problems requiring selection of a single correct answer from a list, a very commonly used testing format, a single line on the scoring form is used to indicate the correct choice as is the case for numerical problems such as the one in Figure 6.

Figure 6 represents an algorithmic numerical question. The masses and tensions in the figure are randomly labeled. There is a choice of two figures, interchanging the forces. Thus there are 288 different versions for the figure affecting the calculation. Furthermore the numerical values of the masses and forces are also randomized. The student marks the correct choice out of an ordered list of eight values, with associated units indicated in the text. The format is similar to that of the problem in Figure 2, except that there is no problem number, but rather the answer's choices are numbered. This becomes clearer in the next example shown in Figure 7.

This problem has several statements each of which requires one of three attributes as an answer: Rises, Falls, Unchanged. Several lines in the scoring form are automatically assigned for the answer, with both line number and lettered bubble choice indicated to minimize ambiguity. The problem shown in figure 4 would, as part of an examination, automatically be formatted in a way similar to that of Figure 7. It also illustrates the flexibility in writing questions as well as the single bubble per line on the scoring form when using LON-CAPA.

4 pt Rank the perimeters of the plane geometric figures, from shortest to longest. (Shortest has rank = 1. If two perimeters are equal, they get the same rank and the following rank is skipped, e.g., 1 2 3 3 5 6)

▷ Square with sides of length 32 mm.
25. A 1 B 2 C 3 D 4 E 5 F 6

▷ Right triangle with hypotenuse of length 42 mm and one leg of length 24 mm.
26. A 1 B 2 C 3 D 4 E 5 F 6

▷ Regular hexagon with side of length 24 mm.
27. A 1 B 2 C 3 D 4 E 5 F 6

▷ Rectangle with sides of lengths 42 mm and 26 mm.
28. A 1 B 2 C 3 D 4 E 5 F 6

▷ Equilateral triangle with sides of length 38 mm.
29. A 1 B 2 C 3 D 4 E 5 F 6

▷ Circle with diameter of 40 mm.
30. A 1 B 2 C 3 D 4 E 5 F 6

FIGURE 8
 RANKING PROBLEM IN LON-CAPA EXAM FORMAT

Figure 8 shows a problem requiring students to rank items as a result of a calculation. Ranking can also be used for items on any ordering scale by assigning numerical values to each item.

Essay questions are hand-graded. A boxed area is provided on the exam paper for a student to enter the essay. The system automatically reserves the appropriate number of lines on the scoring form for scores to be entered by the teaching staff. That same format is also used for questions where the answer is an equation. The first part of the question in Figure 9 shows an example of a formula

6 pt Derive a formula for the straight line which, at $x = 2$, is tangent to the curve below.

$$y = 5 + \left(4x - \frac{2}{x}\right)^2$$

23. Leave blank on scoring form

3 pt Calculate the slope of a straight line perpendicular to the tangent line above.

24. A -1.27×10^{-2} B -1.59×10^{-2} C -1.98×10^{-2}
 D -2.48×10^{-2} E -3.10×10^{-2} F -3.88×10^{-2}
 G -4.84×10^{-2} H -6.06×10^{-2}

FIGURE 9
 FORMULA QUESTION IN LON-CAPA EXAM FORMAT

problem. The second part of that problem is a numerical problem to help emphasize the numbering of answers rather than of problems for the scoring form.

In the online response to a formula question, students use / for divide, * for multiply, ^ for power, and () in formatting their answer. The format is the same as they would use on a typical calculator, and they are given feedback if the answer is not properly formatted. To grade the problem on an examination, the teaching staff has access to the particular student’s answer since the coefficients in the equation vary among students. Hand grading by the staff when such a problem is on an examination allows for variation in writing down a formula. Note the line numbered 23 in Figure 9 which students are instructed to leave blank on their scoring form. The grader just marks that line on the student’s scoring form to enter a grade for the formula derived.

DISCUSSION

Personalized examinations have eliminated cheating as a concern in giving examinations. Students realize that they have different numerical values in algorithmic problems and that the order of questions and wording of statements varies. Complaints by students that other students are cheating, a fairly common occurrence, has disappeared. Preparing and processing individualized examinations is streamlined and requires little additional work. Short quizzes, typically given near the end of the lecture period have proven a strong incentive for students to attend class, even though the impact of the quiz grade on the final score can be very small. Some instructors use such quizzes to get students to pay close attention in class by informing students that the quiz questions will be related to that day’s discussion topics.

Mid-term examinations can also conveniently be used as formative assessment tools by allowing students the opportunity to earn partial credit by reworking a second version of examinations [2]. This functionality is greatly facilitated by the system, and would probably not be an option without information technology as it would require a large amount of additional work, especially in large enrollment classes. Essentially all students avail themselves of the opportunity, in great part no doubt due to incentive given. Typically 25% of missed points are recovered by solving all the problems correctly. A student scoring 60% on a mid-term exam and getting 100% on the version following the exam has now a grade of 70%. Even though participation is optional, essentially all students avail themselves of the opportunity. Note that the ‘correction versions’ are also personalized and are generated so that they represent a different version of the original exam. Just as when doing homework assignments students enter answers on-line, have access to hints, can participate in discussion forums, and get help from the instructional staff.

An interesting impact of this “Partial Credit by Corrections” is that the number of students complaining

about the difficulty and/or unfairness of the exam drops considerably after they get the correct solutions. A much more common response is “It wasn’t that hard! I should have got it!” Another added benefit is that there is little need for the instructor to discuss the examination in great detail in lecture.

The printing of a large number of individualized examinations can take several hours, but the process is automated and has proven very reliable. Students taking the examinations turn in both the printed exam and the scoring form so that a complete record of each student’s work is kept. The turn around time in the scoring office has been shorter than for other type of examinations. No post-processing or analysis is done in that office. The data file, which contains an image of the entries on each student’s scoring form, is electronically transmitted to the system’s server, typically within 24 hours. Processing that file takes only a few minutes so that students can get feedback on their performance rather quickly.

Printed exams provided by the system would be needed even if examinations were administered on-line, especially for the final examinations, as any breakdown in the network technology would cause serious problems. Finally we note that the system provides numerous statistical measures of student performance, including assessment of misconceptions in problems where randomization and random labeling makes such data difficult to obtain [5].

ACKNOWLEDGMENT

We are indebted to the many professors who have used personalized examinations and quizzes and given us the feedback to improve the method.

REFERENCES

- [1] Kashy, E., Thoennessen, M., Tsai, Y., Davis, N. E., and Wolfe, S. L., “Using Networked Tools to Promote Student Success in Large Classes”, *Journal of Engineering Education*, ASEE, Vol. 87, No. 4, pp. 385-390 (1998)
- [2] Tobias, S., Raphael, J., “The Hidden Curriculum”, contribution by E. Kashy and D. J. Morrissey, pp 165-167, Plenum Press (New York 1997)
- [3] Kashy, D A, Albertelli, G., Kashy, E., and Thoennessen, M., “Teaching with ALN Technology: Benefits and Costs”, *Journal of Engineering Education*, Vol. 90, No. 4, October 2001, pp. 499-505
- [4] Kortemeyer, G., Bauer, W., Kashy, D., Kashy, E., and Speier, C. The LearningOnline Network with CAPA Initiative. IEEE Frontiers in Education Conference Proceedings, Vol. 31, 2001, p1003. The software is free, licensed by MSU under GNU General Public License. <http://lon-capa.org>
- [5] Albertelli, G., Minaei-Bigdoli, B., Punch, W.F., Kortemeyer, G., and Kashy, E., Computer Feedback in Computer-Assisted Assignments, ASEE/IEEE Frontiers in Education Conference, Boston, MA, Nov 6-9, 2002.