INDIVIDUALIZED INTERACTIVE EXERCISES: A PROMISING ROLE FOR NETWORK TECHNOLOGY*

Deborah A. Kashy[1], Guy Albertelli[2], Guy Ashkenazi[3], Edwin Kashy[4], Hon-Kie Ng[5], and Michael Thoennessen[6]

Abstract *— For several years we have made broad use of network technology in large undergraduate physics courses. We have found consistent evidence that this technology has had a positive impact on learning. A major task in this project has been to develop questions and problems that can be used on individualized homework assignments and exercises. Our goals have been to promote collaboration among students, discourage students from mindlessly copying each other's work, and guide students away from the plug-in-formula problem solving approach and towards concept-based understanding of the subject.*

In this paper we illustrate and discuss a variety of numerical and conceptual questions that we have developed. We use data from our Fall 2000 course to assess how the different types of problems impact student achievement. Results indicate that individualized interactive exercises, especially those in applet format, show the highest correlation with students' achievement. However, these problems are also the most difficult to prepare. The importance and value of expending that effort is discussed.

Index Terms: On-line Assignments, Numerical Problems, Conceptual Exercises, Individualized Applets.

INTRODUCTION

Beginning Fall 1995, network technology has been used in essentially every aspect of a large, 500-student, calculus-based physics lecture course. The course is a requirement for most students majoring in mathematics, science, and engineering, and success in the course is often a condition for pursuing such majors. Indeed, on many campuses students have dubbed the course a "weed-out" course. Assessment and evaluation of student performance over several years has shown increases in student success rates with technology [1]. This use of technology is often referred to as establishing an Asynchronous Learning Network (ALN). Successful aspects of our use of ALN as well as some significant problem areas encountered have recently been discussed [2].

This paper focuses on an issue central to most if not all uses of network technology in education – the nature and quality of the exercises that are assigned to students. These exercises are at the core of every aspect of our use of ALN technology. Students are given weekly homework assignments in which the problems are individualized. Students enter their solutions on-line, and receive immediate feedback on correctness as well as help in the form of hints or links to useful material when the instructor has included such help. The discussion forum we have provided is linked directly to the questions in the assignments. In addition, we have created a learning center where students can interact with each other as well as with teaching staff. The software system used in the course, CAPA (Computer-Assisted Personalized Approach), supports a broad variety of question types including conceptual, numerical, and essay problems [3]. The first two types constitute the majority of the assignments. A few essay questions are also part of assignments. The essays are submitted on-line and evaluated by teaching staff. The system facilitates the task of grading the essays by highlighting keywords and recording grades automatically. It also allows the instructor to send e-mail feedback to students.

In this paper we describe a range of problem types that we have developed and used. The specific problem types that will be discussed include mathematical problems, traditional numerical problems, conceptual problems with and without random labeling, numerical problems with random labeling, and interactive applet problems. After illustrating these problem types, we discuss how effective they seem to be in terms of student learning based on data gathered in the Fall of 2000.

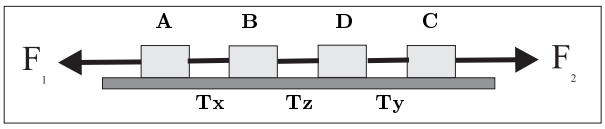
SAMPLE EXERCISES

A total of 260 questions were assigned as homework problems for Fall 2000. Below we give examples of the various question types with our classification of their characteristics. Note that we will not discuss essay questions as only two such problems were assigned during the Fall of 2000.

To appreciate these questions from the perspective of students, we urge readers to propose a solution on their own. Answers to all questions are listed at the end of this section.

* Supported by the Andrew W. Mellon Foundation, the Alfred P. Sloan Foundation, and the National Science Foundation.
[1] Department of Psychology, Michigan State University, East Lansing, MI 48824, kashyd@msu.edu
[2] College of Natural Science, Michigan State University, East Lansing, MI 48824, albertel@msu.edu
[3] Science Teaching Center, Hebrew University, Jerusalem 91904, Israel, guy@fh.huji.ac.il
[4] Department of Physics & Astronomy, Michigan State University, East Lansing, MI 48824, kashy@nscl.msu.edu
[5] Department of Physics, Florida State University, Tallahassee, FL 32306, ng@phys.fsu.edu
[6] Department of Physics & Astronomy, Michigan State University, East Lansing, MI 48824, thoennessen@nscl.msu.edu

0-7803-6669-7/01/$10.00 © 2001 IEEE **October 10 - 13, 2001 Reno, NV**
31st ASEE/IEEE Frontiers in Education Conference



The questions are presented below in the same format in which they appear on the printed assignments.

A number of questions assigned in the Fall of 2000 dealt essentially with mathematical skills which students should have acquired before enrolling in the course. We have examined data from this particular course for several years now, and we have found that mathematical skills are very important for success in the course. The mathematical problems were assigned mostly at the beginning of the course. Question 1 below is a typical example. The numerical coefficients in each equation differ among students, subject to the constraint that the paths never cross.

1. [2pt] The coordinates ... two objects ... along the x-axis are given below as funtion of time t. x_1 and x_2 never have the same value. Calculate the value of ... when the ... are nearest to each other.

$$x_1 = 2\ .0t$$

$$x_2 = -33.74 + \ .0t - 8.0t^2$$

The second question shown is typical of the numerical problems which constitute the majority of homework assignments. The text of the question is the same for all students but the numerical values differ for each student.

2. [1pt] A 87.6 ... kg (weight = 192.7 lb) is walking towards ... patio at 1.40 m/s (3.13 mph) when, at ... recognizes ... He ... in 120.0 ms. Find the magnitude of the average force ... hand.

The third problem illustrated is also a numerical problem, but has a figure associated with it. The order in which the values of the masses A, B, C, and D are given in the text is identical for all students. However, the labels identifying the four masses A, B, C, and D, as well as the three tensions Tx, Ty, and Tz, are randomly located for each individual student. Thus, for problem 3 students must use the picture of the physical situation in order to solve the problem because the solution is specific to their figure.

3. [1pt] Four blocks ... surface.

The blocks are connected by thin strings with tensions **Tx**, **Ty**, and **Tz**. The masses of the blocks are **A**=8.00 kg, **B**=2.00 kg, **C**=8.00 kg, **D**=9.00 kg. Two forces, F_1=94.00 N and F_2=13.00 N ... the masses ... Assume that friction between the masses and the surface is negligible and calculate ... **Tz**.

Coding questions with randomized labels like question 3 is more complex than the more basic numerical problems. The task is facilitated by pre-coded templates. Note that the figure and individualized labels are displayed for each student on both the printed assignment and on the Web.

In addition to the numerical problems, we have also developed a range of conceptual questions to assist students in developing a broader understanding of the material. Below is a typical example that deals with the Doppler effect.

4. [1pt] John ... He ... the frequency ... horn ... If ... hears ... several ... is ... general ... (Give ALL correct answers, i.e., B, AC, BCD...)

A) Both ... having ... same ...

B) Both ... having ... different speeds.

C) John is ... towards ...

D) The distance between John and the horn is increasing with time.

E) ... cannot ... direction.

F) Both ... directions.

For half the students, the pitch heard in the question above is 270 Hz rather than 300 Hz, and that while all students see the same general concepts, those concepts are presented in a different order and may be worded differently.

Question 5 below is another conceptual question about the Doppler effect. Its content which represents data is displayed in graphical form and the various curves are randomly labeled. Thus for some students the highest observed frequency curve is labeled "5", as is the case below, but for others it is "1" or "2" or some other value.

5. [2pt] Amy, standing near a straight road, records the ... sound ... along ... The recorded frequency as function of the car position along the ... is graphed.

For each statement select T-True, F-False, G-Greater ... equal ...

A) The highest ... is ... '**5**'.

B) The ... is ... '**2**'.

C) The ... '**1**' is ... to that ... '**4**'.

D) When ... way, the ... frequencies ...

E) Initial ... '**1**' ... distance ... in '**4**'.

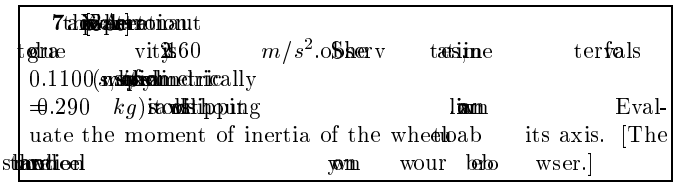
F) The ... '**1**' is ... to that ... '**3**'.

The random labeling makes it much more difficult for students to trade answers without some potentially valuable collaboration taking place. Students helping one another must look at each other's figures because the answer for any given statement depends on the particular labeling. Note that in both questions 4 and 5, the correct answers to all parts must be entered or else the computer simply responds "no" without any indication as to which part is incorrect. This forces students to review their answers, confirming the part they considered correct and working further on areas of uncertainty.

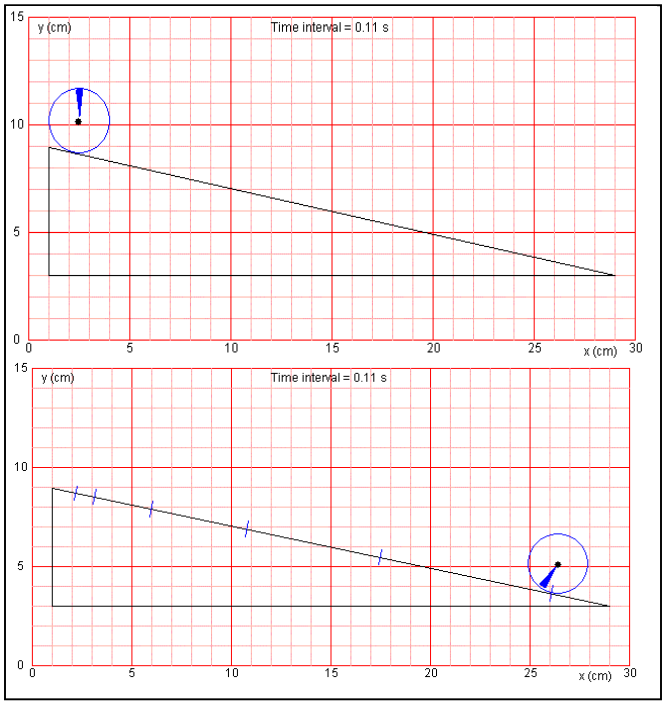
The five different curves on the graph shown in problem 5 also form the basis for a numerical problem assigned at the same time, i.e., to calculate the speed of the car for one of the curves. Its solution requires reading data from the graph so that there are 4 possible versions. While the computer was used to randomly label the curves, it did not generate an individualized graph for each student. We are currently developing interactive applets to generate such graphs and pass their randomized parameters through CAPA to evaluate the students' responses. This will increase the number of versions into the hundreds, which is typical of most individualized numerical problems used.

In the conceptual problem shown below (problem 6) the physical components rather than the data are randomly labeled (which was the case in question 5). The essential aspects are (1) that a "massless, frictionless" pulley only changes the direction of the tension, and (2) that bodies accelerate according to the total (net) force acting (Newton's Second Law). These ideas were discussed and several in-lecture demonstrations were carried out because these basic concepts give students considerable difficulty. The three tensions and two masses are randomly labeled. There are several versions of each statement but each statement addresses the same concept and the order in which statements are presented varies among the students.

The individualized applet problems included a variety of accelerated linear and rotational motions in one and two dimensions. The text of one of these is shown in question 7 below.



The applet is illustrated below, where its initial and final states are shown. As it rolls, the wheel leaves marks at fixed time intervals at the point of contact. Students may replay the motion at will, and each student's applet has a different set of parameters. Thus the applet is an experiment, requiring students to collect the data with some accuracy and then to make use of their understanding of the concepts to solve the problem.



The reader may view applets at:
http://capa4.lite.msu.edu/devolibrary/Links/FIE.html.

Answers to sample exercises:
#1 22.25 #2 1.02x10^3 N #3 64.0 N #4 ABCF
#5 TTEFLL #6 GELEEF

RESULTS

In this section we describe how students performed on the various problem types in Fall of 2000. We examine characteristics of the problems themselves (e.g., success rates and number of attempts before a correct solution is derived) and we attempt to evaluate how effective the various problem types seem to be in promoting understanding of the subject matter. Our index of general understanding is student performance on the comprehensive two hour final exam.

Table I shows the mean performance on the six problem types, the average number of tries, the hours from initial entry of answers until a correct answer is entered, and the correlation between performance on the problems and performance on the final examination.

TABLE I

MEANS FOR CHARACTERISTICS OF THE SIX PROBLEM TYPES AND CORRELATIONS OF PERFORMANCE WITH FINAL EXAM

Question type	Performance in percent	Number of tries	Hours until correct	Correlation with final*
Mathematical	90	2.5	5.5	0.45
Numerical	94	2.1	3.6	0.39
Conceptual	96	2.6	5.2	0.29
Random label Numerical	91	3.0	6.1	0.41
Random label Conceptual	92	3.9	8.8	0.42
Applet	63	5.0	9.9	0.48

* All have $p < .001$. Performance in percent is an average of the number of each problem type a student was able to solve. Number of tries is the average number of solutions students entered, and Hours from first try until correct is the average length of time between the first computer entry and the computer entry of the correct response.

Two points are quite clear from these data. First, students successfully completed the vast majority of the problems assigned – across all problem types the average was 87.7 percent. Second, the Applet problems clearly stand out from the other problem types in that students only solved an average of 63 percent of them. The Number of tries and Hours until correct columns echo this pattern, with students trying the applet problems more frequently on average than the other problem types, and relatedly, taking more time until successful completion. Notably, the random labeling problems appear to be somewhat more challenging than standard format problems. For both conceptual and numerical problems, the random labeling versions required an average of 1 additional try until students solved them.

The final column in Table 1 shows the correlations between students' average performance on a particular problem type, and their score on the final exam. Not surprisingly, these correlations indicate that students who performed better on the homework tended to score higher on the final exam, and this relationship seems to be slightly stronger for Applet problems.

One significant characteristic of the Applet problems was that a fairly substantial number of students (15 percent) never even tried to solve them. This directed us to the question of who those students were and how they performed in the class. For each of the seven Applet problems, we created a variable that was 0 if they never tried the problem, 1 if they tried the problem but never solved it, and 2 if they solved the problem successfully. Then we averaged over the seven problems. If the average was less than 1.0, we classified students into a group labeled "Rarely tried". If the average was 1.0 or greater but less than 2.0, we classified students as "Tried but sometimes failed". Finally, if the average was exactly 2, we classified students as "All correct". The number of students falling into these categories is included in Table II, along with the means on the Final exam for each group. A simple ANOVA test of whether these three means differed yielded an $F(2, 457) = 69.61$, $p < .001$. Clearly those who solved the applet problems were most successful on the final, followed by those who at least tried, followed finally by those who put in little effort.

TABLE II
FINAL EXAM MEAN PERFORMANCE AS FUNCTION OF TRYING AND SOLVING APPLET PROBLEMS

Student Action	N	Overall Mean	Mean (equating for covariates)
Rarely tried to solve	72	47.0	56.1
Tried but sometimes failed	209	62.5	62.1
Solved all correctly	108	75.7	70.3

Note: Covariates included the combined ACT scores; cumulative GPA up to, but not including this class; pre-existing knowledge of physics as measured by the Force Concepts Inventory [4]; and the number of times a student was absent from class.

Drawing a conclusion about the effectiveness of thinking through and solving the Applet problems is premature, however, because the three groups almost certainly differ on a host of variables related to performance on the final exam. That is, there are likely to be large differences in ability, effort, and commitment between these groups. We attempted to address this issue statistically by including measures of general academic potential, the ACT combined, academic performance (GPA from all courses prior to this one), physics knowledge prior to participation in

the class (measured on the first day of class using the Force Concepts Inventory [4]), and number of absences from lecture (number of quizzes missed out of 34 total) as covariates in the analysis. Even equating students on each of these important covariates, there were significant mean differences on the final exam as a function of performance on the Applet questions [$F(2,82) = 16.98$, $p < .001$]. The means equating for the covariates in Table II are a statistical forecast of what the means on the final would have been if the three groups were equivalent in their academic potential, academic performance, prior knowledge of physics, and number of absences.

The correlations between the final exam and the various problem types in Table I indicate that performance on all problem types relates to success on the final. To determine whether some types of questions were particularly helpful, we conducted a multiple regression analysis including performance on each of the problem types as predictors. Also included as predictors were the student's cumulative GPA, combined ACT, Force Concepts Inventory score, and number of absences. This analysis was rather poor because success on the mathematical, conceptual, and numerical problems (with and without random labeling) was very highly correlated – causing a multicollinearity problem. Even in this analysis, however, success on the applet problems was uniquely predictive of success on the final exam. To pursue this finding, for each student, we averaged over the problem types (excluding the applet problems) and computed the percent of all homework problems solved (other than the Applet problems). We then included success on the Applet problems, success on all other homework problems, cumulative GPA, combined ACT, Force Concepts Inventory score, and number of absences as predictors in a regression predicting the final exam score. Thus this analysis estimates whether there is something unique about performance on the Applet problems over and above performance on other homework problem types (as well as other measures of academic aptitude).

TABLE III

MULTIPLE REGRESSION RESULTS PREDICTING FINAL EXAM SCORES A FUNCTION OF VARIOUS MEASURES

Measures	Standardized Coefficient	t-value
Applet performance	.21	3.95**
Other HW Performance	.13	2.47*
Combined ACT	.05	1.37
Cumulative GPA	.28	6.69**
Force Concepts Pre-test	.14	6.39**
Number of absences	-.16	3.51**

Note. * $p < .05$. ** $p < .001$

Table III contains the results from this regression. It shows that success on the Applet problems represents a unique aspect in learning relative to other homework problems. The relative value of the two types of problems can be evaluated by comparing the standardized coefficients. These standardized regression coefficients are essentially partial correlations that remove the effects of every other predictor in the model. Thus, controlling for homework performance in general and various measures of academic success, success on the Applet problems provides a unique learning experience with respect to performance on the final exam.

DISCUSSION

An interesting perspective concerning individual cognitive performance on the various problem types is Bloom's taxonomy of educational objectives [5]. This taxonomy suggests that there are six cognitive levels involved in problem solving: recall, comprehension, application, analysis, synthesis, and evaluation. These levels are hierarchical such that performance at each level depends on performance on the previous ones.

Recall questions require students to provide or identify previously learned information in the same way it was presented to them. Comprehension questions require conversion of that information into a different format, such as translation between different representations. Some of the simpler conceptual problems likely fall into these two categories, which may account for the lower correlation between the conceptual problems and performance on the final exam shown in Table I.

Application questions require the use of previously learned information in new and concrete situations to solve problems that have single or best answers. Many of our numerical questions fall into this category. Although previous research shows that students can use algorithmic knowledge to solve application problems without having an understanding of the underlying concepts [7], the correlation between performance on the numerical problems and the final exam suggests that this may not be the case, since students who did well on those problems tended to do better on the final. The aim of conceptual and applet questions was to discourage such algorithmic approach. Our measurement of a strong correlation between students performance on conceptual and numerical problems on the final exam, $r = 0.63$, $p<.001$, in agreement with earlier results [1,2], indicates some degree of success.

Moving up Bloom's taxonomy, analysis questions require the breakdown of a problem description, design, or situation into its constituent elements such that the relations between these elements and the organizing principals are made explicit. The Applet problems we used fall into this category. Some of the more complex randomly labeled conceptual questions may also be in this category. We would suggest that the unique predictive power of

performance on the Applet questions is due largely to the fact that students who solved these problems were operating at an analysis level – and thus had achieved a higher level of understanding of the course material.

Finally, synthesis questions require applying prior knowledge and skills to produce a new or original whole and evaluation questions require judgment of the value, accuracy or appropriateness of a solution. For the levels of synthesis and evaluation, it seems that open-ended essay questions are the most appropriate. One of us (EK) has for several years read and graded numerous such essays from students in the course, and has concluded that only a few students reach these levels in the subject.

Our experience using network technology has sensitized us to the importance of the social interaction aspect of learning, and we have found that we can use network technology to facilitate social interaction among students, both through the network and face-to-face [1]. Table I shows that success in solving the randomly labeled conceptual problems is more predictive of final performance than regular conceptual problems. We believe that part of the difference stems from the difference in communication level utilized during collaboration in solving these problems.

Consider a case in which two students are working on problem 6, and student A states that "T_z plus T_y is equal to T_x". Student A has considered the physical problem, but his words alone carry no physical meaning, just a relation between three mathematical variables. However, if both students had the same labels in their problem sets, this sentence would carry enough information to correctly match item (D) in the problem, and student B would not have to consider the physical meaning behind the mathematical relation. This is not the case if the labels are randomized. In order to deliver the same information, student A would have to rethink his own problem in terms of physical concepts in order to translate his statement for student B's labeling. Student A would have to say "The sum of the two downward tensions on the pulley is equal to the upward tension", or at least point out the relevant tensions in student B's diagram. Student B would have to consider the physical meaning behind A's words in order to apply the information to his own problem. Therefore, both students would have to actively manipulate physical concepts in their minds rather than just words, operating at a higher cognitive level. At this level communication catalyzes the individual construction of meaning - Active participation in a discussion forces the students to form their own meaning of the concepts in their minds before they can verbally articulate these meanings. This view of social interaction is part of the educational theory of social constructivism, put forward by Vigotsky [7], who stated that "thought is not merely expressed in words; it comes into existence through them."

CONCLUSIONS

The results presented indicate that it is indeed important to devote the considerable resources required in developing individualized interactive exercises as they appear to enhance both individual cognitive performance and social interaction. This task can be considerably reduced by better sharing and communication of such resources, as we hope will be the case in the newly established LON-CAPA collaboration which is a project to create a networked resource pool [8].

ACKNOWLEDGMENT

The support of the Andrew W. Mellon Foundation, the Alfred P. Sloan Foundation, and the National Science Foundation is gratefully acknowledged.

A large number of students over the years have given us feedback as they have struggled with numerous challenging problems. Many gave constructive feedback, others expressed great satisfaction, both in words and body language when CAPA answered YES to their entry, and more than just a few have expressed frustration at both the amount of work and its difficult nature. We thank them all; their comments have helped us greatly.

REFERENCES

[1] Kashy E., Thoennessen M., Tsai Y., Davis N.E. and Wolfe S.L., "Using Network Technology to Promote Student Success in Large Classes", *Journal of Engineering Education*, October 1998, pp 385-390.

[2] E. Kashy, G. Albertelli, M. Thoennessen, Y. Tsai, and D. A. Kashy, "ALN Technology on Campus: Successes and Problems" *Proceedings of the 2000 Frontiers in Education Conference*, IEEE Catalog Number: 00CH37135, ISBN: 0-7803-6424-4, Session S1D (2000)

[3] http://capa.msu.edu/ The system software is free, Licensed by Michigan State University under GNU-General Public License.

[4] Hestenes, D., Wells, M., and Swackhamer, G., "Force Concept Inventory", The Physics Teacher, Vol. 30 No. 3, 1992, pp 141-151. (New version downloaded from the Web)

[5] Bloom Benjamin S. and David R. Krathwohl, (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals, by a committee of college and university examiners. Handbook I: Cognitive Domain*. New York, Longmans, Green, 1956.

[6] Vigotsky, L.V., *Thought and Language*. Cambridge, MA. MIT Press, 1962.

[7] Sawrey, B.A. "Concept learning versus problem solving: revisited". *Journal of Chemical Education*, Vol. 67, (1990). pp. 253-254.

[8] Kortemeyer G., Bauer W., Kashy, D.A., Kashy, E., and Speier, C. *"The Learning-Online Network with CAPA Initiative",* FIE2001 paper #1003, See also http://www.lon-capa.org/