Teaching with ALN Technology: Benefits and Costs *

DEBORAH A. KASHY
Department of Psychology
Michigan State University

GUY ALBERTELLI
College of Natural Science
Michigan State University

EDWIN KASHY
Department of Physics and Astronomy
Michigan State University

MICHAEL THOENNESSEN
Department of Physics and Astronomy and
National Superconducting Cyclotron Laboratory
Michigan State University

ABSTRACT

The implementation of Asynchronous Learning Network (ALN) technology in a large on-campus course over several years is reviewed, and recent data concerning both educational and cost effectiveness are presented. Even with higher course standards for success, student performance on examinations has improved, a larger fraction of students achieve the goals of the class, and the proportion of students who excel has increased. Female students benefit even more than their male counterparts. The level of communication and interaction among students has also increased dramatically, with mostly positive (but some negative) effects. Data concerning cost effectiveness indicate that the technology can reduce costs, but perhaps more importantly, it can increase the quality of education without increasing costs.

I. INTRODUCTION

This paper briefly summarizes the principal results obtained with the use of CAPA (Computer Assisted Personalized Approach) as an ALN tool. The main question considered is how the use of this tool has impacted student performance, but we also present preliminary results on how that use has affected costs associated with the courses.

CAPA is a network tool designed to facilitate and improve the way assignments, quizzes, and examinations are provided and graded in large enrollment courses, and was first implemented for a 90-student class in 1992 [1]. The CAPA system provides immediate feedback to students (and instructors) on conceptual understanding and correctness of solutions and it also includes a variety of statistical and course management features [2]. The very positive initial student response as well as the interest of instructors in other fields provided the impetus to continue to develop and expand the system's capabilities. It was licensed at over 50 institutions before becoming free software in June 2000, provided by Michigan State University under GNU General Public License [2].

II. BACKGROUND

To assess the educational impact of ALN, both subjective and objective measures have been used. They include (1) surveys of students (2) instructors comments, and (3) examination performance and course grade distributions.

Surveys of students indicate that students' attitudes have been and continue to be highly positive, and students give the system high grades for helping them learn and understand [1,3-6]. Instructors also have expressed satisfaction and often enthusiasm [7]. Investigation of exam performance and grade distributions has demanded some control on the performance standards used and grading methods employed. Briefly, an independent instructor evaluated the difficulty of examinations, and grades are based on an absolute grading scale [8]. With fixed scale assignment of grades, exam performance and grade distributions can be compared across sections of the course. We note, however, that evaluating educational effectiveness of new teaching methods represents a uniquely difficult endeavor because there are few, if any, widely used standards that provide a reliable and objective scale on which student achievement can be measured.

Previous analyses of data generated by on-campus physics courses using the CAPA system have resulted in a
number of findings, many of which also have been replicated in other disciplines and at other institutions [3,5,9-11]:

- A majority of students, typically 80%, consider that CAPA helps them learn and understand the course material.
- The technology allows high standards to be implemented while providing students with the opportunity to achieve those standards.
- Time students spend working on assignments and other course requirements has increased by nearly a factor of 2 and approaches the recommended 2-hours outside of class per lecture hour.
- Frequent assignments with firm electronic deadlines keep the course on schedule and help to inhibit the tendency of some students to procrastinate and fall behind.
- Allowing multiple tries on assigned problems with no penalty is highly motivating; most students strive to get all the work done correctly.
- There is a high level of interaction among students and between students and staff. A smaller teaching staff can do more for (and with) students. Reassigned staff can provide a greater level of Socratic interaction with individual students.
- Scores on examinations show a substantial increase, even with higher standards and harder problems.
- Controlling for standards, a larger proportion of students succeed, and an even greater proportion excel in the course, with female students benefiting even more than males.
- With the computer doing the grading, the instructor is viewed by students as a mentor rather than a judge.
- Developing and testing materials well adapted to the technology, especially material that is designed to improve conceptual understanding, is costly in both time and effort.
- Interactions and collegiality has greatly increased among faculty sharing their experience.
- Faculty satisfaction is high in spite of the increase in the level of effort required.

III. CAPA FEATURES

The CAPA system consists of several applications. The system generates unique assignments for each student and students enter their responses online. The system provides students with immediate feedback concerning correctness, as well as specific information on a wide variety of errors that are likely to occur: units, significant digits, and various formatting errors that prevent a proper interpretation of the student’s answer. It also includes a readily accessible discussion forum in which students can interact with one another as well as with the teaching staff.

For the instructor, the CAPA system provides a broad set of pre-coded templates and examples that facilitate the coding of a variety of problem types. These include:

- Numerical problems for which each student gets randomly assigned values for the variables.
- Problems in which students select all the correct statements from a list of randomized statements.
- Matching problems in which a list of randomized statements must be matched with appropriate conclusions.
- Problems asking students to identify features in a diagram, with both the labels and the list of features randomized.
- Problems that require a graphical solution.
- Problems that require students to rank a set of items based on quantitative or qualitative measures.
- Problems with an answer consisting of several parts (and).
- Problems with more than one correct answer (or).
- Open-ended and essay questions that are read and evaluated by the instructor using key-word highlighting by the computer.
- Problems with expressions as answers where students enter symbolic formulas.
- Individualized applets as a basis of interactive questions
- Questions for which sound and/or speech is the medium.

We have recently examined the relative effectiveness of a range of these types of problems by examining how well success on the problem types predicts success on the final exam [12]. Our results indicate that, as would be expected, success on homework problems predicts success in the course. Notably, success on individualized interactive applet problems provides a unique predictive ability over and above success on the other problem types. These applet problems tend to be much more costly in terms of resources needed to create them, but such efforts may be well worth while if students learn more through solving them.

A list of the features of the three instructor modules in the CAPA system is given in Table 1.

IV. NEW RESULTS

A. Educational Impact

In this section we report some new results concerning the impact of ALN technology on student learning. To see how technology has changed student performance, data from a physics course, Physics for Scientists and Engineers (Phy183), during Fall 1999 and Fall 2000 is compared with that of previous years. This is a 500-student introductory course, typically taught in two lecture sections.

Technology was used in essentially every aspect of the course. Individualized midterm examinations were administered in class. After the exams students were given the opportunity to redo the exam online, but with different
and 2000 and resulted in 90% average attendance, a quizzes were given throughout the semester in both 1999 class meetings. Numerous unannounced individualized examinations were a combination of both summative and improved when they redid the exam. Thus midterm numerical values and different versions of randomized conceptual problems. As an incentive, exam grades were then computed by factoring in how much the students improved when they redid the exam. Thus midterm examinations were a combination of both summative and formative assessment [13,14]. Lecture time was made more active with participatory exercises included during half the class meetings. Numerous unannounced individualized quizzes were given throughout the semester in both 1999 and 2000 and resulted in 90% average attendance, a remarkably high value. In addition to examinations and quizzes, individualized homework assignments were given and students were able to enter their responses online and obtain immediate feedback concerning the correctness of their responses. When homework assignments were answered incorrectly, students were allowed to reevaluate their answers and try again. Students were allowed up to 20 tries for each problem. A web-based discussion forum that is part of the CAPA system provided a means for students to share their thoughts and questions with other students as well as with the instructor and teaching assistants. One question we studied is whether students who succeeded in solving numerical problems also showed evidence that they understood the underlying conceptual material. In particular, we were interested in determining whether our students were learning at a level above a purely algorithmic “plug-and-chug” (i.e., finding a formula to plug in variables and grinding out the answer) approach. In 1999 40% of examination questions were conceptual and in 2000 46% of exam questions were conceptual, reflecting the importance we placed on understanding scientific concepts. Students worked on similar questions as part of their homework assignments. The questions were designed to help students appreciate the concepts underlying quantitative numerical solutions. Some questions were designed to lead students to resolve misconceptions on their own by confronting them with contradictions, while others illustrated concepts from a variety of situations [15-19]. The overall correlation coefficient between performance on conceptual problems and computational problems, summing over all problems on the midterms and final exam, was $r = .70, p < .001$ for 1999 and $r = .72 p < .001$ for 2000. These strong correlations suggest that the students who succeeded on the numerical problems also tended to succeed on the conceptual problems. The strong relationship between conceptual understanding and numeric problem solving suggests that our initial emphasis on creating templates in CAPA for a variety of conceptual problem formats was a good design choice. Several of these features are found in other similar systems [20-26]. As noted above, there have been several indications that CAPA allows a greater number of students to achieve the goals of the course, and simultaneously, provides an environment in which a greater proportion of students can excel. Figure 1 compares the 1999 and 2000 grade distributions to previous years by showing the evolution of the distribution of grades starting in 1992 and illustrates the large change associated with the implementation of ALN technology. The 1992–1994 histogram represents the grade distribution when the course was taught in the traditional manner. 1995 was a transition year (not shown) when the first ALN was implemented. The proportion of students with grades of 2.5 or above that year was essentially unchanged from 1992–1994 reflecting in part higher standards and our lack of experience in ALN use [8].

### Table 1: Three CAPA modules for the instructor.

<table>
<thead>
<tr>
<th>Module</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. QUIZZER:</strong></td>
<td>Multifaceted editing tool for preparing homework, quizzes, and examinations</td>
</tr>
<tr>
<td>- Prepares materials in three formats: ASCII, HTML, LATEX</td>
<td></td>
</tr>
<tr>
<td>- Each student receives unique questions and problems</td>
<td></td>
</tr>
<tr>
<td>- Over 180 pre-coded templates to facilitate creation of numerous types of questions</td>
<td></td>
</tr>
<tr>
<td>- Allows printing of text and graphics in a compact, efficient manner</td>
<td></td>
</tr>
<tr>
<td>- Due dates can be set for individual sections independently</td>
<td></td>
</tr>
<tr>
<td>- Includes a timed entry option for use with take-home quizzes and exams</td>
<td></td>
</tr>
<tr>
<td>- A simple transformation allows conversion from homework style to exam style</td>
<td></td>
</tr>
<tr>
<td>- Provides the range of answers for a question across all students in a class</td>
<td></td>
</tr>
<tr>
<td>- Efficient assembly of existing problems from problem libraries</td>
<td></td>
</tr>
<tr>
<td><strong>2. MANAGER:</strong></td>
<td>Course management and statistical analysis tools</td>
</tr>
<tr>
<td>- Provides distribution of grades for an assignment</td>
<td></td>
</tr>
<tr>
<td>- Instructor can examine number of attempts made by students for each problem</td>
<td></td>
</tr>
<tr>
<td>- Analyze answer patterns to detect misconceptions: correlations between items, degree of discrimination and degree of difficulty</td>
<td></td>
</tr>
<tr>
<td>- Course summaries for individual students can be generated for advising purposes</td>
<td></td>
</tr>
<tr>
<td>- Interprets and grades output of scanned forms where the pattern of correct responses varies</td>
<td></td>
</tr>
<tr>
<td>- Can send semi-personalized e-mail to students based on performance</td>
<td></td>
</tr>
<tr>
<td><strong>3. GRADER:</strong></td>
<td>Additional grading tool that supplements online self-grading by students</td>
</tr>
<tr>
<td>- Allows instructor to grade subjective answers such as essays efficiently</td>
<td></td>
</tr>
<tr>
<td>- Provides answers for individual student’s assignments for hand-grading</td>
<td></td>
</tr>
<tr>
<td>- Allows a problem to be excused for an individual student, section, or class</td>
<td></td>
</tr>
</tbody>
</table>
With ALN technology, there is a clear shift away from the traditional bell-shape in the 1996–1998 years as well as in 1999 and 2000. The proportion of students excelling (i.e., earning a grade of 3.5 or 4.0) is 32% in 1996-1998, 32% in 1999, and 36% in 2000. These values are substantially higher than the 20% seen in 1992-1994 with a traditional lecture course. It appears that with the ALN technology used in this course, motivated students are able to overcome deficiencies in preparation through hard work. Perseverance is rewarded in the sense that students who arrive at incorrect solutions initially, can potentially solve all homework problems after correcting their errors and misconceptions.

In 1996–1998, 78% of the Phy183 students achieved the course goals (i.e., had grades of 2.5 or above). This represents an 18% increase over 1992-1994, so that each year approximately 90 more students achieved the required level of understanding. In 1999, 70% of the students achieved the goals. This represents a decrease from the 78% average in the previous 3 years. However, in 2000, 77% of the students had grades of 2.5 or above. Thus, the decrease in 1999 may represent a random fluctuation.

An alternative explanation for the dip in 1999 suggests that it may represent a real decrease in student achievement. An enterprising student developed an elaborate web discussion forum where students could get answers and formulas, often with little understanding, thus defeating the goals in the design of most individualized numerical and conceptual problems. A recent analysis shows that students’ use of that web site is negatively correlated with examination performance, \((r = -0.35, p < .001)\), i.e., students who used it more tended to score lower on midterm exams, quizzes, and the final. This web site uses technology to promote “plug-and-chug” problem solving, which is quite the opposite of learning and understanding [27]. In contrast the correlation for students using the discussion site provided with the course was positive [28]. These results were described to the students in 2000, and the warning may have been sufficient to neutralize the negative impact of the student-run web site.

To respond to the new challenge posed by the student-run web site, we have been working to develop new problem formats that make copying without understanding much more difficult, while not increasing the actual complexity of our problems. The new formats include individualizing the labeling of figures in numerical problems, providing data in histogram or graphical form, and individualized applets in which the students’ actions are transmitted to CAPA for analysis. Versions of such a problem for two different students are shown in Figure 2. Values and labels can differ as well as their positions on the diagram. In this case, the position of the masses and the identification of the angle both vary from student to student. Communications among students for these types of problems is likely to be instructive on either web site since no single plug-in formula solves this relatively straightforward problem. Thus students who post solutions will have to present an explanation of the physical situation in the problem rather than simply giving a formula. Our recent analysis of the effectiveness of a variety of problem types indicates that these new formats represent an improvement over the traditional versions [12].

A somewhat different issue relating to educational effectiveness is whether the effects of ALN technology differ for male and female students. A unique opportunity to examine this question is provided by an earlier physics course in which the first semester of the course, Phy231, was taught using traditional methods and the second semester of the sequence was taught using CAPA [4]. A total of 267 students participated in both parts. Figure 3 shows the grade distributions for male and female students, and one can see a shift indicating that female students’ grades tended to show greater improvement during the second semester (using CAPA) relative to male students’ grades. The gender impact is more clearly seen by plotting the difference as shown in Figure 4, where the women show a larger change towards the higher grades.
2. [2pt] A 4.30 kg beam has a length 1.30 m and is suspended in a horizontal position as shown. There are 10 equally spaced attachment points, 13.0 cm apart with three masses hanging from the beam. A thin cable attached 13.0 cm from the end makes an angle of 53.0° with the wall as shown.

![Diagram of beam with masses and cable](image)

The masses are N = 8.00 kg, O = 6.00 kg, P = 3.00 kg. Calculate the tension in the cable.

2. [2pt] A 3.90 kg beam has a length 1.20 m and is suspended in a horizontal position as shown. There are 10 equally spaced attachment points, 12.0 cm apart with three masses hanging from the beam. A thin cable attached 12.0 cm from the end makes an angle of 35.0° with the wall as shown.

![Diagram of beam with masses and cable](image)

The masses are N = 4.00 kg, O = 8.00 kg, P = 5.00 kg. Calculate the tension in the cable.

Figure 2: Two versions of the same problem.

To better judge these data, we compared the average grade change for this sequence with one in which both semesters were taught using traditional methods (N = 428). With CAPA the average grade change for women was greater than that for men by 0.16 (on a 4.0 scale) [29]. A similar observation was made in a microeconomics course. Male students’ grades in the transition from traditional to computer enhanced rose from 2.63 to 2.73 while those of female students rose from 2.36 to 2.72, with a preliminary analysis pointing to higher homework grades as the reason [30].

It could be argued that the greater improvement in final grades for women is simply the result of women earning higher overall homework grades. Although that may be part of the story, we wanted to investigate whether there was more going on. Note first that for Phy183 the average homework scores of women, 92% in 1999 and 91% in 2000, were not very different from the 90% for men in both 1999 and 2000. To investigate whether there were differences between men and women beyond homework performance, we looked at the exam performance over the course of the semester for men and women. Our analyses indicated significant interactions between gender and time for both
years. Figure 5 shows the women's average exam scores relative to those of men for the two years.

For 1999 men had an average score on Exam 1 (given approximately four weeks into the semester) that was 18% higher than women, a statistically significant difference. At Exam 2, seven weeks into the semester, men again outperformed women, but their score was only 6% higher, a marginally significant difference. At Exam 3 and the final exam, the gender gap has virtually disappeared with women's average score essentially equal to that of men.

The pattern of exam performance over time for men and women in 2000 supports the 1999 findings in that they also indicated that women initially performed more poorly than men and that they caught up. At Exam 1, men outperformed women by about 11%, again a statistically significant difference. In 2000, unlike 1999, that difference disappeared at Exam 2, but a marginally significant difference was evident for Exam 3 such that men outscored women by about 7%. Consistent with 1999, in 2000 men and women did not differ in their Final exam performance. Perhaps if we can understand the causes of the gender effect on exams, we may be able to better address the needs of other learners who have had less success in science courses.

We have also studied factors correlated with performance on the final exam for Fall 1999 and Fall 2000 and found them consistent with earlier results [8]. Correlations with quizzes (1999: \( r = .66 \); 2000: \( r = .68 \)), midterm exams (1999: \( r = .72 \); 2000: \( r = .72 \)) and homework (1999: \( r = .36 \); 2000: \( r = .46 \)) are all strong and statistically significant (all p’s < .001). Attendance in class is also an important factor relating to performance on the final, with number of absences negatively correlated to final exam performance (1999: \( r = -.37, p < .001 \); 2000: \( r = .42, p < .001 \)). To allow for the fact that these correlations could be driven by intelligence level (i.e., smarter students do better on homework and better on the final), we recomputed these correlations partialling out the student’s composite ACT scores. The partial correlations were very similar to those presented above, and all remained statistically significant.

### B. Cost Impact

Does ALN technology reduce costs in education? Preliminary data from an ongoing study of costs in university education indicate that ALN technology can be used to reduce costs under some circumstances [31]. The key variable seems to be how instructors approach homework assignments and other faculty or teaching assistant responsibilities.

Our experience with ALN technology suggests that it can reduce costs when faculty and teaching assistants are involved in grading assignments and keeping records of student performance. That is, when educational staff members are responsible for repetitive, time consuming tasks. In the Physics for Engineers course (Phy183) we have greatly reduced costs by removing recitation sections, many of which were taught by university faculty. These faculty were responsible for leading the sections, grading, and record keeping. A similar result has occurred in our large introductory chemistry courses in that prior to implementing the technology, a large number of teaching assistants were used as graders for student assignments. Removing most of the instructors’ grading responsibility has an additional benefit: They have more time to interact with and teach their students.

A much broader impact of technology comes not from actual cost savings, but rather from improving the students’ educational experience while maintaining current costs. In disciplines for which large problem libraries have been developed (e.g., physics and chemistry), instructors can generate online assignments simply by selecting appropriate questions for their students. If such libraries are not available, however, an initial investment of faculty and teaching assistant time is necessary.

We would also like to note that the technology can and has been used in disciplines outside of the physical sciences (Food Services, Microbiology, etc...). For example, Introductory Psychology courses are typically taught in very large sections, (i.e., ranging from about 200 to over 1000). Sheer logistics generally limit requirements in these courses to several multiple choice examinations. Having students write brief “thought” papers during the class may be a desirable component of the course, but even logging in whether or not students complete such papers is a time-consuming task. CAPA allows students to answer open-ended questions, and then automatically links responses to the students’ records. Instructors have also used the system’s ability to support conceptual questions to develop weekly homework assignments that ask questions based on reading assignments. In a large introductory astronomy course, this appears to have helped motivate students to keep up with the reading assignments as indicated by the forum postings giving the location of relevant sections in the text [32].

![Figure 5: Relative difference between female and male performance on exams in Phy183.](image-url)
V. CONCLUSIONS

There are a number of studies that have reached the conclusion that technology has made “no significant difference” in student education [33]. Our experience and our data suggest that technology can have a profound impact on learning if it is used in a way that capitalizes on its unique ability to “interact” with students, provide them with immediate feedback, and facilitate interactions among students and between students and teaching staff.

One of the most significant challenges of implementing ALN technology in the classroom is the preparation and testing of questions and problems. Coding high quality, thought provoking problems is time consuming, and poor debugging of new problems can cause substantial frustration on the part of students. Libraries of problems and questions, as well as a broad range of animations and simulations are available. A similar situation exists for chemistry, and such materials are being developed for other disciplines. Several publishers now have a large fraction of the problems in their physics texts coded for CAPA. However, the effort required when first implementing an ALN in a new discipline can be very large in spite of great progress in facilitating the use of technological tools. This is especially true when the new opportunities offered by technology are used to do more than deliver traditional materials in a new and efficient method.

In assessing educational effectiveness of ALN technology, one may need to differentiate between technology-mediated courses and technology-enhanced courses. We have demonstrated a continuing improvement in student achievement when network technology is used to complement and enhance on-campus courses. More students succeed and excel. Prompt feedback and increased student time on task rank high among contributing factors. One issue that we need to continue to address is the negative impact of student-created “cheating” web sites that allow students to nominally succeed on assignments in the absence of actual learning. This difficult issue requires a solution that does not infringe on students freedom to communicate. Another challenge is the need to develop improved discipline-specific standards. Such standards would be of great help in assessing educational effectiveness more objectively, and could also lead to a better understanding of how educational materials can be developed or adapted to better take advantage of the technology.

ACKNOWLEDGEMENTS

We wish to acknowledge the support we have received over the years from many colleagues at MSU and at other institutions. In particular we thank Prof. Burks Oakley II of the University of Illinois for his early interest in our work and his continuing support. We also wish to thank the Alfred P. Sloan and Andrew W. Mellon Foundations as well as the National Science Foundation for financial support.

REFERENCES

2. http://www.pa.msu.edu/educ/CAPA
http://msuvmall.msu.edu/imc/
9. Mader, K., G. Peaslee, “CAPA4.6 and Discuss at Hope College”, CAPA as ALN Teaching Tool Workshop, Michigan State University, Feb. 1999


32. Sherrill, B.M., private communication.