Innovative instruction or technological fetishism: 
An evaluation of education technologies at community colleges

Terrence Willett © 2001

Abstract

From computer adaptive testing to distance learning to computer assisted lecturing, technology has begun permeating our instruction. While the rapid infusion of technology into education has critics, possible benefits include increasing access to education to underserved populations and enhancing learning outcomes. This paper provides a synthesized evaluation of education technology at Cabrillo and Gavilan Community Colleges using examples including distance learning of Spanish, computer mediated math instruction, computer adaptive placement testing, and an experimental pilot testing recall of content delivered in traditional lecture style versus a lecture mediated by computer generated slides. Overall, the impact of technology has had mixed results in improved performance indicators. However, most evaluations are hampered by a lack of controls. These findings suggest that the adoption of instructional technologies be preceded with a review of relevant research and implemented in consultation with the campus research staff to maximize the explanatory value of evaluations.

Introduction

From computer adaptive testing to graphing calculators, distance learning to web based instruction and computer assisted lecturing, technology has begun permeating our instruction. The purported benefits of these adaptations include increasing access to education to underserved populations and enhancing learning outcomes (Hall and Barker 1995; Reeves 1998). There exist great possibilities to use technological innovations as superlative visualization and auditory aids or as infinitely patient tutors who will tirelessly and uncritically drill students in a highly individualized manner. Web publishing of course syllabi, notes, assignments, and listserves allow students 24-hour worldwide access to essential course materials. Students can perform library research from home and colleges can conduct course placements with greater accuracy in a shorter time with instant feedback at a lower cost.

The rapid infusion of technology into education has its critics who worry that educational quality may be compromised in some cases (Nissambaum and Walker 1998; Rintala 1998; CFT 2001). Specifically, they indicate the danger of the ideological shift from education as a public duty to education as a market. This concern heightens when a technology is closely tied to a particular private company whose sole objective lay in increasing profits derived from the sale of a product rather than improving education. Further, there exist fears among labor activists that technology could be inappropriately used to replace teachers or support staff. Finally, these new technologies are not equally available to all groups.

What is needed is an evaluation of these new techniques to test their efficacy compared with traditional methods and to ensure they are at least as effective and that they do not create an unmitigatable bias against certain groups. Ideally, a technological enhancement would lead to improved access, retention, or skills acquisition or more accurate and convenient course placement over traditional delivery methods. This report provides summaries of studies performed recently at Cabrillo and Gavilan Community Colleges on the Central Coast of California and provides a synthesis of the impact of technology in education and suggestions for improving instructional technology evaluation.
**Distance Learning of Spanish: Improving Success or Access?**

To evaluate distance learning efforts in Spanish at Cabrillo College, comparisons were made of retention, success, and persistence rates between Spanish telecourses (Destinos) and traditional classroom Spanish. Data were obtained from 5155 students who took Spanish 1 or 2 between Spring 1996 and Spring 1999. All beginning Spanish courses in general had significantly lower retention and success rates than the second level of Spanish. Destinos rates are mostly below classroom rates although the gap narrows in more advanced courses. Overall Destinos persistence was initially higher than classroom courses but has shown a significant downward trend over time of 5.5% per term and is now below classroom persistence rates ($R^2 = 0.906, p = 0.003$).

One goal of Destinos was to reach out to students who would not otherwise be able to take Spanish because of lifestyle constraints. Comparing Destinos 1A to classroom Spanish 1, we find that a higher proportion of Destinos students were female, white, over 40, native English speakers, expecting to be employed either full time or not at all, have an advanced degree, and have “update job skills” or “educational development” as a goal. There did not seem to be a difference in Destinos versus classroom Spanish enrollment by geographic region. It appears that Destinos may differentially attract older professional and homemaking students whose home and work demands inhibit physically attending classes. As students self-selected whether they took the telecourse or the lecture format and those two groups did not appear equivalent demographically, a firm conclusion on the effect of this telecourse on student success and retention cannot be made.

However, because Destinos students seemed to differ from classroom students, it may be that success, retention, and persistence rates may not be complete or appropriate measures of effectiveness. It may be that simply the attraction of new students who would otherwise be unable to take Spanish is the desired outcome.

**Mathematics Placement: Computer Adaptive versus Pencil and Paper Test**

For placement exams at Cabrillo College, a study compared our traditional paper and pencil test, the Mathematics Diagnostic Testing Project or MDTP, to a computer adaptive test (CAT), ACCUPLACER by the College Board. CAT tests take advantage of new technology in the context of Item Response Theory (IRT) to create potentially a more efficient and individualized assessment (Wainer 1990). Classical test theory arose out of the standardization of assessment tests that accompanied increased societal complexity and the desire to assess the skills of large numbers of people. While practical to implement, these types of tests suffer from unmet assumptions such as parallelism (obtaining the same score from the same examinee on multiple administrations) and equality of variance of errors of measurement between examinees. In addition, for comparison purposes, examinees are all given the same test items or at least tests shown to be of comparable difficulty. These tests are generally constructed to be of moderate difficulty and poorly differentiate students of very low or very high ability.

Item Response Theory responds to these difficulties by positing a latent ability, $\theta$, that can be assessed given items of known difficulty, $b$ (Lord 1980; Hambleton and Swaminathan 1985, Wainer 1990). IRT also includes parameters to account for item discrimination, $a$, and guessing, $c$. A CAT test using IRT uses far fewer questions than a traditional test and so can reduce the resources needed for testing and enhance access for examinees. In addition, this method is theoretically more accurate as it is not based upon the restrictive assumptions of classical test theory such as parallelism and homogeneity of examinee measurement error.

We compared ACCUPLACER, a CAT test, with MDTP, a paper and pencil (PnP) test by double testing 52 students. Order of test administration was randomly determined so that roughly half took the CAT version first and roughly half took the MDTP first. All students took the elementary algebra level of the CAT test. Due to an error in test administration, students were
allowed to self-select which level of MDTP to take. 38 took the Algebra Readiness test, 14 took the Intermediate Algebra test.

ACCUPLACER scores showed a significant linear relation to math placement level determined by MDTP tests ($R^2 = 0.614$, $F(1,50) = 82.093$, $p < 0.005$). It appeared that the CAT test could replace the paper and pencil test, all other aspects being equal. However, ACCUPLACER failed the math faculty’s content review and was not used. On the other hand, the faculty became aware of CAT tests and their possibilities and are exploring other products.

**Computer Assisted Instruction in Math**

For several semesters, Math departments at Cabrillo Community College in Aptos, California and Gavilan Community College in Gilroy, California taught computer assisted arithmetic and elementary algebra courses using Academic Systems. The intent of introducing computer assisted instruction (CAI) was to increase success rates in these courses. Such hopes have some foundation as CAI has been related to increases in Math and Reading scores for low-achieving 5th graders (Weller, Carpenter, and Holmes 1998) and increased higher order cognitive reasoning in geosciences instruction (Renshaw and Taylor 2000). However, a meta-analytic review of CAI studies by Fletcher-Flinn and Gravatt (1995) found that studies that control for the effect of different instructors and course materials did not show a benefit or a detriment resulting from CAI. They suggest that CAI appears at least equivalent to traditional instruction and that other evaluation factors such as time savings or the value of immediate feedback to the student should be included in assessments of this new technology. The evaluation of Math CAI at the two colleges primarily focused on student success, retention and persistence.

**Cabrillo College Math CAI**

This study employed two main research strategies. The first pooled student outcomes for 638 enrollments in basic skills Math from Spring 1998 to Spring 2000 and 12,959 enrollments in Elementary Algebra from Fall 1992 to Spring 2000 and compared overall success and retention between instructional modes in their class and in the next math class in the sequence. The second tracked a cohort of 141 basic skills Math students enrolled in the Fall of 1998 through Spring 2000 see how progression through the mathematics sequence related to instructional mode for a particular group of students. This was the first semester CAI was used with 69 students receiving CAI and 72 receiving traditional instruction.

Demographics of students were in general similar between the two treatment groups enhancing confidence in the comparison despite the lack of random assignment to instructional modes. However, instructors were different between the two modes of instruction, which may provide a confounding factor.

CAI basic skills math students did not consistently show enhanced performance in basic skills math and exhibited lower success and retention in their subsequent attempt at elementary algebra than other students. Elementary algebra CAI students had similar success and retention in elementary algebra as compared to lecture students but showed improved performance in intermediate algebra as compared with other students. CAI students in the 1998 cohort were significantly less likely to persist into Elementary Algebra but were proportionately equal in enrolling in Intermediate Algebra as compared to traditional lecture students. Those lecture arithmetic students who enrolled in Intermediate Algebra succeeded at significantly higher rates than CAI students (Figure 1, Table 1).
Figure 1. Success rates of the 1998 Cabrillo arithmetic cohort as they progress through the curriculum (note that only the Intermediate Algebra is significantly different).

Table 1. Success rates of the 1998 Cabrillo arithmetic cohort for arithmetic and subsequent classes.

<table>
<thead>
<tr>
<th>Arithmetic Mode</th>
<th>Arithmetic Success Rate</th>
<th>n</th>
<th>Elementary Algebra Success Rate</th>
<th>n</th>
<th>Intermediate Algebra Success Rate</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Systems</td>
<td>73%</td>
<td>69</td>
<td>38%</td>
<td>45</td>
<td>29%</td>
<td>14</td>
</tr>
<tr>
<td>Lecture</td>
<td>85%</td>
<td>72</td>
<td>26%</td>
<td>58</td>
<td>70%</td>
<td>10</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>3.16</td>
<td></td>
<td>1.7</td>
<td></td>
<td>4.14</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>0.08</td>
<td></td>
<td>0.19</td>
<td></td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>
Gavilan College Math CAI

Between Spring 2000 and Spring 2001, 1000 students enrolled in Arithmetic, Pre-Algebra, Elementary Algebra, and Intermediate Algebra with 803 in the lecture mode and 197 in the CAI mode of instruction. Demographics of students were similar between the two modes. This similarity was indicated by using a Classification and Regression Tree (CART) to attempt to predict instructional mode using the demographic variables of disability, age, race, gender, educational status, gpa, employment hours, educational goal, and academic standing as classifiers. Only 13% of CAI students were correctly identified by the best CART model. As these variables could not accurately identify CAI students, it implies that the students are may be roughly equivalent between the two modes of instruction.

The only significant difference between instructional modes appeared in Arithmetic where retention rates for lecture students was higher than for CAI (Table 2).

Table 2. Success and retention rates of Gavilan students by class level and instructional mode.

<table>
<thead>
<tr>
<th></th>
<th>Success Rate</th>
<th>Retention Rate</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate Algebra</td>
<td>42%</td>
<td>65%</td>
<td>79</td>
</tr>
<tr>
<td>Lecture</td>
<td>48%</td>
<td>72%</td>
<td>460</td>
</tr>
<tr>
<td>Elementary Algebra</td>
<td>56%</td>
<td>74%</td>
<td>27</td>
</tr>
<tr>
<td>Lecture</td>
<td>45%</td>
<td>70%</td>
<td>154</td>
</tr>
<tr>
<td>Pre-Algebra</td>
<td>67%</td>
<td>100%</td>
<td>3</td>
</tr>
<tr>
<td>Lecture</td>
<td>46%</td>
<td>67%</td>
<td>70</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>44%</td>
<td>67%</td>
<td>88</td>
</tr>
<tr>
<td>Lecture</td>
<td>45%</td>
<td>80%</td>
<td>119</td>
</tr>
</tbody>
</table>

a = n too small for significance test
b = significantly differ from each other at the 0.05 level

As noted in the Cabrillo results, differences between instructors are a potential confounding factors. One term at Gavilan, the same teacher facilitated both a lecture and CAI course providing a natural experiment. The differences found for success and retention were not significant (Table 3).

Table 3. Success and retention rates for two different modes of arithmetic taught by the same instructor.

<table>
<thead>
<tr>
<th></th>
<th>Success Rate</th>
<th>Retention Rate</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Systems</td>
<td>32%</td>
<td>61%</td>
<td>28</td>
</tr>
<tr>
<td>Lecture</td>
<td>48%</td>
<td>71%</td>
<td>31</td>
</tr>
</tbody>
</table>

Neither rate differs at the 0.05 level

Math CAI Conclusions

At both colleges, interviews with instructors revealed that differences existed in how the teacher used CAI in their classroom. Some maintained a more traditional lecture format with the computer used as a supplement while others acted more as facilitators with the software providing the initial instruction. While there may be an optimal method for using Academic Systems that may or may not have been implemented by these instructors, the results presented here are the consequence of how this software is actually used in the field. Given this, CAI as taught through Academic Systems may be more appropriate for elementary algebra students than for basic skills students at Cabrillo and for Gavilan no clear advantage or disadvantage is apparent.
**The Board versus the Screen**

One academic department at Cabrillo is requiring that its instructors use computer generated slides for lecturing rather than chalkboards. This recommendation is based upon the assumption that this new method will enhance the learning experience. Studies comparing computer slide delivered content to traditional methods such as the chalkboard are not abundant and have not found significant differences in outcome performance measures (Ahmed 1998; Bowe et al. 2000; Szabo and Hastings 2000). Szabo and Hastings (2000) indicated that computer slide lectures may have a higher entertainment value than traditional lectures and for that reason could enhance memory recall as students are more engaged during class. To test the possibility that the medium of computer generated slides would result in increased memory recall, a pilot experiment was conducted to aid the design of a possible larger study.

**Pilot Methods**

A group of 8 zoology students were presented with the definitions of 18 biological terms and 6 “GRE” words. The biological terms were those that were to be on an upcoming test. Students were told that they were participating in a memory pilot test that was also acting as a review. Two groups of six definitions were written on a white board in blue ink and two groups of six definitions were presented on PowerPoint on slides with a white background and blue lettering. All definitions were read aloud. Groups were alternated between the board and PowerPoint. A coin toss determined that the first group presented was written on the board, the next group on PowerPoint, the third group on the board, and the last group on PowerPoint. PowerPoint slide printouts were provided as well as blank notepaper and pencils prior to the presentation. After a 5 minute study period and a 1 minute distraction period, students were given 15 minutes to answer 12 multiple choice recognition questions and 12 fill in recall questions. Assignment of terms to question type was random with the constraint that equal numbers of board presented and PowerPoint presented terms were tested with each question type. Participants were debriefed afterwards. Most knew that a main factor was the method of presentation but none indicated they were aware that having to take notes or not was the other main factor.

**Pilot Results**

The only two words (homology and inchoate) missed in the recognition format were presented in PowerPoint. Six of the twelve recall words would have lost points if graded on a test. Four words received markedly lower recall scores and of those three were presented on PowerPoint.

**Pilot Discussion**

While the sample was too small for generalization, it does suggest direction for future research. As expected, the recognition portion of the test saw much higher accuracy than the recall portion. Further, the majority of errors were made on terms presented in PowerPoint format. The two main differences between the PowerPoint and board presentation were the length of time the words were presented and more importantly, that students did not have to take notes on PowerPoint presented words as they were presented with printed notes. Theories of learning would predict that the elaboration resulting from the act of writing the definitions would lead to higher recall for those words.

In designing a larger study in light of these indications, one must decide explicitly what is being tested. Instructors can easily enhance gross memory by promoting attention to subject matter through use of color, sound, writing, and repetition. However, what is the impact on the learning of complex concepts? What are the differences in content delivery between text based subjects such as literature, visual subjects such as anatomy, or abstract subjects such as math that are more or less disposed to computer generated slides? Should considerations of convenience, adaptability, and time cost be considered along with student outcomes? Is the right question what is the impact of PowerPoint on student outcomes or what is the impact of PowerPoint on teaching methods? It
may be that the quality of the instructional method and curriculum far outweigh the medium of presentation in importance and should actually receive the focus of research efforts.

**Unsung Costs of Instructional Technology**

Evaluations of instructional technology focus on improving learning for students. While this should be the primary goal, other important considerations are often ignored. For example, many Math instructors have embraced the introduction of the graphing calculator in their classes. They often use the graphing calculator to make the study of functions more visual and to incorporate experimental problem solving. But what is the cost of this tool? To students, these machines can be more expensive than the textbook. Further, the manufacture of electronic devices in general consumes energy, resources, and creates toxic waste products. While high tech manufacturing has a clean visage, those who live in the overlapping toxic plumes of Silicon Valley with periodic groundwater contamination have a different view (www.svtc.org). Also, greater use of computers in classrooms increases a school’s energy consumption, which lately in California means a significant expenditure on electricity that could be spent on other instructional materials or staff.

While these are primarily matters for communities and government agencies to address, we all should be mindful that purchasing and using more technology has costs beyond what comes out of the school’s budget. When we increase the presence of technology, we may be asking students to spend more money, we increase the school’s energy bill, and we add to the toxic waste stream. For these costs to be worthwhile, the benefits should be significant and validated.

**Summation and Future Directions**

Currently, the main benefits of instructional technology do not primarily appear to be with improving student learning outcomes. Its advantage seems to lay more in the areas of improving access, saving time, providing immediate and individualized feedback, and enhancing visually dependent presentations. While the introduction of instructional technologies has not so far shown unequivocal superiority over traditional methods in some cases, we may even fear that some new technology is an impediment to learning. For instance, just as TV has long been reviled as a killer of reasoning and reading skills, some indict PowerPoint as intrinsically detrimental to quality presentations and critical thought (Nunnberg 1999; Stewart 2001). These concerns are primarily anecdotal and in any case, as noted by Reeves (1998), technology is “only a vehicle for content and pedagogy”. Technology does not remove the responsibility of an instructor or a student to create an engaging and effective learning situation. Instructional technology, whether it is a piece of chalk or a graphing calculator, can be used skillfully or poorly.

To fully evaluate these technologies, better controlled research is needed with a more expansive set of outcome variables. Too often, new technology is introduced into the curriculum with little or no planning (Hall and Barker 1995). Researchers are then faced with basing conclusions on the presence of non-random assignment of conditions, grading differences between instructors, curriculum and testing differences between sections and over time, and other confounding factors. Ideally, a campus researcher should be involved in the implementation of a new technology so that these factors can be controlled to a greater extent and make substantive evaluations possible. Further, outcomes other than just success and retention should be explored such as increased access to underserved populations.

Instructional technology can be beneficial in some situations but it is neither a panacea nor a replacement for inspired teachers and responsible students. Adoption of new technologies should be preceded by a literature review of the technology, detailed plans for implementation and evaluation, and a full accounting of all associated costs.
References


California Federation of Teachers. 2001. Globalizing Technology in Education. Perspective. 32.4: 3.


Contact:
Terrence Willett, Director of Research, Gavilan Community College, Gilroy CA 95020
twillett@gavilan.cc.ca.us www.gavilan.cc.ca.us/research

printed on 100% post-consumer waste recycled paper