I. Introduction

Upon receiving the 1978 Millikan Award for his notable and creative contributions to the teaching of physics, Alfred Bork predicted: “We are at the onset of a major revolution in education, a revolution unparalleled since the invention of the printing press. The computer will be the instrument of this revolution… By the year 2000 the major way of learning at all levels, and in almost all subject areas, will be through the interactive use of computers.” [Bork 1978] Although the future has not lived up to Bork’s optimistic forecast, computers are used in numerous ways in the classroom to teach physics. In traditional college environments, we see that computers have been used in lectures, where their use ranges from displaying collected data of demonstrations to fostering active student participation by facilitating voting in class. In laboratories and recitations, computers can have various functions, from simulating complex phenomena to making data acquisition fast and more reliable. Other environments outside the classroom include online homework systems and programs that diagnose students’ understanding and deliver individual feedback.

Despite the computer’s wide-spread presence in numerous educational environments, their use is just beginning to be researched in the college classroom. This paper surveys researched educational practices that utilize the personal computer to teach introductory college physics in traditional classroom environments, namely lecture, recitation/laboratories, and supporting activities outside the classroom. From these research studies, it is apparent that how the computer is used has a significant impact on its pedagogical effectiveness. The computer is by no means a productive tool by nature. The Physics Education Research (PER) community is studying how computers are used in numerous environments, and researchers have identified an assortment of characteristic features of computer-based activities that describe how the computer can be used as an effective pedagogical tool. Ron Thornton and David Sokoloff [Thornton 1990a], while studying their microcomputer-based laboratories (MBL), formulated 5 characteristics of computers that they believe are responsible for their success:

1. students focus on the physical world
2. immediate feedback is available
3. collaboration is encouraged
4. powerful tools reduce unnecessary drudgery for students
5. students understand the specific and familiar before moving to the more general and abstract

To this list, three more features of computers have been added by Edward Redish (6) [Redish 1997] and Noah Finkelstein (7 & 8) [Finkelstein 2005]:

6. students are actively engaged in exploring and constructing their own understanding
7. computers can make visible the models that are useful for forming concepts
8. computers can constrain the students in productive ways
Throughout this paper, these eight characteristic features will be used to underline why some computer-based activities have been shown to be productive educational practices and why others have not. A distinction should be made—features 1, 3, 5, and 6 are generally considered to be ideal features of any education practice [Redish 2003], while features 2, 4, 7, and 8 are specifically coupled to computer-based activities. Thus, the research does not separate the computer from the educational practice associated with it. It should be noted that this list is not complete, nor does it describe all the features that a computer affords. For example, a computer can be used as an ottoman or a doorstop, but these features do not contribute to the pedagogical effectiveness of an activity.

Because of the limited scope of this survey, certain applications of computers will have to be omitted, such as the use of PDA’s, course websites, multi-media textbooks, Physlets, Mathematica, and distant learning. While these applications have been used to teach introductory college physics, they have not been as formally studied as other practices have. Additionally, the focus of this paper is narrowed by only including studies which center on introductory college physics in traditional classroom environments. Other environments outside this domain, such as Workshop Physics [Laws 1991] [Laws 1999], SCALE-UP [Beichner 1999], Studio Physics [Wilson 1992] [Wilson 1994], and K-12 (for example, see [Linn 2000]), are excluded from this survey.

II. Computer-based activities in lecture

For quite some time, the lecture environment has remained unchanged—students sit passively and listen to an instructor’s speech or observe a particular demonstration, which are commonly believed to help students develop a better understanding of the underlying concepts. However, various studies have shown that having students passively observe lecture demonstrations does surprisingly little to help students understand the material or change their naïve conceptions [Roth 1997] [Crouch 2004]. Instead, research has shown that student learning can be significantly improved by having students more actively engaged in lecture (for example, see [Hake 1998], [McDermott 2001], [Redish 2003], and references therein).

One such method employed in lecture to have students take on a more active role is called Interactive Lecture Demonstrations (ILD), developed by Ron Thornton and David Sokoloff [Sokoloff 2001]. This practice uses inexpensive probes that can be connected to a computer through an interface box to measure various quantities, such as temperature, position, and velocity. Data are displayed in digital and graphical formats on the computer screen as the measurements are being taken. After two sequences of ILD’s on kinematics, students demonstrated a better understanding of the material after the intervention, as measured by a subset of questions on the Force and Motion Conceptual Evaluation (FMCE) [Thornton 1998]. Students attending a traditional lecture in a noncalculus introductory physics course scored an average of 35% on acceleration questions and an average 70% on velocity questions. After these same students attended two ILD sequences on kinematics, their averages substantially jumped to 80% on acceleration and 90% on velocity. However, no control group was used in this study—we do not know if students would have improved similarly by some other method. Additional studies on ILD’s have demonstrated similar results [Sokoloff 1997] [Thornton 1997a].
The first five characteristic features in our list were developed with microcomputer-based laboratories (MBL’s) in mind. ILD’s are the same tools, but used in lectures instead. Despite the difference in name, the same five features are utilized by ILD’s as well. Again, it is noteworthy to point out that some of these features (2 and 4) are intrinsic to computer-based activities, while the others (1, 3, and 5) can be features of any educational practice.

Not every application of computers in lecture utilizes the same set of features. Personal response devices (commonly known as “clickers”) were designed to replace other, simpler techniques (such as raising one’s hand or using colored cards) used when an instructor implements ConcepTests [Mazur 1997] [Crouch 2001] in lecture. Although the students use small, wireless, infra-red devices to answer multiple-choice questions, a personal computer is used in the process to receive the data, display it instantaneously in a useable format, and record these data. Such features and tasks, like instantly displaying the exact results from a round of voting, were not possible with large populations of students using simpler techniques. Clickers utilize features two through six and do so in different ways compared to ILD’s. For example, the type of feedback that is available (feature 2) with clickers allows the instructor to instantly assess the class’ understanding, and the students can gauge their understanding compared with the rest of the class.

One study which compares the use of clickers in lecture to an equivalent physics lecture that did not use clickers in an introductory, calculus-based physics course for engineering honors students [Reay 2005]. On a common question placed on a midterm, 93.5% of the students who used clickers throughout the semester answered the question correctly, compared to 82.5% of the students who never used clickers.\(^2\) Although students with clickers did slightly better, the results were not statistically significant. Additionally, student attitudes towards clickers are highly dependent on the means of implementation. This same study conducted surveys given at the end of the course, where students were asked to rate statements from a 2 (strongly agree) to -2 (strongly disagree). When a single clicker was given to a group of 3 or 4 students, the mean for the statement, “I like using the voting machines,” was 0.84. When every student was given a clicker, the mean jumped to 1.30. Although the study did report this and other interesting findings on how clickers should be implemented, the authors did not show that clickers were more effective than traditional methods to deliver concept tests.

**III. Computer-based activities in recitations and laboratories**

Previously, we saw how ILD’s and clickers were assessed to study their impact on students’ conceptual understanding of introductory physics. Some of these same tools (such as the sensors and computers associated with ILD’s) that were employed in lecture can also be implemented in a recitation or laboratory environment, where students can directly interact with the equipment.

These tools, first known as the Microcomputer-Based Laboratory (MBL), were developed by Ron Thornton and David Sokoloff [Thornton 1990a]. Just as in the ILD’s previously discussed, data from the sensors are displayed in digital and graphical formats on the computer screen as measurements are being taken.

In the study mentioned previously [Thornton 1990a], the authors observed improvements in conceptual understanding in kinematics by students in both algebra- and calculus-based physics courses who use MBL’s. One
particular noncalculus, introductory physics course at the University of Oregon consisted of 3 lecture sections and a separate, voluntary laboratory course. Thus, one group, the MBL students, attended lecture and participated in the associated laboratory course; the other group, the non-MBL students, attended only the lecture and did not participate in any laboratories. Since the two groups of students are spending different amounts of time studying the subject, this raises an issue with their research method—will an improvement by the MBL students be a result of the extra time spent on the subject, or to the pedagogical effectiveness of the curriculum? This issue will be addressed later on by Redish.

Both MBL and non-MBL students were randomly distributed in all three lecture sections. To assess students’ conceptual mastery of kinematics, the authors gave students a pre-test before any relevant lectures on kinematics and before the MBL students performed any of the MBL laboratories.³ After all students attended lectures on kinematics and the MBL students participated in two different MBL activities on kinematics, all students were given a post-test with nearly identical questions from the pre-test.

On three velocity questions given in the pre-test, the average for the students was 54%. The post-test average on these three velocity questions for the MBL students improved to 82%, while the average for the non-MBL students was 50%—not significantly different from the pre-test results. This finding agrees with other studies that observe students gaining little from lecture [McDermott 1984] [Halloun 1985a] [Halloun 1985b]. Other studies within this piece [Thornton 1990a] demonstrate similar trends with MBL.

In Thornton and Sokoloff’s study [Thornton 1990a], one may question whether the observed success of MBL’s is simply due to the added instructional time. Redish is able to address this concern by controlling for both the instructor and the amount of time spent on instruction [Redish 1997]. In one course, the instructor spent 3 full hours in lecture teaching instantaneous velocity using MBL equipment. In associated hour-long recitation sections, neither MBL’s nor tutorials were used; graduate teaching assistants reviewed textbook problems. In a similar course given later, this same instructor dedicated only one hour of traditional lecture to velocity and students participated in a hour-long tutorial session using MBL. Students were assessed using multiple-choice questions provided by Thornton and Sokoloff [Thornton 1990a], the Force Concept Inventory (FCI) [Hestenes 1992], and long-answer exam questions. On five velocity questions given during an examination, the average for all five questions for the class without MBL’s was 64.2%, while the class that implemented MBL’s with tutorials had an average of 85.6%. The authors claim this demonstrates that MBL’s with tutorials can play a significant role in improving students’ understanding of velocity. The extra time on instruction may not be the reason for the improvements Thornton and Sokoloff observed using MBL. Other studies have shown similar findings to the ones specifically addressed above [Sokoloff 1995] [Thornton 1997b] [Wittmann 1999].

Thus far, none of the applications discussed have utilized the remaining two characteristic features of computer-based activities: computers can make models explicit (7) and they can constrain students in productive ways (8). We will now examine how computer simulations achieve these two goals.

Some studies have examined students using a computer simulation, compared to students using equivalent real laboratory equipment. For example, Noah Finkelstein et al. [Finkelstein 2005] compared students using real
light bulbs and wires to students who used a simulation, entitled the \textit{Circuit Construction Kit} (CCK)\textsuperscript{4} [Perkins 2005], that explicitly models current flow in DC circuits.

In a second-semester, algebra-based introductory physics course, the authors compared two groups of students: students who used real lab equipment to complete a DC circuits laboratory and students who used CCK to complete the same laboratory. After the students completed the DC-circuits lab, \textit{all} students performed a “challenge” using \textit{real} equipment where groups had to build a circuit, analyze it, and write down their individual responses on a worksheet. Note that the students who used CCK to complete the lab had no formal experience with the real laboratory equipment prior to the challenge.

The challenge worksheets were scored for accuracy on a scale from 0 (demonstrates no knowledge) to 3 (demonstrates a complete and correct solution). The average score for the CCK students was 1.86, while the real-equipment group was 1.64—a statistically significant difference (p<0.03). The authors claim the success of the simulation may be due to the fact that the simulation makes current flow in DC circuits explicit (feature 7) and constrains the students in productive ways by discouraging them to “mess about,” which may be deemed as unproductive (feature 8). Additionally, features 2 and 3 are utilized by this particular simulation.

It should be noted that not all studies that compare students’ use of a computer simulation to other methods have observed similar results. In a study conducted on a calculus-based, introductory physics course, Richard Steinberg compared students in three different weekly recitations—two recitations used a computer-based tutorial with an associated simulation (Class A and B) and one of the recitations utilized an equivalent paper-based University of Washington-like tutorial (Class C), all of which were on air resistance [Steinberg 2000]. Two forms of quantitative data were collected to assess the students’ conceptual mastery: pre and post results on the FCI and common exam questions. The FCI gains for Class A, B, and C are 0.32, 0.25, and 0.32, respectively. The figure below shows the results from common exam questions given to the students—none of the results are significantly different. During a nearly identical follow-up study conducted in a subsequent semester, the authors observed virtually identical results.

It seems that the computer simulation in Steinberg’s study utilizes the same set of characteristic features as CCK in the Finkelstein study. Despite Steinberg making certain physical models explicit in his simulation, no difference in conceptual understanding was observed between students using the simulation and students using a paper-and-pencil based tutorial. We may speculate on the reasons for these observations. As with all
instructional tools, how one implements it can have a significant effect on its effectiveness. Particularly, the curriculum surrounding a computer-based tool is often strongly coupled to its impact on students. Another possible explanation is that the paper-and-pencil based tutorial is already optimized for student learning, and the addition of a computer simulation to this tutorial will not further benefit student learning. Thus, what one compares a computer-based activity to has an effect on its perceived success.

In other studies on how students learn from interactive computer programs, we see some that fail to teach students basic concepts. One such study looked at how 10 first-year college students taking a noncalculus introductory physics course interacted with each other while using an instructional unit entitled, *Movement: The Physics of Bodies in Motion* [Yeo 2004]. The authors’ research methods are questionable, since a control group of students was not created in this study. The computer program was developed to help students learn concepts on mass, speed, force, stability, inertia, and momentum. There are 25 sequential screens that have various activities for the students, including reading text, making calculations, and analyzing data and video clips. A significant portion of this tutorial has students study projectile motion by analyzing the physics in the context of Olympic long jumpers in motion.

In a series of questions designed to evaluate the students’ conceptual understanding, the authors noted that the number of correct responses did not meet their expectations. For example, on a question involving the path of a long jumper, no student was able to correctly identify the location where the jumper’s vertical velocity is a maximum. Unfortunately, more detailed results are not presented in this paper.

Based on numerous observations and student interviews, the Yeo et al speculate that the computer program did not constrain the students in productive ways. Specifically, students were free to move on to further subjects without demonstrating an appropriate understanding. Although the program did utilize the first characteristic feature of computers by presenting the physics in a real-life setting, it appears the program did not appropriately utilize features 2, 5, 6, and 8. This is not to suggest that a successful computer application must utilize many, if not all 8 features. As the authors noted, it seems that certain structural features of the program and its implementation could have been improved.

**IV. Out-of-classroom computer-based activities and class support**

Beyond lectures and laboratories, some students spend a great deal of time studying for physics courses. The personal computer has reshaped this environment outside the classroom—instead of having students work out problems on paper and turn them in to be graded, students can now work out their answers and submit them via programs that attempt to diagnose students’ naïve conceptions. Various online homework systems have been developed, each with its own features, but the bulk of these applications utilize features 2, 3, and 6.

One such system, known as *Computer-Assisted Personalized Approach* (CAPA), is a web-based computer program that creates individual homework assignments for students [Kashy 1993]. Students enter in their answers online, and CAPA gives immediate feedback and hints on students’ responses. Students enter in their answers in either a numerical or multiple choice format (free-responses or figures are not supported by CAPA). The authors
report that during its first use in an introductory college physics course for non-science majors, student support of the new homework system was positive. On a self-reported survey given to students towards the end of the term, approximately 82% of students indicated they spent more time on CAPA assignments compared to other graded homework.

In her Ph.D. thesis, Andrea Pascarella performed a study on the impact of using CAPA in a calculus-based, introductory physics course on students’ conceptual understanding of Newtonian Mechanics and problem-solving abilities [Pascarella 2002]. Students were initially assigned to either CAPA or traditional homework. Halfway through the semester, the students switched to the other homework format. Using the FMCE and common exam questions as an assessment tool, Pascarella observed no significant difference in performance between students who used CAPA and similar, paper-based homework assignments. However, student understanding is just one of many aspects that a researcher can study. One can also study how students view particular classroom tools. It was noted that students with expert-like views of learning believed that traditional homework was a better learning tool, while students with non-expert-like views of learning felt that CAPA was a better learning tool.

In a similar study that used a slightly different online homework system (called WebAssign®), we see analogous reports of researchers observing no significant difference in conceptual understanding between paper-based homework and online homework [Bonham 2001]. While these systems may have logistical benefits (instructors no longer have to grade homework by hand), it seems that they have little impact on improved student understanding.

Other studies, however, have claimed to observe significant gains in conceptual understanding when comparing online and paper-based homework. In one study [Cheng 2004], the authors studied students in a large introductory, calculus-based physics courses over a two-year period. Instructors taught using either interactive engagement (IE) or non-interactive engagement (NIE) methods and utilized either ungraded, paper-based homework (UHW) or graded, online homework (OHW). It should be strongly noted that this study is not determining if online homework is as beneficial as similar paper-based homework. The authors not only changed the mechanism by which homework is delivered (i.e., paper versus computer), but also the homework pedagogy (i.e., ungraded versus graded homework). It should not be too surprising that graded homework is more beneficial for student learning than ungraded homework.

The online homework system was developed at the University of Texas at Austin, and is similar to CAPA. Using the normalized gain of the FCI as a measure of student learning, the group using UHW had a gain of \( g = 23.9\% \), while the group who used OHW had a gain of \( g = 50.9\% \), regardless of whether IE or NIE was utilized in lecture. This difference is statistically significant \( (p < 0.001) \). The table below displays the normalized gain on the FCI for all four groups.

<table>
<thead>
<tr>
<th></th>
<th>IE</th>
<th>NIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHW</td>
<td>20.5% (±3.2)</td>
<td>25.7% (±2.0)</td>
</tr>
<tr>
<td>OHW</td>
<td>58.2% (±2.1)</td>
<td>40.4% (±2.0)</td>
</tr>
</tbody>
</table>
The online homework systems we have discussed thus far are mainly used to replace paper-based homework. Another computer application exists that attempts to tutor students by providing specific and customized feedback on a student’s progress. These tutoring systems emphasize different characteristic features that the computer affords when compared to online homework systems. For example, a greater emphasis is placed on features 2 and 8, since extensive feedback is provided as students are carefully guided through a series of tailored questions. Generally, feature three is not utilized, since collaboration is not encouraged with these tutoring systems.

Researchers have been developing such tutoring systems for quite some time. In the late 1980’s, Lillian McDermott studied a computer program known as Graphs and Tracks that attempted to identify students’ common beliefs of motion graphs [McDermott 1990] [Grayson 1996]. One such modern tutoring system, known as Cybertutor (also called Mastering Physics), developed by David Pritchard et al. at MIT [Morote 2003] [Pritchard 2002], is a web-based computer program that acts as an interactive personal tutor to introductory physics students. The program presents students with multi-part problems, which may require free-response, numerical, symbolic, or fill-in-the-blank answers. Cybertutor can give students hints on how to do certain problems, or offer simpler sub-problems upon request from the user.

In two related studies involving Cybertutor / Mastering Physics [Warnakulasooriya 2004] [Warnakulasooriya 2005], the authors had two groups of students taking a calculus-based, introductory physics course on Newtonian mechanics solve two related physics problems using Mastering Physics—one group solved the pair of problems in one order, and the other group solved the two problems in reverse order. Thus, the first problem in a pair of questions for one group was the second problem in a pair for the other group of students. The idea behind this experiment is to see if there is a difference in the amount of time it takes the two groups to solve a given problem in a particular pair. If the two groups solve the same problem in an equal amount of time, then one may conclude that the remaining problem in the pair will not aid students in solving subsequent, related problems. If the group who solved a particular problem second took less time to do so, then you can conclude that the first problem in the pair primed them for the second.

The latter case was observed by the authors—the group of students who solved a problem second in a given pair was able to solve it in 15% less time on average when compared to the group who solved the same problem first. From this fact, the authors conclude that problem solving transfer is taking place when students use Mastering Physics. However, it remains to be seen if this result is unique to Mastering Physics or if one could observe similar results with paper-based homework. Secondly, one could use this same strategy to see if there is a difference in time to completion with a pair of unrelated problems. If there is a significant difference between the time to complete a problem in reverse order in an unrelated pair, then the author’s claim is severely weakened. This control case was not investigated in this particular study.

Additionally, Pritchard has tried to use Mastering Physics to predict students’ final course grades. Using a variety of information collected by Mastering Physics on how students solve problems, such as correctness, time to completion, and number of hints given, the authors were able to devise an algorithm that allowed them to get a correlation coefficient of 0.634 between Mastering Physics scores and final course grades [Pritchard 2005].
Although this is generally considered to be a high correlation coefficient when attempting to predict final course grades, it remains to be seen if the same algorithm can be used for other courses.

V. Conclusion

This paper surveys the research literature that examines the various computer applications in the three traditional learning environments used to teach introductory physics at the college level: lecture, recitation/laboratory, and other supporting environments outside the classroom. Some of these applications have yielded successful results, such as an increase in student understanding of the material, while others seem to be no more effective than traditional methods.

A number of research studies presented in this survey utilized questionable research methods, either making unproven claims, not employing appropriate control groups, or a mixture of problems. One possible explanation is that these computer-based applications are often developed by the same people who research them, and there is an inherent desire to show that such tools are successful. Another possibility is that these particular researchers employ questionable research methods due to inexperience in the field of education research. PER is a fairly young research field compared to others, and as with all new fields, it will take some time for the community to establish various norms, expectations, and standards of practice.

When these applications that utilize the personal computer are successful, one may ask whether it is the computer itself or the pedagogical materials associated with the computer application that is ultimately responsible for the tool’s success. From the few research studies that demonstrate some success, it seems that both aspects are responsible. The 8 characteristic features frequently referred to in this survey are not innate features that any computer possesses—they are features that a computer-based, instructional tool may utilize and are identified by the research community as the characteristics that are responsible for the success of certain computer-based activities.

Researchers are just beginning to discover the potential that certain computer applications have at helping students with their conceptual mastery of physics. Other qualities of our students, such as their problem-solving ability, knowledge retention, and beliefs about science, are also affected by computer applications, but have not been studied in detail. Additionally, some applications have not been studied in all possible learning environments, such as the use of computer simulations in lecture. The year 2000 has come and left, and Alfred Bork’s prediction seems to be premature, but the personal computer will continue to be used more to teach students physics, regardless of its proven or unproven effectiveness.
Bibliography:


---

1 Updated versions of the FMCE are available at: [http://physics.dickinson.edu](http://physics.dickinson.edu)

2 Note that all students in this study were presented this particular question in lecture, not just the students who used clickers.

3 It should be noted that only 2 of the 3 lecture sections were given the pre-test. Since MBL and non-MBL students are randomly distributed in all 3 lecture sections and the 2 sections that were tested showed no significant difference, the authors claim that the pre-test scores for MBL and non-MBL students should be similar.

4 CCK was developed by the Physics Education Technology project (PhET): [http://phet.colorado.edu](http://phet.colorado.edu)

5 [http://webassign.net/info/](http://webassign.net/info/)

6 The program has evolved over the years, and the name was changed from “Cybertutor” to “myCybertutor.” The current version now utilizes “Mastering Physics.”