00B025I

00B025I
ACH-DC01-45

THE ROLE OF INFORMATION TECHNOLOGY IN STUDENT LEARNING ASSESSMENT

Report Submitted to:

Dr. Arthur Heinricher & Dr. Ed Y.H. Ma

National Science Foundation

By

Valerie Sanders

David Valliere

Anita Wong

In Cooperation With

Peggie M. Weeks PD & Dr. Lee Zia

National Science Foundation - Division of Undergraduate Education

December 19, 2000

Table of Contents:

Abstract

This report, prepared for the National Science Foundation (NSF) Division of Undergraduate Education (DUE), contains details from our investigation and evaluation of Information Technology (IT)-based Learning Assessment Tools in undergraduate education. Twenty-two tools were identified and evaluated to determine their most effective functional use. Recommendations for further studies are included.

Authorship

Executive Summary	David Valliere, Valerie Sanders
Introduction	David Valliere, Valerie Sanders
Literature Review	David Valliere, Valerie Sanders, Anita Wong
Methodology	David Valliere, Valerie Sanders
Results and Discussion	David Valliere, Valerie Sanders, Anita Wong
Conclusions	David Valliere, Valerie Sanders
Recommendations	David Valliere, Valerie Sanders

Acknowledgements

Completing this report would not have been possible without the help of many individuals. We would like to take this opportunity to thank all of these people that have made this project what it is today. First, we would like to extend thanks and appreciation to our liaisons, Ms. Peggie Weeks and Dr. Lee Zia, for their endless contributions. They offered us guidance and shared their wisdom and experience throughout these past 14 weeks. We would also like to thank our faculty advisors, Professors Arthur Heinricher and Yi Hua Ma, for their tireless efforts and scrupulous attention to detail, it was these efforts that have molded our project to the current state. We would also like to extend great appreciation to Professor Wes Jamison for preparing us for this experience. Wes successfully transformed engineering (and Management Information System) students into social scientists in only seven weeks. Working for the National Science Foundation has been a pleasure for us, everyone has been outstanding and willing to help us in every possible way. To the Program Directors (James Lightbourne, Herbert Levitan, Joan Prival, Elizabeth Teles, Gregory Tait, Diane Jones, Myles Boylan, Joan Gosink, Andrew Bernat, and Terry Woodin) who have helped us throughout these seven weeks: Thank you! Finally, we would like to thank Kathleen Valliere and Lucy Schuman for their help in finding information and offering advice. Once again, we want to extend our greatest appreciation to all of the people mentioned and everyone else that helped us with this project throughout the past 14 weeks.

Executive Summary

As technology advances, colleges and universities are developing new and innovative ways to assess student learning. The purpose of this project was to investigate and evaluate information technology-based learning assessment tools. IT-based learning assessment tools are tests using a computer. Our goals were to generate a list and to evaluate the effectiveness of the IT-based tools that are currently being used in undergraduate education. The tools we discovered will be a part of a growing list for the National Science Foundation (NSF) to add to. Evaluating the effectiveness of the tools will aid in informing NSF of what is useful and provide information to Program Directors on tools that they may consider for future support.

In order to identify tools that are presently being used, we conducted informal, unstandardized interviews with ten Program Directors in the Division of Undergraduate Education in NSF. We discussed the tools which some of these program directors were supporting and asked if they knew of any other available tools. We also searched on the Internet to learn more about the tools we encountered and contacted the developers of the tools or professor(s) that use the tool. We conducted email surveys and four semi-standardized phone interviews. We also "interviewed" the tools by using the tools as students would completing a homework assignment or a test, observed all the features the tool had to offer and explored their capabilities.

The list we created is not exhaustive, but it represents the types of tools that are being used in undergraduate education across the country. Our list includes 22 tools, consisting of nine that are primarily homework graders, 12 primarily testers, two tutors,

two labs, and one essay grader. Some of the tools have more than one function and thus are counted in more than one category. Some of the main features of the tools are: immediate scoring and feedback to students and to professors, providing correct answers when students answered problems incorrectly (tests); providing individualized problems to each student; allowing students to try problems over if they get them wrong (homework); and explanations of answers (tutors).

Since these tools were developed for assisting in ongoing assessment, we believe that IT-based learning assessment tools would be more effective in lower level classes, introductory physics, calculus, and algebra for instance. However, IT-based learning assessment tools can be effective, if used as a supplement to other assessment functions in more advanced classes. As technology advances new generations of assessment tools could be developed to assist and assess student learning.

We recommend that the National Science Foundation continue to make additions to our IT-based learning assessment tools matrix and set up future projects to continue research in this field. One potential project should be to evaluate these tools and to develop criteria for evaluating the effectiveness of the tools. The effectiveness of each function, for example the effectiveness of receiving hints on homework. The tools should be evaluated individually due to the many differences in functionality and variations in their purposes. The NSF should also consider supporting tools that did not fall under the categories of homework graders, tutors, and tests.

Chapter 1 Introduction

Information technology (IT) is a term meaning applying computer systems, including networking and telecommunications, generally to store and send out information (FOLDOC, 2000) (Cambridge, 2000). IT has grown tremendously over the last decade. It has affected the ways that students learn and the ways that their knowledge is assessed. It has had an impact on schools in many areas including learning, research and testing. Learning assessment is the judgment of the knowledge of facts obtained or understanding of ideas (Cambridge, 2000). IT-based learning assessment tools are devices for testing students' learning (through homework, tutors and exams) that utilize computer systems.

The National Science Foundation: Division of Undergraduate Education (NSF: DUE) is interested in this project to determine what IT-based learning assessment tools are currently being used at the undergraduate level. NSF would also like to know whether the tools are actually effective in student learning assessment. DUE supports programs that enhance the quality of instruction in the diverse institutions of higher education, that is, two- and four-year colleges and universities (NSF: DUE, 2000). This project is one step in exploring information technology-based learning assessment tools. Other educators may consider the findings useful when they contemplate using IT-based learning assessment tools in their own courses or when improving and developing new tools.

There are various types of IT-based learning assessment tools. One example is a computerized test, which is simply a standard test taken on a computer. Another example is a computer adaptive test, where the computer scores the test as the student takes it and

chooses the next problem according to the student's answer to a previous question. If the problem was missed, the computer would present a question of a lesser difficulty and if answered correctly a more difficult problem would be given.

Our goal for this project was to investigate and evaluate information technology-based learning assessment tools at schools around the country by utilizing phone and face-to-face interviews along with a review of the existing literature to accomplish our tasks.. We obtained professors' views on the effectiveness, simplicity and dissemination of these tools. Part of our goal was to generate a list of IT-based learning tools that are currently being used in undergraduate education in the US.

We talked to various professors and developers of the tools to discover their views on how technology is affecting learning assessment. We also used the information obtained from using the tools ourselves to determine how the tools operated and whether these IT-based learning assessment tools were preferred. We wanted to organize this information so it was easy to read and understand. To accomplish this, we made categories that depicted characteristics that were similar between the tools.

Chapter 2 Literature Review

2.1 Effective Assessment

Assessment of student learning is the process of collecting data to determine what students know. Information technology learning assessment tools start with traditional testing devices, such as exams, quizzes and group discussions, and adapts these devices to new technologies like computers and the Internet. Techniques serving assessment include interviews, questionnaires, and of course tests (Perrone, 1991). Traditional assessment tools such as these are used as the focal point for assessing student learning.

According to the American Association for Higher Education (AAHE) Assessment forum (1997) there are nine "Principles of Good Practice for Assessing Student Learning". One of the principles is "assessment works best when it is ongoing not episodic" meaning improvement is enhanced when assessment involves an associated series of activities carried out over a period of time (AAHE, 1997). An additional principle notes "assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time" (AAHE, 1997). This implies that assessment should involve a variety of approaches to include those that entail actual performance, so it reveals not only what a student knows but also what he or she can do with that knowledge. For example, a professor may believe that a paper, which is due at the end of the term, is an effective assessment of the students' learning throughout the term. If the student is required to submit various sections of the paper throughout the term, then he or she can receive feedback from the professor on the sections. So reiterating this principle, assessment is an on-going process and is not as effective as it could be when it is sporadic. This approach could effectively

produce a better paper for the student, and the professor has a chance to focus the student's attention on the topics that really need to be discussed in the paper. Therefore, for assessment to be successful, whether it is IT-based or not, these principles should be incorporated into the usage of the assessment tools. These are some parameters for making an assessment tool effective for the professors as well as the students.

The following points are some that Carl David of the University of Connecticut (1996) notes the ideal examination should include:

1. Both the student and examiner feel that what was being measured was indeed measured. If a students are being tested on derivatives, then they should feel that they has demonstrated their knowledge on derivatives and equally the professor should feel that the test that was given has accurately measured the students' knowledge in derivatives.
2. The exam should be reproducible within realistic bounds year after year (the assessment should basically be the same each time). There should not be much fluctuation in they way that exams are administered.

Whether or not IT-tools are really effective in higher education is questionable. According to the findings published in the Journal of Engineering Education (JEE, 2000), technology in the classroom is most effective in the special education settings and "tends to decrease in the order of elementary schools, secondary schools, and colleges." However there are ways to determine if the assessment tool is actually effective in assessing students' learning. "The most frequent way that evaluators determine whether performance is good enough is by using comparisons. This method can illustrate the advantage of using the tool, if there are any at all. You can compare students in and out of

the innovation" (Baker, 1999). Baker continues to explain that following up on students and tracking their performance over time is a good way to determine the effectiveness of a tool over the long-term.

2.2 Four Generations of IT- Based Learning Assessment Tools

Computerized educational measurement is the process of combining educational assessment and computer resources together to form IT-based learning assessment (Linn, 1993). Educators have identified four generations associated with information technology based assessment tools. The four generations reflect how well educational measurement and computing resources have been integrated. The first generation, defined as *computerized testing (CT)*, is simply a computer administered a test. This allows students to take the exam during time periods outside of class. Computerized testing is usually restricted to multiple choice or true or false questions. The second generation is called *computerized adaptive testing (CAT)*. The main difference between CT's and CAT's is the computer's ability to tailor the difficulty or content of the next question based on the response to the current question. This means that if a student is answering the questions quickly and correctly, the computer will increase the difficulty of the following questions. The third generation of computerized assessment is called c*ontinuous measurement*. This generation uses tools to profile the students' learning abilities and assesses their current level of achievement. The computer uses calibrated measures based on the content of a course to track and estimate the student's level of achievement in the course. Finally, the fourth generation of IT based assessment is called

intelligent measurement. The computer acts as a tutor providing feedback on what questions the student answered correctly and incorrectly and explains the correct answer. The tool also identifies for the student areas in which he or she needs more study, areas which have been mastered, as well as other advice that will help the student achieve (Linn, 1993).

2.3 More than Exams

Although standard exams are the most common forms of assessment, they are not the only method of assessing student learning. For example a mechanical engineer's ability to design a machine that completes a given task is an assessment of all of the previous physics, material science and mechanics classes that the student has completed. As time goes on, more professors are utilizing IT based assessment tools, so tests may not be the only form of IT-based learning assessment. This project, an example of the WPI IQP (WPI, 2000), is an assessment of student learning that could not be done on a computer. Another example is at Frostburg State University where an Internet-based chemistry is set with remote locations for distance education (FSU, 2000).

2.4 Examples

As colleges and universities incorporate IT-based learning assessment tools into courses, schools push to develop tools that are more effective in aiding student learning. Some of these IT-based assessment tools include, computerized testing, computerized adaptive tests, and on-line tests. At James Madison University, computer-based tests are

utilized in general education, tests measuring information-seeking skills, technology proficiency, and communication skills (JMU, 2000).

TOEFL is a good example of a first generation computer based testing system. TOEFL evaluates the English proficiency of people whose native language is not English. The program was introduced as a computer-based test in July 1998 in many parts of the world. It combines many of the same question types as the traditional paper-based test with new question types that can be offered on the computer. There are four sections on the test: Listening, Structure, Reading and Writing. In the Listening and Structure sections, users are able to change the answer as often as they like until a final choice has been made within the time limits. Once users move on to the next question, they are not allowed to go back to an earlier question.

Just in Time Teaching (JiTT) is a second-generation test used at Indiana University-Purdue University Indianapolis (IUPUI) in the Physics department (JiTT, 1999). This tool allows students to complete exercises, without any time constraints. At IUPUI, students complete preparatory problems via the Web, which are due a short time before class begins. The results of the assignments then help the professors adjust and organize the lessons (JiTT, 1999). "The computer adjusts the characteristics of the administered items to match the proficiency level of each examinee" (JMU, 2000). Thus, students with high proficiency levels will receive more difficult problems and students with a lower proficiency receive questions of lesser difficulty.

Dr. David Pritchard, of the Massachusetts Institute of Technology, is working on a project entitled CyberTutor. CyberTutor is a fourth generation IT-Based learning assessment tool that is currently being used in physics courses at the Institute. For

educators, the platform offers course management tools and the ability to create, deliver, and grade customized interactive homework assignments and exams. CyberTutor presents each student with a problem, provides hints or simpler sub-problems at the student's request, grades the student's responses immediately, and tells the student if the answer has deficiencies. Instead of having questions with only multiple choice or numeric answer, CyberTutor has special Java applets that analyze expressions, word strings, student-drawn vectors (with the mouse), and student-drawn curves (MIT, 2000).

2.5 Advantages and Disadvantages

There are many perceived advantages to information technology-based learning assessment tools. One commonly alleged advantage is instant feedback. Students are given the results of a test immediately. They have the opportunity to go over any wrong answers without having to wait for professors or TA's to be available. This is also a advantageous because the errors are still fresh in their minds. Some believe that it is often more educationally effective if a student receives feedback as soon as possible (Thelwall, 1999).

IT-based tools can save students and professors time. Linn believes that computerized testing (CT) significantly reduces errors made when correcting exams, and that doing away with answer sheets increases the speed of test taking (Linn, 1993). Johnson of Penn State believes that there is efficiency in using computerized testing (Penn State, 2000). It saves time because instructors can reuse questions from old tests and they would not have to spend time writing additional questions.

Computerized testing has advantages in scoring and reporting scores as well. It can eliminate the time required for manual grading. A sub-score can be reported for each

section of the exam. The computer can calculate sub-scores and composites immediately (Linn, 1993). CT provides immediate reporting of scores and often aids in human interpretation of scores. Therefore, feedback from the computers can include what answers are correct, a running total score, and information on how many questions are left as well as remaining testing time (JMU, 2000). The results can also come in the form of charts and graphs for the student to see which types of questions they get right and what topics they should spend more time studying.

Cheating can be reduced or eliminated by computer-administered test (Thelwall, 1999). In some tests, questions are selected randomly so that every student has a test that is different. This allows the institution to reuse the same tests every year. Penn State utilizes a system where photo IDs are checked at the door of the room and names are matched to those on the class roster. Students are not permitted to bring in paper or take paper out of the testing facility; they are given scratch paper to use during the exam to work out problems by hand if needed (Penn State, 1997). This would prevent students from writing down equations and other information to cheat on their exam. Students may attempt to memorize the questions and answers, but after awhile realize that it is pointless. So there are no 'fraternity tests' passed on from student to student. The tests do not all have to be administered at the same time. So there is also flexibility in the scheduling of the tests.

Another advantage of computerized testing is their ability to present questions in new and possibly more realistic ways. For a physics examination, "a complex text description on a series of static diagrams may be required to present an item concerning a particular property of motion" (JMU, 2000). Using CT, the motion might be more

clearly depicted using a brief video clip demonstrating the physical motion. Tests having audio items (where students would use headphones) as well as visual items can be an attractive feature.

For some tools, professors are limited to the type of questions they can ask on computerized tests. In most cases the tests are made up of multiple choice and true/false questions. However now, numeric questions are available, where a number is typed in; fill in the blank; short answer; and even essay questions. There are few tools that allow students to demonstrate their understanding of material via a writing sample and have the computer grade it (IMEJ, 1999).

CATs offer extra time for students who need it, hence reducing one source of test anxiety (ERIC Digests, 2000). With CATs, students are given unique tests because the questions are assembled interactively as the individual is tested. The proficiency of a student is determined while he or she takes the test, based on answers to previous questions. Test questions are compiled in a large database and classified by level of difficulty. These questions are then "scanned and the one determined to measure the candidate most precisely in the appropriate test plan area is selected and presented on the computer screen" (National Council, 2000). The Graduate Record Exam (GRE) was introduced as a computerized test in October 1992. The GRE's employ a CAT where the computer uses the information from preceding questions answered to generate the next question. If an item is missed, the computer will give a simpler question. Then as questions are answered correctly, the test questions will increase with difficulty. According to The Educational Testing Service (ETS) "Each computer-based test section meets pre-established specifications, including that the types of questions answered

correctly and incorrectly and including the difficulty levels are taken into account in the calculation of the score. Therefore, it is appropriate to compare scores of different test takers even though they received different questions…" (GRE, 2000).

There are arguments that the elimination of answer sheets can eliminate some traditional mistakes, for instance, accidentally skipping over an item in the test booklet but not on the answer sheet or failing to completely erase an answer. Computerized test examinees are able to focus on one item at a time without having to deal with the 'anxiety' of the questions to follow (Linn 1993, 376).

Computerized testing produced improvements in problems associated with test-administration procedures. This testing has precise control of displayed items. Instructions like "Stop, Put your pencils down" are not necessary because the program would automatically stop after the allotted time. With CT, there are no paper copies for answers to be stolen, copied, or otherwise misused. Password and security protection block unauthorized access to testing materials.

When administering assessments via the Internet there are some setbacks. Using remote test administration may increase the number of individuals who can access the assessment services but, test administration from remote locations reduces the control over the testing environment and could compromise the standardization of the test (Sampson, 350). In a computer facility with a professor or TA, he or she would control actions like talking among students and the use of notes during the test. Some tests do not allow students to use certain test-taking strategies. Students are unable to look through all the questions on the exam before beginning, which is a strategy that many utilize. More importantly, the National Center for Fair and Open Testing notes that some

computerized tests do not permit students to go back and check over their work after completing a question (FairTest, 1999). Even though the students can check it right after doing the problem, they are unable to go back and correct their work after going onto the next one or two problems having realized they made a mistake.

Computerized adaptive tests are set up so that all the information needed in selecting the items can be summarized in one to three parameters. Thus, CAT's may not be appropriate for some subjects and skills, for example, psychology (ERIC, 2000).

Another disadvantage that students at many schools may face is a lack of resources. While many colleges and universities have computerized testing programs, many do not have an adequate facility to hold enough computers for all the students that need to be tested. In addition, the institutions just may not have a sufficient quantity of computers with the proper software. Software must either be developed or purchased from a commercial software vendor. There is no guarantee that off-the-shelf software will do exactly what the institution wants it to do. (JMU, 2000).

There are some specific advantages and disadvantages that are associated with assessment tools that are used for grading homework. These tools were initially designed to save time for the professors while giving the students more homework practice. Since, "assessment works best when it is ongoing not episodic" (AAHE, 1997), it is not realistic for a professor teaching a class of five hundred students to assign a lot of homework, assess the students' progress, and return it to the students' in a timely manner. Homework graders make this possible. Other advantages have developed as a result of using these tools for learning assessment. According to Edwin Kashy (1999), of Michigan State University, "frequent assignments with firm electronic deadlines keep the

course on schedule and help inhibit the tendency of some students to procrastinate and fall behind." Another advantage that has developed, as a result of using these tools, is the ability to allow students to try a problem out and rework it if their answer is wrong without having to wait a period of time for their work to be checked. This option motivates students to get all of their work done correctly, therefore learning the material (Kashy, 1999). Also, with the computer doing all of the grading, the students see the professor as more of a mentor than a judge. Therefore making them more likely to ask questions and take advantage of the professors knowledge. Although these tools may save professors time by alleviating the need to grade homework, Kashy (1999) explains that homework sets are often time consuming and difficult to prepare. This is especially true with tools that can generate multiple versions of the same question.

Information technology-based learning assessment tools have their advantages and disadvantages as well. IT-tools give instant feedback to students and sometimes to professors, save time, are able to present questions in more realistic ways using video clips, Java applets, etc. In contrast, some tools do not allow students to employ certain test-taking skills, IT-tools are not effective for some subjects and skills, and using these tools in a class may present problems for some schools due to a lack of resources.

Chapter 3 Methodology

The purpose of this section is to provide a detailed explanation of exactly what we did to complete our project. This section includes our processes for creating and administering our interviews as well as methods for compiling and analyzing data. The purpose of this project was to investigate information technology-based learning assessment tools at the undergraduate level around the country. We compiled a list of the tools and evaluated the effectiveness of the tools. We used face-to-face and phone interviews as our primary source of information, we also used as many tools as we could, and we supplemented this information with literature found both in print and on the Internet.

3.1 Question Design for the Interviews

The Social Exchange Theory (SXT) is an element of the foundation of social science, which states that whenever someone is asked a question they do a quick cost-benefit calculation in their head before answering. This calculation helps them determine whether it would be in their best interest to answer the question and answer honestly, or if the risk would be too great. So in order to ensure that we do get answers, and honest ones, it is essential to make the benefits outweigh the costs. At first the only potential cost we discovered, for our interviewees, was their time. This was the main factor to overcome in trying to get professors and National Science Foundation Program Directors to talk with us. They were all very busy people and we therefore had to convince them that our research would effect them on some level and that their input would only help further development of IT-based learning assessment tools. They might have felt that

their participation in this research was a waste of their time. However, being professional educators, any possible improvement in education is something that is important to them. As we talked to the various professors and developers of the IT-based tools, we learned that another cost for these people was in discussing possible disadvantages of the tool and in providing some statistical data that might show the ineffectiveness of the tool. We emphasized the fact that our project was a starting point for NSF because we were compiling a list of the tools as well as attempting to determine new directions for research in IT-based learning assessment tools. We continued by saying that this information would be maintained at the National Science Foundation and one day professors would be able to obtain this information for their own needs.

Once these costs were balanced, the next most important factor in the success of our research was designing an effective questionnaire. According to Dillman and Salant (1994, page 77), the issues that should be considered when developing the questions are:

- How specific should the questions be?
- Will the questions produce credible information?
- Will the respondents be able to answer the questions?
- Will the respondents be willing to provide the information?

Berg (1998) further explains factors that we need to consider when developing our questionnaire. They include the following:

- Specific ordering of questions
- Phrasing
- Level of language
- Education level of respondents
- Cultural traits and age of respondents

There are four types of questions that are used when developing an interview, focus group, or survey: *essential questions, extra questions, throwaway questions, and probing questions*. Essential questions "exclusively concern the central focus of the study" (Berg, 1998, 65). These types of questions are designed to get specific information from the respondent. Examples of essential questions that we used in our interviews were:

What do you feel makes an effective assessment tool?

and

How does this tool meet those criteria?

These types of questions can either be grouped together in the instrument, as was the case with these two questions, or scattered throughout.

Extra questions are "roughly equivalent to certain essential ones but worded slightly differently. These are included in order to check on the reliability of responses or to measure the possible influence a change of wording might have" (Berg, 1998, 66). Throwaway questions are used to develop a rapport with the subject. The throwaway questions that we used were better classified as discussion questions. We would start the interview by discussing the tool that we were investigating. This developed a good rapport with the interviewee because this was a topic that they were very familiar with and it also got the conversation off on the right foot immediately. These types of questions may be used to avoid starting an interview cold. These questions generally have nothing to do with the topic but get the subject talking freely and make the atmosphere more comfortable. In our case, one throwaway question that we used was:

What was your initial motivation for creating this tool?

This question got the respondent thinking about their tool and the response that is given discusses how much better the tool can make the lives of the students and professors, therefore starting the interview off on a positive foot.

Throw-away questions are also used if the interview has entered a sensitive area. Because our interviews were semi-standardized (which is explained in detail in Section 2.4) the interviewer could digress on topics and ask questions that were not on the question sheet. If at any point the conversation got cold, or the subject seemed disinterested, the interviewer would ask a question from the list and thus get the conversation moving again. Finally, probing questions are used to draw the subject out more. Questions like "Could you tell me more about that?" are examples of probing questions. "Their [probing questions] central purpose is to elicit more information about whatever the respondent has already said in response to a question" (Berg, 1998, 67).

When using these four types of questions to develop our list, there are problems that can occur. One problem that can occur is a "double-barreled question." A double-barreled question is one that actually is asking two different questions. These types of questions are not good when interviewing because the subject will most often only answer one of the questions or ask for the question to be repeated.

These four question types fall into two categories depending upon how they are asked. Informational questions are direct and will allow obvious answers, while attitudinal and behavioral questions are sometimes obscure and often the answers are not obvious to the subject. This means that the questions evoked thought and the answer that is given is based upon the amount of trust the subject has in the researcher. Therefore, we had to use a "funnel approach" to get truthful answers.

The funnel approach involves beginning with general questions, then moving into domain (area of interest) questions, and finally into specific questions. So, the placement of our questions was most interesting first, followed by most difficult questions and lastly the most sensitive questions. Through this process we tried to gain the trust of the interviewees, so when we asked the specific questions we could presume that the individuals would answer our questions. Our questions may not be considered sensitive, but, for example, asking for data that proves a certain tool to be effective may cause some emotion. If the data shows that the tool is ineffective, this data is still necessary for our research but the professor may not want to disclose the data. Many professors may believe the tool they developed and used to be effective, but they do not have substantial evidence to verify this. Therefore we needed to approach questions like these delicately. The question: "*Have you found any disadvantages to this tool?*" needs to be approached by first asking the professor if he or she has seen any problems in other similar tools. This is a domain question and gets the subject thinking about possible disadvantages with tools in general, but we haven't specifically made them think that their tool has these problems.

When developing our set of questions we thought about what information we wanted to obtain. We tried to come up with questions that would draw out answers and give us data that was relevant to our project. As stated earlier, the first question receives the most scrutiny. We did not want to start with a question such as *"Have you found any disadvantages to using the tool?"* Starting with a question like this could have put the interviewee on the defensive. Therefore, as mentioned earlier, our first question was "*What was your initial motivation for creating this tool*?" When asking about the

disadvantages of the tools we wanted to stay away from implying that the tools indeed had some problems. So we worded our question, *"Have you found any disadvantages to this tool?"* to find out if there were problems and what the problems were.

Inter-subjectivity or misinterpretations, was very important for us when it came to analyzing the content of the interviews with National Science Foundation Program Directors and professors. We had to be certain that the questions we asked were actually what the participants heard. Also, we had to be sure that what the participants were saying was what we were hearing. This is important because if the interviewee interprets the question in a different way than we were, the data received may not be as helpful to us. Because the people that we were interviewing are so busy, it was difficult to ask them to clarify an answer later. Therefore we had to get it right the first time.

To ensure that the questions for the questionnaire and interviews were designed properly, we used the Total Design Method (TDM) extensively. According to Berg (1998), the TDM is a good method for implementing Social Exchange Theory and ensuring the subject understands the questions that are being asked. To accomplish this, we used Quality Circles. Quality Circles are made up of the following steps (repeated until the pretests are effective):

- Mapping every aspect of the procedures
- Identifying the weaknesses
- Correcting the weaknesses
- Pre-testing
- Repeat the process

After constructing our initial set of interview and questionnaire questions, we performed preliminary tests with each other and with our liaisons. The advantage of pretesting with our liaisons was that both were former college professors. Although our liaisons have

never created an Information Technology-based Learning Assessment Tool, they have been responsible for funding the development of many such tools. Therefore they have done research on them and understand the objectives of these tools. Pretesting on Dr. Lee Zia and Peggie Weeks enabled us to gain feedback from people that could relate to those we interviewed. Once we made the corrections to the questions based on our feedback and our liaisons' feedback, we completed another pre-test with the same subjects. The second test determined if the weaknesses found in the first round of pretests had been corrected and if any new weaknesses had arisen. Fortunately, it seemed that we had removed all misinterpretations and poor question ordering after the first pretest, therefore making the second pretest easy and efficient. For example, one of the initial questions was:

What do you think is the most valuable component of the tool, if there is one?

This first version limited the respondent to describing only one feature of the tool and introduced a somewhat negative twist at the end. This question was replaced by

What sort of advantages could you foresee for the students and faculty?

The new question allowed for the respondent to give more information and elaborate on the properties of the tool which make it effective. The final versions of the questions used are included in Section 3.2.

3.2 Interview Questions

1. What was your initial motivation for creating this tool?
2. What sort of advantages could you foresee for the students and faculty?

3. What concerns did you have when creating this tool? For instance, roadblocks like costs, or technology? How have you been able to get around these?
4. What disadvantages do you feel are associated with Information Technology-based Learning Assessment Tools?
5. Have you found any disadvantages to this tool?
6. What do you feel makes an effective learning assessment tool? Does your tool meet these criteria?
7. How have you evaluated the effectiveness of the tool? For example, assessment of the tool like surveys or statistical data comparing students grades.
8. Have you received any feedback from the students concerning this tool?
9. Does this tool lend itself to other disciplines? If so, where?
10. Have there been any incentives for the faculty to use this tool? For instance, is there a support team to help the faculty with implementing it into their classes effectively, workshops, or tutorials?
11. Do any other schools use this tool or an adapted version? Which ones?
12. Do you know of any other assessment tools that are being used at your institution or elsewhere?

3.3 Interviews with Tools

Interviews are generally only reserved for human-human interaction. However, in our project we decided to stretch the conventional interpretation of an interview to include human-computer interaction. One of the best ways to determine if a tool is effective is to use the tool. For this project we did just that. The "questions" that we

asked were slightly different than the ones found in Section 2.2. We asked questions concerning the features, and ease of use. We were interested in learning if the student would have the opportunity to change an answer at any time, and if the tool graded the problems instantly. We logged into the tools exactly as a student would. Fortunately, most of the tools were available on the Internet for anyone to use. Once using the tool, we tried all of the functions that the particular tool offered (feedback, grading, etc.). We often submitted answers that did not make sense in order to see what feedback was available and the extent of information a student would receive. All of this information was used exactly as we used the information gained from the interviews, and was analyzed in the same manner. The techniques are explained in Section 3.6.

3.4 Frame for the Interviews

The frame defines the exact population of people that we were interested in gaining information from or about. Our exact population included eleven Program Directors from the National Science Foundation, and five professors who either developed or were involved in the development of a tool. The number of Program Directors was determined in the middle of A-term after our liaison, Peggie Weeks, sent an email to Program Directors throughout the National Science Foundation. This email requested information from the Program Directors concerning IT-based learning assessment tools. The number of professors contacted was limited by time constraints. As our list of tools grew we knew that we could only talk to a fraction of the professors that were represented by the list. In order to have sufficient time to analyze all the interview data, we set the deadline of December 1, 2000 as the last day that we would schedule interviews with professors.

3.5 Interviews

Phone interviews and face-to-face interviews were our main methods for gathering data for this project. An interview is a conversation with a purpose (Berg, 1998). The purpose is to gather data that is relevant to the topic that is being researched. There are three types of interviews: *standardized, unstandardized and semi-standardized.* A standardized interview is rigidly structured and the interviewer asks the interviewee questions that have been predetermined (Berg, 1998). A standardized interview is used when the interviewer believes that the questions he or she will ask are comprehensive enough to gather all of the necessary data. An unstandardized interview is the opposite of a standardized interview, meaning that the interviewer has no prepared questions. The "interviewers must develop, adapt and generate questions and follow-up probes appropriate to the given situation and the central purpose of the investigation" (Berg, 1998, 61). Finally, a semi-standardized interview is a combination of the other two methods. In this method, the interviewer has a series of predetermined questions but also allows the interviewee to digress on a topic if it is relevant (Berg, 1998). This method also allows for the interviewer to use probing questions to draw the interviewee out on a topic.

We used a semi-standardized interview. The questions were developed, and refined using the Total Design Method. However, if the person digressed on a topic that was relevant, we allowed this to occur and tried to probe further. This happened quite often when conducting our interviews. The professors were interested in discussing the advantages of their tool as well as their motivation for creating these tools.

One person conducted the interview and one person took notes. Both the person conducting the interview and the person taking notes came prepared with clipboards, plenty of paper, a copy of the questions, and a few extra pens. During our phone interviews we asked all of the professors if it was all right to record the interview for transcription purposes. All of the subject agreed, but were all interested in reviewing the areas of our report that included their quotes so that they could make sure that we didn't quote them incorrectly.

When conducting the interview, we followed the "Ten Commandments of Interviewing" as defined by Berg (1998):

- Never begin an interview cold (Use small talk to begin the interview and make the interviewee more comfortable and talking freely.) We may would start the interview asking how the person was doing and then go on to explain what we're doing. The people we talked to were free to say what they please and then we went on to ask for permission to tape the interview.

- Remember your purpose (Keep the subject on track.)

- Present a natural front (Ask questions as if they just 'popped' into your head.) After the interviewee answered a question that may have been on the list we would often ask them another question to further explain on the topic or to get additional information.

- Demonstrate aware hearing (Offer the subject non-verbal responses to show that you understand and are paying attention to what they are saying.)

- Think about appearance (Business attire is most appropriate.) This is important, though it did not apply to us because we conducted phone interviews.

- Interview in a comfortable place (where the subject feels comfortable.) When we talked to the program directors we went to their offices, so we hope they were comfortable there.

- Do not be satisfied with monosyllabic answers.

- Be respectful.

- Practice, Practice, and more Practice!

- Be cordial and appreciative. (Thank the subject and answer any questions that they might have.)

At the completion of the interview, the person taking notes left the room and did not discuss anything with anyone. He or she transcribed all of the notes into a record of the events that occurred in the interview.

The person conducting the interview wrote down key words or 'buzz words' throughout the interview. After the interview was completed, these short notes were transcribed into a record of the interview as well. It was imperative that these notes were transcribed immediately after completing the interview so that nothing was forgotten. All of these notes were then left alone until it was time to analyze the data.

In order to get useful data we had to produce relevant questions. Creating good questions was vital to these interviews to ensure that they flowed well and the content of

the answers was appropriate. This was explained in greater detail in Section 2.1 of this methodology.

3.6 Content Analysis

Once we conducted these interviews we analyzed the data obtained. Content analysis is the systematic analysis of results of interviews (Berg, 1998). Content analysis includes the counting of seven elements to find similarities between what interviewees say, which are:

- Words
- Themes
- Characters
- Concepts
- Paragraphs
- Items
- Semantics

In content analysis, *words* are the smallest unit counted (Berg, 1998). For example, a person may mention the same adjective frequently when discussing a topic. A *theme* is a simple sentence (Berg, 1998). A reoccurring simple sentence can depict a general feeling a person has. For instance, if a professor states "studies show that students who use the tutors do better," ten times in a twenty minute interview, it may be possible to ascertain that this professor is an advocate of the Tutor function on IT-Based Learning Assessment Tools. *Characters* are people, we would count the number of times a specific person or persons are mentioned (Berg, 1998). An *item* is a whole unit, for example a book, a letter, or even an in-depth interview (Berg, 1998). *Concepts* are words that are group together into an idea. This method is a more detailed way of counting words (Berg, 1998). "Words such as crime, delinquency, kiting, and fraud might cluster around the

conceptual idea of deviance" (Berg, 1998, 232). *Semantics* relate to the number and type of words as well as how affective the words are. Affective words are words that provoke some kind of emotional response for a person, generally negative (Berg, 1998). We used the extensive transcriptions of our interviews to find these seven items as well as analyzing trends that we discovered from the formulation of our list of tools. Words that we heard frequently during our interviews were *advantages*, *homework*, *time*, and *cost*. Content analysis is important for our project for a number of reasons. Half of our project is to evaluate the effectiveness of Information Technology-based Learning Assessment Tools. Content analysis of an interview with developers of WebWork yielded the phrase "immediate feedback" nine times. We can make the assumption that immediate feedback is something that is important to be included in a tool, and therefore the effectiveness of it needs to be explored.

For our analysis, we counted words and themes more than anything else. In our process of coming up with a list of elements for an effective assessment tool, we looked at the various descriptions of the tools found on web sites and in paper-based literature. We would compare, for example, what was said for the homework tools, noticing most of them told the students what answers were right and wrong. Therefore, we put this down as a component for graders to be effective.

3.7 The Matrix

Our objective for this project was to investigate and evaluate Information Technology-based Learning Assessment Tools. While investigating these tools, various characteristics of the tools that were similar and different became apparent. These characteristics were collected and organized in a spreadsheet, which we call "The

Matrix". Each row represents one tool that we found during our investigation. Each column heading represents one piece of information pertaining to the tools. These came from our previous research, as well as information that we found during our interviews with people and the tools. A complete glossary of the Matrix column headings and explanations for why they were chosen is found in Appendix A. In Chapter 2, Literature Review, we discussed the *Nine Principles of Good Assessment (AAHE, 1997).* One of the principles was that assessment is best when it is ongoing and not episodic. Tools that have the ability to grade large amounts of work in very short periods of time save professors a lot of time and allow them to assign more work than normally. We created one of the Matrix headings as *Grader* to encompass this functionality. This is effective because now the professor can make the assessment ongoing instead of episodic. Another heading that is found on the Matrix is *Generation.* In the Literature Review in Chapter 2, Dr. Robert Linn (1993) defined four generations of Information Technology-based Learning Assessment Tools. These generations represent broad categories that all tools can fit into.

These headings were then used to then make comparisons between the tools and helped to determine what makes a tool effective. All of this data and the results that were found are located in Chapter 4.

Chapter 4 Results and Discussion

Chapter 4 introduces the results of our research for the past 14 weeks. It discusses our observations from the investigation of IT-Based Learning Assessment Tools and the effectiveness of these tools in undergraduate education.

4.1 Summary of Tools

Table 1, found below, depicts the categories that each tool fits into. As you can see some tools encompass multiple categories. The majority, about 35%, of the tools can be classified as tools that test students.

Table 1: Categories of Tools

Tool	Homework	Tutor	Test	Other
CyberTutor	X	X	X	
OWL	X	X		
JiTT			X	
Electronic Student Portfolio				X
WebWork	X	X		
Accounting Tutorial		X		
LabView			X	X
Computer Assisted Test (Chem)		X	X	
The Quiz CGI	X		X	
IIL		X		X
WhizQuiz			X	
Teaching Engineering Graphics		X	X	
SALG				X
COMPASS			X	
Mallard	X		X	
WebAssign	X			
QuizSite	X		X	
CAPA			X	
Gateway Test			X	
Visual Calculus		X		
Online Homework and Quizzes	X	X	X	
Intelligent Essay Grader				X
Total	8	9	13	5

Based on Linn's definitions of the four generations of IT-Based Learning Assessment Tools, 80% of the tools that were revealed in our investigation can be classified as Generation one tools. In contrast, only about 6% can be classified as Generation 2 tools, and no tools are classified as Generation 3 tools.

Table 2: Generations of Tools

Tool	Generation I	Generation II	Generation III	Generation IV
CyberTutor				X
OWL				X
JiTT	X			
Electronic Student Portfolio				
WebWork	X			
Accounting Tutorial	X			
LabView				
Computer Assisted Test (Chem)	X			
The Quiz CGI	X			
IIL				
WhizQuiz	X			
Teaching Engineering Graphics	X			
SALG				
COMPASS		X		
Mallard	X			
WebAssign	X			
QuizSite	X			
CAPA	X			
Gateway Test	X			
Visual Calculus				
Online Homework and Quizzes				
Intelligent Essay Grader				
Total	**12**	**1**	**0**	**2**

The following table (Table 3) compares the Generation that the tools are classified in with the general use of the tools.

Table 3: Distribution of Tools by Generation and Use

	Generation I	Generation II	Generation III	Generation IV
Homework Graders	6	0	0	2
Tutors	4	0	0	2
Tests	8	1	0	1
Other	1	0	0	0
Total	19	1	0	5

4.2 Description of IT-Based Learning Assessment Tools

During our research on information technology-based learning assessment tools we investigated 22 tools. These tools include homework graders, on-line exams, tutors and other various tools. Following are descriptions of the tools that we found. Our descriptions include information that we gathered from literature, information on the tools' website, from talking to professors and developers of the tools, and our own experiences in using some tools.

4.2.1 COWculus

"Calculus on the Web" or COWculus, is the name of a program currently being used at Temple University. Students use COWculus to practice calculus problems. Students can either practice or login and have their scores recorded. They are able to check their answers before going on to the next problem. They can also get help if they need it (Temple, 2000). There is an expression interpreter that turn the equation in a more simplified form, which may help the student understand better by allowing the student to see the simpler parts that comprise the complex problem. If questions are missed, the student is given a second chance to complete the problem correctly. A student's answers are analyzed using Maple, a computer algebra system capable of analyzing symbolic mathematical expressions (Temple, 2000).

4.2.2 Accounting Tutorial

The program at West Virginia University has incorporated an online tutorial into their accounting classes. Students are able to select the questions they want to answer. If the answers are incorrect then the correct answers are given. The basic format of the tutorial includes a discussion of the topic, which are brief containing important concepts and helpful hints. There is also a function called *Contact Tutor* that allows students to e-mail a tutor with a question pertaining to a given topic. The *Question* section contains multiple-choice questions that are relevant to the topic being studied. The *Answers* section defines the letter corresponding to the correct multiple-choice answer. Students can click on this letter which leads them to an explanation of why this is the correct answer (WVU, 2000).

4.2.3 Electronic Student Portfolios

At the Virginia Commonwealth University (VCU), engineering students participate in a program referred to as *Electronic Student Portfolios* (VCU, 2000). In this program (1999-2000 school year), 63 engineering students design and develop their own Internet web sites that contain certain required information as well as information that they wanted to include. Information contained in one student's portfolio was, for example, the student's year of graduation, major, electronic version of their resume, a reflective essay, email address, relevant courses, an example of a lab report, and various essays relevant to their major. These portfolios are then evaluated by members of the

VCU Industrial Board for content of the documents and development of the web site. The Industrial Board is comprised of people outside of the University, including CEOs of various companies (Tait, 2000). The Industrial Board evaluates the students through the use of a scoring rubric provided by VCU. The rubric is comprised of five *Learning Outcomes* that has a group of questions to be scored on a scale of zero to three (three being the best). The *Learning Outcomes* are *Laboratory*, *Design* and *Innovation*, *Written Communication*, *Interdisciplinary Approach*, and finally *Life-Long Learning and Global Issues* (VCU, 2000).

4.2.4 Interactive Internet Laboratory

Bernard Mohr, of Queensborough Community College (QCC), has been involved in developing a device that allows electrical engineering students to complete lab work outside of the lab. The lab is called Interactive Internet Laboratory (IIL). In his proposal to the National Science Foundation, Mohr explains, "what has been lacking in the distance-learning arena are hands-on laboratory experiences" (Mohr, 1999). He continues to say that many community college students work during the day and have to travel great distances to attend classes, therefore making scheduling classes difficult especially classes with labs. "For them a distance-learning laboratory would make full time enrollment possible and perhaps accelerate their graduation" (Mohr, 1999). ILL combines software and hardware components to produce a lab environment. The hardware component is a data acquisition device called "e-LAB", and the software component, referred to as webLAB, interfaces the e-lab portion with a computer and the

Internet. (Shown as Figure 1 below) Not only is this system convenient for students, but it also has certain advantages over traditional lab work. For example, when students are using the tool, there is a web page shows only one portion of the lab. In order to move on to the next question, the student must finish the current question completely and correctly and answer questions to ensure that comprehension of the topic is also mastered. There are online tutorials that aid the student in completing the lab. The system also gives feedback to the professor, alerting him or her to the questions that gave students more difficulty. This helps the professor to focus on these topics during class time to reinforce the concepts.

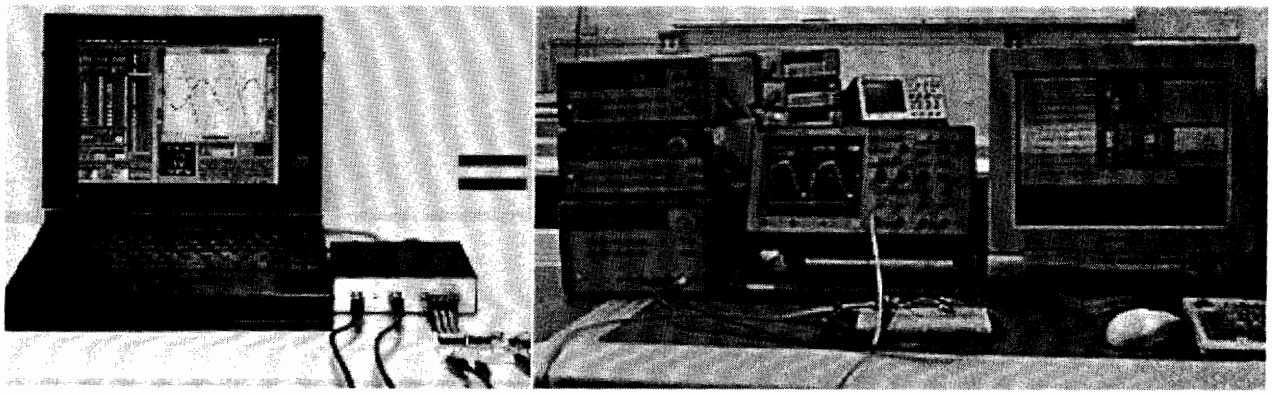

Figure 1: Photograph of the e-LAB system

One of the main disadvantages to this tool is the cost. Because of the technology and hardware required in making one of these devices, the cost for just one unit is $6500. With a large portion of a community college's students having full-time jobs, the cost of owning enough of these IILs to help benefit the class is enormous.

4.2.5 Labview

Frostburg State University (FSU) utilizes an Internet-based chemistry set for distance education. The experiment is not a simulation, but the students control equipment, located in remote locations, via their computers. The web interface is used to collect data, to obtain interactive technical support and background information, and to display and analyze the results. This project deals with the issue of practical experiences in Web-delivered courses by providing students with remote access and control of real equipment (FSU, 2000). The experiment does, however, involve equipment that is unavailable in most undergraduate laboratories. The Labview program has a simple graphical interface that is used to control the apparatus onsite. (As shown in Figure 2 on the next page) The interface is used during the assessment stage in order to compare the performance of students accessing the experiment online and offline (FSU, 2000). There is a Web camera that collects pictures of the equipment during the process to be used for data collection and helping students monitor the reaction's progress. The experiments are performed on an instrument panel that recreates the actual look and feel of the equipment. There is a panic button on the panel that commences a conferencing connection with the system's caretaker via NetMeeting during working hours and by email at other times. The panel is used to observe and interpret student activity, identify and correct conceptual problems, and to prevent actions that could damage the apparatus (FSU, 2000).

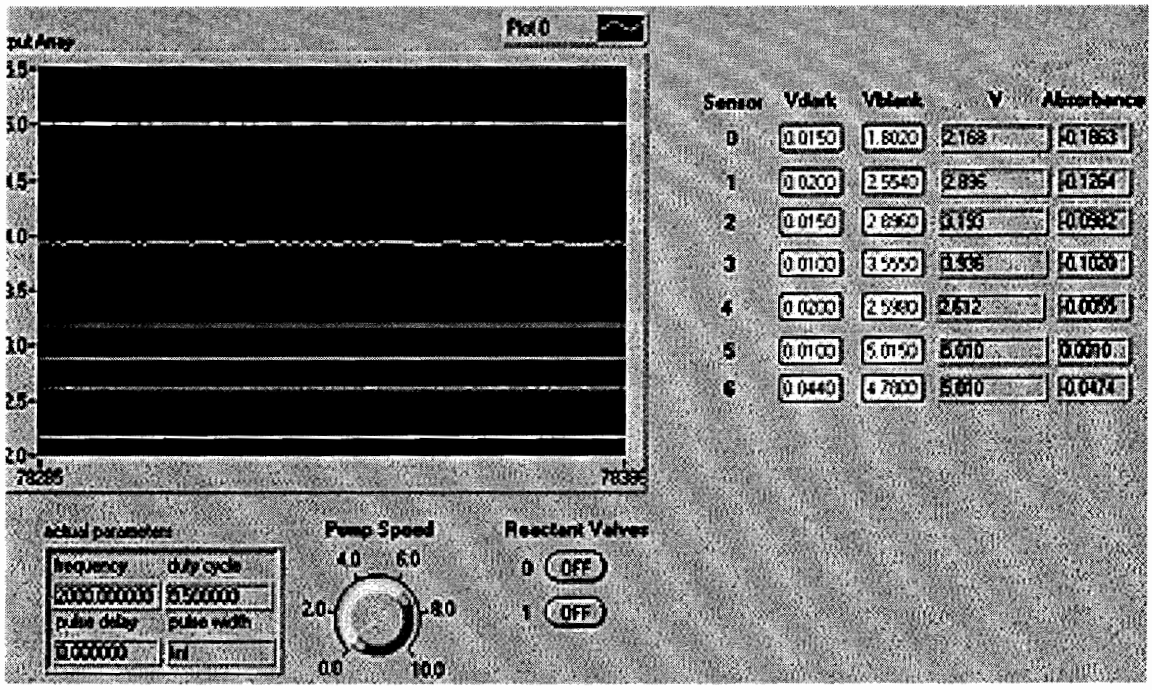

Figure 2: Labview Monitor image

4.2.6 WhizQuiz

Virginia Polytechnic Institute has a quiz on the web called *WhizQuiz*. The quiz contains multiple choice questions, true/false questions, and drop down list of answers, where the users choose from the list of choices. Drop downs are used when there is a list consisting of 1 to 2-word choices, where a popup menu will be produced. WhizQuiz also supports questions that require exact numeric answers or answers with maximum and minimum acceptable values. Answers are graded according to the key provided by the quiz author. The results are displayed with correct answers. This tool supports images and also provides help links for students. There is also a password protection, which prevents the grading of the quizzes but does not stop users from looking at the questions (Vtech, 2000).

4.2.7 Enhancing the Teaching of Engineering Graphics

There was a multimedia instructional CD-ROM/web page developed in response to tutorial videotapes for an engineering graphics class at the University of Texas. Currently this project, entitled "Web-Based Learning: Enhancing the Teaching of Engineering Graphics", and consists of an integrated web site with links to hours of tutorial movies, lecture presentations of all class lectures, web-based games and interactive quizzes. The lectures are accessed through links to Microsoft PowerPoint presentations. Lotus Screen Cam movies present lab tutorial material. (As shown in Figure 3 below) Games and quizzes are Java Applets created in Java Script 1.2 code and displayed on the web through HTML files. The project requires a little more than 2 Gigabytes of computer storage (IMEJ, 1999).

Movies can be broken into segments that allow students to go forward or backward to the start of each segment. The size of a typical 10-minute tutorial is approximately 4 megabytes of computer storage.

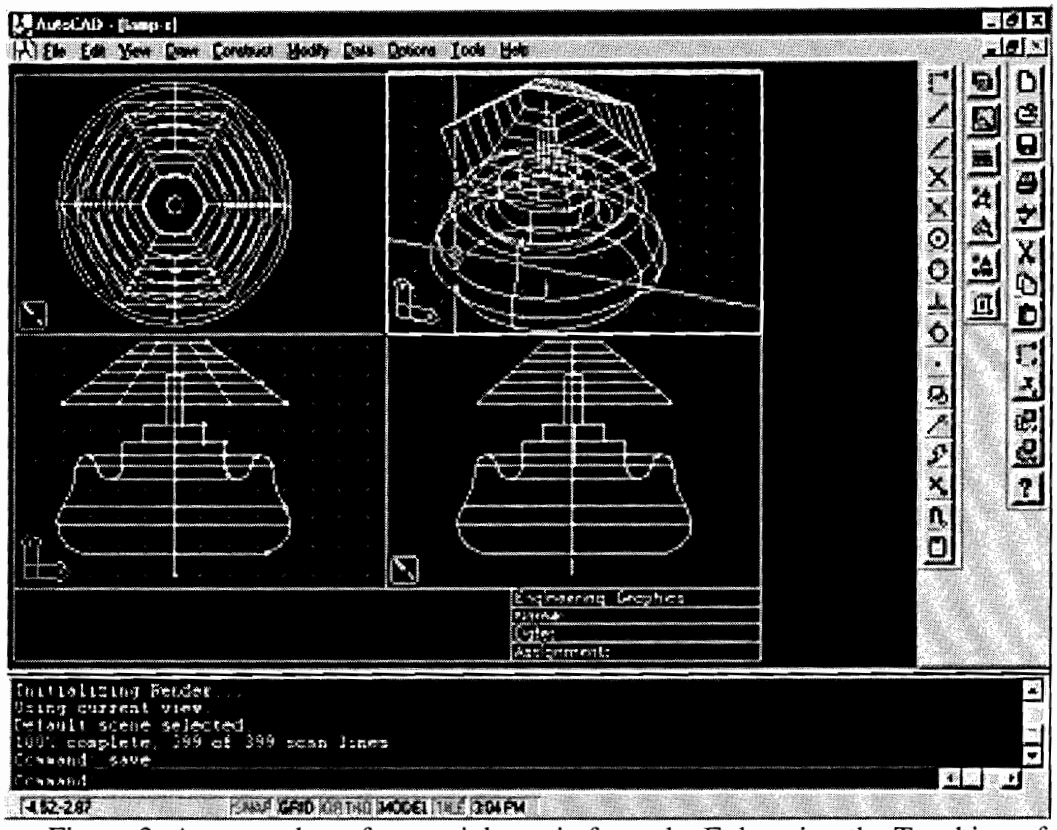

Figure 3: A screenshot of a tutorial movie from the Enhancing the Teaching of Engineering Graphics program

There are true or false and multiple-choice style questions, with the answers stored in memory. When the user selects the *Grade* button, their answers are compared to the array containing the correct answers. As shown in Figure 4, questions missed are displayed in an "alert window" (IMEJ, 1999).

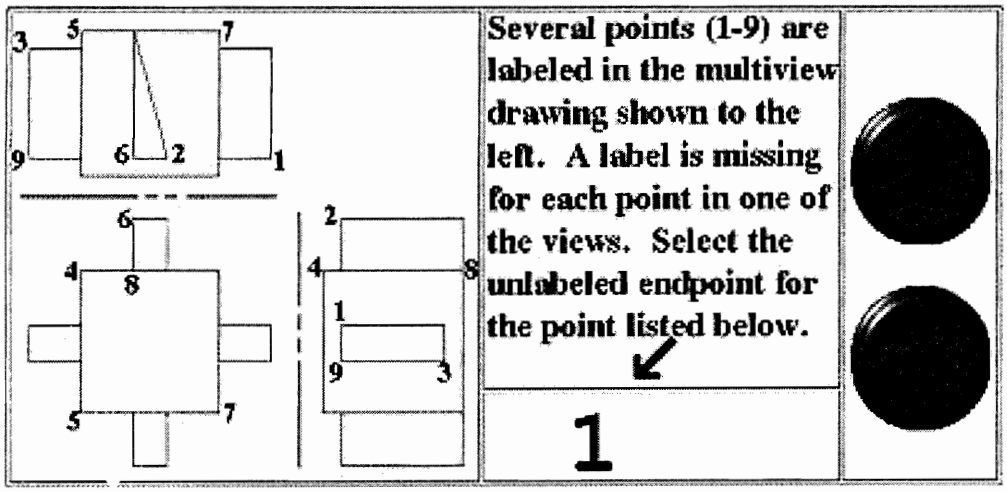

Figure 4: A JavaScript game that uses image maps to aid students with graphic visualization techniques

4.2.8 Just in Time Teaching (JiTT)

Assessment doesn't have to be restricted to the students' learning, but it can also be applied to the teachers and their courses. "Trying to read eyeballs, it's tricky," explains Dr. Myles Boylan, a Program Director at the National Science Foundation, "[professors] can't wait until the exam to find out how the students are doing." Boylan is explaining the problems that many professors have determining how a course is going. It would be beneficial for professors to gain feedback from their students throughout the course, not just when the first exam is given.

Gregor Novak, professor of physics at Indiana University-Purdue University Indianapolis, developed a software tool in the early 1980s that allowed students to solve physics problems and complete exercises at their own pace on computers (Rozycki, 1999). Novak explained that although the computer was very patient, it lacked one key item. "We know from cognitive learning studies that group interaction is a key ingredient in enhancing learning. The computer-assisted learning programs of the 1980s,

including mine, lacked that one ingredient. There was no sharing of ideas with peers or teachers" (Rozycki, 1999). Novak, with help from Andrew Gavrin and Evelyn Patterson, developed a program called *Just-in-Time Teaching* (JiTT) to address the issues of using computer-assisted learning with the necessary teacher and peer interaction. Students taking Physics classes with the JiTT program in place, complete a short quiz briefly before attending class each day (Rozycki, 1999). The results of the quiz do not affect the students' grades, but are used as a benchmarking tool for the professor to base the lecture on. If the results show that there is a lack of understanding in a particular area, the professor knows what to focus on for that day's lecture. This approach also keeps the professors from moving the class at an unreasonable rate, whether it is too slow or too fast. This program has shown considerable success in practice, "of those surveyed after two semesters of JiTT courses, 92 percent preferred the approach to a standard course" (Rozycki, 1999). JiTT isn't the only program that gives teachers feedback on how the course is progressing, SALG found in section 4.2.9.

4.2.9 Student Assessment Learning Guide (SALG)

SALG, Student Assessment Learning Guide, is a web-based program that allows students to log in anonymously and give feedback on a particular course that they are enrolled in. SALG includes 46 sample questions that the professors can choose from to comprise the questionnaire that the students will see when logging onto their course. Professors have the opportunity to adjust the questions, even add and remove questions throughout the semester. The results are returned to the professors in the form of charts

and graphs. The professors also have the option of seeing the raw data and developing their own conclusions (SALG, 2000).

4.2.10 Mallard

Mallard is a Web-based, interactive quizzing program. It can be used as a quiz or as a homework grader. Mallard offers the following features:

- Blank-Provides a plain input blank for the student to type in text.
- Multiple Blank-Provides a user-specified number of blanks for the student to type text into.
- Arith-Provides a blank for the student to type in a number or a mathematical expression.
- Multiple Choice.
- True/ False.
- List-Provides a single blank in which the student types a list of items, similar to Multiple Blank (Mallard, 2000).

Several versions of the same question can be made but it is not recommended that questions of multiple parts be made. In Mallard answers are case-sensitive and punctuation also matters determining whether a question is right or wrong. When the answers are graded, feedback that has been prepared by the developer is given.

4.2.11 The Intelligent Essay Assessor (IEA)

The Intelligent Essay Assessor (IEA) is a set of software tools designed to evaluate the quality of essay content. The text analysis underlying the essay-grading scheme is based on Latent Semantic Analysis (LSA) (IMEJ3, 1999). LSA methods focus

on the conceptual content, the knowledge conveyed in an essay, rather than its style, or even its syntax or argument structure. In order to assess essays, LSA is trained on domain representative texts, which include, textbooks, articles or writing samples that students would encounter while learning in that area (IMEJ3, 1999). This allows the tool to have a semantic representation of the information contained in the subject matter.

This tool can also detect whether an essay is very different from its domain. In this case, the computer would 'flag' it for human evaluation (IMEJ3, 1999). Essays that are flagged are ones that are off topic or contain unusual syntax. The program is also set up to detect plagiarism. When instances of plagiarism are directed to the instructor for additional grading considerations.

The essay grader can do a simple grading that takes about five seconds or can be set to grade in an extended time of 20-30 seconds. The IEA then gives the student a score from zero to one hundred and gives feedback to explain the grade.

4.2.12 WebWork

WebWork is an internet-based system for generating and delivering homework problems. It was developed in 1996 at the University of Rochester, and built upon the Web-based program, CAPA, developed at Michigan State (Rochester, 2000). The tool is now used for pre-calculus, calculus and physics classes. WebWork is web-based allowing students to access it from any computer. Since many students review old assignments to prepare for an exam, the server that WebWork is housed in can accommodate large number of simultaneous connections. The tool allows the students to try a problem as many times as he or she wishes before the due date. One of the key

benefits to using WebWork is immediate feedback. This enables the students to correct mistakes while the problems are fresh in their mind (Rochester, 2000). Students are encouraged to seek help elsewhere from fellow students, from the TA's or from the instructor (either in person or via e-mail). Educators at University of Rochester believe it is beneficial to students to seek help from humans when they are experiencing difficulty (Gage, 2000). With the use of individualized problems, students cannot cheat in the sense of copying each other's answers but are able to collaborate on the methods of solving a problem. (As shown in Figure 5)

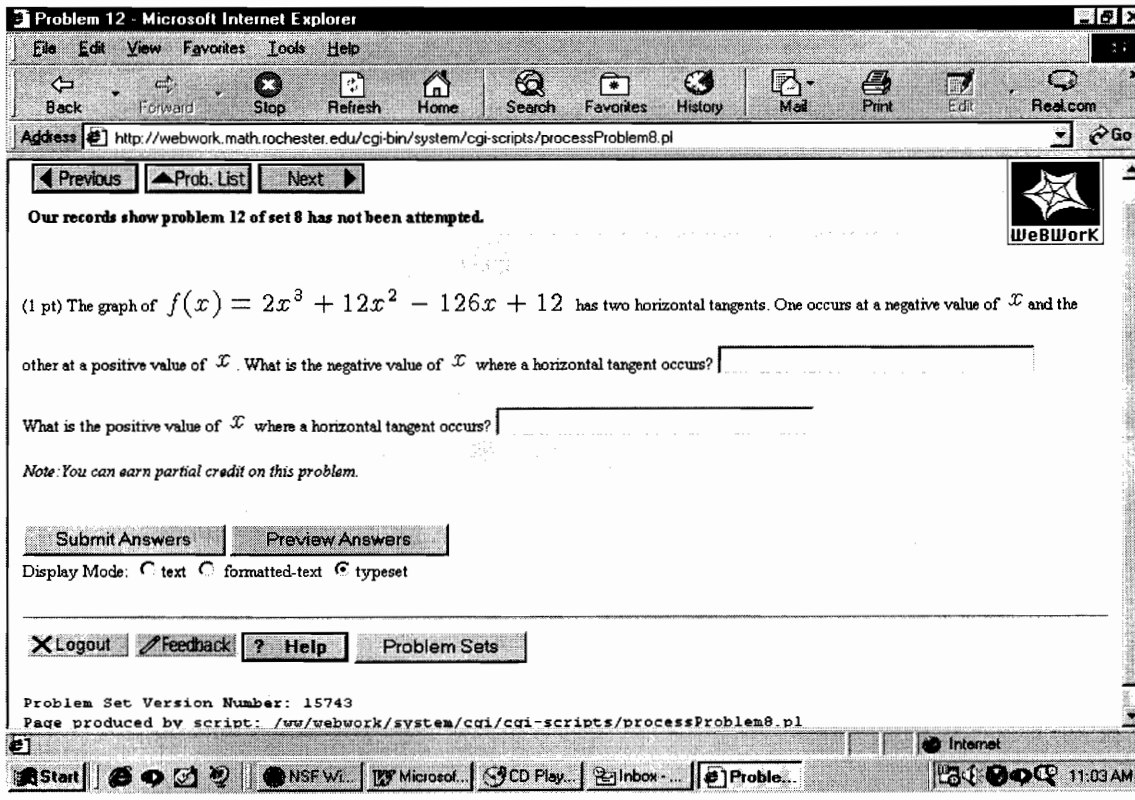

Figure 5: Sample WebWork problem

There is a mechanism that shows the instructor the same problem a particular student is using to answer questions better, as well as showing the professor the current progress of the students (Rochester, 2000). Notice, in Fig. 5, the *preview* button that

shows the student what his or her answer looks like. Students are able to press the *feedback* button and contact the professor or a TA to receive more assistance or to send comments. The *help* button links to a page that gives an explanation of all the functions on that particular page (i.e., submitting answers, the *preview* button, *feedback* button, etc.).

WebWork grades homework problems and gives immediate feedback. Partial credit is given on problems that have two part answers. Developers of this tool, at this point, are unclear of the advantages that the hints would have for the learning process and therefore have decided to leave them out.

4.2.13 Visual Calculus

When Visual Calculus was first being developed in 1996, the developers' goal was to be able to show professors how technology (computers in particular) could be used in the teaching of calculus. What resulted from this is an online tutor for students enrolled in pre-calculus and calculus classes at the University of Tennessee, Knoxville, TN.

Visual Calculus combines a tutorial, discussion, and drill problems. The tutorial contains a detailed explanation of calculus concepts. The program is basically an online textbook. Students complete the problems by hand (using paper) and check their work on Visual Calculus. The discussion areas provide more detailed explanations of the various aspects of the calculus concepts. These aspects are displayed to the student using Adobe Flash technology, and are displayed with words, in easy common language to avoid confusion, and accompanying animated pictures to reinforce the concept. The discussion

area continues to give examples of the concepts in symbolic or algebraic terms, numerical terms, graphical terms, and finally verbal terms in an effort to help students with different learning styles. The drill problems are simply problems that could be found in a textbook and are all on one web page in a list. Next to each of the problems there is a link to a solution for that particular problem.

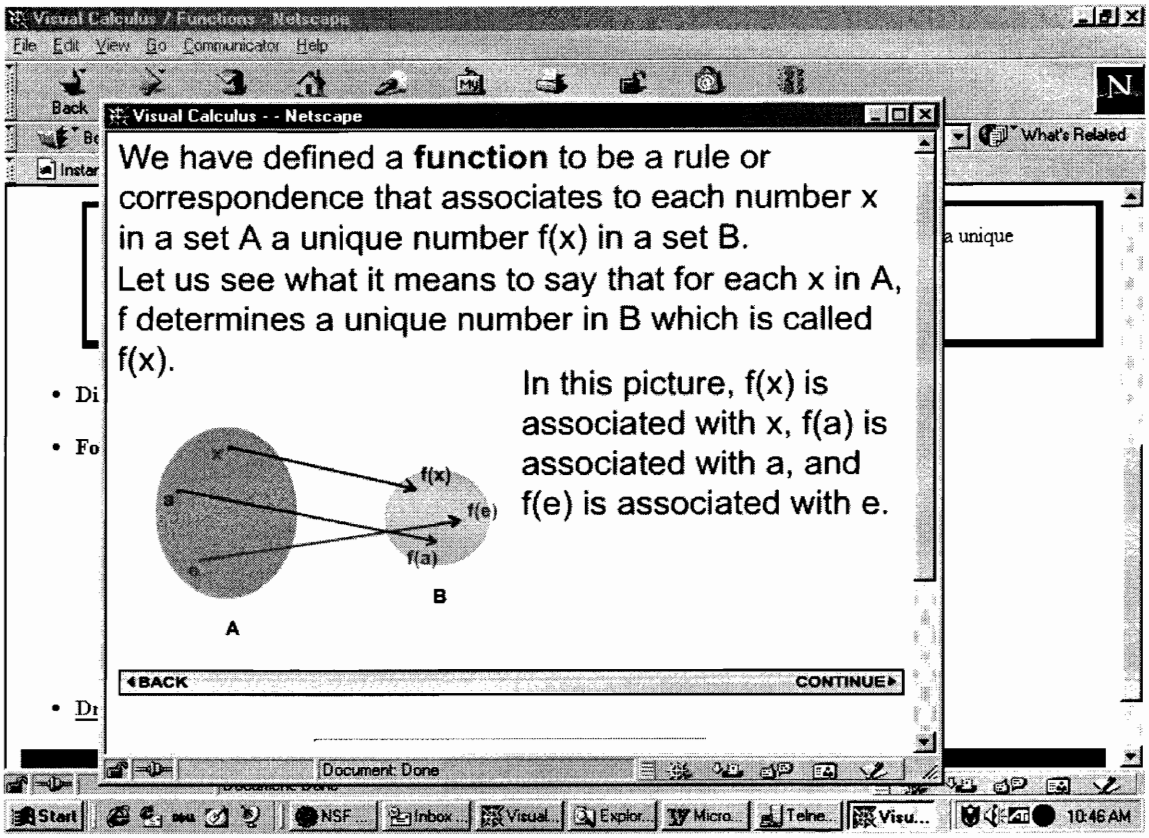

Figure 6: Example of Visual Calculus's Tutorial features

Visual Calculus is housed on a web server, and easily accessible to all students at all times of the day or night. The two main features that we feel this tool is lacking are an interface that allows the students to answer the questions through the computer, and a feedback component (as shown in Figure 6). The students need to complete the problems by hand on paper and then they can check their answers against the solutions, but if they

get an answer wrong there isn't any feedback to help the students know where they went wrong.

4.2.14 CAPA

CAPA is an Information-Technology-based Learning Assessment Tool that allows teachers to create quizzes and tests for their astronomy, biochemistry, chemistry, mathematics, physics, botany, accelerator physics, human food and nutrition, family and child ecology, and computer science courses. The questions can be presented in a variety of different ways including multiple choice, short answer, numeric answer, and matching. CAPA can even support questions that require a graphical answer, and can pose questions using sound and speech as the medium (Kashy, 2000).

Students are given instant feedback and relevant hints and may correct errors without penalty prior to an assignment's due date. The system keeps track of students' participation and performance, and records are available in real time both to the instructor and to the individual student (CAPA, 2000).

There is a discussion portion on CAPA that allows students to interact with a faculty member. CAPA allows the exams to be timed, which allows students to take these tests at home. To create an exam the professors put together a template of questions that the CAPA system can then use to create similar but different questions so that no two students get the same exam. The 'template' would be a question that the professor made up and then the computer would just insert different numbers or variables to make the questions slightly different.

One of the main advantages to this tool, as well as other IT-based learning assessment testing tools, is their ability to save time for the professor. If a professor is teaching a course for a class of over a hundred students, it becomes extremely difficult to administer quizzes on a weekly basis. However, as we have noted earlier, assessment is best when it is ongoing and not episodic. Although this is desirable in a large class it isn't feasible without a tool like CAPA. According to Dr. Kashy (2000), of Michigan State University, in a class where CAPA was used to give unannounced individualized quizzes, resulted in an average attendance rate of ninety per cent.

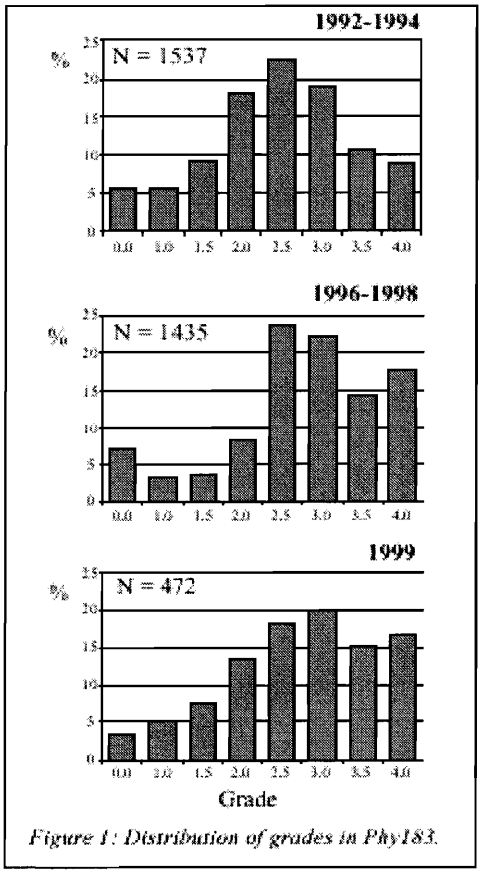

Figure 1: Distribution of grades in Phy183.

Figure 7: Charts the grade distribution in Physics 183

Figure 7 compares the years (1996-1999) when Phy183 was taught with the use of CAPA (use becoming more extensive each year) to the years (1992-1994) when the same

class was taught using the tradition methods (Kashy, 2000). These classes had an enrollment size of roughly 500 students. In the years when the CAPA wasn't used, only 60% of the class or 300 students obtained the passing level. In the years 1996-98, 78% of the class or 390 (as shown in Figure 7).

4.2.15 Computer Assisted Testing (Chemistry)

Professor Carl David, from the chemistry department at University of Connecticut uses a Web-based system for assigning homework and testing the students in his physical chemistry class. As seen in Figures 8a and 8b, the following features are offered:

- A link to get an explanation of the problem given
- Drop down lists offered to select correct numbers
- Students are able to send comments via e-mail to the professor
- After submitting the answer, the tool tells the student whether his or her answer was right or wrong and allows the student to look at the correct answer if they choose

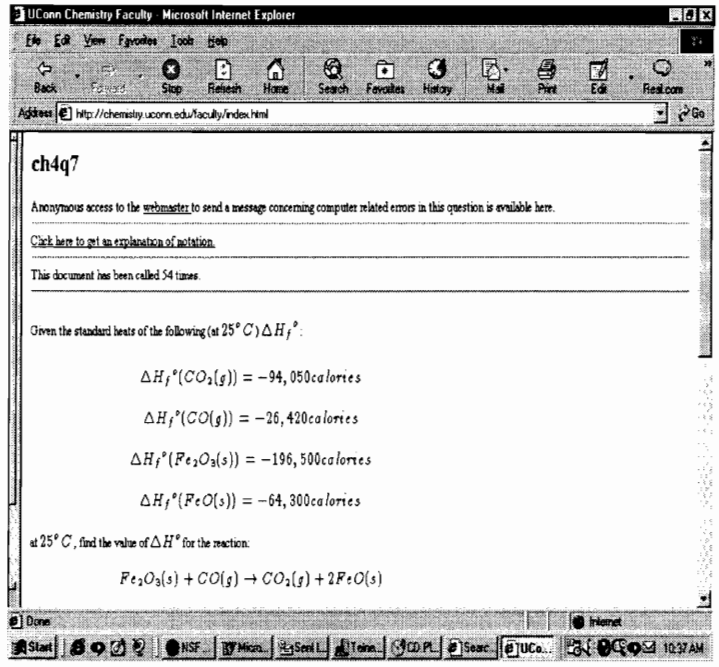

Figure 8a: Sample physical chemistry problems from CAT at UCONN

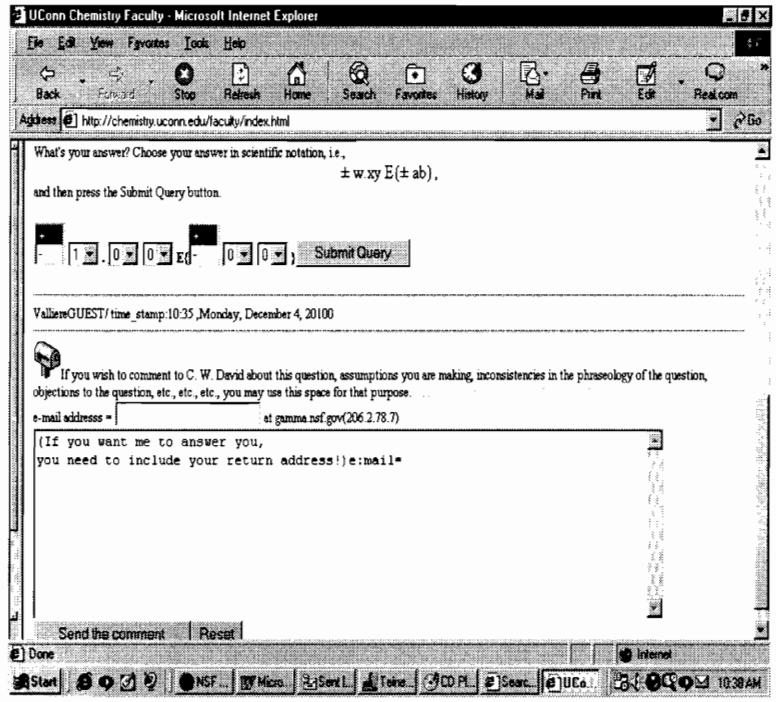

Figure 8b: Sample physical chemistry problems from CAT at UCONN

4.2.16 Quiz CGI

Quiz CGI is a perl script that administers quizzes. Instructors create text files, which assigns values to different variables. Multiple choice and short answer are the most common questions used. Students' answers to short questions are compared against answers instructors have given.

4.2.17 WebAssign

WebAssign is used to deliver and grade homework assignments over the Internet. This tool is used at several schools, including North Carolina State University, Massachusetts Institute of Technology, Northeastern University and Rensselaer Polytechnic Institute. Assignments are generated with different numerical values in order to individualize the questions. More than 1000 students are using it at North Carolina State University. WebAssign offers multiple choice, multiple select, numerical and fill in the blank questions. There are options that allow the professor to provide hints or detailed solutions to questions. When professors allow students to resubmit their homework, they are able to edit their latest version and submit the assignment again. Students are able to review their submissions by following a link on their assignment summary, which is used for that purpose.

4.2.18 QuizSite

QuizSite was developed by BEST (Bureau of Evaluative Studies and Testing) and is used at Indiana University at Bloomington, ID. "QuizSite was created specifically for Indiana University. Over five thousand students are currently using this tool. It was not

designed to be modified for use at other universities." (IUB, 2000) It delivers homework and exams via the World Wide Web. QuizSite allows for timed exams where professors specify a start and end time for accessing an activity, however for the activity to be timed, all students must start at the same time (IUB, 2000). QuizSite is unable to show the remaining time left on the exam. This may cause a problem because there are not any restrictions to the location where the tool can be accessed. QuizSite also provides immediate feedback to both instructor and students. Multiple-choice, matching, numeric, and fill-in the blank questions are scored automatically but the instructor grades essay and short answer questions.

4.2.19 COMPASS

COMPASS is a computerized adaptive test used for placement at Ivy Tech State College. It is used to test writing skills, reading comprehension and knowledge of mathematics (Numerical Skills/Pre-algebra, Algebra, College Algebra, Geometry, and Trigonometry) (COMPASS, 2000).

4.2.20 Derivative Gateway Test

The Derivative Gateway test was developed to assure that students could take derivatives quickly and accurately. Students are allowed to take the test as many times as is necessary to pass it, however only one attempt per day is allowed, and they must pass it by a certain deadline. Individualized problems are presented to each student. There is a penalty for not passing the test by the deadline, which is a reduction in the final course grade by one-third of a letter (Gate, 2000). If the test was not passed on the first

attempt, then all other attempts must be made in one specific math lab. When this happens, there is usually a backup of students waiting to use the tool scoring is not neither completed nor received immediately (Gate, 2000).

4.3 Categorizing the Tools

Throughout our research we have come across a variety of information technology-based learning assessment tools. These tools range in function and the disciplines in which they can be used. As we investigated these tools, we realized that each tool had its own objective as well as characteristics. We had to categorize them based on similarities between these objectives and characteristics. As we mentioned in the Literature Review section of this paper, Linn (1997) has defined four generations of IT-based learning assessment tools. We used Linn's generations as broad categories and then created our own more specific categories that all fall within the generations. As seen in Table 2, most of the tools we encountered were generation I tools.

We categorized them as Homework Graders, Tutors, and Tests. The generations that Linn described can be applied to all of the tools that were found within each of these categories. These categories represent the main types of assessment tools that we have identified during our investigation, but we have also discovered tools that did not fall into any of these three categories (as shown in Figure 10).

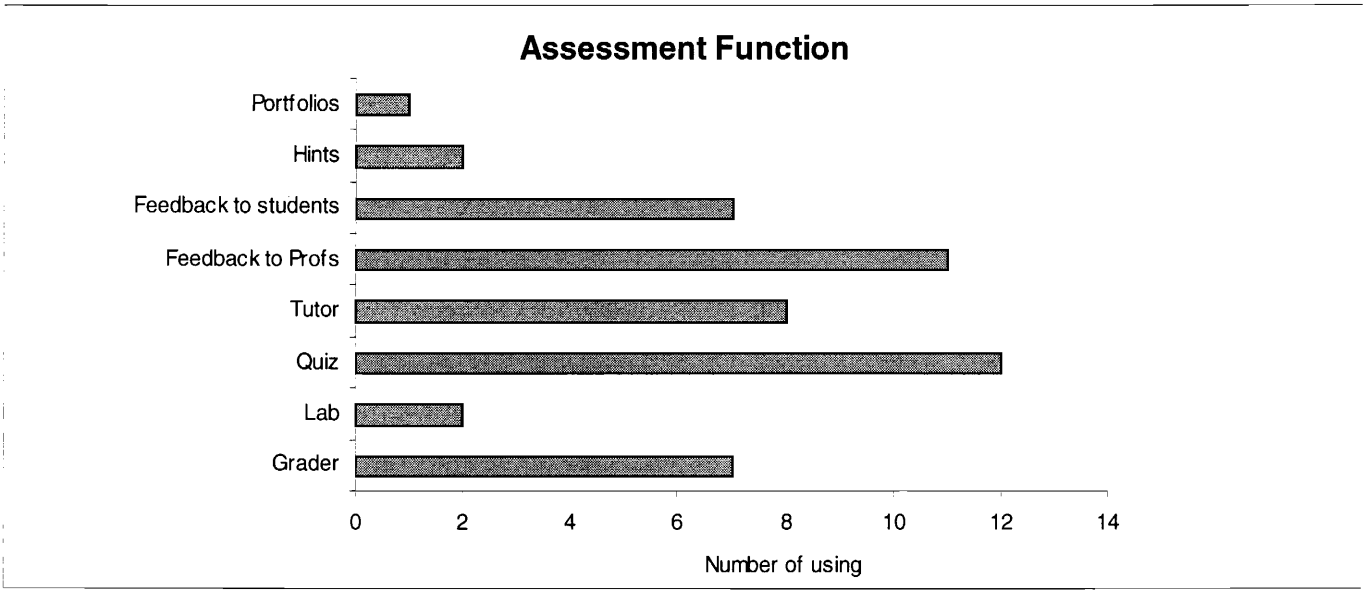

Figure 10: Main functions that IT-Based Learning Assessment Tools possess

4.3.1 Homework Graders

The purpose of a homework grader is to administer and grade professor generated homework sets. From literature, interviews and our own opinions, we have selected the elements a homework grader should have in order to be effective. The following criteria for effective homework graders are features of the tool or functionality that we feel help make these tools better. The first criterion is that there should be unrestricted use of the tools. Students need to be able to use the homework tool whenever they want and not just when homework has been assigned. For example, high number of students at University of Rochester used WebWork right before an exam. Therefore, it would have to be able to handle that many people using the program at once. According to Dr.

Arnold Pizer from the Department of Mathematics at University of Rochester, "[There are] just about a thousand students in Mathematics and I think this semester about two hundred students in a Physics course [using WebWork]" (Pizer, 2000). In order to avoid cheating, there should be a large number of varying questions on the same concept. If a student gets a question wrong, he or she should receive another question, similar but not identical, on that material. If two students were working on the same material then "they can copy each other's methods, but they can't copy each other's answers" (Gage, 2000). (Dr. Michael Gage is a Mathematics professor from University of Rochester.) In other words they would study in-groups, which many students do in college. To allow students to work together in-groups, a homework grader will need to have a database of questions that is quite extensive. Another feature that the tools should have is the ability to interpret expressions; for example 1+2= 3 is the same as 2=3-1. The tool would work this way for mathematical expressions, but this would also come in handy for short answer and essay questions. Students often mistype words and expressions or fail to break their answers down to simpler components. If a professor or teacher's assistant was grading a paper of student with messy handwriting, he or she might try to make out what the student meant and give credit. Therefore if a computer is to replace human for grading, the computer needs to be capable of interpreting messy handwriting and give partial credit. For WebWork, "One of the features that has been added to the program, recently, is an answer pre-viewer so that when a student types in an answer, they can see what the answer looks like when it's in an appropriate way" (Pizer, 2000). According to the *Nine Principles of Good Practice for Assessing Student Learning,* assessment works best when it is ongoing and not episodic. With a computerized homework grader,

professors can assign more homework to their students. There are students who will do homework to learn the material, but then there are many students who will not complete the homework because it is not being graded and counted towards their final grades. With the grades being counted, more students will work on the homework and spend more time on it, to ensure a good grade. Along with grading, there should be "special grading". By special grading we are suggesting that students who complete the homework by using a number of hints should not receive the same grade as a student who completes the homework without the use of hints. If the homework were meant to assess the students' knowledge of concepts, it would not be fair to count them the same. Of course, students should be allowed to try the problems over and over until they get them correct and use hints if needed, however when the total score is being calculated, the number of attempts made and hints used should also be counted and graded accordingly. One feature that would be beneficial to professors is to have some type of statistical feedback to them. In this case, when a student has completed their homework assignment, the professor would get the score along with a timestamp, the number of questions missed and what material the student needs more work. This would allow the professor to tailor the lectures to the students' needs.

4.3.2 Tutors

The elements of effectiveness for tutors are much like the ones for homework graders with only a few differences. The main difference would be the feedback feature. Though computers can not fully replace humans, the tutors should be able to tell the students what problems they got right or wrong and why. Ideally, the software should be able to

recognize where you made your error and reply with a detailed explanation of the concept pertaining to your error. The computer should be able to supply hints (or simpler sub-problems). Since the tools are tutors, the two components of grading (homework grades counting towards the class and scaled scoring) mentioned previously would not be a factor. The rest of the elements, however, are important for an effective tutor. Among those, the unrestricted use and the accessibility of the tutor are very important.

4.3.3 Tests

There is an assortment of features that quizzes should have to assess student's knowledge effectively. The tests should offer more than multiple choice and true or false questions. There is a wide range of questions that could be used, such as: numeric, fill in the blank, drop down list, short answers and essay. (Incidentally, short answer and essay differ in the length and elaboration of the answer. Essays are longer and require the student to elaborate extensively on a given topic.) These should be taken advantage of and a variety of questions should be given on a concept. This would cut down on cheating. We, as students, believe one of the more important features is to allow students to see all of the questions on the exam and go back to previous questions to change answers if necessary. The entire exam should be submitted after the student is satisfied with the answers, and given the chance to check over their work. After the test has been submitted, there should be instant scoring telling the students which questions they got right and wrong. The test should be able to interpret equivalent expressions the same way the homework grader should.

4.4 Graphs and Commentary about Tool use

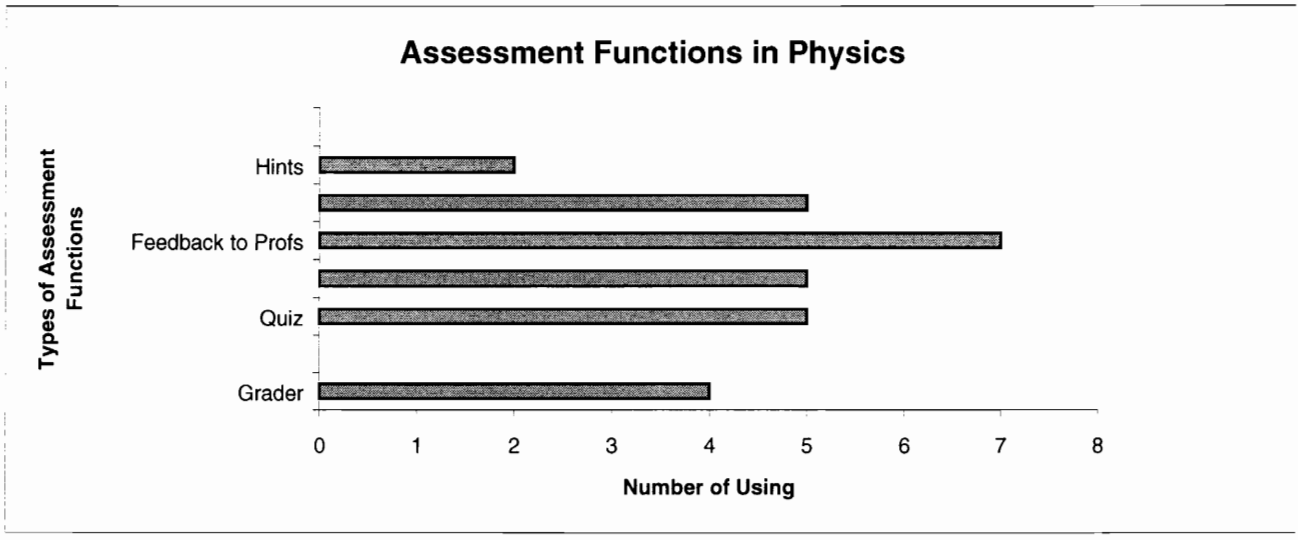

Figure 11: Questions Physics tools use

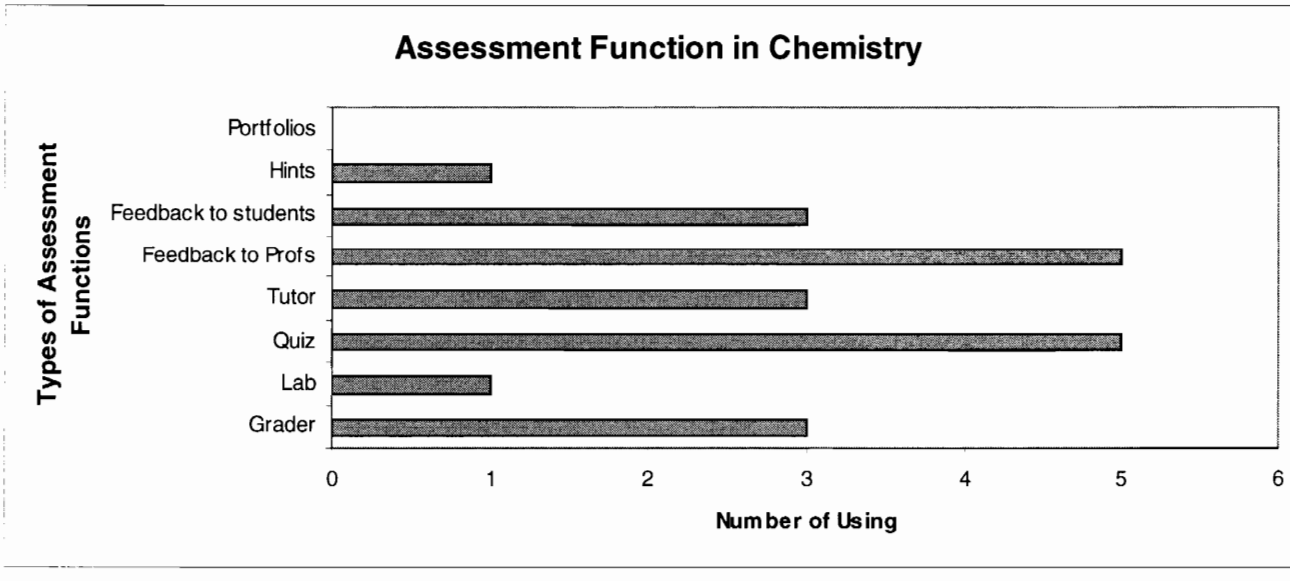

Figure 12: Questions Chemistry tools use

We wanted to compare the disciplines to the types of questions that are being offered to see if there is any correlation. However, there are various reasons why the disciplines use the questions they do. Multiple-choice questions are the most widely used

questions. This could be because they are the simplest to develop and grade. The tool does not have to be 'intelligent' and one letter signifies the correct answer. True or false are also very popular for the same reason. We were surprised to find how many were using short answer and fill in the blank questions. It may be more complex to develop a tool that grades short answer problems than multiple choice and true/false. The short answer and fill in the blank questions require the student to figure out the answer without having the option of guessing (as shown in Figures 11 and 12).

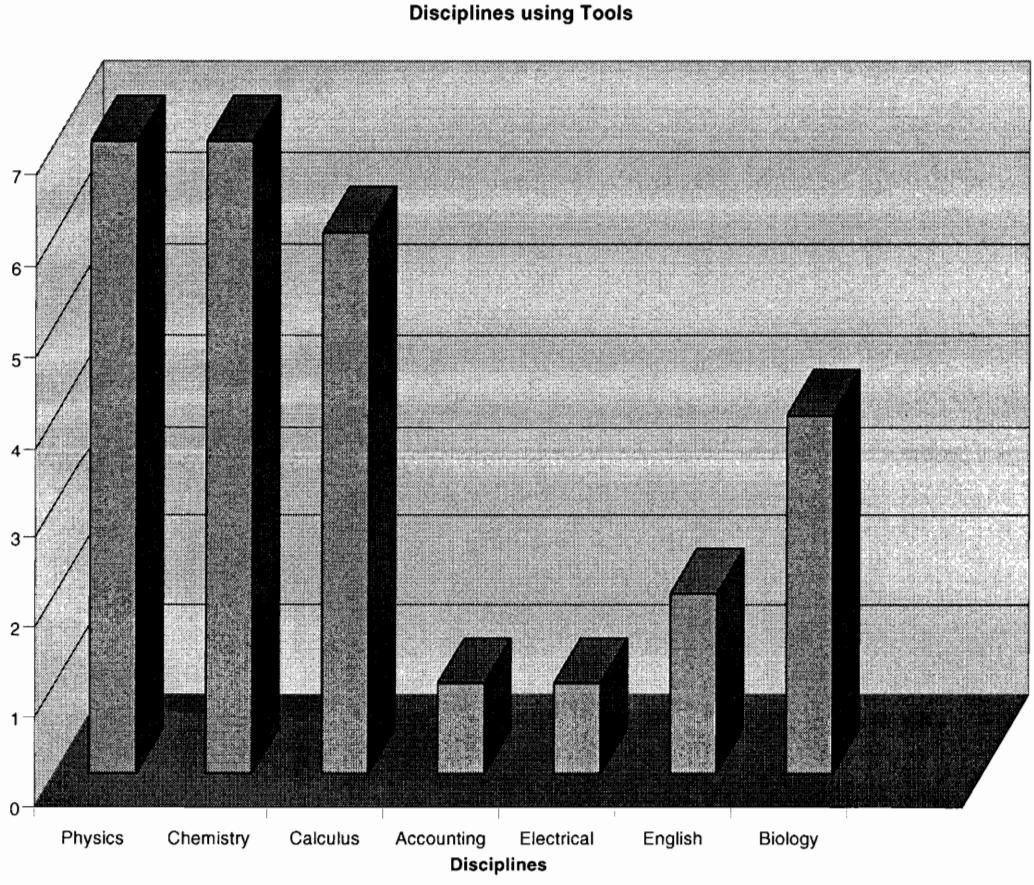

Figure 13: Range of Disciplines that use IT-Based Learning Assessment Tools

The tools that are used for humanities classes as well as the math and science classes tend to have more options on the question styles. So there may not be a big correlation between each discipline and question styles used, but rather a comparison in what questions are used in the humanities versus the sciences.

Of the tools that we found, Figure 13 depicts the disciplines that have IT-based learning assessment tools. As one can see from Figure 13, the majority of the tools lie in the Physics, Chemistry and Calculus disciplines. One reason for this is that the multiple choice and true or false questions are widely used.

4.5 Explanations

4.5.1 "Drill and Skill"

"Drill and Skill" is a phrase used to describe a homework assignment that requires a student to complete a plethora of questions that are all very similar. Similar means that the questions all relate to the same concept. The objective of "Drill and Skill" is to train students to complete a specific type of problem, but not necessarily understand the concept. An example of "Drill and Skill" is completing 50 derivatives as a Calculus I assignment.

The majority of the IT-Based Learning Assessment Tools that we have encountered focus on this type of learning. The students are supplied numerous questions all relating to the same concept. This is the motivation behind a tool like *WhizQuiz*.

4.5.2 "Enrichment"

"Enrichment" is an extension of an activity or lesson. It is an application of a skill that is has been learned by a student. The "Drill and Skill" may teach the student

how to use the skill, but it is enrichment that teaches the student when and why to use the skill. Relating back to the derivative assignment, an enrichment exercise would be applying the knowledge the student has to a problem involving the rate of change as a gas tank is emptied. This is the motivation behind tools like *Interactive Internet Laboratory*, which like any lab is designed to apply a student's learning.

Chapter 5 Conclusions and Recommendations

5.1 Conclusions

IT-based learning assessment tools are finding their way into colleges and universities around the country. From our investigation of IT-based learning assessment tools, the following conclusions are made:

1. IT-Based Learning Assessment is ongoing and not episodic therefore, IT-Based Learning Assessment tools promote continuous rather than sporadic assessment.

2. IT-based learning assessment tools lend themselves better to "Drill-and-Skill" style assessments than "Enrichment" assessments. In "Drill-and-Skill", the students complete many similar problems as a way of learning a particular skill. This style of learning is conducive to lower-level math and science classes.

3. Since the grading performed by most IT-Based Learning Assessment Tools is based on the final answer, it is impossible for students to receive partial credits on the tests. We regard this as a disadvantage of the IT-Based Learning Assessment Tools.

5.2 Recommendations

We have developed the following recommendations to the National Science Foundation are made.

1. The NSF should continue to study the effectiveness of IT-Based Learning Assessment Tools and develop criteria for evaluating these tools. A project should be organized to determine the effectiveness of each function of the tools. For example, do computer-provided "hints" help students learn more from homework? This could be determined by administering a test to two groups of

students. One group of students should use the IT-based learning assessment tool with the hints and the other not using the hints. Then the same test should be administered again to both groups of students.

2. The NSF should consider supporting research dealing with *Unique Tools*, meaning tools that do not fit into the three categories (Homework Grader, Tutor, and Test) that we created.

3. The NSF should expand the *matrix* that we have started. The matrix, as it stands now, is not exhaustive but rather a complete representation of the types of tools that are presently being used in undergraduate education. We recommend that NSF support additional projects to enhance the list, and make it accessible on their web page. The Matrix contains invaluable information that professors around the country could refer to when deciding to use these tools in their courses.

Appendix A – National Science Foundation's Mission

The National Science Foundation promotes and advances scientific progress in the United States by competitively awarding grants for research and education in the sciences, mathematics and engineering.

Available Online: http://www.nsf.gov

Appendix - B

Name of Tool (Origin of	Unique Features	Who's Using the Tool	Primary Discipline	Secondary Discipline	Other Disciplines	Style of Questions	Generation
CyberTutor (MIT)	-provides hints or simpler sub problems at the student's request - provides individualized feedback about performance - java applet allow to draw the curves and vectors on web	MIT	Physics			Multiple/Numeric Answer; Mouse drawn vectors	IV
OWL (University of Massachusetts at Amherst)	- can create variations of questions based on templates developed by professors - Intelligent Tutoring	University of Massachusetts at Amherst	Chemistry	Physics	Geosciences, Art History, Biochemistry	Multiple Choice, Short Answer	IV
JiTT (Just in Time Teaching) (Indiana University, Purdue University-Indianapolis)	- prior to every class students complete Web based "WarmUp Exercises". - assist the professor in adjusting and organizing the class	Indiana University, Purdue University-Indianapolis, United States Air Force Academy	Physics		Astronomy, Meteorology, Geology, Psychology, Chemistry, Biology, Genetics, Mathematics, Business Management and Ethics	Multiple Choice, short answer	I
Electronic Student Portfolios (Virginia Commonwealth University)	-Engineering students design and develop their own Internet web site that contains certain required information. - reviewed by the "Industrial Board" at VCU and graded based on a scoring rubric.	Virginia Commonwealth University	Engineering			Web-Based Portfolio	
WeBWorK (University of Rochester)	-produces similar but individualized problems for each student -can handle different forms of the same answer	University of Rochester, John Hopkins University, Indiana University, SUNY Stony Brook, Ohio State, University of Virginia,	Precalculus & Calculus	Physics		Multiple Choice, True/False, Short Answer	I

Name of Tool (Origin of Development)	Assessment Function								Developme nt Stage	Contact
	Grader	Lab	Test	Tutor	Feedback To Profs	To Students	Hints	Portfolio		
CyberTutor (MIT)	X		X	X	X	X			1.5 years, since 1999	David E. Pritchard - Physics Professor MIT dpritch@mit.edu
OWL (University of Massachusetts at Amherst)	X			X		X			1.5 years, since July 1999	William J. Vining - Department of Chemistry, UMASS Amherst vining@chem.umass.edu
JiTT (Just in Time Teaching) (Indiana University, Purdue University-Indianapolis)			X		X				5 years, since Fall of 1995	Gregor M. Novak - Department of Physics, Indiana University - Purdue University Indianapolis gnovak@iupui.edu
Electronic Student Portfolios (Virginia Commonwealth University)								X	2 Years	Gerald E. Miller - Department of Enigineering, Virginia Commonwealth University gemiller@gems.vcu.edu
WeBWorK (University of Rochester)	X			X	X	X			4 years, since fall of 1996	Douglas C. Ravenel - Chair, Department of Mathematics, University of Rochester doug@ravenel.net

Name of Tool (Origin of	Unique Features	Who's Using the Tool	Primary Discipline	Secondary Discipline	Other Disciplines	Style of Questions	Generation
Accounting Tutorial (West Virginia University)	-answer section defines the correction - allows student to click on a letter response that takes them to	West Virginia University	Accounting			Multiple Choice	I
Labview (Internet-based chemistry set) (Frostburg State University)	-interactive technical support and background information and to display and analyze the results - panic button on the panel commences a conferencing connection	Frostburg State University	Chemistry			Interactive Lab	
Computer Assisted Homework/Testing - Chemistry (University of Connecticut)	-Freshmen Chemistry Examination online - refenece is provided	University of Connecticut	Chemistry I	Physical Chemistry		Multiple Choice	I
The Quiz CGI (University of Connecticut)	-instructor creates a text file containgin questions that are then sent to the server where the quiz is created -the instructors' answers are given to the student for comparison	University of Connecticut, Connecticut Community Colleges; University of Rhode Island	chemistry, math, physics, pharmacy			Multiple choice, short answer	I
Using Distance Learning and Remote Access Technology for an Interactive Internet Laboratory (IIL) (Queensborough Community College)	-can be transported anywhere -DVM that is connected to the laptop that is interfaced with a laptop using software called webLAB -allow students do the labs off campus at their own convenience	Queensborough Community College	Electrical Engineering			Interactive Lab	

Name of Tool (Origin of Development)	Assessment Function								Development Stage	Contact
	Grader	Lab	Test	Tutor	Feedback To Profs	To Students	Hints	Portfolio		
Accounting Tutorial (West Virginia University)				X					2 year, since spring of 1998	Paul D. Melton - Department of Business, West Virginia University
Labview (Internet-based chemistry set) (Frostburg State University)		X							2 years, since 1998	Dr. Frederick Senese - Department of Chemistry, Frostburg State University fsenese@frostburg.edu
Computer Assisted Homework/Testing - Chemistry (University of Connecticut)			X	X	X				5 year, since fall of 1995	Carl W. David - Department of Chemistry, University of Connecticut david@uconnvm.uconn.edu
The Quiz CGI (University of Connecticut)	X		X		X				1year, since Fall of 1999	Andrew Depalma - Professional Staff, Department of Mathematics, University of Connecticut Andrew.Depalma@uconn.edu
Using Distance Learning and Remote Access Technology for an Interactive Internet Laboratory (IIL) (Queensborough Community College)		X		X	X					Bernard Mohr - Department of Electrical & Computer Engineering, Queensborough Community College 718 281-5240

Name of Tool (Origin of	Unique Features	Who's Using the Tool	Primary Discipline	Secondary Discipline	Other Disciplines	Style of Questions	Generation
WhizQuiz (Virginia Polytechnic Institute and State University)	-master log file for all quizzes for use by the webmaster, which can be used for obtaining access statistics on various quizzes	Virginia Polytechnic Institute and State University, University of	English, Information Technology			Multiple Choice, True/False, Numeric	I
Web-Based Learning: Enhancing the Teaching of Engineering Graphics (University of Texas)	-integrated to hours of tutorial movies, lecture presentations, web-based games, and interactive quizzes -students have the opportunity to change answers until all of the answers are corrected	University of Texas	Engineering			Multiple Choice, True/False	
Student Assessment Learning Guide (SALG) (University of Wisconsin)	-allowing students to log in anonymously and give feedback on a particular course -professors can select or edit the questions - statistical information of the class is reported to the professor	University of Wisconsin	SMET			Rubric, Short Answer	
COMPASS (American College Testing)	-entrance exam, used to test students' proficiency in: Writing Skills, Reading and Mathematics	American College Testing, Ivy Tech State College, Columbus State Community College,	English	Math		Multiple Choice	II
Mallard (Universtiy of Illinois-Chicago)	-offers multiple versions of the same question -developer can generate quizzes as HTML-forms -supply conditional feedback prepared by the developer	Universtiy of Illinois-Chicago	Electrical Engineering			Blank, Multiple Blank, Arith (type in numerical expression), Multiple Choice, T/F, List(student types a list of items)	I
WebAssign (North Carolina State University)	- automatically provide an answer key after the assignment due date -questions and problems originate from student's textbook which can be modified, created a new, or used directly from the database -students can work a problem multiple times until they get the correct answer	North Carolina State University, King College, LeTourneau University,Marshalltown Community College, Massachusetts Institute of Technology, Merrimack College, Milwaukee School of	Biology, Business, Chemistry, Engineering, Math and Physics			Fill in the blank with numerical number or single word (units are given)	I

Name of Tool (Origin of Development)	Assessment Function								Development Stage	Contact
	Grader	Lab	Test	Tutor	Feedback To Profs	To Students	Hints	Portfolio		
WhizQuiz (Virginia Polytechnic Institute and State University)			X		X				3 years, since 1997	Information Systems and Insect Studies Lab, Virginia Polytechnic Institute and State University whizquiz@vt.edu
Web-Based Learning: Enhancing the Teaching of Engineering Graphics (University of Texas)			X						4 years	Stephen Crown swcrown@panam.edu
Student Assessment Learning Guide (SALG) (University of Wisconsin)					X				1 year, since 1999	Dr. Susan Millar, University of Wisconsin smillar@engr.wisc.edu
COMPASS (American College Testing)			X						4 year, since Dec 1996	COMPASS compass@act.org
Mallard (Universtiy of Illinois-Chicago)	X		X						2 year, since 1998	Donna Brown djb@uiuc.edu
WebAssign (North Carolina State University)	X				X	X			2 years, since the fall of 1998	WebAssign North Carolina State University webassign@ncsu.edu

Name of Tool (Origin of	Unique Features	Who's Using the Tool	Primary Discipline	Secondary Discipline	Other Disciplines	Style of Questions	Generation
QuizSite (Indiana University)	-evaluations via the web -a variety of reports automatically generated for the instructor	Indiana University	Math, History			multiple-choice, fill-in, numeric, and essay	I
CAPA (Michigan State University)	-numerous templates of problems that encourage students to collaborate -keeps track of students' participation and performance		Physics	astronomy, biochemistry, chemistry, mathematics, physics, botany, accelerator physics, human food and nutrition, family and child ecology, and computer		Multiple Choice, True/False, Short Answer, Essay	I
Derivative Gateway Test (University of Michigan)	-computer-generated quiz that tests students' ability to complete derivatives quickly and accurately -Student's can complete as many		Calculus			Short answer functions	I
Visual Calculus (University of Tennessee, Knoxville)	-explains concepts through the use of interactive tutorials - supplies "Drill" problems to be completed on paper and the solution can be viewed		Pre-Calculus, Calculus			Problems are given, students complete by hand	
Online Homework and Quizzes (University of South Carolina)	-use jave animation to supplement student's homeworks and quizzes -several topic of tutorials with tables and charts		Chemistry			Multiple Choices	
Intelligent Essay Assessor (New Mexico State)	-scores essays -detects plagiarism	New Mexico State University University of Colorado	Psycho-linguistics	Biology, Psychology, History		Essay, Short Answer	

Name of Tool (Origin of Development)	Assessment Function								Development Stage	Contact
	Grader	Lab	Test	Tutor	Feedback To Profs	To Students	Hints	Portfolio		
QuizSite (Indiana University)	X		X		X	X			3 years, since the spring of 1997	BEST—Franklin, best@indiana.edu.
CAPA (Michigan State University)			X		X	X	X		Since Fall of 1992	Edwin Kashy, Department of Physics and Astronomy, Michigan State University kashy@nscl.msu.edu
Derivative Gateway Test (University of Michigan)	X		X						5 year, since the fall of 1995	Bob Megginson, Department of Mathematics, University of Michigan
Visual Calculus (University of Tennessee, Knoxville)				X					Grants awarded in 1996	Lawrence S. Husch, Department of Mathematics, University of Tennessee
Online Homework and Quizzes (University of South Carolina)	X		X	X		X			2 years, since fall of 1998	USC Department of Chemistry, apotter@chem1.usc.edu
Intelligent Essay Assessor (New Mexico State)						X			Since fall of 1997	Peter Foltz New Mexico State University pfoltz@crl.nmsu.edu

Appendix C – Glossary of Matrix Headings

The purpose of this section is to give a detailed explanation of what each heading on the matrix means and why we decided to choose it. The purpose of our matrix was to provide a starting point for the National Science Foundation to continue research on what kinds of tools are out there. As we discovered the some of IT-based learning assessment tools that were being used at schools throughout the United States, we determined that a matrix would be the best way to represent them. Although every tool is different and the objectives behind their development vary, the matrix allows us to organize focus on the components of the tools that are similar as a basis for comparison. The following is a list of these matrix headings that are compared within the matrix.

- **Name of Tool (Origin of Development)**: This is the first matrix heading and it simply gives the name of the tool (and any acronym or nickname it may have) and the institution that was responsible for the development of it.
- **Unique Features**: The purpose of all of these tools is to assess a student's learning in a particular course that they are currently enrolled in. The methods for which they accomplish this task are similar throughout all of the tools. However, each tool has it's own special attribute that makes it different from the rest of the field. This feature is something that we like to call "thinking out of the box". This means simply that the developer thought of something that goes beyond the typical methods for assessing the students that sets this tool apart from the rest. For instance, this can come in the form of different question types or better student feedback.
- **Who's Using the Tool**: Because many of these tools are new, finding information that assesses the effectiveness of the tools is difficult at times. However, one good

way of determining the effectiveness of an IT-based learning assessment tool is to see how many other schools have adopted it. If there are fifteen institutions using the tool then this is a good indication that this tool is effective. On the other side of the coin, just because only one school is using a tool doesn't necessarily mean that the tool is ineffective. This can be because the tool is still in development.

- **Primary Discipline**: This section of the matrix indicates the discipline that the tool was originally designed for. This section can be used to compare what disciplines are developing and using the tools the most. It may also indicate in what disciplines the IT-based tools can be used most effectively.
- **Secondary Discipline**: This section defines the next discipline (if any) decided to adapt the tool into their curriculum. The more widespread a tool is can be an indication to the effectiveness of it.
- **Other Disciplines**: Like the secondary discipline, this section indicates other disciplines that the tool has adapted itself too. If there are not any other disciplines that are using the tool this doesn't necessarily mean that the tool is ineffective but rather maybe the tool is too specialized to work anywhere else.
- **Style of Questions**: IT-based learning assessment tools are generally confined to the use of multiple choice and true/false questions. However, this is changing as technology advances. This section is a chance to see the variety of styles that are used.
- **Generation**: Dr. Robert Linn defined "generations" that IT-based learning assessment tools fall into. Greater discussion of exactly what each generation means is discussed in the Chapter 2.

➢ **Assessment Function**: As you can see, this section has been broken down into various subsections. The purpose of this section of the matrix is to define the objective of these tools. Because these tools are all so different, we need to focus on the major objectives of them for comparison reasons.

- **Grader**: A grader is a tool that scores the work a student has done, immediately following the completion of the assessment. Meaning that a human doesn't need to review the answers given by the students when the assessment is complete.
- **Lab**: This indicates a tool that is used for students to complete lab work. Not to complete a lab report but to physically complete the lab.
- **Quiz**: This indicates an exam, whether it is a quiz or a test.
- **Tutor**: A tutor is a tool that gives not only poses questions for students but offers them information to ensure that they learn as well as get the question correct.
- **Feedback to Profs**: "Profs" is our abbreviation for professors. Tools that have this functionality give information back to the professors for a variety of reasons. Some of the reasons include the professors' ability to tailor the lectures to the right knowledge level or statistical information about the performance of students on a given exam.
- **Feedback to Students**: This indicates that the tool gives information back to the students when a question is answered incorrectly. The information can come in the form of topics that the students need to study more

carefully, or information as simple as the units that they used were incorrect.

- **Hints**: Does the assessment tool give the students hints or simpler sub-problems if they request them?
- **Portfolio**: This section indicates that the nature of the program is an electronic portfolio. Something that is created and updated by the student and then put on display through the use of technology.

- **Development Stage**: This section deals with how long the tool has been used. This is yet another indication to the effectiveness of a tool; if it has been used for twenty years then there is a possibility that it is a very effective tool.
- **Contact**: This section is simply a person to contact for more information about this tool. Sometimes this is the developer or the project manager.

Appendix D – References

Tetreault, Donald. (1998). "Educational Technologies: What are they and what are the Costs?" University of South Carolina [Online]. Available: http://www.bellsouthcorp.com/bsf/technology/educational_technologies.htm. Accessed on September 3, 2000.

Bennett, Randy Elliot. (June 1998). "Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing." Educational Testing Service [Online]. Available: http://www.ets.org/research/pic/bennett.html. Accessed on September 3, 2000.

U.S. Department of Education. (1998). An Educator's Guide to Evaluating the Use of Technology in Schools and Classrooms[Online]. Available: http://www.ed.gov/pubs/EdTechGuide/preface.html. Accessed on September 3, 2000.

Ysseldyke, J., & Salvia, John. (1998). Assessment. Boston: Houghton Mifflin Company.

Linn, Robert. (1993). Educational Measurement. Phoenix: The Oryx Press.

Albright, Michael, & Graf, David. (1992). Teaching in the Information Age: The Role of Educational Technology. San Francisco: Jossey-Bass Inc.

Breivik, Patricia Senn. (1998). Student Learning in the Information Age. Phoenix: The Oryx Press.

Wiggins, Grant P. (1993). Assessing Student Performance: Exploring the Purpose and Limits of Testing. San Francisco: Jossey-Bass Inc.

AAHE Assessment Forum. (1997). Learning Through Assessment: A Resource Guide for Higher Education. Edited by Lion F. Gardiner, Caitlin Anderson, and Barbara Cambridge. Washington DC: American Association for Higher Education.

Gray, Matthew. (1995). "Measuring the Growth of the Web: June 1993 to June 1995." Massachusetts Institute of Technology [Online]. Available: http://www.mit.edu/people/mkgray/growth/. Accessed on September 5, 2000.

National Commission on Testing and Public Policy. (1990). "From Gatekeeper to Gateway: Transforming Testing in America." Chestnut Hill, MA: NCTPP.

Glaser, Robert, & Silver, Edward. (1994). "Assessment, Testing, and Instruction: Retrospect and Prospect." Los Angeles: National Center for Research on Evaluation, Standards and Student Testing.

Gumport, Patricia, & Chun, Marc. (1999). "Technology and Higher Education: Opportunities and Challenges for the New Era." Stanford, CA: National Center for Postsecondary Improvement.

ERIC Clearinghouse on Information and Technology. (1993). "Alternative Assessment and Technology." Syracuse, NY: ERIC [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/365312.html. Accessed on September 22, 2000.

Office of Educational Assessment. (2000). "Data Processing." University of Washington [Online]. Available: http://www.washington.edu/oea/. Accessed on September 22, 2000.

Pritchard, David E. (2000). "CyberTutor." Massachusetts Institute of Technology [Online]. Available: http://cybertutor.mit.edu/. Accessed on September 22, 2000.

Boston University. (2000). "Comparison of WebCT and CourseInfo." Boston University [Online]. Available: http://www.bu.edu/webcentral/research/courseware/. Accessed on October 6, 2000.

TOEFL. (2000). http://www.toefl.org/cbabttfl.html. Accessed on October 6, 2000.

Thelwall, M. (2000). "Computer-based assessment: A Versatile Educational Tool." Computers & Education, vol. 34, 37-49.

Sampson, James P. Using the Internet to Enhance Testing in Counseling. Journal of Counseling & Development, vol. 78, 348-356.

Johnson, R. Neill, & Enerson, Diane, & Plank, Kathryn. (2000). "Computerized Testing Roundtable." Penn State University [Online]. Available: http://www.psu.edu/celt/largeclass/comptest.html. Accessed on October 6, 2000.

James Madison University. (2000). "The Basics of Computer-Based Testing and Assessment. JMU [Online]. Available: http://www.jmu.edu/assessment/cbtbasics.htm. Accessed on October 11, 2000.

Berner, R. Thomas. (July 1993). Using Computers To Teach Journalism: What Some Students Think Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, Volume 1, Number 3.

Educational Resources Information Center (ERIC). (February 2000). "Computerized Adaptive Tests: Eric Digest No. 107" [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/ed315425.html. Accessed on October 11, 2000.

National Council of State Boards of Nursing. (2000). "Computerized Adaptive Testing (CAT) Overview." NCSBN [Online]. Available: http://www.ncsbn.org/files/nclex/cattop.asp. Accessed on October 20, 2000.

Graduate Record Examinations. (2000). "How does the Computer-Based General Test Work?" GRE [Online]. Available: http://www.gre.org/cbttest.html#howwork. Accessed on October 20, 2000.

Berg, Bruce L. (1998). Qualitative Research Methods for Social Scientists. Needham Heights, MA: Allyn & Bacon.

Salant, Priscilla & Dillman, Don. (1994). How to Conduct Your Own Survey. New York: John Wiley & Sons, Inc.

FairTest: The National Center for Fair & Open Testing. (2000). "Computerized Testing: More Questions than Answers [Online]." Available: http://www.fairtest.org/facts/computer.htm. Accessed on October 20, 2000.

David, Carl W. "New Ways To Test Science from: The College Board Review, No 119 Spring 1981." http://www.chem.uconn.edu/~cdavid/diatribe/CAT.html May 30, 1996. Accessed on November 4, 2000.

Kadiyala, Madhavi, Crynes, Billy L. A Review of Literature on Effectiveness of Use of Information Technology in Education. Journal of Engineering Education Pg. 177-189 April 2000

Crown, Stephen W. Web-Based Learning: Enhancing the Teaching of Engineering. (IMEJ) Interactive Multimedia Electronic Journal of Computer-Enhanced Learning February, 1999

Rozycki, William Just-in-Time Teaching (JiTT) Research & Creative Activity, April 1999 Vol. 12, No.1

Kashy, E., Albertelli, G., & Thoennessen, M. (1999). "ALN Technology on Campus: Successes and Problems." Michigan State University.

Baker, Eva. (1999). "Technology: How do we Know it Works?" National Center for Research on Evaluation, Standards and Student Testing [Online]. Available: http://www.ed.gov/Technology/TechConf/1999/whitepapers/paper5.html. Accessed on November 4, 2000.

Accounting Tutor (2000), West Virginia University, [Online] Available: http://www.access.wvu.edu/class/acctutor/accthome.htm Accessed on November 8, 2000.

Blackboard (2000), [Online] Available: http://www.blackboard.com. Accessed on November 8, 2000.

COMPASS (2000), COMPASS/ESL of ACT, [Online] Available: http://www.act.org/compass/overview/index.html Accessed on November 8, 2000.

Computer-Assisted Personalized Approach (CAPA) (2000). Michigan State University [Online] Available: http://capa4.lite.msu.edu/homepage/ Accessed on November 8, 2000.

Computer Assisted Homework/Testing (2000), University of Connecticut Chemistry Department, [Online] Available: http://www.sp.uconn.edu/~cdavid/chem263 Accessed on November 9, 2000.

CourseInfo(2000), [Online] Available: http://www.courseinfo.pitt.edu Accessed on November 9, 2000.

CyberTutor (2000), MIT, [Online] Available: http://cybertutor.mit.edu Accessed on November 9, 2000.

Electronic Student Portfolios (2000), Virginia Commonwealth University, [Online] Available: http://www.vcu.edu/mdcweb. Accessed on November 10, 2000.

Gateway Test (1999) University of Michigan Department of Mathematics, [Online] Available: http://www.math.lsa.umich.edu/courses/115/gateway.html Accessed on November 10, 2000.

FLAG- The Field Test Learning Assessment (2000), National Science Institute Education, University of Wisconsin Madison, [Online] Available: http://www.wcer.wisc.edu/nise/cl1/flag/tools/Tframe.asp Accessed on November 11, 2000.

Intelligent Essay Assessor (2000), New Mexico State University, [Online] Available: http://www-psych.nmsu.edu/essay/ Accessed on November 10, 2000.

JiTT (2000), Indiana University, [Online] Available:
http://webphysics.iupui.edu/jitt/WhatOVR.html

LabView (2000), Frostburg State University, Chemistry Department, [Online] Available:
http://imej.wfu.edu/articles/2000/2/06/index.asp
Accessed on November 10, 2000.

Mallard (2000), University of Illinois, Department of Electrical Engineering, [Online] Available: http://www.ews.uiuc.edu/Mallard/. Accessed on November 12, 2000.

OWL (2000), University of Amherst, [Online] Available:
http://www.math.umass.edu/~knightly/m127/owl.html
Accessed on November 12, 2000.

PIVOT (2000), MIT Physics Department, [Online] Available:
http://curricula2.mit.edu/pivot/. Accessed on November 12, 2000.

The Quiz of CGI, University of Connecticut (2000), [Online] Available:
http://www.sp.uconn.edu/%7Evirtclas/quiz/documentation/
Accessed on November 12, 2000.

QuizSite, Bureau of Evaluative Studies and Testing (2000). Indiana University at Bloomington (IUB), [Online] Available:
http://www.best.indiana.edu/quizsite/faq.html#general
Accessed on November 12, 2000.

WebAssign (2000), North Carolina State University, [Online] Available:
http://wasnet03ws.physics.ncsu.edu/info/
Accessed on November 12, 2000.

Web quizzes (1998), University of South Carolina Department of Chemistry, [Online] Available: http://michele.usc.edu/105b/105bfall98/
Accessed on November 12, 2000.

WebCT (2000), [Online] Available: http://about.webct.com/. Accessed on November 12, 2000.

WebWork (2000), University of Rochester, [Online] Available:
http://webwork.math.rochester.edu/docs/
Accessed on November 12, 2000.

WhizQuiz (2000), Virginia Polytechnic Institute and State University, [Online] Available: http://www.math.vt.edu/Cyber1114/jalessio/whizquiz/format.html
Accessed on November 18, 2000.

WPI (2000), Worcester Polytechnic Institute, [Online] Available:
http://www.wpi.edu/Academics/Projects/intro.html
Accessed on November 18, 2000.

Student Assessment Learning Guide- SALG (2000), University of Texas, [Online]
Available: http://www.wcer.wisc.edu/salgains/instructor/
Accessed on November 18, 2000.

SUNY (2000), SUNY Learning Network State University of New York,
[Online] Available:
http://www.sln.suny.edu/admin/sln/original.nsf/504ca249c786e20f85256284006d
a7ab/bd9a90d415c32956852563f6007a1297?OpenDocument
Accessed on November 18, 2000.

Visual Calculus (2000), University of Tennessee, Knoxville, Math Department, [Online]
Available: http://archives.math.utk.edu/visual.calculus/
Accessed on November 18, 2000.

Appendix E – Bibliography

FairTest: The National Center for Fair & Open Testing. (2000). "Computerized Testing: More Questions Than Answers [Online]." Available: http://www.fairtest.org/facts/computer.htm. Accessed on September 5, 2000 – December 10, 2000.

Thelwall, M. (1999) Computer-based assessment: A Versatile Educational Tool. Computers & Education

Johnson, R. Neill, Enerson, Diane M., & Plank, Kathryn M. (1997) "Computerized Testing Roundtable" [Online]. Available: http://www.psu.edu/celt/largeclass/comptest.html. Accessed on September 5, 2000 – December 10, 2000.

Sampson, James P. Using the Internet to Enhance Testing in Counseling. Journal of Counseling & Development

James Madison University. (2000). "The Basics of Computer-Based Testing and Assessment" [Online]. Available: http://www.jmu.edu/assessment/cbtbasics.htm Accessed on September 5, 2000 – December 10, 2000.

Berner, R. Thomas. (1993) "Using Computers To Teach Journalism: What Some Students Think." Interpersonal Computing and Technology: An Electronic Journal for the 21st Century July 1993, Volume 1, Number 3.

Educational Resources Information Center (ERIC) (2000) Computerized Adaptive Tests: Eric Digest No. 107 [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/ed315425.html February 2000.

National Council of State Boards of Nursing, Inc. (2000). "Computerized Adaptive Testing (CAT) Overview." [Online]. Available: http://www.ncsbn.org/files/nclex/cattop.asp. Accessed on September 5, 2000 – December 10, 2000.

Educational Testing Service. (2000). "The General Test" [Online]. Available: http://www.gre.org/cbttest.html#howwork. Accessed on September 5, 2000 – December 10, 2000.

National Science Foundation: Department of Undergraduate Education 2000. "Mission" [Online]. Available: http://www.ehr.nsf.gov/ehr/due/about/mission.asp. Accessed on September 5, 2000 – December 10, 2000.

Worcester Polytechnic Institute (WPI). (2000). "About the WPI Projects Program." [Online]. Available: http://www.wpi.edu/Academics/Projects/intro.html.

Accessed on September 5, 2000 – December 10, 2000.

Blackboard, Inc. (2000). "Customer List." [Online]. Available: http://company.blackboard.com/customers.cgi?SELECT=26. Accessed on September 5, 2000 – December 10, 2000.

Free On-line Dictionary of Computing (FOLDOC). (2000). "Information Technology." [Online]. Available: http://wombat.doc.ic.ac.uk/foldoc/foldoc.cgi?query=information+technology&action=Search. Accessed on September 5, 2000 – December 10, 2000.

Cambridge Dictionaries On-line. (2000). "Definition: Assess." [Online]. Available: http://dictionary.cambridge.org/define.asp?key=assess*1%2B0. Accessed on September 5, 2000 – December 10, 2000.

Cambridge Dictionaries On-line. (2000). "Definition: Information Technology." [Online]. Available: http://dictionary.cambridge.org/define.asp?KEY=inform*1%2B5&x=74&y=5. Accessed on September 5, 2000 – December 10, 2000.

University of Massachusetts at Amherst (UMASS). (2000). "OWL Feature Highlights and Benefits" [Online]. Available: http://ccbit.cs.umass.edu/owl/pages/features.html. Accessed on September 5, 2000 – December 10, 2000.

Tetreault, Donald. (1998). "Educational Technologies: What are they and what are the Costs?" University of South Carolina [Online]. Available: http://www.bellsouthcorp.com/bsf/technology/educational_technologies.htm. Accessed on September 5, 2000 – December 10, 2000.

Bennett, Randy Elliot. (June 1998). "Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing." Educational Testing Service [Online]. Available: http://www.ets.org/research/pic/bennett.html. Accessed on September 5, 2000 – December 10, 2000.

U.S. Department of Education. (1998). An Educator's Guide to Evaluating the Use of Technology in Schools and Classrooms[Online]. Available: http://www.ed.gov/pubs/EdTechGuide/preface.html. Accessed on September 5, 2000 – December 10, 2000.

Ysseldyke, J., & Salvia, John. (1998). Assessment. Boston: Houghton Mifflin Company.

Linn, Robert. (1993). Educational Measurement. Phoenix: The Oryx Press.

Albright, Michael, & Graf, David. (1992). Teaching in the Information Age: The Role of Educational Technology. San Francisco: Jossey-Bass Inc.

Breivik, Patricia Senn. (1998). Student Learning in the Information Age. Phoenix: The Oryx Press.

Wiggins, Grant P. (1993). Assessing Student Performance: Exploring the Purpose and Limits of Testing. San Francisco: Jossey-Bass Inc.

AAHE Assessment Forum. (1997). Learning Through Assessment: A Resource Guide for Higher Education. Edited by Lion F. Gardiner, Caitlin Anderson, and Barbara Cambridge. Washington DC: American Association for Higher Education.

Gray, Matthew. (1995). "Measuring the Growth of the Web: June 1993 to June 1995." Massachusetts Institute of Technology [Online]. Available: http://www.mit.edu/people/mkgray/growth/. Accessed on September 5, 2000 – December 10, 2000.

National Commission on Testing and Public Policy. (1990). "From Gatekeeper to Gateway: Transforming Testing in America." Chestnut Hill, MA: NCTPP.

Glaser, Robert, & Silver, Edward. (1994). "Assessment, Testing, and Instruction: Retrospect and Prospect." Los Angeles: National Center for Research on Evaluation, Standards and Student Testing.

Gumport, Patricia, & Chun, Marc. (1999). "Technology and Higher Education: Opportunities and Challenges for the New Era." Stanford, CA: National Center for Postsecondary Improvement.

ERIC Clearinghouse on Information and Technology. (1993). "Alternative Assessment and Technology." Syracuse, NY: ERIC [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/365312.html. Accessed on September 5, 2000 – December 10, 2000.

Office of Educational Assessment. (2000). "Data Processing." University of Washington [Online]. Available: http://www.washington.edu/oea/. Accessed on September 5, 2000 – December 10, 2000.

Pritchard, David E. (2000). "CyberTutor." Massachusetts Institute of Technology [Online]. Available: http://cybertutor.mit.edu/. Accessed on September 5, 2000 – December 10, 2000.

Boston University. (2000). "Comparison of WebCT and CourseInfo." Boston University [Online]. Available: http://www.bu.edu/webcentral/research/courseware/. Accessed on September 5, 2000 – December 10, 2000.

TOEFL. (2000). http://www.toefl.org/cbabttfl.html. Accessed on September 5, 2000 – December 10, 2000.

Sampson, James P. Using the Internet to Enhance Testing in Counseling. Journal of Counseling & Development, vol. 78, 348-356.

Johnson, R. Neill, & Enerson, Diane, & Plank, Kathryn. (2000). "Computerized Testing Roundtable." Penn State University [Online]. Available: http://www.psu.edu/celt/largeclass/comptest.html. Accessed on September 5, 2000 – December 10, 2000.

James Madison University. (2000). "The Basics of Computer-Based Testing and Assessment. JMU [Online]. Available: http://www.jmu.edu/assessment/cbtbasics.htm. Accessed on September 5, 2000 – December 10, 2000.

Berner, R. Thomas. (July 1993). "Using Computers to Teach Journalism: What Some Students Think." Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, Volume 1, Number 3.

Educational Resources Information Center (ERIC). (February 2000). "Computerized Adaptive Tests: Eric Digest No. 107" [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/ed315425.html. Accessed on September 5, 2000 – December 10, 2000.

National Council of State Boards of Nursing. (2000). "Computerized Adaptive Testing (CAT) Overview." NCSBN [Online]. Available: http://www.ncsbn.org/files/nclex/cattop.asp. Accessed on September 5, 2000 – December 10, 2000.

Graduate Record Examinations. (2000). "How does the Computer-Based General Test Work?" GRE [Online]. Available: http://www.gre.org/cbttest.html#howwork. Accessed on September 5, 2000 – December 10, 2000.

Berg, Bruce L. (1998). Qualitative Research Methods for Social Scientists. Needham Heights, MA: Allyn & Bacon.

Salant, Priscilla & Dillman, Don. (1994). How to Conduct Your Own Survey. New York: John Wiley & Sons, Inc.

Kadiyala, Madhavi, & Crynes, Billy L. (April 2000). "A Review of Literature on Effectiveness of Use of Information Technology in Education." Journal of Engineering Education.

Russell, Michael, & Haney, Walter. (January 2000). "The Gap Between Testing and Technology in Schools." The National Board on Educational Testing and Public Policy: Statements, vol. 1, number 2.

Gumport, Patricia J., & Chun, Marc. (1999). "Technology and Higher Education: Opportunities and Challenges for the New Era." National Center for Postsecondary Improvement, Stanford University.

Raskin, Robin. (October 2000). "Resuscitating Education." Family PC Magazine.

Foltz, Peter W & Laha,, Darrell & Landauer, Thomas (February 1999). "The Intelligent Essay Assessor: Applications to Educational Technology" Wake Forest University. TIEA[Online] Available: http://imej.wfu.edu/articles/1999/2/04/index.aasp Accessed on September 5, 2000 – December 10, 2000.

Woolf, Beverly Park & Hart David & Day, Roberta & Botch, Beatrice & Vining William (March 2000). "Improving Instruction and Reducing Costs With a Web-based Learning Environment" International Conference on Mathematics/Science Education & Technology (MSET 2000).

Crown, Stephen (February 1999). "Web-Based Learning: Enhancing the Teaching of Engineering Graphics" Wake Forest University. ETTOEG[Online] Available: http://imej.wfu.edu/articles/1999/2/04/index.aasp Accessed on September 5, 2000 – December 10, 2000.

Mohr, Bernard (1999) "Using Distance Learning and Remote Access Technology for An interactive Internet Laboratory" Technology Instruction for the 21st Century, Queensborough Community College

"Alternative Assessment and Technology. ERIC Digest" ERIC Clearinghouse on Information and Technology Syracuse NY. [Online]Available: http://www.ed.gov/databases/ERIC_Digests/ed365312.html Accessed on September 5, 2000 – December 10, 2000.

Russell, Michael & Haney, Walter, (January 2000) "The Gap between Testing and Technology in Schools" The National board on Educational Testing and Public Policy, Peter S. and Carolyn A. Lynch School of Education, Boston College.

Bennett, Randy Elliot (June 1998) "Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing" Education Testing Service. [Online] Available: http://www.ets.org/research/pic/bennett.html Accessed on September 5, 2000 – December 10, 2000.

Gumport, Patricia & Chun Marc (1999) "Technology and Higher Education: Opportunities and Challenges for the New Era", National Center or Postsecondary Improvement, Stanford University School of Education

Kadiyala, Madhavi & Crynes, Billy (April 2000) "A Review of Literature on Effectiveness of Use of Information Technology in Education", Journal of Engineering Education.

Ogilvie, Craig (1999) "Effectivenes of Different Course Components in Driving Gains in conceptual Understanding", Department of Physics, MIT

Heinecke, Walter & Blasi, Laura & Milman, Natalie & Washington, Lisa (1999) "New Directions in the Evaluation of the Effectiveness of Educational Technology", Curry School of Education, University of Virginia.

David, Carl (May 1996) "New Ways to Test Science from: The College Board Review, No 119 Spring 1981", The College Board Review, No 119 Spring 1981. [Online] Available: http://www.chem.uconn.edu/~cdavid/diatribe/CAT.html Accessed on September 5, 2000 – December 10, 2000.

Baker, Eva (1999) "Technology: How Do We Know It Works?", National Center for Research on Evaluation, Standards and Student Testing, UCLA.

Russell, Michael (2000) "It's Time to Upgrade: Tests and Administration Procedures for the New Millennium", The Secretary's Conference on Educational Technology 2000, [Online] Available: http://www.ed.gov/Technology/techconf/2000/russell_paper.html Accessed on September 5, 2000 – December 10, 2000.

Russell, Michael (1999) "A Follow-up Study Comparing Performance On computer and On Paper", Education Policy Analysis Archives, Boston College, [Online] Available: http://epaa.asu.edu/epaa/v7n20 Accessed on September 5, 2000 – December 10, 2000.

Noval, Gregor & Patterson, Evelyn (1998) "Just-In-Time Teaching: Active Learner Pedagogy with WWW", IASTED International Conference on Computers and Advanced Technology in Education, [Online] Available: http://webphysics.iupui.edu/JITT/ccjitt.html Accessed on September 5, 2000 – December 10, 2000.

Rozycki, William (April 1999) "Just-In-Time Teaching", Indiana University, [Online] Available: http://www.indina.edu/~rcapub/v22n/p08.html Accessed on September 5, 2000 – December 10, 2000.

Kashy, E & Thoennessen, Albertelli & Tsai, Yihjia & Kashy, D.A. (October 2000) "ALN Technology on Campus: Successes and Problems", 30th ASEE/IEEE Frontiers in Education Conference.

ERIC, Educational Resource Information Center (2000) "Alternative Assessment and Technology. ERIC Digest. ",ERIC Clearinghouse on Information and Technology Syracuse NY
[Online] Available: http://www.ed.gov/databases/ERIC_Digests/ed365312.html
Accessed on September 5, 2000 – December 10, 2000.

Accounting Tutor (2000), West Virginia University, [Online] Available:
http://www.access.wvu.edu/class/acctutor/accthome.htm
Accessed on September 5, 2000 – December 10, 2000.

Blackboard (2000), [Online] Available: http://www.blackboard.com
Accessed on September 5, 2000 – December 10, 2000.

COMPASS (2000), COMPASS/ESL of ACT, [Online] Available:
http://www.act.org/compass/overview/index.html
Accessed on September 5, 2000 – December 10, 2000.

Computer-Assisted Personalized Approach (CAPA) (2000). Michigan State University [Online] Available: http://capa4.lite.msu.edu/homepage/
Accessed on September 5, 2000 – December 10, 2000.

Computer Assisted Homework/Testing (2000), University of Connecticut Chemistry Department, [Online] Available: http://www.sp.uconn.edu/~cdavid/chem263
Accessed on September 5, 2000 – December 10, 2000.

CourseInfo(2000), [Online] Available: http://www.courseinfo.pitt.edu
Accessed on September 5, 2000 – December 10, 2000.

CyberTutor (2000), MIT, [Online] Available: http://cybertutor.mit.edu
Accessed on September 5, 2000 – December 10, 2000.

Electronic Student Portfolios (2000), Virginia Commonwealth University, [Online] Available: http://www.vcu.edu/mdcweb
Accessed on September 5, 2000 – December 10, 2000.

Gateway Test (1999) University of Michigan Department of Mathematics, [Online] Available: http://www.math.lsa.umich.edu/courses/115/gateway.html
Accessed on September 5, 2000 – December 10, 2000.

FLAG- The Field Test Learning Assessment (2000), National Science Institute Education, University of Wisconsin Madison, [Online] Available:
http://www.wcer.wisc.edu/nise/cl1/flag/tools/Tframe.asp

Accessed on September 5, 2000 – December 10, 2000.

Intelligent Essay Assessor (2000), New Mexico State University , [Online] Available: http://www-psych.nmsu.edu/essay/
Accessed on September 5, 2000 – December 10, 2000.

JiTT (2000), Indiana University, [Online] Available: http://webphysics.iupui.edu/jitt/WhatQVR.html
Accessed on September 5, 2000 – December 10, 2000.

LabView (2000), Frostburg State University, Chemistry Department, [Online] Available: http://imej.wfu.edu/articles/2000/2/06/index.asp
Accessed on September 5, 2000 – December 10, 2000.

Mallard (2000), University of Illinois, Department of Electrical Engineering, [Online] Available: http://www.ews.uiuc.edu/Mallard/
Accessed on September 5, 2000 – December 10, 2000.

OWL (2000), University of Amherst, [Online] Available: http://www.math.umass.edu/~knightly/m127/owl.html
Accessed on September 5, 2000 – December 10, 2000.

PIVOT (2000), MIT Physics Department, [Online] Available: http://curricula2.mit.edu/pivot/
Accessed on September 5, 2000 – December 10, 2000.

The Quiz of CGI, University of Connecticut (2000), [Online] Available: http://www.sp.uconn.edu/%7Evirtclas/quiz/documentation/
Accessed on September 5, 2000 – December 10, 2000.

QuizSite, Bureau of Evaluative Studies and Testing (2000). Indiana University at Bloomington (IUB), [Online] Available: http://www.best.indiana.edu/quizsite/faq.html#general
Accessed on September 5, 2000 – December 10, 2000.

WebAssign (2000), North Carolina State University, [Online] Available: http://wasnet03ws.physics.ncsu.edu/info/
Accessed on September 5, 2000 – December 10, 2000.

Web quizzes (1998), University of South Carolina Department of Chemistry, [Online] Available: http://michele.usc.edu/105b/105bfall98/
Accessed on September 5, 2000 – December 10, 2000.

WebCT (2000), [Online] Available: http://about.webct.com/
Accessed on September 5, 2000 – December 10, 2000.

WebWork (2000), University of Rochester, [Online] Available:
http://webwork.math.rochester.edu/docs/
Accessed on September 5, 2000 – December 10, 2000.

WhizQuiz (2000), Virginia Polytechnic Institute and State University, [Online]
Available: http://www.math.vt.edu/Cyber1114/jalessio/whizquiz/format.html
Accessed on September 5, 2000 – December 10, 2000.

WPI (2000), Worcester Ploytechnic Institute, [Online] Available:
http://www.wpi.edu/Academics/Projects/intro.html
Accessed on September 5, 2000 – December 10, 2000.

Student Assessment Learning Guide- SALG (2000), University of Texas, [Online]
Available: http://www.wcer.wisc.edu/salgains/instructor/
Accessed on September 5, 2000 – December 10, 2000.

SUNY (2000), SUNY Learning Network State University of New York,
[Online] Available:
http://www.sln.suny.edu/admin/sln/original.nsf/504ca249c786e20f85256284006d
a7ab/bd9a90d415c32956852563f6007a1297?OpenDocument
Accessed on September 5, 2000 – December 10, 2000.

Visual Calculus (2000), University of Tennessee, Knoxville, Math Department, [Online]
Available: http://archives.math.utk.edu/visual.calculus/
Accessed on September 5, 2000 – December 10, 2000.

Andrew Depalma 11/14/00

University of Connecticut 9:20AM

Conductor: Dave

Noter: Valerie

QUIZ CGI

Mr. Depalma was into web development and graphics support. The idea/want for creating the tool came from the faculty that there was a need for it. Quiz CGI is used for Drill and Practice, it isn't built rich with content. To have a rich-with-content tool would take a lot of time and would be expensive. There was no true cost for U. Conn in developing this tool b/c Unix is used as the platform and the support staff needed was already there. Depalma and other have begun to investigate other tool- smart/intelligent tools that would be able to teach students. When CGI was first formed there were only multiple choice questions and then there were short essay questions over time. Some concerns that Depalma had were if people would actually use the tool and justify what he and other created; also if there was enough printed material on the tool. When the tool was first developed several years ago, tutorials were offered in the classes for faculty. Some disadvantages were the availability of the tools, errors that seem to creep through- where faculty would say the answer is one thing and the computer say another.

Depalma notes that as far as effective assessment goes, the quiz is informal and insecure and only used to evaluate the user's skills. He further says that typically on-line assessment is rich in content. For them to be, it would be quite costly. He noted the Jasper Series at Vanderbilt University having a rich tool. He uses the example that if a lab is to be given to test the students' skill/knowledge it is best that the lab be given in a laboratory where the student can use real equipment. Any other way would just be seing if the student could simulate.

There were in some pre and post tests given in the beginning use of the tool. Depalma says that the results were predictable. The Web wasn't really as accessible several years ago as it is today and that back then the users just took the quiz for the novelty of taking it- because they could. He noted that there really wasn't a difference in the scores from the quiz and the paper test- so that "doesn't say much for technology", he

added. He also mentioned their ability to make last minute changes to the multiple choice tests on the CGI that would be easier than the paper. Furthermore on the quiz it was possible to make it so that there were drop down lists where a selection had to be made and that a series of correct answers had to be made just to get one question correct.

Students and faculty both thought the tool was "spectacular" and both groups gave ideas in further formatting of the quiz. Currently the tool is scattered throughout disciplines. Depalma said that about 85% of the tool is availabe on ShrinkWrap. U.Conn uses WebCT. He also noted that there was a decrease in user population of the tool but an increase in user population for WebCT. Earlier this year (mid-July) he said there like 1.5 million hits on the Quiz. U. Conn has distribted this tool to local area schools such University of Rhode Island, and some Connecticut Community Colleges.

Now they are looking to support various functions on-line, for example: a plagiarism detector, artificial intelligence application, chat, and data accessing. WebCt has it so that they can build things that can't be built with other platforms. WebCt is built to handle volume- all the students have accounts.

Using WebCt to build application upon is almost hassle for them because a new version comes out all the time, they would have to constantly rebuild/update their applications. For that reason, they use Unix. One new thing they have come up with using WebCt and writing application is a nutrition awareness game.

Phone Interview with Professor David Pritchard from MIT
Wednesday November 15, 2000 11:30A

Valerie Sanders = S.
David Pritchard =P
David Vallerie = V

P: Hello
S: Hello, May I speak with Professor Pritchard?
P: Hi, is this Valerie.
S: Yes this is
P: Good morning
S: Good morning, how are you?
S: I guess before we will start, I'd like to know do you mind us in tape recording this conversation? It's just for our notes. And your name will appear in out report and things.
P: OK that's fine. Just doing the transcript.?
S: Yes, sir
P: Okay well you can send the transcript to me for checking
S: That's fine.

S: Just to start, for CyberTutor, did you take part in the creation of the tool?
P: Yes, I did. But I didn't do any of the programming...that was done by my son,
S: OK
P: I also wrote or was in charge of the people that were writing the problems... help me to write the problems
S: OK
S: What was your Initial motivation for creating this tool?
P: Well, my Fundamental motivation was that I think a lot of time is wasted on homework, and I want to able to give some tutorial assistant to the students while they are doing the homework.
S: Did you see any advantages for the faculty using this? Other than just saving time.
P: For the faculty I think the major advantage is, well it's two-fold. The first is, for say people who are teaching in a recitation mode or just using the library problems, they can see where students are having difficulty... they can look at the problems and the program will display, well, compare how many students got the right answers to how many got the wrong answers. For this in particular problems, how many requested the hints, and how many students requested all hints and still unable to solve the problems and therefore requested the solution, All that information is presented to the instructor and the instructor can know which part of problem to discuss in class.
S: So for the CyberTutor, is it just for each class, or is it for all?
P: Well, the basic way CyberTutor works, you have to write or somebody has to write a library of problems. And then one step library exists, other teachers can use it to make homework assignments, selected from those problems for their students.
S: OK
P: Now I didn't say for another thing and that is for the author problems, the feedback from student is very valuable because you learn for instance where you don't have

enough hints, where a lot of wrong answers were given, in facts the program does more than just tell you how many wrong answers, it tells you of the various wrong answers how many wrong answers of each type were given. So may I backup one little step here, CyberTutor mainly emphasizes three form responds not multiple choices, but single word, fill in the blank, and analytic expectation which is probably what we use, far and way the most. And then some simple applets that the students can draw curves on or students can draw vector diagrams on, so in all of these, students don't pick one of the five multiple choices, but they don't understand the problem or if they get a wrong answer and they submit the wrong answer, so the idea is maybe the author can figure out what the students are thinking, what they are doing wrong by looking at the wrong answer say how would I get that answer

S: OK, how are the students able to get feedback, is it through CyberTutor or is it in another way through classroom?
P: The students get feedback from CyberTutor, and they get the feedback from… right now what we are doin is, we have certain number of generic answer checkers. For instant let's say that the answer contains three terms A+B-C, the student might submit answer in like A-B-C, so that will be grade wrong, because the solution is A+B-C. But, the computer would then compare most students answers, take both the student answer and the solution and change every negative sign to the plus sign, and then both of them would be A+B+C, and at that point it would say you seem to be on the right track, but you should check over your signs
S: OK
P: It could make statement of that, call them generic to the students. In addition, after we've run the problems on one class, the computer displays the most frequently answers the wrong answers in order to most frequently to be given and less frequently to be given, and you can look however many you have patience for but typically about three or four will, maybe half of all the responses will be three or four different wrong responses, and then you can figure out what you would say to a student who gave each of those wrong responses, maybe it's like, don't you think that acceleration should be depend on the mass m two, that's not in your answers.
S: You mentioned students being able to give feedback to the professors. Do they give the feedback though CyberTutor or is it classroom?
P: No, there is an opportunity for the students to send comments to the professor, and the comments are basically, I would say a major of three kinds. The Most trivial is…well, I just push the answer button by mistakes and I don't want to lose half of the point whatever it is for the hint I didn't need. OK, the second one is that they have a particular question for the TA who is online but not all the time just a few hours of each night of three days before the assignment is due. And they can say please help me, I have the following question. TA actually responds to them by email. The third thing is, they can say student will say I am really frustrated by this problem, I just didn't understand this in following, I kept submitting this answer which I thought would be graded correctly but it wasn't.
S: OK. Have you found any disadvantages or setbacks from using the tools?
P: Hmm, well, I would say first for the tools will have limitation it is difficult deal with the problem, which it would solve in one of two different ways. You can only basically…

easily give a set of hints to the students to do in one way. Also there is open-ended problem not good for CyberTutor, of course there have been frustrations when our server was inadequate. As far as students are concerned, the students didn't like it when the server didn't work, but in the second term, students recommended it to subsequent students just over two to one and I think as we use it more, learn what the students rating are, fix the problem when they given the wrong answers, I believe that percentage students use this learning tool will go up.

S: Could you tell me or do you have a figure of about how many students use CyberTutor?
P: Yeah we have been used in classes here for a last year and half, that range from one-hundred to one hundred eighty students.

S: OK, I'm going to switch gears here. What do you feel makes an effective assessment tool?
P: OK, so now we are switching from the tutorial role to the assessments?
S: We are looking at assessment tools but we are also looking at computer assist learning like CyberTutor can help to learn and it's a grader also, so we are also looking for the grader, assessments like quizzes.
P: So, OK, fine. Well, I get .. what was your exactly question again?
S: How do you determine the effectiveness of CyberTutor, like have you assessed it in any way, do you have any statistical analysis?
P: Oh, I see what you mean. Yes,we did one thing last semester, we gave the force concept inventory test to the student at the beginning and the end of the term, the course was a conceptual assessment of understanding of kinematics and Newton's Laws. Then we correlated the various components of the course.. CyberTutor for work, we had written homework, we had class participation and recitations and we have some tutorial in recitations. And so we correlated the grade the student got in these various things against their improvement in the force concept inventory test. What that showed was that the CyberTutor was the winner in terms of increasing the effectiveness...increasing the student's score in the conceptual tests, it was little bit more effective working out the problems in group, it was twice as effective as the written homework and improving the students' ability on this particular conceptual tests on force concept.

S: We'd be very interested in the data, would it be possible for you to send me some of the data?
P: Yes, I can do that. I have your email address?
S: Yes, you do.
P: These surveys can always be attacked , because we did not really do any blank testing with Group A and Group B, and then there is a question of what goes into the CyberTutor grade, is a complicated formula of how many did you get right and how many hints did you ask, you have to lose a little credit, where you will lose a lot more credit if you ask for a solution to the hints, if you will lose all credits if you ask for the solution for the main part of the questions. There are also numbers of optional problems and the students were graded in the ways about three quarters of them got one hundred, but we gave a little bit of extra credits, maybe up to 105 points gave to the people who

did twice as much work. So, in general CyberTutor's grade is proportional how much work student did. And in fact, the homework grade most of the students get all the problems they turned in right, so the homework grade is also proportional to how much problems students did.

S: For the CyberTutor, does the homework that they do on CyberTutor, does it count toward their class grade? Or is it just a tutor?
P: No that counts as the grade of the class. I don't remember what's the formula was but it counts as the same as written homework, each count about like 10-12 %, and there also test and final exam grades.

S: Is MIT the only school using CyberTutor? Have any other school adopted the tool?
P: We just working on it, getting the problems up and right now I would say the assignment system, the software for making the assignment is not good enough to just be used by other people, hopefully next semester we will have few other places trying it. There are some people express interests, but we just didn't get up to that stage with it yet.

S: OK, do you have couple names of school that they are interested in?
P: Well, I have couple names of professors that are interested. But they haven't agreed to use it, so I don't want to go in to the exactly who they are.
S: Perfectly understandable.

S: Have you worked with any other tools before CyberTutor?
P: I haven't worked with them, but I have looked them over and in fact it was my originally intention to use CyberProfs, which I thought was about the most sophisticated of the programs like the WebSign or WebCT, where you could compose your problems and administer them. I though the CyberProfs was the best but when we went to license the thing, MIT wanted to make a few changes in the license agreement and they never responded, so we start to write our own. Well, that's... period.

S: Other than, CyberProfs, do you know any other tutor or assessment tools that schools are using?
P: Hmm, well there is Andys, there is blackboard, there is also a free one in Texas University, I forget the name, I have a big list that I looked at one or another, I would say most of the well known things are what I would generally classify as administrative programs. That they're designed to grade a lot of homework and take that load off and do it in a way of assessable students, and provide immediate yes or no answers to the student which is important, then they don't have wait for the solutions to been given out a week later. There also some other programs, for example, like the Freebody, which I think is distributed by the APS, which is more in the favor like CyberTutor tutorial problems of programs that help students with one aspect of one another. And that's the more flavor CyberTutor is aimed at. I think Freebody is probably is most widely distributed, it seems to be more used in high school, it's a fairly simple program.

S: Well that's all the questions that we have for you at this time. We definitely will give you the transcript after we finish transcribing our conversation and we are looking forward to receiving the data from you.
P: Assume you type it up then send me electronically one and I would get it the most efficiently way to go.
S: We will do then.
P: OK, I guess I want to make one more final comment.
S: Oh, please.
P: As far as assessment tools, I think that the amount of information that you have about for students, something like CyberTutor is immense amount that you get compared to the examination.
S: I guess what type of information?
P: No, just the information of performance of the students. On a typical kind of homework, you want to get, you know you may assign three point of each one, so that's how much you will know how's the student does know that three problems. On the other hand if you want to look at the student do on CyberTutor, you know which hints the students needed to solve the problem, you know how long it took them to solve the problem, you know how many wrong answers the student gave, you know which particular wrong answer the student gave, some of the wrong answers are indicative of lack of skills in trigonometry, or solving simultaneous equations. So you have all this information from the students, and so ultimately, it seems to me that you should be able to assess students much more accurately with something like CyberTutor where you have so many more interactions than you do with the homework.

S: When we write our final report, would you mind being quoted on any of this?
P: I don't, I have to look over the transcript before I would do that you know because when you talk you things little more sloppily or a little more ostentatiously, majority to other people, and I don't want to interact their side, so I have to be a little careful about.
S : Totally understandable. OK, thank you for your time.
P: OK, fine, tell me a little about your survey.
S: About our survey. What we are doing which is completing the project for our school, and we're just looking at all the Information Technology learning based assessment tools, tutor, graders, that are out there forming a huge list. This is a stepping stone for the National Science Foundation, we will pull together as many examples as we can come up with.
P: I see. How many I have mentioned that you have?
S: I don't know we have any of those.
P: Are you just doing physics?
S: No, any… all disciplines
P: OK
S: Is the CyberTutor only use for physics? Or use it in any other disciplines?
P: It could be in other disciplines. There is a Professor here in math that has started to write a problem library for differential equations. But so far it hasn't been used , we don't have enough problems to accept.
S: OK
V: But it's easily adaptable?

S: Yes, someone have to do the work of writing those problems, and my estimate is that the it takes to write the problems is about the same amount of time to write the solution on the same processor. However, before you give a problem electronically which have to check it over very very carefully to make sure the wording is as unambiguous as possible. So to check the absolutely certain that the answer is correct or there are any other reasonable answers can also graded correctly. We almost spent an hour and half to check over each problem after written, it might take 3 or 4 or 5 hours to complete if there are some computer art work then it will take a long time, then also after the problems are given, you want to go over the wrong answers and respond, make a response to them, and that will take another hour per problem. And for a typical course, you need something in the order of one hundred…one hundred fifty problems.
V: It gets pretty time consuming.
P: I think it is less time consuming than a textbook, order of time to take to write a book of homework problems and solutions in a word processor. Maybe take up to thirty to fifty percent more effort than that. But when you look at the scalability, and what it actually does, I think it's worthwhile.
V: So do you feel it is a more quantitative test? Being an electronic test you will get more data out of the test then you would from the traditional pencil and paper test?
P: If you're using it in tests, I think you do, because what you would essentially find is that…say your standard a problem was split into four parts, so you'd grade those four parts whether the answer was right or wrong. On CyberTutor for the students who couldn't get it right, they will be told that they didn't have it right and then they might just start using some hints, and then they might eventually be able to get it right, but you could watch them in that process to see, kind of pin where their level of knowledge extended to.
V: Let me just say, I'm Dave, Valerie's partner. I didn't mean to jump in there, I just had a question there at the end
P: That's okay, What's your last name?
V: Valliere
P: Valliere?
V: Valliere,
P: OK, its French
V: Yep. I think that's all the questions that we have.
S: Do you have any more questions for us?
P: Nope, I didn't I just hope you will send me a copy of report, cause I am sure I will think up some stuff that I haven't found.
S: We will do that.
P: OK Fine.
S: Thank you for your time.
P: OK Bye Bye
S & V: Bye Bye.

Interview Transcription: Dr. Bill Vining **November 14, 2000; 1:30PM**
Department of Chemistry, UMASS Amherst
OWL

Valerie: Did you take part in the creation of the tool?
Vining: um, peripherally so.
Valerie: peripherally?
Vining: Peripherally. There's essentially a few different parts to OWL. There's the main system. The main homework system. Which I didn't have a lot to do with. But there's two sets of fairly large enhancements. One is Discovery exercises and the other is intelligent tutors.
Valerie: uh huh.
Vining: I had a lot to do with the latter two parts of it.
Valerie: ok
Vining: but not a lot to do with the original main system. But I'm on the main team that was doing it.
Valerie: What would you fell are some of the advantages for the students and the faculty for using the OWL tool?
Vining: Well, the advantage, let me start with the students.
Valerie: ok.
Vining: The advantage to the students is that, what the system does, what the main philosophy behind it is what's called Mastery Learning. Which is, where people can work on something and if they get it wrong, they get help.
Valerie: mmmhmmm
Vining: but then they also have to go back and do it again. And the biggest problem with doing that with homework is if you hand grade the homework, then someone has to grade it over and over and over again. Which is not realistic when you have over 1200 students.
Valerie: Ok.
Vining: So that's not a good thing. When you have it on a computer-based thing, what happens is the computer gives them a question, and then they get it wrong and it tells them how to do it and then the next time they already know the answer. So what the OWL system does is it very elaborately, um, constructs questions on the fly. So if a student gets something wrong it tells them in great detail how to do that problem, but then they have to go back and do it again but it gives them a different problem.
Valerie: Oh, Ok.
Vining: It offers and unlimited, infinite number of questions that it can give them based on the way that we construct the assignment. So students can work on a problem over and over and over again until they have really learned it. When they get it right they get just as much credit as the people that did it right in the first place. So the mastery learning thing is the main advantage to the students because they can keep trying it and trying it until they finally know it. From the instructors point of view, what that does well there's two things. One is, students, as with everybody, will do their homework more if it is graded than if it is not. You know if you have something that is not graded it sends the message that it is not as important. So it offers us a way of assigning homework to a large number of students without actually having to do the grading. That's the main advantage to the faculty. There are a couple other advantages, one that you can um set deadlines with um, as you like, individual instructors keep the students up with the work. Lots of time students won't do any work until right up until the exam is about to happen. Using the system you can give more assignments like one assignment a week or even two or three assignments a week, to make sure that the students work continually throughout the semester. That's particularly good for the instructor so that the instructor can keep tabs on that sort of thing.
Valerie: Right. Can you tell us some of the disadvantages of the tool for either students of faculty?
Vining: Um, well, um, one disadvantage, which isn't too hard to work around, is that sometimes the computer equipment doesn't work reliably, so that the chemistry department server that was serving the system the central database holder of everyone's grades will sometimes clog up and stop working for a little while. Never like losing grades or anything like that, but sometimes when people have something due, and the system won't work, sometimes it's our fault or something on campus or sometimes it's the worlds fault, that clog the system up, so sometimes people have something due and the system doesn't work and everybody gets mad. What you need to do then is just be fairly flexible in changing deadlines when that occurs and make that clear to the students that that's ok. That's the only disadvantage, when things go wrong with the computer stuff.

Valerie: yeah, ok. When you were in the process of creating the two components that you worked on, were there any concerns that you had, I guess, for instance like, costs and technology?
Vining: Costs or technology to the students you mean, or,
Valerie: yeah or just to the school. Like for instance, would you use Unix or would you use another platform?
Vining: Oh I see, well what we basically did was we used the tools that made the most sense for the job that we were doing. So most of the programming was done on relatively inexpensive Windows types of computers. And that wasn't very hard. And it ended up that that wasn't really much of a concern. We had um the financing um came in part from the NSF, who you are apparently working with, part from the Department of Education called FIPSI there and they fund higher education projects, as well as a fairly large amount of money from the university because the system actually saves the university a great deal of money because we don't have to pay people to grade. We were always concerned about how we spent the money that we were spending, but it was never really a problem.
Valerie: Um, ok, I guess getting away from the specific tool for a little bit, what do you feel makes an effective assessment tool?
Vining: an assessment tool for assessing how the students are doing, you mean?
Valerie: Yeah.
Vining: yeah it's actually pretty good for that. One of the things we have studied and that we track fairly well is the students who do all of their assignments do better on their exams. So we have a very good correlation of that for both chemistry and physics now. We expect to see the same things in other things as well. For some of the other things like enhancements like the intelligent tutors we have seen fairly close alignment being students completing the assignments using the tools and how well they do on exam questions. So, so they're getting grades on their homework, and the people that are doing well on their homework grade, homeworks, are also doing well on their exam grades. So it seems to be working out well that way, yeah. But I think it is good for that.
Valerie: Ok. Are there… Are there any other disciplines that use this tool other than physics and chemistry?
Vining: Yeah um, let me think, yeah there's a fairly large variety now but they use it if different amounts. I'm trying to remember who they are at the moment. Um Geosciences is using it, um, I think Art History is using it, Biochemistry is using it, there are a few others, I could send you something about that if you want. There's five or six others. Chemistry and Physics are using it more elaborately, than the others because we have been working with it more and we have built up a database of questions. So, but other folks are using it quite a bit. Oh, but the other thing it's used for on campus is there's a group on campus that is in charge of the kinda like OSHA regulations, the Health and Safety Department and the Health and Safety department uses OWL system to do all of their safety training on campus for thousands of employees.
Valerie: Ok getting to the faculty, are there any incentives for the faculty to use the tool, are they very receptive to using the OWL assessment?
Vining: um, I'm not sure about it is physics because I don't work in that department, but in chemistry it's all very, um, there's no problem, people really like using it. So there actually isn't really any need for incentive because they actually can see that their students are doing better. As we've made new things, what's happened, like these tutors for instance, what's happened is that we've done studies that show that the students who use the tutors do better and as soon as we show that to the faculty they start assigning those as well. So there hasn't been any problem with that at all. And one of the things that has happened is, in chemistry the initial use was in the general chemistry classes, the first year chemistry class, and now what's happened is the second year chemistry class, the organic chemistry class, have also started using it. Have started making databases of questions as well. So the fact that students are doing better with it has made the faculty start to expand what courses it gets used for in chemistry. So there's no problem with that at all.
Valerie: Are there any tutorials or any support staff that is used or is needed for…
Vining: um, most of the…Ah…Well…Ah, there are those things, but they are not very much needed. People tend to just be able to get on and start their instructions on the software itself, that makes it pretty easy to use. There is also an online manual for it, but in addition to that there's a whole, um, there's two ways that we offer support. One is that, in the system itself there is a messaging system, where students can send messages asking for help and then they will get email messages back telling them you know you know what to do. The other thing is in chemistry in particular, I know this exists else where at other schools, we have a resource center where students can come and do their work at the resource center. And

at the resource center there are people trained there to particularly help students that are having problems, so there's a place for them to come where there are computers and there are people that are trained to help. And that's staffed from like noon till midnight, all work days and Sunday nights. So, we have places that people can come get help if they need it.
Valerie: You just mentioned other schools, are there any other schools that use OWL assessment?
Vining: Um yeah, there are I think 6 other schools that are using the chemistry version of OWL. I don't know that if they are using it in other disciplines or not, but in the Chemistry discipline, Umass Dartmouth, Umass Lowell, Umass Boston, St. Anselems, CCRI, and I think there's another one but I can't remember.
Valerie: Have you used any other tools other than OWL?
Vining: I personally haven't but the other people in chemistry have. They used to use a system named PLATO, which still exists but not too many people use it anymore, it's been around like 20 years. So they used to use this system named PLATO but they don't use it anymore. It was very crude by modern standards, but at the time it was quite good. So, ah, so ah I personally haven't used anything else.
Valerie: Ok, I guess I have one last question, do you know of any other assessment tools that are being used a any other colleges and universities?
Vining: Yeah, there's a few of them. One is called Kappa. Which I believe comes out of the University of Illinois, but it might be the University of Indiana, it's one of those mid-west schools. Kappa is it's name, another one is called Web Assign. Which comes from North Carolina State University. Those are the two I know most of.
Valerie: Ok, well those are all of the questions that we have, we would like to thank you for your time.
Vining: Thank you, Valerie; I look forward to reading your report.
Valerie: We will be sure to send a copy of it to you.

WebWork- University of Rochester
Interview- 11/20/00, 9AM
Dave- David Valliere, *Arnold*- Arnold Pizer, *Mike*- Mike Gage, *Vicky*- Vicky Roth, *Val*- Valerie Sanders, *Ravenel*- Doug Ravenel

Dave: We were wondering if we could talk to you a little about WebWork today. If you don't already know, we're doing a research project for the National Science Foundation. We're investigating and evaluating information technology-based learning assessment tools. And after we've done some research on it, we feel that WebWork is a good example so we wanted to get some more information about it. Our first question is what was your initial motivation for creating this tool was.
Arnold: Part of the motivation was that we knew that it was important it was to grade homework in calculus, but it was impossible for us to do that by hand with the large # of students involved, we did not have the resources to grade homework. This is a way that we can do it,
Mike- And a second reason was that, we were involved we were involved in teaching a joint course in Physics and they were using something called Kappa- which essentially has the same philosophy as WebWork and when we heard about Kappa we thought that was a great idea. Kappa basically is a system similar to WebWork, you get immediate feedback on the validity of your answers. Technically it's different, but the philosophy is the same.
Dave: The second question is: What sort of advantages do you perceive for the students and faculty using WebWork- is it just that it's faster and easier for the faculty so that they can grade a lot of calculus problems.
Arnold: There's several advantages, what you mentioned is one advantage and that is probably the thing that grabs people first, especially like chairmen responsible for getting a lot of homework graded. The second thing is, I don't know Mike you might want to take this- it has a lot of advantages for the student
Mike: The most important advantage is that students get immediate feedback when they've made a mistake so it really changes the way they do their homework. So typically what would happen is a student would answer a question and then…well some students try and figure out whether the answer is reasonable but a lot of students- you know once you've got some answer, while its up to the T.A of the professor something like before the tend to respond…so this gives immediate response. I don't know if you've checked the characteristics of WebWork, its important to understand it gives immediate response to a question. It tells you whether the answer is correct or not; it doesn't tell you what the correct answer is. And you get to try the same question again. This is not like a series of quizzes in which if you don't do well on that you get a new quiz and start on that-this is much more the homework model, you make a mistake, your homework isn't right- you work on it and get the right answer
Dave: Yeah, Does it give the students any feedback on if they got the question wrong, where to go look where to study more or what the concept was they were missing?
Mike: It really focuses on the homework, we pretty consciously tried not to create a computer instructed learning system- that's a pretty difficult thing. So this really meant to work with recitations and professors teaching classes and things like. It is possible to put hints; it's even possible to key the hints into the kind of answers the student got. But, by

and large we haven't explored this very far. In my opinion this is actually a good thing, while its easy to make some kinds of guesses to what the students get wrong and give elementary hints, but often hints do almost as much harm as good. We try to encourage students who are having difficulty to go talk to a human…T.A or professor or something like that.
--Let me just make a small point and then I think we can get Vicky R in here. As an example of usefulness of WebWork, we found that this semester that the busiest that are system has been is right before common exams in our large Calculus courses. Students are really using the system to study for their exams. I'm going to turn it over to the secretary.
Secretary: Hold on one moment
Vicky: Hi this is Vicky Rauss, sorry to be joining late, we got a little confused about getting me hooked up.
Dave: I'm Dave Valliere and Valerie Sanders and Anita Wong are here and we're all from WPI. We were just talking about the advantages for students and faculty for using WebWork.
Vicky: How far did you get in this conversation?
Mike: I gave the discussion about immediate feedback
Vicky: Maybe you could give me a little background on like is it decision-tending, you're looking to try something out
D: Us, we're students from Worcester Polytechnic Institute, we're doing a research project for the National Science Foundation and what we're doing is investigating evaluating information technology-based learning assessment tools. Basically what they just want is for us to find as much as we can about the different tools that are out there and we came across WebWork during our research and we figured that it was interesting tool and we wanted to learn as much as we could
Vicky: And you folds have already tried out the web site?
Dave: Yep
Vicky: Okay good, perhaps you've already heard that we've done evaluations based on 3 different formats, sort of triangulate what we find. We did a survey of students, online survey, we've done observations of students using WW and we've done interviews-students who have used WW, faculty members who have used WW and a few TA's
Dave: Are the results of these different things are they available on-line?
Vicky: These things haven't been published-we have them kind of halfway ready to go and if there are specific questions that you've got, we have an article in the midst of preparation if there are very specific questions that you've got we would be happy to answer that. I think in addition to the two immediate adv. we saw coming, we anticipated-kind of a hypothesis coming into this that students would like the immediate feedback and faculty and TA's would like to relieved of the burden of paper and pencil testing. Something that we've discovered at our institution that may or may not hold true and other institutions is that students would, for example, would often times turn to other students for help to work their way through problems. Which was from our perspective a desirable goal. They would turn to other students and make contact to a faculty member
Dave: Okay, I'm going to change gears really quickly. Getting back to the feedback that you were just talking about are there any specifics…you said you did doing a survey, was that survey of the opinions of the students?

Vicky: Yes, it was a survey of whether they like it or not we also were taking a look at how they were using WW. Like whether they used it in their dorms or whether they used it the night before- those kinds of questions. We were also trying to see if we could see some connection between patterns of use and anticipated grades. Also trying to see whether or not WW is being used different by students at different levels of the Calculus course. You know like if people are in something that would be considered a very entry-level course here... are they likely to use WW different that the students at the higher end. Roughly off the top of my head, without having all the numbers in front of me: students, who are at the top, end to work together more and not less.
Dave: Yeah, Do you have any data that indicates that students' grades are better as a result of using the tool. Perhaps like taking their grades in a class before and then after....
Vicky: Someone's doing their dissertation as part of this project right now. That's going to give us real facts about that. I wouldn't say we're in a position right now to say that grades are- I could show you some comparisons in patterns of use and satisfaction. We haven't done a carefully controlled match kind of study where I would be able to say we controlled all the factors that would let us say that...
Arnold: Some anecdotal evidence along these lines and Vicky and can say if there is any experimental data on this. Because we're suing WW, we grade somewhat differently. In that homework-at least I allow homework to count considerably a greater part of the grade. Before this homework has often not been collected or if it's been collected its been collected and returned late and I would count maybe 5% of the grade. And even when it's collected, it's spot-checked. We just don't have the manpower to grade every single problem for every single student. With WW I know that every single problem has been checked, so I count it a lot bigger apart of the grade somewhere between 20 and 30%, depending on the class. There's probably result a little bit of grade compression and that students who don't do so well on tests, they know that they can work harder on the homework. The other thing- well 2 characteristics that I've noticed is that students like this in some sense and they really feel a sense of control and sense of responsibility that in some sense to get better grades they just have to work harder. And for some reason this is more effective on the homework than on the test. I think the tests they feel the tests are a little bit the laugh of the Gods, the homework they see quite clearly the relationship between better grades and doing their homework. That's one thing. The other thing is that the homework grades tend to be really quite high. You'll frequently find a huge portion of class getting 100% or close to a 100% on the homework, which indicates they've really done a lot homework
Vicky: And that's something that the usage survey confirms that students reported that they would keep working on WW until they got all the problems right. Which is something quite unlikely to happen when people are doing paper and pencil homework because they are not getting the feedback that even lets them know whether they're right. They may stick with it until they've reached their level of saturation but, beyond that there is nothing more they can do unless they have an answer key somewhere.
Dave: Right. How many students actually use WW at your school?
Arnold: We got just about a thousand students in Mathematics and I think this semester about 200 students in a Physics course.
Vicky: And then it's used elsewhere
Dave: At other schools?

Arnold: Yes
Dave: Do you know specifically what other schools using it?
Arnold: We can start giving you a list. Indiana has several thousand students...Johns Hopkins, Ohio State, Dartmouth...Stony Brook
Mike: University of Virginia
Arnold: University of Virginia. Then there are a number of places like McGill that are using it. Then there is a number that are experimenting with it- I'm not sure how, actively they're using it-Rutgers is using it for certainly a course…Utah is using it, UC Irvine, I think, but I'm know 100% sure
Ravenel: Who else is using our machine, Arnie?
Arnold: Our Machine…Cleveland State…Buffalo State is using it…Rochester Institute of Technology is using it…there's been a little experiment with a few high school students are just starting to play with it
Ravenel: Penn State- Altoona was using it for a while, I'm not sure if they're still using it or not
Dave: Is WW a tool that can just be used for Calculus, Pre-Calculus and Physics or can it lend itself to other disciplines? Like English classes or instance.
Arnold: It can easily lend itself to like Engineering in fact its been used for some engineering courses. It is a tool and I think my feeling is that you should use it for types of problems it is best suited for. And it is certainly best suited for problems that are like would have a numeric...and answer that can be checked with an algorithm. You know an answer, which is a function...asks you what the derivative is, asks you the anti-derivative a function, short answers things like that. If you want its set up in such a way, that if you ask like essay questions, the essay questions can be emailed back to someplace...that's how we do our questionnaires where students basically do the evaluations. At some point they write things...most things are checking off numerical evaluations but they also write essay responses. And those are essentially sent back. Now with those, if you wanted to you could figure out how to grade essay by computer, but I don't think I would want to try that and I don't think that's a really good use of WW.
Dave: Yeah
Vicky: There is something similar that is being used in one of the California schools, unrelated to WW… actually in a conference I stood up and argued quite vealently that. But I would see the nature of WW could go certainly beyond math and physics, I think there is a lot of other applications…that are out there. I just wanted to make a point about the list of schools we just gave you, a year from now… if you folks at NSF or whoever else wanted to come and talk to us about performance outcomes, we should be able show you some information from the U. of Rochester and UVA- kind of a comparison study. Our 1st round of analysis, something that I would argue with NSF is always something that should be done…before you start looking at grades, you start looking at what the treatment is, i.e. how people actually use WW and then you start making some decisions about whether the treatment had an effect
Dave: Right,Okay…
Dave: I'm going to change gears really quickly, have you found any disadvantages using WW…is it difficult for faculty to implement it, problems, or anything along those lines.
Arnold: Well…it's a complex system So setting it up, you need a Web server…you need a lot of tools…so that's certainly something. It takes especially…WW is very flexible, it

allows you write very flexible problems. Now if a professor wants to use problems other people have written at this point at Rochester we've written about 1,300 problems and people at Dartmouth and Indiana and Stony Brook have written problems too. If you're using problems that other people have written, it's not that onerous but if you want to go off and start teaching a course for example a financial math course which I think someone and Virginia did…a course where the problems don't exist- advantage…it's a lot of work to write problems. You have to be very careful in writing problems. It….. Its easy to make a mistake. If you make a mistake with a WW problem…you're publishing that out to 300 students, with a course of 300 students- that can cause some problems. That certainly takes some time. WW is set up so that our students can send you, send the professor, send somebody very easily a question about a problem…by email and depending on…especially if you have a problem that's wrong or very complicated or students don't understand what they're suppose to do, that could generate a lot of email and if you respond to that email...its gonna take some time. Some profs think this is great, and some profs don't care for it. If they don't care for it, they don't have to take it…they can redirect it to someone else. It certainly can take some time.
Mike: I think that if you're expecting to save enormous amounts of time, I don't think you'll be happy with WW as if you're after some other goals. And generally I think that we've been able to do more things with WW…things we didn't even attempt, such as grading everyone's homework. But in the end the amount of work that goes into it is distributed different but is largely the same.
Arnold: Let me make one comment to that. I think that is true from our perspective b/c we never ever graded at U. Rochester. Profs essentially in most large courses… we never graded the homework. We might have had graduate students or somebody like that grading small parts of the homework. But if you're comparing a lot of work using WW the amount of work that would be involved if you actually graded homework then its probably an enormous amount of saving. Profs have made the point that they are spending a lot more time writing homework problems and a lot less grading homework problems.
Vicky: And in some disciplines there is some advantage to grading all the homework. Physics people tell me –that around the country that they and their TA's grade the homework and so if we list that off of them and direct their efforts elsewhere, I that there can be some benefit.
???- But, I still think that largely what tends to happen is that the efforts get directed elsewhere.
Vicky: Right, We hope that it's something more interactive than sitting there with paper and pencil grading homework.
Mike: Exactly and the email feedback is an example of this. So for example b/c a student are at a computer when they have trouble, they send an email to me…I get email and respond to it. I may see fewer students in my office hours although I still see a fair number from there…but I certainly interact with a lot more of them by email, that I use to. So in some sense, I haven't saved myself any time. But I think mostly these interactions have been valuable.
Vicky: Yeah I would think so and if I could speak from the student's perspective from what we observed…about advantages and disadvantages. From years back students talked a great deal about the difficulty in understanding how to input answers as WW has

changed and improved there's much less confusion about those particular types of sets if issues. So I think that's an issue that was there before, that is much less there now. The one concern that I have doesn't really have to do with schools like the U. of Rochester and UVA and any of the other places you had on the list, and that is if this program was rolled out a school with uneven access to computers. We're so use to it here being a 100% wired campus, but in other contacts, in contacts with people who are working at colleges- often times community colleges, where much smaller percentage of their students are routine computer users don't have computers at home, they're not residential campuses, so I would think there whoever would be implementing WW at institutions of that nature would have to think carefully about making sure that there was equal access to computers or at least reasonable sufficient access to computers and email…we just take it for granted here.

Mike: I'd like to make a couple of comments. First thing, I don't know if you're aware of this, WW gives each student a slightly different version of the problem. Something you couldn't do if you were assigning homework any other way...I can imagine. Every student gets a slightly different version, which means they can't copy each other's answers…they can copy each other's methods, but they can't copy each other's answers. Secondly, the feedback feature- the email- it's possible for a student…when a student is using WW, if they have a question, there is a button they can push to send an email to the instructor and that the email that the instructor gets includes a URL that will enable the instructor to see that student's particular problem…the problem that that particular student is working on, which can be helpful if the statement has a specific question about how to do this problem.

Arnold: And another point is on the issue…as Vicky said, in the earlier years students had problems knowing how to type in their answers. I have not encountered any problems like that. I've got about 80 students in a course and I don't recall getting a single complaint, single question about how to input answers. One of the features that has been added to the program, recently, is an answer pre-viewer. So that when a student types in an answer, they can see what the answer looks like when it's formatted in an appropriate way. So for example, when a student types in n1/2+3, when they see it formatted they will understand that what they typed in is 3 and a half, rather than 1/5.

Dave: So it just alleviates some confusion for them, just in case…that's a good idea

Dave: This is not really having to do with WW, but have you worked with any other tools like WW or do you work with any other tools in conjunction with WW?

Arnold: Well we worked with Kappa before we worked with WW, which is certainly similar and we certainly do work with other- well that depends if you mean technological tools. WW wasn't meant to be something to lower or replace all other instructions. We do lectures, we have workshops slash recitation type sessions, office hours- we have lots of things to support WW and I don't think WW would work alone. In fact this is a difficulty of WW that in learning how to use WW, I think one of the things that's important is to encourage students who are having difficulty is not to have the computer help them. The computer tells them whether their answer is right, whether there is difficulties or not, but it really doesn't help them figure out what their doing wrong. I think an important aspect of using WW effectively is to encourage students that if they are having difficulties then they need to go somewhere else than a computer program to

get help and it might be another student, and that's effective in many cases or it might be a T.A or an instructor either directly in person or by email.
Vicky: And I think that is why we were so pleased to see those figures so high on our survey, students are telling us that they turned elsewhere that was their point of intent
Dave: Well I guess I phrased my question wrong, but…what I'm interested in is are there information technology-based learning assessment tools, like WW…like you mentioned Kappa. Do you know if there are any other ones that are being used at the U. of Rochester, or elsewhere?
Mike: Elsewhere there are certainly a lot of them. I don't have firsthand knowledge about there use
Dave: Yeah, We were just wondering if you knew, specifically, of any other ones so that we could have another direction to go in for research.
Mike: Oh, Okay well there's a long list.
Dave: That's what we've noticed
Mike: North Carolina State that's something. I don't know…which ones do you know about?
Dave: We know about quite a few right now, probably 20-25.
Mike: Well you probably know more than I know
Dave: Okay, I guess I just have one last question and this is more general about student assessment…learning assessment tools. What do you feel makes an effective learning assessment tool?
Vicky: I would say there's no such thing as ***a*** tool. It's one of the things that a graduate professor taught me years ago and it has been a steady piece of advice that I return to. There's no one measure, no one survey, one observation, or one interview is going to be able to give you a complete picture. And that's why I used the word triangulate before. I think you need to go in with something that you're trying to test out. You go in with some ideas, goals or questions that you want answered. And then you pick it up and look at it from a number of different directions. So I think that's one thing that's confirmable from a number of different angles. I think another effective approach, rather than tool, might be something that is replicable that when you do it again the next year, you get something similar…assuming the treatment has been the same. It isn't something that doesn't fluctuate according to fluky situations too wildly, that you're able to use it. I think another component of something that's effective is that we could teach it to someone else. And that we could get reliable data from it as well. I don't think effective tools are only related to grades, that is one of the outcome issues you would use. But it might be other things, like, that a program is beneficial across demographic groups, for example. A whole 'nother angle to take on this is that a program, a curriculum innovation or teaching tool that you use is sustained across time.
Mike: I'm trying to remember the correct phrasing of you question. But I wouldn't have described WW, at least the one we use, as a learning assessment tool. We're actually trying to make learning effective…that is to make homework effective and learning. And the assessment part is pretty much secondary for us
Vicky: Maybe I misunderstood the question. I thought we were being asked how were we assessing WW. But perhaps your take on the question, Mike is what the question really was.
Dave: Right, yeah

Mike: Which was the question?
Dave: I think it was more along the lines of what you were saying Mike. I guess its more.
Mike: Vicky's answer I certainly agree with in terms of what you do with assessment, but primarily you could use WW for assessment. Our real motivation was not assessment. It was to make the homework component of what they're doing more effective.
Dave: Okay
Vicky: Good clarification
Dave: That's all the questions that we have, do you have any questions for us?
Vicky: Where is this going? How are you putting this all together?
Dave: We have to do a research project as part of our requirements to graduate from school and we could apply to do it at different locations and we were accepted to do a project at the National Science Foundation. And its just a research project, they haven't- the National Science Foundation hasn't put any efforts into finding out what's out there in terms of these types of tools and they wanted to have a stepping stone. And I guess what they plan on doing is having more projects along these lines in future years, but they needed some place to start. So what they want us to do is figure out what's out there, so they have an idea to help professors and institutions know that there are tools like this out there at their disposal. So that's basically where we're coming from.
Vicky: Are there faculty members that you're working with?
Dave: We're working with some Program Directors and we also have 2 faculty members from our school down here advising us.
Vicky: I see, what department are they in?
Dave: One is in the math department and the other one is in the chemistry department
Mike: What division of NSF are you working with?
Dave: The division of Undergraduate Education
Mike: Who's the program director there?
Dave: Peggie Weeks
Mike: Okay that's somebody I don't know
Dave: And Dr. Lee Zia
Mike: Oh okay, well Lee certainly knows a lot about WW
Dave: We have talked to him about this and other tools
Mike: When you're done with this, your report, would it be available to us?
Dave: Sure
Mike: We would like to receive a copy of it
Vicky: Are you folks acquainted with Judith Miller in the biology department at WPI
Val: Yes, I had her for a class freshmen year
Vicky: Because there are things that she's done that I think- her basic idea about some things to do with workshops, which correspond with some of the workshop ideas we're working with here at Rochester. I felt all along has had some applicability to incorporating WW or WW-like tool, as well. And I don't know whether Judy has thought about something like this. I just saw her at a meeting and wish I had time to sit down with her and talk to her. I would like to pull her in on the loop.
Vicky: Anything else that we should talk about?
Dave: That's all the questions that we have, I just want to thank you very much for you time, this morning
Vicky: Sure

Mike: Your welcome
Arnold: Well Thank You
Dave: We'll be sure to send you a copy of the project when it's all done
Mike: Okay, well good luck on it
Dave: Thank you
Vicky: bye-bye
Mike: Bye
Dave: Bye
Arnold: Bye

Interview with Edwin Koshy – Michigan State University
K – Edwin Koshy, Department of Physics and Astronomy, Michigan State University
V – David Vallerie
S – Valerie Sanders

V: OK, I was wondering for the first off, is it OK by our tape recording interview is for transcription purposes so that we don't miss anything.
K: No problem.
V: Alright, good, thanks you. I was wondering we can talk a little bit about CAPA.
K: OK, go ahead. What would you like to know.
V: Ah, first for, can you give us a brief explanation what it is? We know it is information based learning assessment tool, um, but we didn't find too much information about it, that's why I was wondering if you can go like an overview of it.
K: Well, listen, ah, I can, we have an publish paper and journal and several publication on it.
V: OK.
K: One is American journal of physics that was back in 1996, then that was journal of education,
V: OK
K: Also, we recently have a paper of frontier educational conference. Those kinds of system involved and a pretty good details, I would send you a copy of some of that issues.
V: OK, that will be great.
K: OK, but if you would like to know CAPA what is about, it is essentially a complete factor course management, which include homework, quizzes, examinations.
V: OK
K: Teachers which offer the personal life which meaning each student has its own version of it. It has formative assessment and summative assessment. Formative assessment which means you gonna test and learn it, summative basically means testing not time to teach. But the emphasis are the formative, so we have a lot of work involving and the INC have piles of questions which we have, as impact we hope we can teach the student's concept, instead of just plug into the formula then get number.
V: Alright.
K: Although we have those also tools as necessary.
V: Does it give the feedback to student or mean by base on what they put in the answers?
K: As example, the things like the formation of answers, if the answer doesn't have the right format, the guide will tell to get into it correctly, if you have wrong unit, it won't standardize but it will tell you it is incorrectly, then you need to fix it. If your answer is wrong, I won't tell you what is wrong with the answer, because that point is very important, if you did that it completely kill the idea of problem solving, because asking the problem solving is to find what did I do wrong, if somebody else tell you what did I do wrong, then you haven't solve anything. So as the purpose of other system, we guide the student toward after he had put the answer. At the end of the semester we all know the students can do it because they do it again until the answer went over. And that explain they can do it very well. So basically the system gave you the feedback of your answers, it also have 12 different kinds of problem questions, including subjective questions. Subjective question of course are like short essay, but those are coordinate by the factors and the system only help use one of the work that submitting by the system and convening processes and automatically to content and to student's file. And you can get feedback by emails and so on.
V: So is it CAPA generate the statistical information for the professors?
K: Absolutely. That's one of the main thing, it has complete statistical package that tell you broad and individual in details how does the class going. And that feedback is essential for both students and professors, because for the students have to see how they are doing, and what they need to work on. For the professors they can tell for the class in whole what particular problem what question, which many students are missing, which we need to explain the concept to them.
V: Along those lines have, have you done any evaluations of CAPA performances as an effective assessment tools for student?
K: Absolutely, we found out the paper of journal of education about the best award, we get the award of American society of education, I think the best of engineering education of 1999, I think it based on our assessment, on the fact we

able to set up the standard which were high, and student significantly battle, not because of the tool, but like how are we using the tools, so the tool itself can be good or bad. It just like a hammer, it can be good or bad, you can't see it, if a carpenter use it will be good, if we are using it, it may not be; It can have many many option of how we going to use the tool, for example, I can do it in quiz by electoral, electoral normally sometimes maybe once a week, or maybe once for every other week. I physically has two a week, it could not possibly do this by classified student, that system to handle all the work.
V: Have you found out any disadvantages to CAPA?
K: There is always a number of students, let's say 5-8% student, who really work very hard and still don't get it. For those students seem a negative thing in sense to work, these students got a way of little work and they don't like it. On the other hand, a number of students they work significantly more, they like it very much, typically the student evaluate the whole, it used to be they used according to university, if you ask them do they like it, it's only 50 and 50. If you ask them do they love, it's more like 80 and 20.
V: You mentioned that there are other disciplines using it now?
K: Sure, it can be using it in a big wave, physics, human ecology, food services, microbiology, chemistry, psychology, statistics, lots and lots of disciplines in lots of universities.
V: Lots of universities.
K: Indiana there are about 15 universities, in Canada about 8 or 9 ..
V: Are there any intensive start using for the staff to get to use the tools? Is it easy? Are there any tutorial stuff for them to setting it up or is it very easy ..
K: It has evolved over the years, it is much easy now then 2 years ago, because we work very hard for the improve the interface. For the student it will always be easy, first it for we work on to make that accessible for students, and then over the years, we work on the structure side to make it better and better, in these days we will gain the public license, do you familiar with Linux?
V: Yea
K: Well, Linux is the new general public license kind of software, meaning open source software, we have to charge license over $600-3000 for the software. Now we adopt it and open source free software. Of the public license I think it simulates lots of people to interest in it. Now we are also join another group of Michigan State, we also have the learning new system. New system is called "long capa". Basically it is a online learning network, with capa, it has all the capa ability that we have in capa and include many other features. It has ability to share materials like course platform and references.
V: How old is CAPA?
K: CAPA was first invented in 1992, way began with the browser, began with telnet, classes began to develop and switch to web. Back to several year ago we have both telnet and web, and later the student got much more into the web, this year is the first year that we essentially we didn't use telnet, up until last year we still have significantly use of telnet, but this year is full web.
V: Have you mentioned it you use it also for quizzes in your classroom? Do the students have the opportunities if they don't necessary go in to use CAPA to reinforce something they learn in class or study for exam like ahead of time, do they have the opportunities
K: Yes, they do what they can do is very easily, to have a set of homework problem which assign for though the year, you can have both problems available for the student, in the version which is different from the originally, so the student can work on, and because it is the same files. The system provide the different version depend on the seat that you chosen, so there is no problem on these factors, you can make lot lot of work available. In fact the people at home will use CAPA do successfully in time for them to prepare the examination. Basically putting very large amount of all materials for practices, and when we assign homework for example, if the student miss the first time, he doesn't get penalize and he gets the chances to correct himself with finite number of time, it is very useful because sometimes you may need time for get it all, so it doesn't penalize you. In the mean time, you also get a chance to learn it. A lot of thing maybe you miss the first time, should get 60%, second time will be 40%, so forth and so, but is very negatively because it means to expect the student get it in the first time, the student shouldn't be in your course, they should be taking more advance classes. We expect students to miss, to have a very high stand of the class. When they miss a question they will have a chance to correct it.
V: Let me change in here real quick, just take about assessment information technology learning assessment tools as a whole outside of CAPA, what do you feel to make an effective assessment tool?
K: First thing you have to make sure is assessing is to doing work, so you make sure the first person who is doing the work is the one who accessing. Second thing is, make sure the standard of assessment are high of the flag and the goal of the course. So we need that, the third thing is what we study we publish in journal

we have professor not involve in the study but responsible for the difficulties for the exams, maintain the difficulties over couple years, what have to be traditionally in that course, the course that I am talking about, which is the Calculus and physics based for Engineering major. So these are some of thing that maybe aspect for assessment, then you got to make sure the tool that you are assessing with, the content between materials, reasonable easy, medium hard, and hard. All of them should have a reasonable measurement. All very hard and the score are low, or all very easy and score are high, then you are not measuring that much. So we need some reasonable fact with examinations use for assessment.
V: Oh, that's all the questions that we have, thank you very much for your time.
K: Are you familiar with our homepages?
V: No, no we haven't got to there yet.
K: Yea, you should try to log in sometimes
V: Is it under Michigan State University?
K: You can go to google.com to search engine to find it as you like.
V: OK, great. Now we are reviewing the research project and part of it now we are trying to evaluate effective of assessment tools are. Is it possible that we can quote any of your assessment information today?
K: Yes, you can. Assessment has three components, basically we use student survey, OK, to see how hard they react, and establish the college students pretty good challenge learning, that's one aspect. Second assessment instructor survey, how do the student and instructor's aspect, use the system and how do the student performing. The third one we use is the examination performance, these three components are patience of our assessments. And all the people seems dramatically increasing exam scores, they are all possibly increasing types, especially class and student feedback, finally the students mentioned learning battle, by lose the majority say yes. It indicate the educational impact, anyone by himself may not be odd, but it combines it will be in a good case.
V: Great
K: And you can call me any time. Want me to send any of the papers so you can look that up by yourself.
V: Sure, that will be excellent.
K: I will send you the latest one.
V: Great, thank you very much, appreciate it. Can you send it electronically or do you want to send it as
K: I can send it as attachment as you mean, though email,
V: Do you have my email?
K: dvalliere@nsf.gov
V: Yea, that's it. Do you have any questions for us?
K: Well, I will expect the result come out from your study.
V: We will be done on our research on Dec19-20th, so we can send you a copy if you are interested in it.
K: You have my email address?
V: Yea.
K: Appreciate to getting your copies.
V: OK, thank you very much.
K: OK, bye.