

Education research using web-based assessment systems

Scott W. Bonham, Aaron Titus, Robert J. Beichner and Larry Martin
to be published in Journal of Research on Computing in Education, Fall 2000

Abstract

The growth of the Internet, and in particular the World Wide Web, is already influencing the way science is taught and will undoubtedly do so to greater extent in the future. One important facet of this is the development of web-based assessment and testing systems. These systems also provides a valuable new tool to the Education Research community; a tool that combines the ability of multiple-choice diagnostic tests to handle large numbers of subjects with some of the greater flexibility and additional information that other methods offer. On the other hand, some of the particular strengths of this tool also give rise to some unique disadvantages. In this paper the strengths and weaknesses for education research are discussed, and some suggestions for its use are presented.

Introduction

The explosive growth of the Internet is making available radical new means of communication that affect such diverse areas as business, entertainment and education. While older methods of accomplishing tasks continue to be used, the Internet offers unique advantages. In areas of education it offers a medium that has the potential to be more responsive to students, to encourage greater participation in their own learning, and to give greater access to different sources of information than traditional methods offer (Brooks, 1997; de La Beaujardiere et al., 1997; Khan, 1997). In this paper the focus is on the marriage of computerized testing systems with the World Wide Web to produce web-based assessment and testing systems (WATS). These systems are on the verge of becoming widely adopted at the university level in physics and other courses. A list of many of the WATS known to the authors may be found in the appendix. At least one of these systems, WebAssign, is also currently used in several dozen secondary schools. In the past, new types of technology, such as audio and video recording, have given rise to new methods of conducting education research. In the same way, the coming wide spread use of WATS will open up possibilities for educational research. Surprisingly, this area has not received much attention in the literature (Campbell, 1997).

The adoption of WATS for student work in physics is being driven in part by the promise of reduced grading load and the provision of more immediate feedback to students. More

comprehensive overviews of the use of such systems can be found elsewhere (Goldberg & Salari, 1997; Titus, Martin, & Beichner, 1998; Zhao, 1998). The various systems listed in the appendix differ greatly among themselves, ranging from interactive course material that an individual professor has created for his students to comprehensive systems incorporating databases, chat rooms, and multimedia. Some of them are systems developed by groups within universities for in-house use, while other university-developed systems are available for use by others. Commercial offerings come from some companies where the system is their main product, and from other companies—in particular textbook publishers—where the system is provided as a service to their customers. This paper focuses on comprehensive systems in which password-authenticated homework web pages are delivered and graded by a central server.

The system developed and used at North Carolina State University (NCSU) is described at this point as an example of a comprehensive WATS (NCSU, 1998). WebAssign uses a Sybase database in which homework questions are stored, assignments are organized, and grades are recorded. Class rolls and student information are downloaded from the University's registration and records database. The majority of the questions it contains are standard problems from various physics books—currently eleven different texts are represented in the database. The database includes survey questions and questions from nationally normed physics diagnostic tests (Engelhardt, 1997;

Halloun & Hestenes, 1996; Redish, Saul, & Steinberg, 1998). WebAssign is also being used for courses in computer science, mathematics, business and statistics. Questions present students with buttons for multiple choice (both exclusive and non-exclusive format), text boxes for numerical answers and fill-in-the-blank questions, and/or text area boxes for free-format (essay) questions. The instructor constructs individual assignments by selecting from the available questions, setting starting and ending times for the assignment availability, and setting grading and feedback options. The students log on and are authenticated by the same username and password as their campus computer account. They receive a page generated by Perl script, which lists the currently available assignments and a list of the past-due assignments for which they may now view the answer key. Upon selecting an assignment, the system generates an HTML document containing the questions. An example assignment is shown in Figure 1. The students enter their answers in the form's boxes or select among the buttons and submit the page for grading. The multiple choice and numerical answers are automatically graded and students quickly receive another page with the incorrect answers marked wrong. The answers are in boxes where they can be changed and—if permitted by the instructor—may be resubmitted an unlimited number of times until the assignment deadline. The free response questions are collected for later grading by the instructor or other designated persons. Students may view their current homework grades at their convenience. The instructor may view students' overall scores as well as individual responses. They may also perform administrative tasks for their class such as adding students to the roster and granting individual students extensions on homework assignments. Question formatting is in HTML, so pictures, audio, video, Java applets, and links to other sites may be included. Another feature of the system is the ability to use randomized numbers so that each student receives a different set of numbers for a given problem. In Figure 1 the problem has randomized numbers for the velocity and time, as can be seen by the lighter color of these numbers (on the screen the randomized numbers appear in red). In the physics courses at NCSU, students receive short assignments each class day.

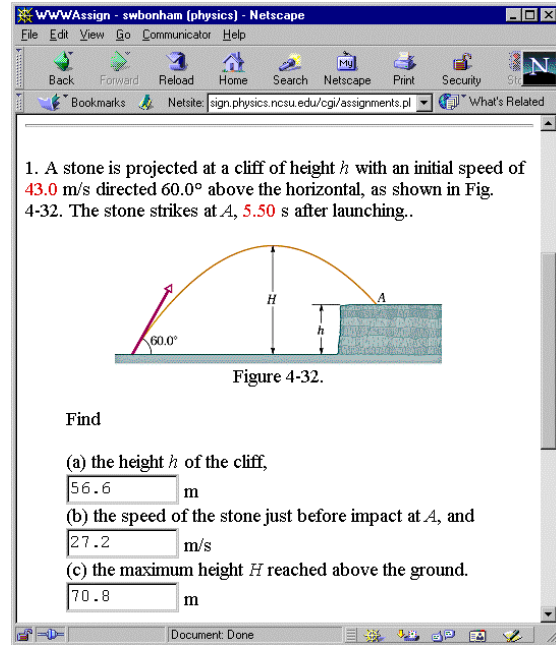


Figure 1: A sample assignment page from WebAssign. This problem is from (Halliday, Resnick, & Walker, 1997), Copyright © 1997, John Wiley & Sons, Inc. Reprinted by permission.

Education research via WATS may be conducted with different degrees of intervention in the learning process. The approach with least intervention is to look at data generated by students completing regular assignments. Essentially, this is passive observation with a WATS. In one sense, this is a limited type of observation; the researcher will only have information about students' interactions with the WATS. However, the computer log files will have a complete, accurate record of this type of information for all students and all assignments, so from another perspective this is a very comprehensive method of observation. A second approach of conducting research is through the inclusion of special exercises in the assignments, which is a limited type of intervention, since this will appear to the subjects as part of the regular work. These could include items from standard diagnostic tests or other questions to probe students' knowledge, and could be either given as separate assignments or embedded in the regular course work. A third approach is through the use of surveys and self-reported data, which represents the greatest level of intervention by the researcher. For example, after working a problem the student could be asked to report in a text box how he/she went

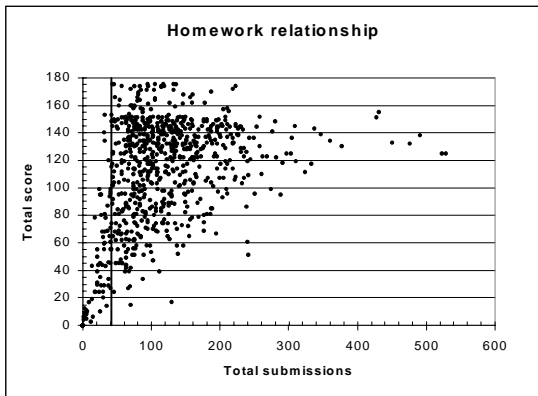


Figure 2: Scatter plot of total homework scores versus total number of submissions for students in the first semester engineering physics course. The vertical line represents (approximately) the minimum number of submissions to have attempted every assignment.

about solving it. The flexibility of the Web environment means that there are many combinations and permutations of these means that could conceivably be used for research.

As an example, Figure 2 presents a result from an investigation via the first research method. This is a scatter plot of student homework scores vs. the number of submissions for the entire semester of the introductory engineering physics course. Each correct answer was worth one point. The vertical line represents the minimum number of submissions to have attempted every assignment. (This number—40—is an approximation; there was some variation in number of assignments between sections.) Dots to the right of this represent students who made enough submissions to have attempted every assignment. The relationship is somewhat surprising; the main effect of increased submissions is elimination of the low-scoring students. Most students, on average,

attempt an assignment two to three times. Significant numbers of submissions beyond this has very little affect on homework scores beyond reducing the numbers of students with low scores.

An investigation via the second approach at NCSU has investigated the influence of multimedia on students' success at solving physics problems. "Video-enhanced" problems were created by supplying a video clip depicting the motion described in a traditional physics problem with the exercise. The WATS was used both to deliver these problems and to gather data on students' performance on the video-enhanced problems compared to the same problems that only had a still picture. Including video with traditional problems significantly affected students' performance on some types of problems but not others (Titus, 1998). This work was followed up by the creation of "multimedia-focused" problems in which information necessary to solve the problem is embedded within a Java applet animation (Christian, 1996-1998; Christian & Titus, 1998). Using the animation, students had to make relevant measurements in order to solve the problem, incorporating elements of a laboratory experiment. The WATS was used to automatically deliver the multimedia-focused problems to students randomly selected to be the treatment group and traditional problems to the control group. As a follow up to this study, students were asked to complete an on-line survey about their experiences with the multimedia focused problems—an example of the third method described above (Titus, 1998).

WATS have the potential to bring significant changes to courses in which they are used. With their power to collect and process data, computers and networks are used to mediate

	Internet setting	Computerized data collection
Pros	flexibility in time and location already used in the class being studied tightly integrated into a regular course	flexibility in how information may be presented the ways in which students may respond individualization of the material every detail of the subject's interactions recorded decreased contact volume of data
Cons	greater variation of the research environment tightly integrated into a regular course complications due to the technology	every detail of the subject's interactions recorded decreased contact volume of data

Table 1: Summary of implications for the use of WATS as an education research tool from its use of the Internet and of computerized data collection.

instructor-student interaction in a very flexible electronic environment (Khan, 1997). These same aspects also provide a new tool for education research, both for evaluating this new method of teaching and for broader issues. The suitability of WATS for education research arises naturally from the ability to individually tailor material and responses for each student and its ability to keep track of every activity of that student. In the following sections many of the issues related to the use of web homework systems for education research are discussed. These issues can be separated into two broad categories. First discussed will be how use of the Internet can significantly change the setting in which the subjects work, the type of work done, and related implications. Second is a discussion of the data that can—and cannot—be obtained by this method, the type of data and its quantity. Following this, the relationship between this new tool and other existing research tools is briefly discussed and suggestions about using this tool are made.

The settings of the work

A significant advantage for education research that WATS offer is the removal of many restrictions on studies. There is much greater flexibility of when and where the subjects may participate, accompanied by a corresponding loss of control of some elements of the study. In general, the students will already be using the system for their class work and the investigation will be integrated with this to a certain extent. However, the heavy reliance on technology also introduces potential complications.

A major characteristic of anything done via the World Wide Web is *flexibility in time and location*. Information can be sent and retrieved from any location in the world at any time of night or day as long as there is access to the Internet. What this means for the education researcher is that his or her subjects can, in principle, work in any location and at any time that is the most convenient. The investigator and the subjects need not be in the same room or even on the same continent. The minimum requirements for a subject to participate in a given location—a computer with Internet access—have become widely available (NSTA, 1998), which reduces greatly the extent and cost

of carrying out large studies in multiple locations. The subjects also have the possibility of working in a familiar setting—such as at home—as opposed to an unfamiliar laboratory. As can be seen, this tool offers a valuable method of looking at realistic learning situations.

Of course this flexibility comes with a cost. The ability of the subjects to work anywhere and anytime necessarily introduces a *greater variation of the research environment*. Some subjects will be alone and others may collaborate with fellow students. Even if instructions are given about this, they might be difficult to enforce. Some may be in a quiet place while others may be surrounded by many distractions. Some may work on it late at night when tired while others log in first thing in the morning. Some students will be very diligent about the work, while others may play at it between browsing through web ‘zines and the latest Dilbert cartoons. If this is too great of a problem, it can be addressed in part by restricting when and where subjects may work. Subjects could be required to work in certain locations, perhaps by restricting the main server to accept requests only from certain IP addresses, such as a specified computer laboratory. To further increase control, these designated locations could be proctored to prevent collaboration and to keep people on-task. Security could be addressed by posting a photo of the student logged in at the top of the screen, which would make it easy to detect cheaters. The server could also restrict the times at which it will accept submissions. As may be seen, WATS offer different levels of flexibility versus control for educational research.

An important aspect of using a WATS is that usually it would be *already used in the class being studied*. Under most circumstances, the investment to set up a WATS and for subjects to learn it would be impractical were it only used for education research. However, once the system has been introduced for class use and the students have become familiar with it, the additional cost to use it for research is small. There is little additional work required for the infrastructure, the database, or training the subjects. Almost all of the work would be in planning, developing the materials and analyzing results, which are the heart of any research project.

One way to recruit subjects would be to ask for volunteers to carry out certain tasks and then reward them in some way for their participation. This could easily be done using web-based evaluation. Much of this process could be automated, should this be desirable. These extra assignments could appear either with the regular homework or separate from it, and could be given with extra instructions and restrictions as needed. However, one of the unique features of using web evaluation as a research tool is the ability to be *tightly integrated into a regular course*. In this way, all of the students in a course, by virtue of doing work for that course, would be participating in the study. This is attractive for several reasons. First, this can provide a large number of subjects, as WATS are particularly attractive for large-enrollment courses. Second, this would provide a less biased sample than the use of volunteers is likely to provide (Campbell & Stanley, 1963; Gay, 1980). Third, students will not be very aware of being studied and observed. This means that the study will happen in a very realistic (albeit high-tech) setting where there is nothing to remind them that the computer at which they are doing their homework is doubling as an observation tool. Even if they are told at the beginning, they will tend to forget. This, of course, raises ethical issues of coerced participation and the rights of the subjects to know. This will necessarily require some sort of oversight to ensure that the rights of the students are not violated and will place restrictions on the design of studies. One restriction will be how control groups are set up so that no students are unfairly penalized. For example, in the study with multimedia-focused problems described above, the treatment group and control groups were reversed every other week, ensuring that all students had equal opportunity (or frustration) of working on multimedia-focused problems. Another restriction on investigations will be that all the material involved is appropriate in both content and difficulty level for the students' goals of learning the material and performing successfully in the course. For example, care must be exercised when including standardized diagnostic exams as part of course material, as many are designed to have average scores around 50%, far lower than typical classroom exams.

The use of computerized systems as the setting for educational research also introduces the potential for *complications due to the technology*. This is particularly true when the investigator and subject have little direct contact—the normal case for WATS. One reason is the potential for bias related to the technical abilities of the subject or the investigator. A subject that is not as knowledgeable or as comfortable is more likely to be intimidated or confused by the task at hand than a more knowledgeable subject is. In most cases, the investigator will not be present to clear up the difficulties, and may never know to what extent the subject's response represents their knowledge of the subject and to what extent their technical aptitude. A poorly designed system or web page could aggravate this. Another potential problem due to the heavy reliance on technology is the effect of technical problems. Again, this is particularly true of WWW systems, where no one has full responsibility for all the technology used. Problems could include: crashes of the main server, crashes of the student's computer, bugs in the program or the assignment, incompatibilities with browsers and Internet service providers (especially when using advanced features like Java), and limited access to the Internet by subjects. All of these issues can negatively affect the validity of the study. Knowledgeable planning and judicious use of advanced features can go a long way to minimizing these problems, but these issues must be taken into account when interpreting the results, particularly if control and treatment procedures use different features of the WATS.

The data that can be collected

In WATS, both the instructor and the students interact directly with the computerized web system. This becomes a powerful filter as to the nature of the interactions and the information available for the researcher. A WATS is very flexible in the content that can be delivered. Computers can easily collect and store every detail of the interaction. However, the nature of human-computer interaction restricts the types of information that can easily be collected.

One of the significant characteristics of the World Wide Web is its ability to be used in a variety of ways. Examples can be seen that range from personal self-expression to

commercial transactions and from public announcements to entertainment. The web context of WATS means that there is great *flexibility in how information may be presented* to students and in the ways they are asked to respond. Nearly anything that can be sent in digital format can be delivered to the students. This information could include text, images, video, sound, animations and simulations. Information could come from the instructor, fellow students, or other third-party sites. This has several potentially major benefits both for instruction and education research. First is the ability to address a greater variety of learning styles (e.g. visual-oriented vs. text-oriented). Second is the “real world” aspect of the problems presented. Third is that a greater level of participation can be required of students. Students could be presented with a video or an animation from which they must obtain information in order to be able to solve the problem. Students could be given a problem in which they must set up and run a simulation in order to determine the answer. Students could be given a problem requiring the use of information from several different web sources in its solution. An example is the multimedia-focused problems discussed above, which combines some aspects of traditional homework problems with laboratory work. This could serve as an alternative or supplement to traditional laboratories, especially when there are distance, time or budgetary constraints on the latter.

Likewise, WATS have, at least in principle, few restrictions on *the ways in which students may respond*. As described above, the current system can present questions in a variety of forms. Students can select among several options, either in an exclusive multiple-choice format or in a non-exclusive checkbox format. These need not be in the traditional form—selecting a region of an image map also works as a multiple-choice problem but in a much more natural format. Students can fill in boxes either with numerical answers that are checked against an internal answer or with a word that is checked against one or more correct answers. These are all automatically graded. Free response (essay) questions can also be asked, which in some cases could be graded automatically, or they could be collected and then analyzed by the instructor or another person. Although currently somewhat

limited as far as automatic grading, submissions in the form of graphical drawings also have great potential. Finally, the web system could accept submission of computer files or code fragments.

The use of the web and custom-generation of pages by Common Gateway Interface protocol (CGI) scripts also allows *individualization of the material* presented to each student. This is currently used in randomizing some of the numerical values students receive. Custom generation of pages can present students with different material depending on their class, group, or any other classification—including treatment and control group membership. Students may be presented with different material or problems depending on other information, such as past history. This is exploited in adaptive testing, where the subjects receive a series of questions, and each one depends on the responses to previous ones (Trentin, 1997). In a well-designed format, this can probe students’ understanding and provide the researcher with many of the benefits of an interview.

One of the greatest benefits for education research is that, by the very nature of the computer-mediated interaction, *every detail of the subject’s interactions* with the system can be recorded. Unlike observations, interviews, written samples and most every other research method, the primary data from web-based systems is already in digital form where it can easily be searched and processed. A log of every subject’s activity on the system—requests for pages and submissions for grading—can be easily maintained. Furthermore, client-side applications such as Java applets can be written so as to send information back to the server about the subject’s every action. In this way, a large amount of detailed information from every subject’s interaction with the system can easily be obtained and processed.

The use of computer-mediated interaction *decreases the contact* between the investigator and the subjects. This has a positive aspect as this reduces the chance of investigator bias and helps to ensure that all students are treated alike. On the down side, it greatly reduces the collection of some types of information. In such a system, the primary data from a given subject is in the form of lists of requests, question id numbers, submission numbers, etc. This is not a

very intuitive form in which to have data. It requires a certain amount of electronic manipulation to even begin to sift out useful information. This form also captures only one facet of the subject's activity—that with the computer. Thus, it would be difficult to capture a comprehensive picture of the learning process by means of a WATS. This tool does not lend itself very well to a qualitative study and often will need to be supplemented with interviews or other means in order to interpret the data. The removal of direct interaction between investigator and subject also removes many channels that might alert the investigator to problems during the study, such as body language. Thus, it may be discovered only after the data is collected—or worse, never realized—that design flaws or technical problems have compromised the study's validity.

Another characteristic that sets this approach apart from many other research tools is the sheer *volume of data* that can easily be collected and processed. This follows from many of the characteristics described above. The use of the web removes many of the geographical and temporal constraints on who could be used as a subject. Generally this tool would piggyback on a system used for classroom teaching, so with little additional overhead cost and one could easily obtain large numbers of subjects through the class enrollments. A large variety of work can be presented to the students, and their every interaction with the system can be recorded. Since this interaction is entirely in digital format and computer-mediated, data collection requires little additional effort to go from ten subjects to a thousand. For the same reason, it is fairly easy to include surveys and other types of feedback along with the problem solving. Being already in digital form, it is also easy to supplement this data with information from other databases, such as the university records department, or track students through all the classes in which they use the WATS. Education research through WATS faces few logistical and economical constraints for carrying out large, detailed studies. This is usually a desirable quality, as larger sample sizes tend to increase reliability of results

The sheer volume of data that can easily be collected of course also has potentially negative aspects. The ease in which WATS generates data means that the focus of research must be as much on the selection and screening of

information as on the acquisition of it. The fact that the investigator can quickly assign additional material means that it would be easy to overlook its cost in terms of workload on the subjects. This could result in annoying the subjects (or worse) and thereby damaging the value of the study. Unfortunately, this is also a type of problem that the investigator may be less likely to notice during the study. Also, any time databases contain detailed information about many people there are ethical issues of privacy and right-to-know, which only become more important as the size increases. There may even be political implications. Will a university continue supporting a research effort that shows substantial differences in performance in class work between gender or racial groups? Finally, the sheer size of the data collected can cause significant problems in interpretation. Most of the data will be uninteresting lists of requests and submissions, and so the researcher must find ways to sort through this and glean that which is interesting. Even so, one is likely to wind up with a large aggregate of facts about each of the subjects, while at the same time the data itself is not very helpful in giving the researcher a good over-all picture of what is going on. This would aggravate the problem of dealing with large amounts of data and could wind up wasting students' time. The volumes of data that can be collected by this tool can be a great boon for many research projects, but by the same token require good understanding and careful planning by the investigator to avoid the pitfalls of dealing with this quantity of data.

Suggestions about the use of WATS as a research tool

WATS offer some unique advantages for use as an education research tool, but they also present a number of significant drawbacks. All of these need to be taken into consideration when considering its appropriateness for a given study and throughout the planning process. The strengths of this tool relate largely to its flexibility, ease of data collection and the closeness it can approach (some types) of real course settings. Closely connected to the flexibility and computational power are the major weakness of this tool: the researcher's reduced control of the experiment, reduced contact with the subjects, and reduced channels

of information. WATS are a tool that has great strengths and significant limitations. Both should be carefully considered in the design of studies employing it.

Here the authors suggest several types of research studies that could benefit greatly from this tool.

- **Large, quantitative studies.** WATS are better suited than any other research tool for many types of large-scale studies. This includes traditional multiple-choice (MC) exams and questionnaires. The latter can easily be delivered via WATS, which can also accept and score answers that are beyond the ability of the MC format (Zandyliet & Farragher, 1997). Since the person directly enters the data, there is no need for optical readers or other data entry methods. A well-designed screen layout is more intuitive than an optical scanning sheet, which could reduce errors in data entry. Already computers are replacing pencil-and-paper exams (Bugbee, 1996) and the use of the Web will further facilitate this. Just as with MC forms, Web questions can be administered either in a proctored environment or at the convenience of the subjects. The one significant drawback is the technology—having the computers for the subjects to work on. However, if the technological infrastructure is already in place for other purposes, this could be even cheaper than paper MC tests.
- **Widely dispersed subjects.** Again, the web is in many cases superior to the other methods of dealing with a geographically separated population, i.e. mailing forms or conducting telephone surveys. It is faster, provides more consistent information and—provided the technological infrastructure is already in place—is cheaper than other means. It would be particularly useful in evaluating distance-learning students using the web; already there are many courses being offered in this medium (CASO, 1998; Collis, 1997; Donahue, 1996; Faulhaber, 1996; Leonard, 1996). Other kinds of studies could also benefit. For example, a proposed diagnostic test could be given to subjects from around the country. A third use would be undertaking a study that involves a narrowly defined but physically scattered group (e.g. professional enhancement resources available

to African-American middle-school science teachers located in rural communities).

- **Monitoring of courses.** For courses that already use Web-based evaluation systems for homework, WATS are the easiest, fastest and cheapest means of monitoring students' progress and getting feedback. Scores, patterns of use, and level of participation are already recorded in a format that can be analyzed and linked to other data. Patterns of use could be tracked over time and the range of scores on individual exercises is easily viewed. Some of this monitoring should take place in courses using WATS to compensate for the loss of contact and direct feedback between instructors and students. Computer scripts could be set up to automatically analyze student answers, looking for patterns that indicate conceptual or other types of difficulties. The system could then alert the instructor and/or student of this pattern, or even automatically direct the student to additional resources.
- **Evaluating the use of WATS.** WATS are very new and are a different approach to the way in which students do homework. The implications of this for their learning is not entirely understood, and is an area that needs to be researched. All WATS courses modify some aspects of the pedagogical process. For example, students using WATS usually receive more immediate feedback on their homework but only their final answers are evaluated, which is a trade-off compared to handing in written work. Although the relative merits have been debated, no studies have been carried out to find out under what circumstances this is a net gain for the students. The flexibility of question presentation and answer collection presents a variety of ways curricula can be changed, leading to similar to those raised above. The WATS themselves will prove to be an important tool in gathering this sort of information, though not necessarily the only one.
- **Evaluating the use of multimedia.** In one sense, this is merely an extension of the above point, but it has even greater potential impact on instruction than the automatic grading. The ability of the Web to deliver diverse information has the potential to completely

change the appearance and use of homework. A great deal of time and effort will be expended on this, but only some of the efforts will really improve students' learning—and some may actually interfere with it. Much guidance needs to be given to these efforts. An important component will be how real students use multimedia in real classes. There is no other tool that can match Web systems for unobtrusively monitoring large numbers of students as they do their work in real situations. This also has potential beyond just evaluating multimedia use in WATS. Java applets written to send back information to the server about how they are being manipulated could be used on many different types of pages. A public site like that of a science museum could have its use monitored by a similar scheme as described here. While other research tools will also be of value, WATS will be an essential element of understanding how students interact in multimedia problem-solving environments.

- **Complex studies.** Sometimes it would be desirable for the course of the treatment or the evaluation to depend on a large number of factors. One may wish to give a student different instruction based on a combination of performance on a pretest and a self-described learning style. Adaptive testing, where the subject receives different questions depending on answers to earlier ones, may be a desirable tool. Computer adaptive testing can validly duplicate paper and pencil testing if it is well designed (Bugbee, 1996; Trentin, 1997; Zandvliet & Farragher, 1997). These type of studies can be easily and flawlessly administered by a well designed WATS, even with the large numbers of subjects that would be required to obtain statistical results from such complex studies. Furthermore, all the information about exactly what sequence that each subject went through would be recorded as part of the log file.
- **Complementing other tools.** As has been suggested above, the particular strengths and limitations of research using WATS are very different from some of the other education research tools. For example, in comparison to collecting student web responses, a video-taped interview about problem solving is far more invasive, more geographically

constrained, harder to do with large numbers and more challenging to analyze. On the other hand, it provides a much richer, personal picture of what is going on than a web system could ever provide. In this way, web system studies and studies with more qualitative tools, such as interviews and observations, can provide strong complements in developing a comprehensive picture of student learning.

This is by no means intended to be a comprehensive list, but rather to provide an idea of the types of education research projects which could benefit from the use of WATS as a research tool. In general, these projects are those where flexibility is a desirable quality, where computational power is important to deal with large or complex studies, and/or where the technology itself is a major focus of the study.

Practical Suggestions

A few practical suggestions on implementing research via WATS are offered here. This is certainly not a complete guide to carrying out such studies but rather represents issues that the authors have encountered up to this point. For the most part, these can be summed up as “plan ahead and be realistic about how the system works”.

- **Know how the students use the system.** WATS allow students to interact with the system in a variety of ways, some of them unexpected. A poll given in the Fall semester of 1997 at NCSU found that over half of the students frequently printed out assignments, presumably to work them out on paper and then enter them into the computer at a later time. By informal or formal means, the researcher needs to obtain a realistic picture of how the students actually go about doing the work.
- **Plan ahead.** Any good research project needs to be carefully planned so that it will achieve its goal, and research via WATS is no exception. What is different is the ease with which one can plunge into research activity without making careful plans, and the mountain of unintelligible data that can be effortlessly collected without a clear picture of the goals.
- **Maintain a good relationship with the system administrator.** Unless the

researcher also happens to be the system administrator, these two people will often have different backgrounds, outlooks and goals. The level of cooperation provided by the system administrator will affect the usefulness of the WATS for research. Since in most cases, the research will be piggybacking on a system originally designed for other purposes, the researcher must make every effort to not be just another demand on the (often over-worked) system administrator.

- **Judicious use of advanced features.** Unless the subjects will always be using designated computers, the HTML for given pages or problems needs to be written with older browsers in mind. For example, Applets written in Java 1.1 will run only on Explorer 4.0 or higher, and at the time of writing there is no Macintosh browser that can run them. Compatibility across platforms also must be considered; fonts used may need to be limited to those available on all platforms (Macintosh, PC and UNIX machines). Similar consideration may need to be given to avoiding the inclusion of large amounts of graphics, in deference to students with slow modems.
- **Minimize interference with the on-going classes.** This is particularly relevant when the researcher is not the instructor. The goal of the instructor and students is their education, and the research must not hinder their progress. One issue that has arisen in the use of WebAssign at NCSU has been confusion among students over whether surveys posted on the web count towards

the total homework score—even after a statement to the contrary was posted at the top of every survey.

- **Be realistic about authentication.** One can never be 100% sure that a given submission came from the student whose name is on it. Students may work together, or might give their password to someone else. This must be taken into consideration; just as with pencil-and-paper work, there is the possibility of copying or submitting work in another person's name.

In conclusion, web-based assessment and testing systems offer the education researcher a new and, in some aspects, unique tool. It offers tremendous advantages including flexibility of setting and format, the ability to collect large amounts of data, and tying in closely with courses. By the same token, computer mediated-communication also reduces the investigator's control of the experimental setting and direct contact with the subjects. This tool can provide some types of studies with great benefits. Careful planning can minimize many of the significant pitfalls of this tool and monitoring that recognizes the unique aspects of WATS. Like any other product of technology, WATS offer education research great benefits when properly understood and wisely used.

Acknowledgments: The authors would like to thank the NCSU Department of Physics for its support for developing a WATS and allowing its use for research purposes, and John Risley and Peg Gjertsen for their work in turning it into a reality. One of the authors (SWB) would like to thank the National Science Foundation for its support of this work through Postdoctoral Fellowship DGE-9714546.

- Brooks, D. W. (1997). *Web-teaching: a guide to designing interactive teaching for the World Wide Web*. New York: Plenum Press.
- Bugbee, A. C. (1996). The Equivalence of Paper-and-Pencil and Computer-Based Testing. *Journal of Research on Computing in Education*, 28(3), 282-299.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally College Publishing Company.
- Campbell, T. (1997). Technology, Multimedia, and Qualitative Research in Education. *Journal of Research on Computing in Education*, 30(2), 122-129.
- CASO. (1998). *Internet University*, [WWW]. Available: <http://www.caso.com/>.
- Christian, W. (1996-1998). *WebPhysics*, [World Wide Web]. Davidson College. Available: <http://WebPhysics.davidson.edu/>.
- Christian, W., & Titus, A. (1998). Developing Web-based Curricula Using Java Physlets. *Computers in Physics*, 12(3), 227-232.
- Collis, B. (1997). Pedagogical Reengineering: a Pedagogical approach to course enrichment and redesign with the WWW. *Educational Technology Review, Autumn/Winter1997*(8), 11-15.
- de La Beaujardiere, J. F., Cavallo, J., Hasler, A. F., Mitchell, H., O'Handley, C., Shiri, R., & White, R. (1997). The GLOBE Visualization Project: Using WWW in the Classroom. *Journal of Science Education and Technology*, 6(1), 15-22.
- Donahue, S. (1996). Programs and Resources in Distance Education. *Journal of the American Society for Information Science*, 47(11), 870-874.
- Engelhardt, P. (1997). *Examining students' understanding of electrical circuits through multiple-choice testing and interviews*. Unpublished Ph.D. dissertation, North Carolina State University.
- Faulhaber, C. B. (1996). Distance Learning and Digital Libraries: Two Sides of a Single Coin. *Journal of the American Society for Information Science*, 47(11), 854-856.
- Gay, L. R. (1980). *Educational Evaluation & Measurement*. Columbus: Charles E. Merrill Publishing Company.
- Goldberg, M. W., & Salari, S. (1997). *An Update on WebCT (World-Wide-Web Course Tools) - a Tool for the Creation of Sophisticated Web-Based Learning Environments*. Available: <http://homebrew1.cs.ubc.ca/webct/papers/nauweb/full-paper.html>.
- Halliday, D., Resnick, R., & Walker, J. (1997). *Fundamentals of Physics*. (5 ed.). New York: John Wiley & Sons.
- Halloun, I., & Hestenes, D. (1996). *Views About Sciences Survey: VASS*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, St. Louis, MO,.
- Hecht, E. (1998). *Physics: Algebra/Trig*. (2 ed.). Pacific Grove: Brooks/Cole Publishing Company.
- Khan, B. H. (Ed.). (1997). *Web-based instruction*. Englewood Cliffs, N.J.: Educational Technology Publications.
- Leonard, D. C. (1996). Using the Web for Graduate Courses in Technical Communication with Distant Learners. *Technical Communication: Journal of the Society for Technical Communication*, 43(4), 388-401.
- NCSU. (1998). *WebAssign*, [World Wide Web]. Available: <http://wwwassign.physics.ncsu.edu/demo/>.
- NSTA. (1998, May/June). Survey Shows K-12 Internet Use Is Up. *NSTA Reports!*, pp. 18.

- Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics courses. *American Journal of Physics*, 66(3), 212-224.
- Serway, R. A., & Faughn, J. S. (1995). *College Physics*. (4 ed.). Fort Worth: Saunders College Publishing.
- Titus, A. (1998). *Investigation of the Effectiveness of Animation on Solving Physics Problems*. Unpublished Ph.D. dissertation, North Carolina State University.
- Titus, A., Martin, L., & Beichner, R. (1998). Web-based Testing in Physics Education: Methods and Opportunities. *Computers in Physics*, 12(2), 117-123.
- Trentin, G. (1997). Computerized Adaptive Tests and Formative Assessment. *Journal of Educational Multimedia and Hypermedia*, 6(2), 201-220.
- Zandvliet, D., & Farragher, P. (1997). A Comparison of Computer-Administered and Written Tests. *Journal of Research on Computing in Education*, 29(4), 423-438.
- Zhao, Y. (1998). Design for Adoption: The Development of an Integrated Web-Based Education Environment. *Journal of Research on Computing in Education*, 30(3), 307-328.
- Zitzewitz, P., Davids, M., Neff, R., & Wedding, K. (1992). *Merril Physics: Principles & Problems*. (1992 ed.). Westervill, Ohio: Glencoe/McGraw-Hill.

Appendix: A partial list of Web-based assessment and testing systems

System	Developer	Web address (http://)
CAPA	Michigan State University	capa2.nsl.msui.edu/homepage/
Companion Websites	Prentice Hall	www.prenhall.com/pubguide/
Course Management System	John Wiley & Sons	www.wiley.com/college/cms/
CyberProf	University of Illinois (Center for Complex Systems Research)	www.howhy.com/home/
Gateway Testing System	University of Nebraska -Lincoln. (Dept. of Mathematics and Statistics)	www.math.unl.edu/webtests/
homework service	University of Texas-Austin (Department of Physics)	hw.ph.utexas.edu/overview.html
interactive problems	University of Illinois (Department of Physics)	webbug.physics.uiuc.edu/courses/ie.html
Mallard	University of Illinois (Department of Electrical and Computer Engineering)	www.cen.uiuc.edu/Mallard/
MicroTest	Chariot Software Group	www.chariot.com
practice quizzes	Addison Wesley/Benjamin Cummings	occ.awlonline.com/bookbind/pubbooks/campbell_awl/
Question Mark	Question Mark Computing	www.qmark.com
TopClass	WBT Systems	www.wbtssystem.com/
Virtual-U	Simon Fraser University	virtual-u.cs.sfu.ca/vuweb/
web evaluation and feedback	University of North Carolina-Charlotte (Department of Physics)	http://www.physics.uncc.edu/cgi-bin/CGI/CGI4edu.html
web quizzes	McGraw-Hill	www.mhhe.com
web-based homework	University of North Carolina-Ashville (Department of Physics)	www.ctl.unca.edu/bennett/Phys222/default.asp
Web@ssessor	Computer PREP	www.webassessor.com
WebAssign	North Carolina State University (Department of Physics)	wwwassign.net/info/
WebCT	WebCT	http://www.webct.com/
WWWAssign	North Park University	www.northpark.edu/~martin/WWWAssign