

Web-based Testing in Physics Education: Methods and Opportunities

Aaron Titus
Larry Martin
Robert Beichner

Introduction

The World Wide Web is impacting education in profound ways. So far, the predominant use of the WWW in teaching has been for finding and distributing information, much like an on-line library. However, as information technology evolves, the Web is increasingly used for more interactive applications like testing. But as with any technological development or media revolution, there is potential for both success and failure. To use the WWW effectively for testing in physics education, we must identify our goals and determine the best way to meet those goals. This paper describes methods of web-based testing, highlights the positive attributes of the WWW for testing, and marks the dangers that threaten its effectiveness.

Methods of Web-based Testing

Testing primarily serves to motivate students to spend time on task. Students will spend that time if it contributes both to their understanding (internal motivation) and to their grade (external motivation). As a result, testing is used for two essential purposes: to provide *feedback* and *evaluation*. Feedback refers to the response regarding a critical analysis of students' work. Evaluation refers to the grading and recording of students' work for the purpose of assessing their understanding of material. These two purposes are not mutually exclusive. A testing instrument, whether it is a homework assignment, quiz, exam, or practice test, can satisfy both purposes to a varying degree. For instance, a practice test is primarily used to provide feedback to students for their self-evaluation. On the other extreme, an exam is primarily used for evaluation.

Evaluation and feedback have different goals and thus have different implementation requirements. Because evaluation is primarily used to record student responses and assign grades, security concerns such as verifying a student's identity, protecting answer keys, limiting access according to a specific time or location, and preventing unauthorized sharing of information need to be considered. The second use, feedback, is used to respond to students' input by providing "correct/incorrect" responses, hints, and solutions or by engaging the student in additional learning activities much like present computerized tutorials. Feedback routines may even be intelligent, where the computer learns the best response for a given input. In most cases feedback and evaluation are usually combined in order to both calculate and record students' progress and provide expanded feedback regarding students' answers.

Modes of Questions

Questions in web-based assessment are typically of four modes: multiple choice, fill-in-the-blank, numerical response, and free-response (e.g. essay). Except for free-response questions, all can be automatically graded [add endnote here about intelligent grading of free-response questions]. However, technology is not necessarily limited to these modes. Questions may use Java applets, video, or imagemaps to directly pass values to client or server grading routines. For example, in Figure 1 the coordinates of a mouse click on an imagemap are recorded and evaluated. (In one sense, the imagemap presents discrete multiple choices to the user---each individual pixel on the screen represents a choice.) This paper, however, will focus on multiple choice, fill-in-the-blank, numerical, and free-response questions.

Figure 2 shows a multiple choice question that uses HTML radio buttons to indicate a student's answer. This question makes use of the "exclusive" nature of radio buttons which disallows the student error of marking more than one answer. As shown in Figure 3, HTML checkboxes may be used for questions with multiple correct answers.

Fill-in-the-blank (Figure 4) and numerical (Figure 5) questions are similar in the way answers are submitted; however, they differ in the way answers are graded. They both use an HTML text input field where the user types an answer. To grade a fill-in-the-blank response, the grading script compares the input, character by character, to the correct answer or a list of possible correct answers. Numerical responses on the other hand are evaluated within a range or tolerance so that small rounding errors within the user's calculation do not produce an incorrect answer. JavaScript can be used to control the data type submitted and thus not allow text when only a number is expected (e.g., to disallow entry of units where units are already specified).

Free-response questions use an HTML text area for students to submit responses. Answers are typically written to a text file or database to be scored later by the instructor. Figure 6 illustrates an essay question on the WWW.

Evaluating Questions

Responses can be scored in three ways, client-side, server-side, or a combination of the two. Which method is best depends on the goal of the assessment. A summary of techniques and recommended uses are listed in Table 1

Method	Exemplary Use
JavaScript	Practice tests Interactive tutorials Front-end processing of responses
CGI and flat text files	Small-scale assessment
CGI and relational database	Large-scale assessment

Table 1: Summary of methods for Web-based assessment

For client-side evaluation, questions and answers are downloaded to the client and scoring is accomplished on the client's computer. Advantages of client-side evaluation are decreased load on the server and immediate feedback. Because answers are downloaded to the client, security questions inevitably arise. As a result, this method is best for an interactive tutorial or practice test. Users' answers and grades can be uploaded after the user has completed all questions, thereby using client-side scripting for scoring students' input and providing feedback and server-side scripting for recording students' grades.

Client-side assessment is accomplished with JavaScript, the scripting language developed by Netscape. JavaScript code is contained in the source of the document and therefore is downloaded with the Web page. Simplistic JavaScript code for scoring questions in Figures 2-5 are shown in Boxes 1-4.

```
<FORM>
Two carts on a horizontal, frictionless track
make a head-on collision. A cart of mass 3M and
velocity 20 cm/s to the right collides
elastically with a cart of mass 4M which is at
rest. After the collision, the cart of mass 3M
<P>
<DL>
```

```

<DT>
<DD><input type="radio"
onClick="alert('Correct')"> moves to the left
<DD><input type="radio"
onClick="alert('Incorrect')"> moves to the right
<DD><input type="radio"
onClick="alert('Incorrect')"> is stationary
</DL>
</form>

```

Box 1: JavaScript code to score the multiple choice question in Figure 2.

```

<FORM>
A collision occurs between two bodies,
m<sub>1</sub>=100 kg and m<sub>2</sub>=1 kg.
Which of the following are true?
<DL>
<DT>
<DD><input type="checkbox"
onClick="alert('Incorrect')"> The magnitude of
the force of m<sub>1</sub> on m<sub>2</sub> is
greater than the magnitude of the force of
m<sub>2</sub> on m<sub>1</sub>
<DD>
<input type="checkbox"
onClick="alert('Correct')"> The magnitude of the
force of m<sub>1</sub> on m<sub>2</sub> is equal
to the magnitude of the force of m<sub>2</sub>
on m<sub>1</sub>
<DD>
<input type="checkbox"
onClick="alert('Incorrect')"> If the collision
is elastic, the kinetic energy of m<sub>1</sub>
<i>cannot</i> change
<DD>
<input type="checkbox"
onClick="alert('Incorrect')"> If the collision
is perfectly inelastic, the total kinetic energy
is conserved
<DD>
<input type="checkbox"
onClick="alert('Correct')"> The velocity of
m<sub>1</sub> after the collision will be
similar to what it was before the collision
<DD>
<input type="checkbox"
onClick="alert('Incorrect')"> If m<sub>2</sub>
is shattered into many pieces due to the
collision, momentum is not be conserved
</DL>
</form>

```

Box 2: JavaScript code to score the multiple choice question in Figure 3.

```

<FORM name=formname>
What type of energy is conserved in an elastic
collision?

```

```



```

Box 3: JavaScript code to score the fill-in-the blank question in Figure 4.

```

<FORM NAME=formname>
Two carts on a horizontal, frictionless track
make a head-on collision. A cart of mass 3M and
velocity 20 cm/s to the right collides with a
cart of mass M which is at rest. The carts
stick together upon collision. Their velocity
after collision is


```

Box 4: JavaScript code to score the numerical question in Figure 5.

Recording students' responses and scores requires a server-side script. The Web browser passes responses to the server where they are processed by a script or compiled program called a CGI (Common Gateway Interface). The CGI can score responses, record desired information, and return feedback to the student; consequently, it can be used for either evaluation, feedback, or a combination of the two.

Server-side scripts are often written in Perl, an interpreted language specifically designed for text processing that is available on all major platforms. However, other languages, such as C, C++, and Java, and platform dependent languages, such as Applescript and Frontier, can also be used. An example (written by LM in Perl) which you may download is at <http://www.northpark.edu/martin/WWWAssign/>.

Server-side scripts often retrieve questions and record responses in flat text files; however, for large enrollment courses, this can be cumbersome. Assigning multiple assignments per week to many students generates large amounts of data. At NCSU, we deliver and grade approximately 3,000 homework assignments per week. To store efficiently a large number of questions, homework assignments, student responses, and grades, we interfaced a CGI script with an SQL (Structured Query Language) database which acts as the backend, or data source. SQL is a database standard for searching and managing information in relational databases[5] and is thus a simple yet powerful tool for large-scale, Web-based assessment.

One advantage of using a database backend is that questions can be stored and retrieved by keyword, subject area, author, and other fields, thus allowing quick generation of assignments. In addition, all assignments, grade reports, and database queries are dynamically generated; that is, no text files are used. Therefore, multiple professors and courses can use the same assessment application without worrying about file structure, non-unique filenames, etc. It seems to be the best solution for department-wide or institution-wide use.

Regardless of the method one uses, the computer (either server or client) can easily handle random selection of questions and numbering of problems, as well as randomization of question content when appropriate. For instance, given a suitable range, WWWAssign will generate random numbers within a question and then calculate the answer based on those numbers (see Box 5). Other functions randomly select an item from a list or select the corresponding item from a different list as shown in Box 6.

(a) Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass $\langle \text{eqn } \$m1=\text{randnum}(1,5,1)\rangle M$ and velocity $\langle \text{eqn } \$v1=\text{randnum}(10,30,10)\rangle$ cm/s to the right collides and sticks to a cart of mass $\langle \text{eqn } \$m2=\text{randnum}(1,5,1)\rangle M$ which is initially at rest. After the collision, the velocity of the carts is

Ans. $\langle \text{eqn } \$v=\$m1*\$v1/(\$m1+\$m2)\rangle$ cm/s

(b) Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass 2M and velocity 30 cm/s to the right collides and sticks to a cart of mass 4M which is initially at rest. After the collision, the velocity of the carts is

Ans. 10 cm/s

Box 5: In (a) is the code which generates the question shown in (b). The $\langle \text{eqn} \rangle$ tag is interpreted by WWWAssign which evaluates the randnum function and returns the result. The answer is calculated based on the numbers generated in the question.

(a) Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass $\langle \text{eqn } \text{pickone}("2","4")\rangle M$ and velocity 10 cm/s to the right collides elastically with a cart which is initially at rest. After the collision, the first cart travels $\langle \text{eqn } \text{picksame}("to the left","to the right")\rangle$ at 5 cm/s and the second cart travels to the right at 2 cm/s. What is the mass of the second cart?

Ans. $\langle \text{eqn } \text{picksame}("15","10")\rangle$ kg

(b) Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass 2M and velocity 10 cm/s to the right collides elastically with a cart which is initially at rest. After the collision, the first cart travels to the left at 5 cm/s and the second cart travels to the right at 2 cm/s. What is the mass of the second cart?

Ans. 15 kg

Box 6: In (a) is the code which generates the question shown in (b). The $\langle \text{eqn} \rangle$ tag is interpreted by WWWAssign which evaluates the pickone and picksame functions and returns the result. The answer is also specified with the picksame function.

Positive Attributes of the WWW for Assessment

Pedagogical improvements. The traditional method of teaching---giving a standard lecture,

requiring weekly or biweekly homework consisting of a few end-of-chapter problems, and then testing students with a few exams per semester---has developed out of the available pedagogic media. The new medium of the internet offers a tool for teaching which makes additional teaching strategies possible. Using automated submission and scoring of assignments, instructors can give students more frequent assignments and more questions on each assignment than is possible with traditional methods, thus increasing the time that students spend studying material, answering questions, and solving problems. In fact, the computer can control the path through the assignment if desired, making the better prepared student's progress more efficient while choosing a more gradual approach for the less able student. Customized hints based on a student's response can target known misconceptions. In this way, the content of feedback becomes just as important as the content of questions.

In addition, the traditional method of collecting homework on a weekly basis (or longer) can now be done daily or after every class. This motivates students to study continuously throughout the week rather than procrastinating until homework is due. Homework is a tremendous motivator for students, and providing assessment on a frequent basis keeps students more focused on material presented in class.

A study by Edward Thomas[3] reports a correlation between repeated execution of "skills" problems and student performance on traditional physics word problems presented on class exams. The group of students benefiting most by the "Precision Teaching" method were those who daily accessed questions by computer. The computer graded their answers and measured completion time. Using Web-based assessment, similar questioning techniques can easily be administered to a large number of students. Even when assignments are more for self-evaluation and practice than for formal testing, records can be made of time on task and correlated with success in the course. Reports of this information to future students can make using the Web-based assignments more palatable.

Web-based assessment may also impact the lecture itself. Student responses can be used during class discussion or recitation for peer-reviewing. Students can then re-evaluate their own answers and strengthen their knowledge before a final evaluation is made by an instructor or grading assistant. LM has used a technique of averaging multiple peer-evaluations of a set of student writings to derive an evaluative score without the necessity of an instructor evaluation.

Decreased administrative effort. Increasing students' time on task and continually administering homework is made possible by automated grading. The server takes over the mundane grading of papers which is often the obstacle to assigning more frequent homework or quizzes. It saves grading time for the instructor (or hired graders) and improves the quality of class time spent in problem solving recitations. Automated scoring also allows the instructor to easily perform item analysis to determine which questions are best predictors of student performance. Instructors can then tailor their assignments, whether homework, quizzes, or exams, to include questions that are best for probing student understanding.

Multimedia-enhanced questions. The Web's capabilities allow questions that include video, animation, simulation, or audio. But merely adding multimedia to traditional problems is not necessarily the best use of technology[4]. Multimedia-enhanced questions should be written differently than traditional questions in order to take advantage of the new medium. Unlike traditional questions, a multimedia-enhanced question does not have to use solely text and static graphics to describe the situation. Rather, it can use motion and sound to illustrate pertinent information. Students can even be encouraged to make measurements directly from the multimedia, thus making the assessment a more engaging process. This type of problem where data necessary to solve the problem are embedded in the multimedia is more appropriately called a "multimedia-focused" problem.

At NCSU, multimedia-focused problems have been developed using Physlets, single-purpose scriptable Java applets, written by Wolfgang Christian (<http://WebPhysics.davidson.edu/>). Physlets can be embedded directly into HTML documents and can interact with the user by employing JavaScript. One example, Animator, displays an object (or multiple objects) according to an equation of motion $[x(t), y(t)]$. It includes control buttons to play, pause, or step through the animation and allows the student to measure the position of any point at any time.

Immediate feedback. With computer-aided assessment, students can receive immediate feedback about their progress. Interactive tutorials and practice questions engage students to test their understanding. At present, students may practice solving textbook problems and look up answers in the back of the book, yet receive no insight into why they miss a question. But on the Web, students can answer questions, gain insight into their errors, and be immediately propelled into curricular material that stresses important points. Surveys given to students indicate that immediate feedback is one of the most appreciated aspects of web-based assessment. For this reason, it is valuable to deliver assignments with immediate feedback and assignments with delayed feedback. The former type encourages students to thoughtfully consider why they missed certain questions and the latter ensures that students fully consider their answers before submitting. The combination of these two types of assignments seem to enforce in students' minds the value of doing homework.

Negative attributes of the WWW for assessment

Inability to see students work. When grading, it is useful for physics instructors to view students' work, check their diagrams, and follow their reasoning. Unfortunately this is not currently possible on the WWW. As grading becomes automated, the instructor may lose some perception into the problem solving processes of students. This may be the largest deterrent to Web-based assessment when used for evaluation. However, for large-enrollment classes where exams are automatically scored and homework is graded by undergraduate or graduate students, this aspect already occurs.

Using a sampling technique, proper problem solving technique can still be encouraged without overloading graders. Students can be required to submit their answers on the WWW and hand in their work on paper. The instructor can randomly select one problem and grade it based on method of solution. The overall assignment grade is then determined by the number correct on the web-based part and the method of solution on one selected question. Thus, students will be motivated to use good problem solving procedures, yet benefit from the advantages on-line assessment provides. Meanwhile, the instructor's time grading papers is decreased..

Technocentric approach to courseware development. A significant concern in the advent of such a powerful communication medium is that instructors and developers will exhibit a technocentric approach to instructional design, thus making decisions based on the capability of technology, not the needs of the learner[2]. It is the responsibility of researchers to help determine how the WWW should be used for assessment, and it's the responsibility of instructors to use the WWW in appropriate ways. **Decisions to use the WWW for assessment must be based on proper learning objectives and not on technology.**

Undetermined effectiveness. The WWW's popularity in education for information delivery, interactive simulations, and assessment has grown; however, little research has been devoted to learning how to use the technology effectively. Researchers should design studies to discover the most effective methods to achieve our pedagogical goals. Instructors must have an open mind and be willing to adjust preferences and ideas as they discover what does or does not work with on-line assessment.

Less variety of questions and grading methods. Automating the grading process eliminates certain types of questioning and grading. However, there are new types of questions that can be delivered on-line that cannot be delivered on a piece of paper.

Security issues. If the WWW is used for evaluation, security issues inevitably arise. How can the instructor be assured that a student is doing his or her own work? In the case of homework, the grade is not heavily weighted, students are usually expected to work together, and password verification may be enough to deter less-motivated cheaters. As a result, security concerns are not so great. And there can be benefits from encouraging students to collaborate.

For exams or quizzes security is a vital issue since these elements often count for a large part of a course grade. In this case, there are sufficient ways of insuring security. One such way is to use randomization. Numbers within questions or the order of choices in multiple choice

questions can be randomized. In addition, techniques from traditional assessment methods in distance education, such as requiring proctors to oversee exams, could similarly be applied to Web-based assessment. Hopefully, the electronic form of the answers may make detecting instances of possible cheating an easier task and thus dissuade students from making the attempt.

Another method of security is to restrict access to questions based on the user's IP address, in which case, students must take quizzes or exams from a certain computer lab, building, or campus. The lab can be monitored, printing can be disabled, and a proctor can check student ID cards upon entry. Students' pictures can even be displayed in the header on the exam so that a proctor can easily identify cheaters. Security in Web-based assessment may not be perfect, yet may be better than what is currently used.

Technological difficulties. When using the WWW for assessment, instructors must realize that technical problems will occur. For instance, the server may be down or server software may need to be updated. Perhaps the instructor wants to use browser specific features or platform dependent plugins which all students are not able to install or access properly. There will always be some technical difficulties. Instructors must plan to handle these difficulties, and it must be clear to students what they should do in the inevitable event of a technical problem. Typically it is sufficient to extend the assignment deadline for individual students who have such trouble.

Beyond technical problems, some students just do not have computer experience and may be anxious about using computers. Fortunately, the Web offers an intuitive interface so that less experienced users can quickly learn to navigate hypertext and submit HTML forms.

Students also need access to the internet. At this time, it is unreasonable to assume that every student owns a computer and has internet access from home. Therefore, a school should have computer labs where students can access the internet on a regular basis.

Presentation difficulties. Although the WWW is improving in its graphic design capabilities, it is not easy to represent mathematical equations, and symbols within text; therefore, they are usually presented as images where each one is created individually and saved in an image format. This is not only a time consuming development process, but since each image is downloaded separately, it results in additional strain on the server. Users with slow modem connections are negatively affected. Therefore, it is best to format equations using HTML tags as much as possible. As an example, one can use subscript and superscript tags for writing equations. Similarly, greek letters are written using the "" HTML tag with the parameter "face=symbol". Also new server side java applets can interpret equations on the fly and generate images of the equations as the page is served to the client.

Students' Perception. Although assessment on the Web offers new opportunities to engage students in learning experiences, students may perceive it to minimize human interaction and ultimately replace the teacher[1]. This is certainly not the goal; nonetheless, some students resist using instructional technology. Instructors must build the trust of students with effective uses of assessment. Then students may come to see the time spent answering questions on the Web as useful for learning. With the increasing use of the WWW in our culture, it is likely that students will become very comfortable with Web-based tools in education.

Passing the test

Educational potential for Web-based assessment is almost unbounded. Unfortunately, it is both potentially wonderful and potentially disastrous for students, instructors, and courseware developers. We must discover effective uses of the WWW for assessment and use it accordingly. To do so, we should use the appropriate technology for the assessment task, be it evaluation or feedback. Simple client-side feedback is efficiently accomplished using JavaScript, whereas server-side evaluation and feedback is done with CGIs. For evaluation of large-enrollment classes, a database backend supplies the power required to manage the large amount of data.

Understanding the technology is not enough. We must understand the benefits of the technology, as well as its limitations. Web-based assessment can be effectively used for

providing unbiased evaluation and feedback on a frequent basis that is immediate, platform independent, multimedia-enhanced, and automated. But new technology brings new concerns. Initial instructional development may be technocentric, with technological capabilities, not learning needs, driving the process. Finally, if we are to depend on technology, we must be prepared when the technology fails. Recognizing these dangers is the first step toward utilizing the positive aspects of technology-assisted education and the appropriate technological methods necessary to achieve effective assessment in physics education.

References

- [1] G.D. Bothun and S.D. Kevan. Networked physics in undergraduate instruction. *Computers in Physics*, 10(4):318 - 325, 1996.
- [2] Lloyd P. Rieber. *Computers, Graphics, & Learning*. Brown & Benchmark, Dubuque, Iowa, 1994.
- [3] Edward W. Thomas, Jack Marr, and Neff Walker. Enhancement of intuitive reasoning through precision teaching and simulations. In *Proceedings of the Conference on Frontiers in Education*, pages 3c3.10 - 3c3.13, 1995.
- [4] Aaron P. Titus. Present PhD research.
- [5] John K. Whetzel. Integrating the world wide web and database technology. *AT&T Technical Journal*, pages 38-46, March/April 1996.

List of Figures

Click on the point(s) on the velocity vs. time graph where the described object reverses direction (this is usually called a turning point).

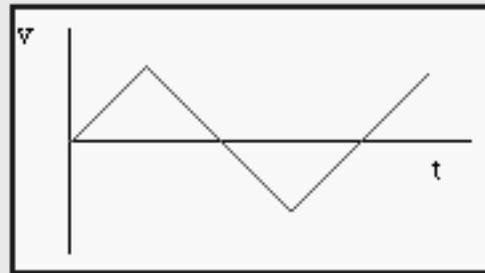


Fig. 1: Imagemap question

Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass $3M$ and velocity 20 cm/s to the right collides elastically with a cart of mass $4M$ which is at rest. After the collision, the cart of mass $3M$

- moves to the left
- moves to the right
- is stationary

Fig. 2: Multiple choice question.

A collision occurs between two bodies, $m_1=100\text{ kg}$ and $m_2=1\text{ kg}$. Which of the following are true?

- The magnitude of the force of m_1 on m_2 is greater than the magnitude of the force of m_2 on m_1 .
- The magnitude of the force of m_1 on m_2 is equal to the magnitude of the force of m_2 on m_1 .
- If the collision is elastic, the kinetic energy of m_1 cannot change.
- If the collision is perfectly inelastic, the total kinetic energy is conserved.
- The velocity of m_1 after the collision will be similar to what it was before the collision.
- If m_2 is shattered into many pieces due to the collision, momentum is not be conserved.

Fig. 3: Multiple choice question with multiple right answers.

What type of energy is conserved in an elastic collision?

kinetic

Fig. 4: Fill-in-the-blank question.

Two carts on a horizontal, frictionless track make a head-on collision. A cart of mass $3M$ and velocity 20 cm/s to the right collides with a cart of mass M which is at rest. The carts stick together upon collision. Their velocity after collision is cm/s

Fig. 5: Numerical question.

In an inelastic collision, kinetic energy is not conserved. What is a possible explanation for the loss of kinetic energy?

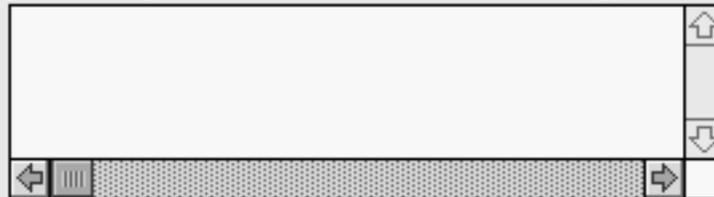


Fig. 6: Free-response question.