

## **Utilizing Usability Testing Methods to Improve OERL**

Judith Fusco, Jeff Huang, and Alexandra Harris  
Center for Technology in Learning  
SRI International

Paper presented at the annual meeting of the American Educational Research Association, April, 2005, Montréal, Canada. Send correspondence to Geneva Haertel, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025. This research was supported by contracts from the National Science Foundation (REC-9912172 and NSB-0353574).

## **Utilizing Usability Testing Methods to Improve OERL**

Usability testing is a well-developed tool for examining Web sites (Rubin, 1994; Schneiderman, 1998), and one that we believe should be utilized by evaluators during evaluations of Web-based resources. A poorly designed site has the potential to prevent visitors from completing tasks, to keep them from finding what they need, and to frustrate them to the point of leaving. A good user interface is at the core of a site's effective implementation; without it, the outcomes of the project with which the Web site is associated cannot be achieved. Throughout the development of the OERL Web site, we have employed usability testing and incorporated many different methods, from standard laboratory usability tests to more in-depth studies of the user experience. We have also conducted user tests of novice and experienced evaluators, NSF principal investigators, and graduate evaluation faculty and their students. From the feedback and results we have refined the navigability, functionality, design, and usability of the site. In this paper, we present the methods employed in our usability tests in an evaluative context, including the profiles of users tested, the iterative testing cycles, topics that were investigated, results from our cycles of testing, and improvements that were made to the site.

### **Utilizing Usability Testing Methods in an Evaluative Context**

What does it mean to evaluate a Web site? Does it mean evaluating the impact the Web site has on its users? Does it mean evaluating the accuracy, quality, readability, and verifiability of the site content? Or does it mean evaluating how people interact with the Web site? While the first meanings seem obvious, and are often addressed by evaluators, the third is usually left to usability engineers rather than traditional evaluators. Jakob Nielsen, a pioneer in the field of Web site design, has identified learnability, efficiency, memorability, error rate, and the user's level of satisfaction as the key elements of usability. These are all issues that evaluators frequently address in the evaluations they design. In fact, in this paper we argue that when an evaluator is charged with evaluating a Web site, an increasingly prevalent occurrence, the evaluator should be concerned with the site's interface and usability, in addition to its impact and content. A usable interface is at the core of a site's effective implementation; without it, the goals of the project with

which the Web site is associated cannot be achieved. A poor, unusable interface will frustrate users, hinder them from finding what they need, cause them to be dissatisfied with the site, or even drive them away, in which case the Web site has little potential impact.

In this paper, we will first describe some of the methods that usability engineers typically employ when they examine a Web site. The methods used will be very familiar to evaluators and thus easily adoptable. Second, we will describe the usability methods used throughout the development and evaluation of the OERL site, and how the use of these methods has impacted the site. Third, we will report what we have learned from usability testing in the context of evaluating the OERL Web site.

### **Usability Methods Background**

Usability engineers typically examine a Web site through some form of user testing. User testing tends to be used as a formative tool, with a cycle of tests intended to expose weaknesses in the system's usability to help inform the design of the system. While usability testing methods have their roots in experimental psychology, they place much less emphasis on quantitative or statistical methods, since usability engineers are interested in users' reactions to the product. In this section of the paper, we will introduce several usability methods that are commonly used to evaluate Web sites at various points in their development. In the earliest phases, testing typically gives the design team information about how to improve the design. As the Web site begins to mature, testing can confirm the design choices, as well as provide feedback that will allow refinements to the site design. Later, usability testing can help determine if design objectives have been met and if the site is working as intended for the audience.

The most basic and common of user testing methods is what is typically called a "user test," which can be used in all phases of Web site development. In this protocol, a usability engineer observes users who are representative of the target audience interacting with the Web site in a computer laboratory setting. In most cases, the user is asked to go through a set of predetermined tasks and to think aloud as he or she performs them.

Thinking aloud indicates how a user is approaching or reacting to the system and can give qualitative data about the user's reactions, including problems interacting with the Web site. A disadvantage of this method is that it is a laboratory-based test; hence, the subject using the Web site is asked to perform decontextualized tasks. This may make the tasks harder or easier to do and may not be representative of a more organic user experience.

In addition to the think-aloud method commonly used in user tests, the usability engineer can follow a question-asking protocol. Instead of waiting for the users to respond themselves, the observer can ask specific questions to prompt for more specific types of responses. A variation of the think-aloud method is the co-discovery method, in which two participants work together. The two participants may have a conversation about the site that helps them to understand it together. For specific quantitative data about users during user tests, the observer can also use a method called performance measurement, which would include collecting information about mouse movement or eye tracking with specialized equipment.

Another group of commonly used usability methods involves a careful inspection of the system. Hiring a professional to do a heuristic evaluation, an evaluation performed by an expert using industry-accepted guidelines, can be a useful way of making sure the system follows basic usability principles. Alternatively, pluralistic walk-throughs can be conducted with individuals from several different disciplines. This method usually consists of a leader and the group of diverse individuals stepping through a task scenario and discussing how the individuals' might interact with the Web site as a group. A group process allows a diverse range of perspectives on usability issues to be exposed.

Two related usability methods, cognitive walkthroughs and contextual inquiry, are often used in the design phase for a Web site. These methods are meant to help the designers understand the audience and the context in which the system will be used. A cognitive walkthrough requires the construction of task scenarios from design specifications to make the users' goals and purpose for each task more explicit. In order to increase understanding of the context in which the system will be used, contextual

inquiry, involving structured interviewing of the target audience, can be employed. Field observations or ethnographic studies of the context are often used to complement design approaches. In addition, focus groups, surveys, and questionnaires are other commonly used methods for collection of data to build or refine a particular system.

All methods identified thus far are most likely familiar to evaluators. The greatest difference between typical evaluation methods and usability methods is that in the latter case the technology is being evaluated explicitly for the sake of usability. The evaluator who adopts usability methods will need to be sure that he or she has the technological skills to understand what a user is doing or attempting to do when observing the user. If the evaluator does not feel comfortable in this role, we recommend finding a technically minded colleague or graduate student to assist. In addition, we recommend books by Rubin (1994) and Schneiderman (1998), or the Web sites Useit.com, Jakob Nielsen's Web site (<http://www.useit.com/>), and the Usability Toolbox (<http://jthom.best.vwh.net/usability/>) as good reviews of usability methods.

### **Methods used in OERL's Evaluation**

Since OERL was created approximately five years ago, several usability methods have been used to test its design (e.g., cognitive walkthroughs and contextual inquiry). The focus of this paper is not to discuss the methods used in designing the site, but rather to document how the OERL team has utilized usability-testing methods in its formative evaluation of the OERL Web site after it was developed. We have employed both laboratory-based user testing methods and an approach similar to a heuristic evaluation, with a panel of evaluation experts reviewing the Web site. We will describe both of these projects and what we learned from them below. The usability evaluation is one component of the evaluation of OERL. (A second paper [Thurston, Fusco, Javitz, & Smith, 2003] describes the conduct of three surveys that are part of the overall evaluation of OERL, and a third paper describes the methods employed in the expert panel review and the complete findings [Zalles, Trevisan, & Haertel, 2003].)

### **Laboratory-Based User Tests**

*Participants.* Because OERL is a Web site specifically for evaluators, we conducted our user tests on evaluators. Our first user test involved a total of seven evaluators in two rounds of testing. Three evaluators participated in the first round. Based on the feedback from the first three evaluators regarding the Web site's usability, several changes were implemented. After the changes were made, a second round was conducted with four graduate students who were familiar with the evaluation field (novice evaluators are part of OERL's target audience). The evaluators were recruited from SRI International, and the graduate students were recruited from a local university. The graduate students were given a small honorarium and the researchers from SRI International were given an hour of billable time for their participation in the hour-long user test.

*Procedures.* The focus for these two rounds of testing was on how well users understood how the navigation of the Web site worked and how easy or difficult it was for them to complete typical tasks such as searching and browsing information. A script was developed and used with all seven subjects. It was modified slightly before the second round of user testing, so that the script questions accommodated the refinements made to the Web site after the first round of user testing (see Appendix A for the script).

Each participant first answered a series of background questions to determine their evaluation experience and familiarity with computers, and then was asked to do predetermined tasks that demonstrated the use of the site's primary features and resources. The participants were asked to think aloud as they used the site and performed the tasks that were designed to test how easy or difficult it was to navigate the Web site. When users did not spontaneously offer their thoughts, the observer prompted them to do so. The users' movements were tracked directly through video capture of the computer screen for the duration of the test. An audio recording was also made of the participants and the observer while they went through the script.

*Round 1 Findings.* The first round of user tests uncovered some major problems with the site navigation that needed to be addressed. The three participants in the first user test were in many ways "baffled" by how to navigate the Web site and had difficulties with several of the tasks. All completed the tasks, but with varying levels of assistance from

the usability tester. At the end of the test, all three participants indicated that they still were not confident about how the site worked (though one explored the site further and came up with an idea for improving it). Because the results from the first three participants clearly indicated major problems in navigating the site, the first round of user testing was discontinued and improvements based on their feedback were implemented before the second round of testing.

Figure 1 shows a screen snapshot of the home page of the OERL Web site when the first round of user tests began. The site had been developed and organized around three types of evaluation products typically produced as part of NSF-funded evaluations: plans, reports, and instruments. Three tabs representing these three evaluation products were placed along the top of the OERL Web site; when the tabs were clicked, the corresponding artifacts (i.e., reports, instruments, and plans) were displayed. Down the left side of the Web page was a second set of tabs that organized the site into project types (e.g., teacher education, technology, faculty development) and provided access to additional resources such as criteria, glossary, and FAQs. Above the upper tabs and to the left of the side tabs were buttons, each with an arrow, that would take the user to the overview of each section. The function of these arrow buttons was explained in a graphic on the home page, but the graphic was very subtle (the function was almost hidden), so that it was difficult to get to the overview page for each type of artifact. In addition, the target area for clicking on the arrow was fairly small and hard to hit; if the users missed the arrows, they would simply stay on the same page and become confused or concluded that the arrow buttons were decorative rather than functional.

Another confusing feature of the OERL Web site was the representation of the matrix of resources, with tabs acting as row and column headers. This matrix concept was confusing to users, since it was the buttons on the tabs that took users to the overview page of each Web site section, instead of an overview tab in the rows or column headers of the matrix. Thus, it appeared that the matrix was missing one row and one column for these overviews. In addition, when the user clicked one of the upper tabs to enter the plans, instruments, or reports section, that tab would become the left column tab, at the top of the gold side tabs, regardless of its previous position. This “jumping” or

rearranging of the tab order also contributed to the confusion users experienced when trying to use the Web site. See Figure 2 for an example of the tab movement. In Figure 1, the “Instruments” tab is the center tab. In Figure 2 it has become the left-hand tab at the top of the gold tabs on the left.

**Figure 1. The Home Page used in Round 1 of the Laboratory-Based User Testing**

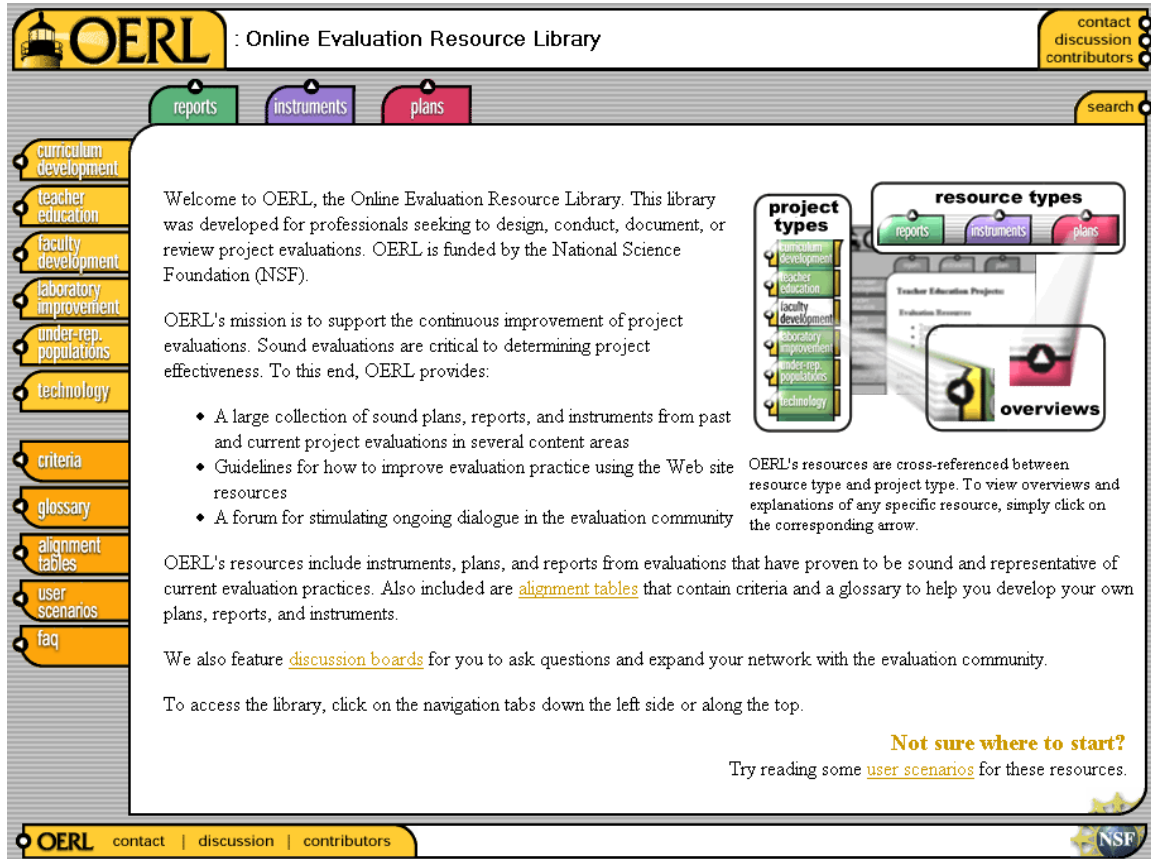
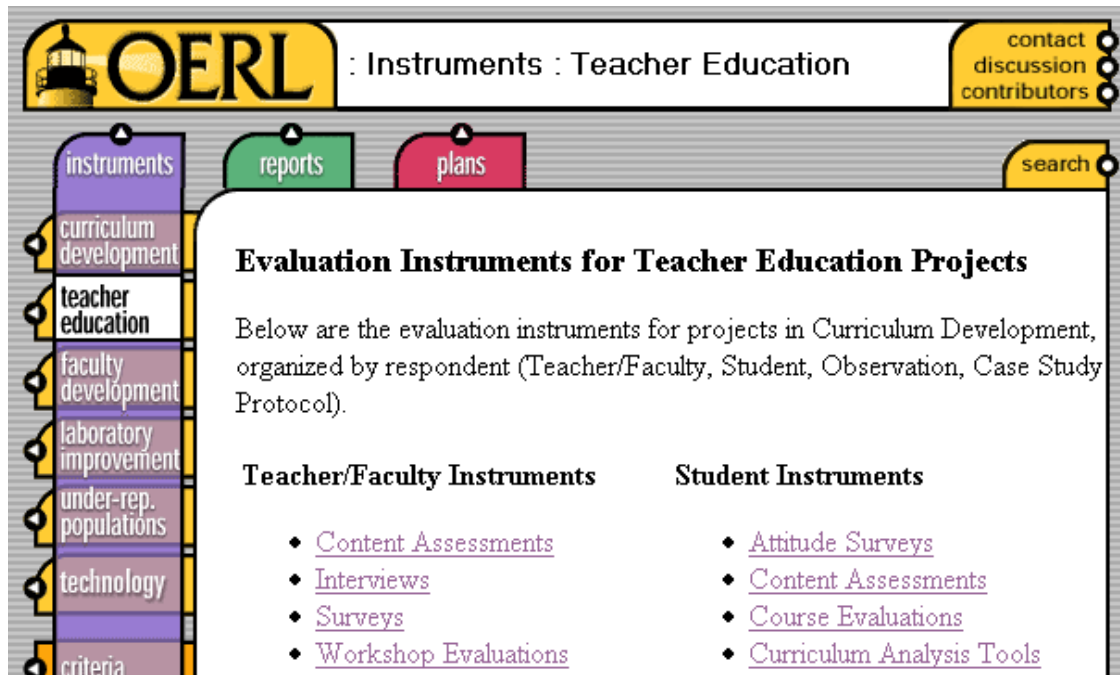




Figure 2. Figure Illustrating the Movement of Tabs



*Changes to the Site after the First Round of User Testing.* Since the site had an “incomplete” matrix, and users had significant difficulty finding the overview pages for both the tabs running across the top of the interface and those running down the left-hand side, we added a row and column to the matrix. Completing the matrix made it easier for people to understand how to get to the overview pages on the Web site. We did this by adding the OERL tab to the top as a column, and adding the “overview tab” on the side. Additionally, the arrow buttons were removed from the tabs, since the matrix now included overview tabs and the arrow buttons were no longer necessary. See Figure 3 to view the changes to the navigational aspects of the OERL home page. Other feedback from the first round of user testing resulted in additional changes:

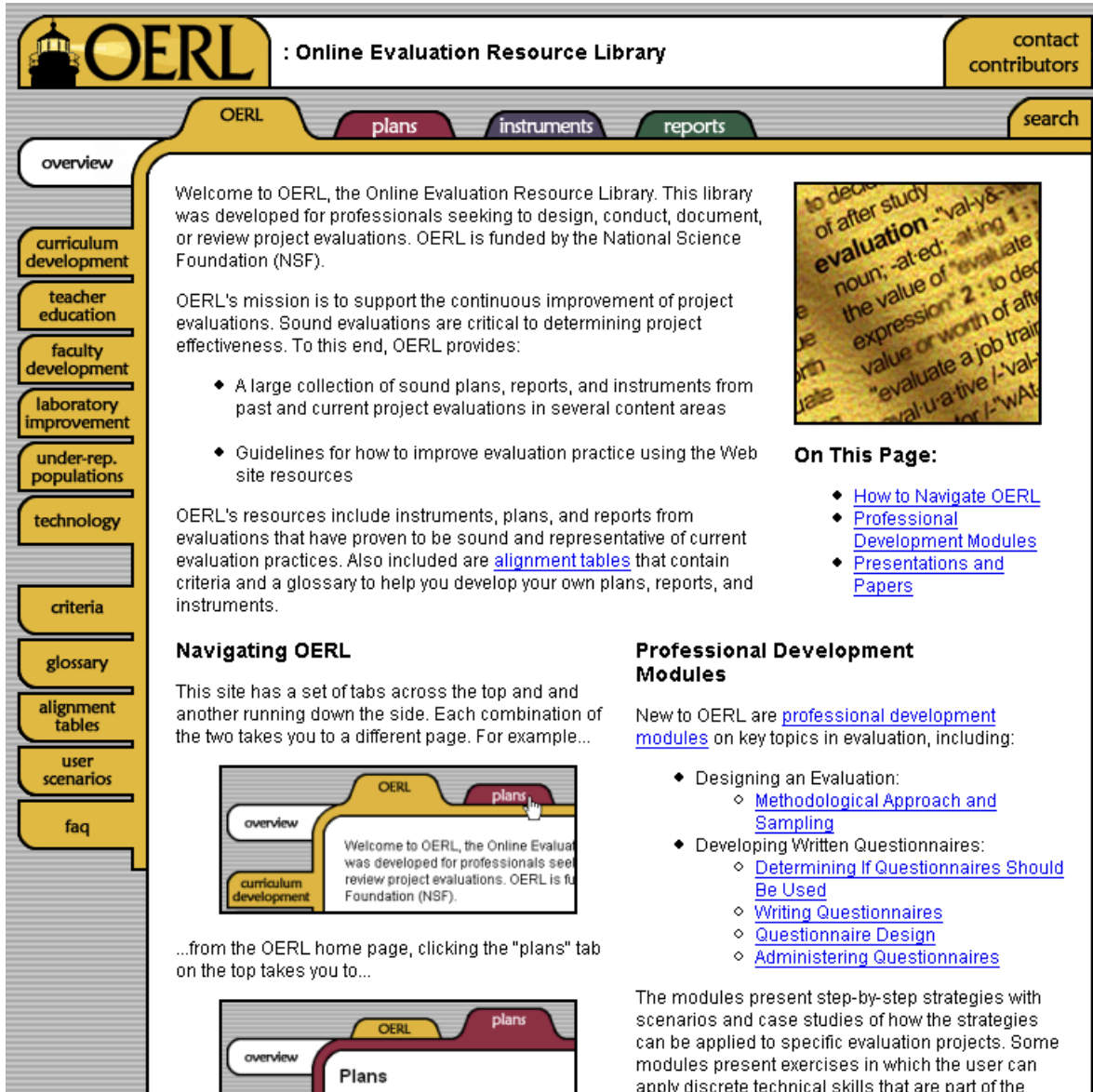
- The colors of the tabs were darkened to give them a more professional look and provide better contrast in color with the text. Earlier feedback indicated that the text in some tabs was difficult to read.
- The OERL additional resource tabs below the project types were made the same color as the project type tabs. The subtle color difference was not commented on by the users, but the Web team felt it contributed to a look of “busy-ness” on the

home page. The spaces between the two sets of tabs and the different shapes of the upper and lower tab sections distinguish the tabs from each other.

- Resource tabs now remain in the same place. Originally, when a tab was clicked, it jumped to become the leftmost tab. This confused users because their cursor arrows were no longer on top of the place where they had last clicked and the order of the tabs along the top had changed. Now, instead of jumping to the left, the tab is raised slightly to signify that it is active.
- Illustrations explaining navigation were included on the home page.
- A Google keyword search replaced the previous keyword search.
- The discussion forum was removed because of low user activity.

In hindsight, the problems users encountered with the original design of the OERL Web site seem as though they should have been obvious when the site was designed. However, Web site designers must make trade-offs and compromises during a design phase and some problems are not obvious because they are often designed to be solutions for other problems. For example, leaving the matrix incomplete was most likely an attempt to reduce the number of columns from which a user would have to choose from. Many first attempts at design are riddled with flaws, which underscores the importance of user testing after a Web site is developed.

Figure 3. The Home Page after Improvements



*Round 2 Findings.* The second round of user testing for the OERL Web site began in December 2002, after the first round of navigational improvements were made. Four graduate students from Stanford University were recruited. Results from the second round of the user-testing study showed that the changes to the site after Round 1 had improved the site's usability. The students were generally able to understand how the site was structured and could browse the plans, reports, and instruments of the site. Although the four participants were unsure of how the Web site worked when they began

the user test, after they had completed a few tasks each spontaneously indicated that he or she understood the layout of the Web site and commented that the Web site's organization was structured like a matrix. In the debriefing that followed the user test, all four participants indicated that the structure of the Web site was a good way to organize the evaluation artifacts (i.e., plans, reports, and instruments) by project type (e.g., teacher education, technology).

With the matrix issue resolved, other issues regarding the organization of information arose in this round of usability testing. First, the presentation of FAQs was confusing. At a minimum, they needed to be reorganized in a way that grouped similar questions together. The presentation of text on the OERL Web site is another topic that was raised. Users felt that some Web pages contained too much text or that the font was too small; users felt that on other pages some of the headings and subheadings needed to be changed to make them more distinctive. Perhaps the most significant usability issue that came to the surface in this round of testing was difficulty with the organization of the search interface—the users found the many choices on the search page to be overwhelming. This issue was explored further with the participants who were part of the heuristic evaluation of the OERL Web site and are referred to as the expert panel; their comments will be discussed below.

*Utility of the Laboratory-Based User Tests.* From laboratory-based user testing, we identified several issues that caused users difficulty, made changes to the Web site to address those issues, and then re-tested with another set of users to see whether the changes had solved the problems. If we had surveyed users of the OERL Web site to determine if they could successfully use the site, we might have learned that there were usability problems, but we most likely we would not have learned what was causing the problems. By bringing users into a laboratory and observing their reactions to the Web site's behavior as they explored the site and attempted to complete the tasks we asked them to do, we learned which specific aspects of the site were causing problems. While the user test was a laboratory-based test, it was valuable in that it helped us learn how to improve the site from a usability standpoint, particularly in regard to navigation.

## **Expert Panel**

A second approach we took to evaluate the usability of OERL was to have several evaluators participate in an expert panel, as in a heuristic evaluation. We did not ask the evaluators to evaluate the site according to guidelines for Web design, but to evaluate OERL's evaluation content and general presentation, based on their own heuristics as expert evaluators and professors of evaluation.

*Participants and Procedures.* The evaluators were all professors who taught graduate-level evaluation. Some had been principal investigators or evaluators of NSF-funded projects. While the main focus of this evaluative exercise was to review the site from the perspective of a professor of evaluation courses (Zalles, Trevisan, & Haertel, 2003), we did ask them to review other aspects of the site such as usability, search features, screen navigation cues (e.g., tabs and color coding), the readability of the text, and the responsiveness of the server. We also asked them to identify any problems they encountered.

The expert panel began their review in December 2002, after the first round of laboratory-based user testing and the resulting changes, and during the same time period as the second round of user testing. The expert panel's findings will be considered in two parts. The panel looked at the OERL Web site as a whole, and then looked at the professional development modules. The professional development modules were a new section of the Web site that the user test did not examine. We will first consider the feedback the expert panel gave on the Web site as a whole first and then discuss their feedback on the professional development modules.

*Findings on the General Web Site.* The panel confirmed what our subjects in the user test demonstrated: that the Web site was navigable and usable. Everyone on the expert panel understood how the site was laid out and had no complaints regarding the organization of the site in terms of its usability. These reviewers did present feedback suggestions on different ways to organize the site (e.g., to have a path through the site that would allow a novice evaluator to identify all artifacts, plans, reports and instruments from a single evaluation).

The largest usability issue that came from the expert panel's feedback regarded the search interface. Like the users in the laboratory-based user test, the expert panel reported that the search function was difficult to use. Originally, the user was given a screen that required significant scrolling and offered up to seven search parameters that the user could use to refine the search, several of which had sub-menus. Figure 4 (parts 1 and 2) shows the original search interface. A simpler search interface was implemented that offers the user fewer search parameters and is fully visible without scrolling on most screen sizes (see Figure 5). The longer, more configurable original search page with all the search parameters is still available under the "Advanced Search" link on the default search page.

Figure 4 (Part 1). The Original Search Interface

### Search for Instruments

Specify the features you're looking for below.

To make **multiple selections** in a select box, press the **Ctrl (Windows)** or **Command (Macintosh)** key and click on the item. Use the same procedure to remove a choice that has already been selected.

**For Instruments from these Types of Projects:**  
 Choose All Project Types:

- Curriculum Development
- Faculty Development
- Teacher Education
- Technology
- Under-Represented Populations

**From Projects funded by:**  
 Choose All Funding Sources:

- NSF Division of Elementary, Secondary, and Informal Education (ESIE)
- NSF Division of Human Resource Development (HRD)
- NSF Division of Research, Evaluation & Communication (REC)
- NSF Division of Undergraduate Education (DUE)
- U.S. Department of Education Technology Innovation Challenge Grants
- Unidentified

**From Projects in these Content Areas:**  
 Choose All Content Areas:

<b>mathematics:</b> <input type="checkbox"/> Choose All: algebra calculus general mathematics statistics	<b>science:</b> <input type="checkbox"/> Choose All: astronomy biology chemistry computational science	<b>social science:</b> <input type="checkbox"/> Choose All: culture economics geography	<b>technology:</b> <input type="checkbox"/> Choose All: general technology robotics
---	---	---	--

unidentified

**From Projects contributed by these Organizations:**  
 Choose All Organizations:

- Algebra Project, Inc.
- American Institute of Physics
- Anonymous
- Arizona State University
- Auburn University

Figure 4 (Part 2). The Original Search Interface (cont.)

**Of these Instrument Types:**  
 Choose All Instrument Types:

<b>Case Study Protocol Instruments:</b> Choose All: <input type="checkbox"/> Case Study Protocols	<b>Observation Instruments:</b> Choose All: <input type="checkbox"/> Classroom Observation Instruments
<b>Parent Instruments:</b> Choose All: <input type="checkbox"/> Parent Interviews	<b>School Staff Instruments:</b> Choose All: <input type="checkbox"/> School Staff Interviews School Staff Surveys
<b>Student Instruments:</b> Choose All: <input type="checkbox"/> Student Attitude Surveys Student Content Assessments Student Course Evaluations Student Curriculum Analysis Tools	<b>Teacher/Faculty Instruments:</b> Choose All: <input type="checkbox"/> Teacher/Faculty Content Assessment Teacher/Faculty Course Evaluations Teacher/Faculty Focus Groups Teacher/Faculty Intake Questionnaires

**On these Instrument Topics:**  
 Choose All Instrument Topics:

Academic Habits  
Academic Preferences  
Attitudes & Beliefs (School Staff)  
Attitudes & Beliefs (Student)  
Attitudes & Beliefs (Teacher/Faculty)  
Background Characteristics & Activities (School Staff)  
Background Characteristics & Activities (Student)

**In these Formats:**

Open-Ended Only     Closed-Ended Only     Open-Ended & Closed-Ended



**Figure 5. The Streamlined Version of the Search Interface****Search for Instruments**

Specify the features you're looking for below.

An [advanced search](#) for instruments is available, allowing you to also search by instrument topic, instrument format, project type, project funding source, and do a more refined search by project content area and instrument type. You can also view a list of [all instruments](#).

**For these Types of Instruments:**  
 Choose All Instrument Types:

<input type="checkbox"/> Case Study Protocol Instruments	<input type="checkbox"/> School Staff Instruments
<input type="checkbox"/> Observation Instruments	<input type="checkbox"/> Student Instruments
<input type="checkbox"/> Parent Instruments	<input type="checkbox"/> Teacher/Faculty Instruments

**From Projects in these Content Areas:**  
 Choose All Content Areas:

mathematics    science    social science    technology  
 unidentified

**From Projects contributed by these Organizations:**  
 Choose All Organizations:

Algebra Project, Inc.  
 American Institute of Physics  
 Anonymous  
 Arizona State University  
 Auburn University

To make **multiple selections** in a select box, press the **Ctrl** (Windows) or **Command** (Macintosh) key and click on the item. Use the same procedure to remove a choice that has already been selected.

In addition, the panel gave some feedback on issues that the users in the laboratory-based tests did not notice, given the nature of the tasks they were asked to do. The panel spent time on the Web site in reviewing it for use with their graduate classes, and interacted with the site as evaluators would when visiting the site for their own purposes. One of the panel members thought that it was difficult to get back to previously viewed material. The panel member thought that the page titles that appear in the browser or in the browser history were not helpful for discerning which page contained information that a user might want to find again after browsing past it on the site. Additionally, the different panel members raised other minor issues; we are trying to discern if the issues are related to individuals' preferences or problems that could impact many users. In general, the panel did not uncover any major new usability issues that the laboratory testers had not already pointed out. The panel's feedback provided evidence that the Web site improvements made after Round 1 of the user tests had improved the usability

of OERL. A third round of laboratory-based user testing would help determine if the issues identified by individual panel members are problems that need to be addressed or the results of individual preferences that do not have a large impact on the majority of users.

*Findings about the Professional Development Modules.* The expert panel was the first group outside of the OERL Web and content team to review the professional development modules. The panel had some suggestions for improvement to the content, but in general, emphasized problems with the interface. Since the expert panel's task was to review the site while thinking about using it with graduate-level evaluation students, their comments were significant. In one of the teleconferences with the panel, we spent a great deal of time answering their questions and taking notes on the problems they had encountered with the professional development modules and on their suggestions about ways to enhance the modules' usability.

Figure 6 shows a screenshot of the professional development modules as the expert panel originally saw them, and Figure 7 shows the improvements made in response to their comments. One problem the expert panel encountered with the original interface was that they could not easily tell what was "clickable" and what was not. For example, the first line of text in the gray box (see Figure 6 and the tabs labeled Introduction; Objectives; Strategy; etc.) is the main tool for navigation through the different sections of the module. The section headings in that gray box did not appear to the panel to be clickable. After considering the problem, the Web team decided that fully developed tabs would better indicate clickability for the main navigation area (see Figure 7). The panel was also confused by the difference between the "Introduction" and "Objectives" sections; those two section headings were combined into one section heading called "Key Topics."

The second row of the gray box is the sub-navigation area, used to move between pages within sections of the module. It was confusing to the panel that the second line started in the middle of the screen and not on the left (see Figure 6). The sub-navigation was re-aligned flush left (see Figure 7).

The intermodule navigation menu (the drop down menu at the top of the page) initially caused some confusion for the panel members because they did not understand why it was there or what its purpose was. Its location on the top right side of the screen (see Figure 6) was thought to contribute to the confusion as it was not conspicuous enough for some of the users. The menu was moved toward the OERL logo to make it more prominent and text was added above it to explain its purpose (e.g., Select a Professional Development Module) (see Figure 7).

In the first iteration, there was a “previous page” and “next page” link on the top and bottom of every page in the modules (see Figure 6). These links were eliminated because they were redundant with the gray navigation bars at the top. Removing them allowed the pages to be less cluttered and take up less vertical space.

Figure 7 also illustrates one additional improvement. The professional development modules have links back to examples from the OERL library of plans, reports, and instruments. When an “R” appears on a page (as it does in Figure 6), it is a link to an example highlighted from the Reports section of the OERL Web site (similarly, “I” is a link to an Instrument example, and “P” is a link to a Plan example). The panel found the letters confusing because they could not locate the legend that explained the letters. The legend was originally located only at the bottom of the main introduction page for all of the professional development modules, along with a great deal of other information. If the users did not notice the legend on the main introduction page, they would have trouble finding the legend after delving into a professional development module. We therefore added a legend to each page that included an example from the professional development module. Figure 7 shows the legend in the top right of the page.

The panel also offered feedback on how the examples were displayed, which resulted in better explanations of the examples, more efficient use of screen space on the example pages, and the addition of a link from the example page back to the point in the module where the example was linked. This link helped to prevent the users from getting lost

(originally, after looking at an example the users had to click the back button repeatedly to return to where they had been in the module).

**Figure 6. The Original Professional Development Modules Interface**

Introduction Objectives **Strategy** Scenario Case Study References

Introduction / **Step 1** / Step 2 / Step 3 / Step 4 / Step 5 / Step 6 / Step 7 / Step 8 / Step 9 / Step 10

<previous page next page >

**Step 1.** Identify assumptions about change (R).

Your first step is to identify the assumptions of change that provide the rationale for the project. After all, the very word "intervention" reminds us that the project is intervening to change something. Defining what the intervention is supposed to change, and how, is a necessary preliminary to deciding how to measure its effectiveness. Ideally, the assumptions are grounded in theory, which means that they have research to back them up.

Different assumptions about the value of the project carry different implications for determining its worth. Sometimes, a project is theorized as providing a new and better solution to a commonly accepted problem. Such projects need to be evaluated on the basis of whether they solve previously defined problems better than alternatives do. Other times, the project is viewed as an opportunity to pursue new goals not previously articulated. An example would be a project that uses new learning technologies in an instructional context that are hypothesized to open up new possibilities for learning that were previously beyond reach of the learners in those contexts. Such projects need to be evaluated on whether they succeed in being the agents of the outcomes they are hypothesized to cause.

Therefore, you need to decide if the intent or intents of the evaluation should be to:

- ◆ Determine effects of the intervention (that is, causal relationships between the intervention and the outcome measures).
- ◆ Allow you to evaluate the strength and uniformity of the intervention across the different intervention participants.
- ◆ Help you explain the sources of variability in the results, provided that:
  - ◇ such an analysis is of interest to the stakeholder, and
  - ◇ additional conditions and factors have been identified that could be exerting an influence on outcomes.

Figure 7. The Second Iteration of the Professional Development Modules Interface

**Step 1:** Identify assumptions about change [\(R\)](#).

**(R)** = report example

Your first step is to identify the assumptions of change that provide the rationale for the project. After all, the very word "intervention" reminds us that the project is intervening to change something. Defining what the intervention is supposed to change, and how, is a necessary preliminary to deciding how to measure its effectiveness. Ideally, the assumptions are grounded in theory, which means that they have research to back them up.

Different assumptions about the value of the project carry different implications for determining its worth. Sometimes, a project is theorized as providing a new and better solution to a commonly accepted problem. Such projects need to be evaluated on the basis of whether they solve previously defined problems better than alternatives do. Other times, the project is viewed as an opportunity to pursue new goals not previously articulated. An example would be a project that uses new learning technologies in an instructional context that are hypothesized to open up new possibilities for learning that were previously beyond reach of the learners in those contexts. Such projects need to be evaluated on whether they succeed in being the agents of the outcomes they are hypothesized to cause.

Therefore, you need to decide if the intent or intents of the evaluation should be to:

- ◆ Determine effects of the intervention (that is, causal relationships between the intervention and the outcome measures).
- ◆ Allow you to evaluate the strength and uniformity of the intervention across the different intervention participants.
- ◆ Help you explain the sources of variability in the results, provided that:
  - ◇ such an analysis is of interest to the stakeholder, and
  - ◇ additional conditions and factors have been identified that could be exerting an influence on outcomes.

In both the general Web site review and the review of the professional development modules, the expert panel members offered excellent feedback from the perspective of users of the OERL Web site. The panel members were not experts on Web interface design, and their main focus was not on the usability aspects of the Web. However, they were target audience users with a great deal of knowledge of the content on the site, and spent a significant period of thoughtful time using and reviewing the site to understand how it might be used in their classrooms. They identified and pinpointed problems that were not initially apparent in the laboratory-based user tests, enabling the OERL team to improve and refine the site significantly as well as identifying new issues that can be examined in future user tests.

## Summary

The methods employed in usability studies are familiar to most evaluators. The implementation of these methods may be a bit different and may require the evaluator to bring in a colleague or consultant initially to clarify technical issues, but the evaluator already has skills in instrument development, observing participants in studies, asking questions, and analysis. Usability methods give the evaluator an important new tool. If an evaluator ignores the usability of a Web site when conducting an evaluation of a project that features a Web site, especially a project designed to engender a community or encourage ongoing use, then a great deal of contextual information will be lost, and the evaluator will not be able to characterize the evaluation results fully. In a formative evaluation, findings from usability studies are important feedback to a project that enable it to refine its Web site. In summative evaluations, usability methods may help an evaluator understand why a Web site that seems to have all the right information, might be rated poorly by users. For example, the Web site may have useful content but is ineffective because it is difficult to use.

The usability methods employed in the OERL Web site evaluation were fairly basic, but yielded a great deal of useful information to the OERL project team. Ultimately, they helped improve OERL's usability. In hindsight, some of the problems and solutions seem somewhat obvious, but even the most careful designers will not be able to design a perfectly usable Web site in the first iteration, as there are always trade-offs to make in the design process. Usability methods serve to highlight problem areas so they can be improved. The laboratory-based usability studies and the expert panel review yielded more information that resulted in direct improvements to the OERL Web site than hundreds of surveys could. Observing a user interacting with the Web site is a way to gain, in a very short time, a new understanding of what the site can give or fails to give.

## References

- Neilsen, J. (2000). Why you only need to test with 5 users. In *Useit.com Alertbox columns*. <http://www.useit.com/alertbox/20000319.html>
- Rubin, J. (1994). *Handbook of usability testing: How to plan, design, and conduct effective tests*. Wiley: New York.
- Schneiderman, B. (1998). *Designing the user interface (3rd ed)*. Addison Wesley: Menlo Park, CA.
- Thurston, E., Fusco, J., Javitz, H., & Smith, N. (2003). *Technological design considerations for surveys of online learners*. Paper presented at the annual meeting of the American Evaluation Association, November 2003, Reno, NV.
- Zalles, D., Trevisan, M., & Haertel, G. (2003). *Expert panel review of OERL's instructional utility*. Paper presented at the annual meeting of the American Evaluation Association, November 2003, Reno, NV.

## **Appendix A. User Testing Script for OERL**





*Navigation User Testing Script for OERL (v. 4.2)*

*User:*

**Job/Company:**

**Speaker:**

**Note-Taker:**

**Date:**

**OS:**

**Browser:**

*Welcome (Time:        )*

<Greet participant. Introduce yourself and note taker. Make them feel welcome and comfortable.>

<Explain that you will be taking notes during this interview. Show equipment.>

Today, you're going to see a Web site called the Online Evaluation Resource Library (OERL) that has been in existence for a couple of years and was recently revised. More revisions and updates are planned for the future, partially based on user input like yours.

This is not a test of your abilities or knowledge; it's a test of the usability and organization of the Web site. Nothing you do is wrong, nor should you feel the need to apologize.

*Background (Time:        )*

*Before we get started, I'd like to ask you a few background questions:*

1. Would you describe yourself as an evaluator? If so, about how long have you been doing work of this nature?
  
  
  
  
  
  
  
  
  
  
2. Are you familiar with NSF programs? If not, what evaluation funding sources are you familiar with, if any?
  
  
  
  
  
  
  
  
  
  
3. Do you ever read or write evaluation reports, plans, or instruments? If so, which and in what context?
  
  
  
  
  
  
  
  
  
  
- 4a. Have you used the OERL Web site in the past? *(If no, circle Never on the next question)*

4b. If so, describe how often you use the OERL Web site:

*Never*    *A few times*    *Once a month*    *More than once a month*

5. Describe how often you use the Internet:

*Once a month or less*    *Once a week*    *Once a day*    *Many times a day*

6. Describe how often you use the Internet for evaluation purposes, excluding literature searches:

*Never*    *Once a month or less*    *Once a week*    *More than once a week*

7. Would you say you are more comfortable using Macintosh or Windows? If you prefer another operating system, which one?

*Macintosh*    *Windows*    *No preference*    *Other* \_\_\_\_\_

8. How comfortable do you usually feel using computers?

*Very uncomfortable*

*Uncomfortable*

*Neutral*

*Comfortable*

*Very comfortable*

### *Getting Started*

*We'll start with your first impressions of the home page, followed by asking you to navigate your way around the site, then ask for your general comments on several topics.*

*You don't need to be an expert on the content presented by the Web site to help us out. We're looking to see if we have provided enough support to allow you to learn about the content and navigate around the site successfully. Today we're focusing on how well the Web site enables these activities, so please comment on what about the Web site helps you or confuses you. Please feel free to mention anything you think could be improved.*

### *First Impressions (Time:       )*

*First we're going to get your first impressions of the OERL site from just looking at the home page.*

Go to <http://oerl.sri.com>.

1. What are your first impressions of the site? These could be about any feature, including the graphics, text, color scheme, organization, usefulness, etc.

2. Just by looking, what do you infer the purpose of the OERL site to be?

*<The purpose of OERL is to support the continuous improvement of project evaluations by providing a library of examples of sound evaluation plans, reports, and instruments to evaluators.>*

3. Just by looking, do you feel the home page gives you a good idea about how to navigate around the site?

*Tasks (Time:        )*

*Now we're going to ask you to start looking around the Web site to find certain pieces of information. It will be helpful if you would share your thoughts and impressions out loud as you have them. Try to find each web page or piece of information on your own, but feel free to ask questions if you get stuck or can't figure something out. If you feel like you need to explore a little bit or spend some time reading content, that's okay. If you don't fully understand something or can't find something, that will be helpful feedback for us to have. There are about 10 of these tasks.*

*For each task, the note-taker should record*

- *What the user did*
- *Anything the user described as helpful*
- *Anything the user described as confusing*
- *Any errors the user encountered in the site (navigation, interface, or content)*
- *Anything else that surprised the note-taker or was otherwise interesting*

1. Find the OERL FAQ (frequently asked questions page).

2. Find student content assessments in Under-Represented Populations.

3a. The next two questions are related: Find the glossary entry for the Project Context report component.

3b. Find a definition for "participant".

4. Find a page that describes what an alignment table is in OERL.
  
  
  
  
  
  
  
  
  
  
5. Find the alignment table entry for Interpretations & Conclusions, a report component.
  
  
  
  
  
  
  
  
  
  
6. Find the document that the criteria for OERL are based on.
  
  
  
  
  
  
  
  
  
  
7. Find out what programs have contributed to Faculty Development projects.
  
  
  
  
  
  
  
  
  
  
8. Find a page that gives you an overview of OERL instruments.





*Follow-up (Time:        )*

*Please describe any additional comments you might have on:*

1. Navigation

2. Organization and structure of the content

3. Search feature

4. Evaluation References (criteria, glossary, alignment table)

5. Help materials (user scenario, FAQ)

6. Anything else?

*Thanks for your participation!*

*(Time:        )*