EXPANSION OF EIA'S WEB SITE USABILITY TESTING PROGRAM

Colleen Blessing, Howard Bradsher-Fredrick, Herbert Miller, Renee Miller, Robert Rutchik
Energy Information Administration

KEY WORDS: Web site, usability testing, think-aloud, participants, energy experts, Kid's Page, students, teachers, links

For its 2002 usability tests, the Energy Information Administration (EIA) changed the focus of its testing in two ways. In addition to asking participants to find specific energy information, EIA also asked participants if they understood the wording on links; where they would find definitions of terms; and where they could find survey EIA survey forms[1].

EIA also changed its recruitment process in 2002. EIA divided its participants into energy experts and non-energy experts. In 2003, EIA tested grade school children and teachers on the site's Kid's Page. This paper focuses on EIA's change in recruitment and testing methodology.

I. Energy Information Administration Usability Testing Program Description

EIA's experience in implementing a usability testing program for its Web site is a microcosm of a wider trend in government and the private sector to develop testing programs as agencies and companies continue to collect and disseminate more information through the Internet. Thus, EIA may have developed its program independently of other agencies and private entities, but its development is not an isolated occurrence.

EIA launched its Web site in 1995 and from 1996 through 1998 the agency analyzed indicators such as number of accessions per month, number of accessions by domain, file accession patterns, and other measures from its Web logs to help improve the content, navigation, and design of its Web site. EIA began to conduct usability testing on the site with actual users in 1999. From 1999 through 2001, EIA asked test participants to find specific information on the EIA site[2], starting their search from the EIA home page. The participants were an unstratified[3] "sample" of 15 to 18 college-educated professional people who did not have to have any energy expertise.

EIA decided to use the think-aloud interview approach for its testing. EIA developed this design from a 1997 Statistics Canada study, ((*Research on the Canadian Statistics Area of the Statistics Canada Website,* (http://www.statcan.ca)) and from some tests of its own. EIA has used this approach for all its testing[4].

EIA has done most of its testing at the Bureau of Labor Statistics (BLS) cognitive lab. (The agency has conducted the remainder in an improvised lab in its Washington offices.) The EIA interviewer and participant would sit at a work-station. The interviewer would have a four-part protocol. The first part was a brief introduction, explaining the purpose of the exercise and a few other things, such as asking for the participant's permission to videotape the session.

The second part was three demographic questions. These were all self-evaluation and recall questions, of which the accuracy of the answers is subject to the vagaries of these types of questions. For example, EIA asked "which of the following best describes your level of expertise using the Internet – Novice, Intermediate, Expert." [5]

[1] The authors want to recognize Mary Spruill of the National Energy Education Development Project and Grace Sutherland, Joelle Michaels, and Stacy MacIntyre of EIA who worked on this project. The authors also want to thank Jenny Zhang, a Carnegie Mellon University student who was a Joint Program in Survey Methodology Intern at EIA during the summer of 2003, who helped us immensely with the tests and data analysis.

[2] For example, "How much coal was produced in Kentucky in 1997?"

[3] This was not a probability sample. Recruits included librarians, students, and researchers.

[4] For a more detailed description on how EIA developed usability testing methodology, see Colleen Blessing, Howard Bradsher-Fredrick, Renee Miller, Robert Rutchik, and Antoinette Ware-Martin, "Cognitive Interviewing: Applications to Evaluating the Energy Information Administration's Web Site," paper presented at the annual meeting of the American Statistical Association (Baltimore, MD, August 1999).

[5] The others were "How familiar are you with the energy industry and energy terms? Not at all familiar, Somewhat familiar, Very familiar." And "Before today, how many times have you used the EIA Web site?"

The exercises in which EIA asked the participants to find specific energy information were the third part of the session and its major focus. The questions were in common English (EIA hoped), instead of industry jargon, to make them understandable to the participants. EIA recorded how long it took a participant to answer each exercise. Participants were allowed a maximum of ten minutes to find the answer to a question. The purposes of the exercises were:

- To observe participants' ability to locate information on the EIA site
- To observe whether participants unknowingly located incorrect information when performing specific searches
- To find out why participants located incorrect information
- To observe the navigation behavior of participants
- To gather the participants' overall impressions of the site and their suggestions for its improvement.

The session would conclude with the participants being asked three "debriefing" questions. For example, "how confident are you that you found the correct information?" The complete session would last from forty to sixty minutes.

The methodology worked. The first several years of testing revealed some basic problems with the EIA site. The major conclusions were that:

- Participants came across lots of energy jargon and other undefined terms. This confused them.
- They had no consistent way to return to the EIA Home Page
- Most participants said there was too much text to read. "Reading text is hard to do."
- Participants often found text reports when they wanted a table. "Expecting a table. It is here somewhere."
- Yet, there were many text tables with lots of horizontal scrolling

In summary, more than half the time respondents did not find the correct information. Often, they would scroll right by it because it was buried in text or tables and/or labeled in terms that confused participants. They would stop and look at the information (or a link) and not be sure that it was the information for which they were searching.

EIA used these and other findings to:

- Improve the design of the EIA Home page and second level pages[6]. (In more recent tests, users have navigated through these pages quicker, and more correctly, than previous tests.)
- Use much less energy jargon on its Web pages
- Write more pages for the Web. Users now find data tables instead of text and find the tables more easily. The tables fit on the screen and the totals and more recent information are at the top of the table. Finally, text is written in bullet points and with links to related information.

II. EIA's 2002 Tests

A. Recruitment

EIA changed the focus of its testing for its 2002 test. The agency moved away from a macro evaluation of the site to a more micro evaluation. From 1999 through 2001, EIA did not stratify its testing participants. For the 2002 tests, EIA stratified its participants into energy experts and non-energy experts. The experts were energy consultants/researchers from Washington DC area "think tanks," colleges, and other institutions. EIA recruited them from attendees at EIA's National Energy Modeling System Conference and recommendations from EIA staff.

EIA decided to test "energy experts" because since EIA's Web site was launched in 1995, the agency's Web statistics consistently have shown that the twenty percent of the site's users who access the site ten or more times a month access between seventy and eighty percent of files accessed on the site. EIA calls these users "frequent users." EIA wanted their feedback on the site.

[6] EIA has redesigned its Web Site four times since 1995. The last redesign was in 2001.

B. Exercises

EIA also expanded the questions to participants about the site. In addition to finding specific energy information, EIA asked respondents if they understood the wording on links on the Home Page; where they would find definitions of energy terms that they did not understand; where to find EIA survey forms; and a few other things. The experts were asked to do tasks on the site that they frequently do for their work. These are examples of what EIA asked participants to do in the 2002 test:

- Find specific information or data – "Which major companies import oil from the Middle East?"
- Find an EIA publication – "Can you find *The International Energy Outlook*?"
- Find the definition to technical terms – "What is bunker fuel?"
- Find an EIA survey form – "Find the Form EIA-860?"

C. 2002 Results

There were differences between the results of the experts and non-experts. To the demographic question "before today how many times have you used the EIA Web site," all the energy experts answered "frequently." For the non-experts, the model category was "a few times."

The overall percent of correct answers to the informational exercises rose from 40 percent in EIA's 2001 test to 53 percent in the 2002 test. The experts found the correct answer for 85 percent (17 of 20) of the exercises, while the non-experts found the correct answer for only 39 percent (18 of 46), a percentage nearly identical to the 40 percent rate from the 2001 tests. The experts were able to find the correct answer more often than the non-experts because of their greater knowledge of the site and the subject matter.

The experts also had comments about the inaccessibility and organization of EIA's data sets. Several said that EIA databases were hard to find. For example, they would be looking at a table or graph in an EIA publication and would want to access and download the data set from which the table or graph was constructed. Yet, they said there was no link from the table or graph to the data set. It was not transparent on the site where the data sets were located.

Others said that some EIA data sets were poorly organized, hard to understand, and inefficient to use. For example, in some there were no column headers to describe the variables and their units of measures. Some did not have years in the rows to tell the user for what year (or month) the record was. There were also no indicators of breaks in data series. These data sets did not have any metadata to tell the analyst anything about the data. One expert suggested that EIA should have a primer on the use of its data sets. The researchers presented these findings to EIA Senior Staff.

III. 2003 Tests

A. Introduction

EIA's 2003 test continued the change in its testing design to focus on a specific subgroup of users that it started in the 2002 test. This time EIA tested grade-school children and teachers on the site's Kid Page, http://www.eia.doe.gov/kids/ . The Kid's Page has been EIA's fastest-growing web product for the past 2 years. Its use has increased from around 19,000 visits in March 2001 to 64,000 visits in May 2003 (and this is down from 73,000 in April 2003). It is now the 3rd most widely used EIA web product.

The Kid's Page, EIA thought, has a simple layout[7]. There is a heading that says "EIA's Kid's Page." Below and to the left of the heading is a cartoon character that EIA calls the "Energy Ant." The cartoon bubble from the Ant says "Hi, I'm your host, Energy Ant. Click on me to learn more about energy." To the Ant's right there is a column, in yellow, with ten information channels that children (and others) can click on.

The overall purpose of the 2003 test was to see if grade-school students and teachers could navigate through the Kid's Page to

[7] The Kid's Page is managed by the Kid's Page Committee, a group representing all EIA offices, and is supported by the National Energy Education Development Project.

find the information for which EIA asked them to look, and see:

- If the links accurately convey to users the information the page contains?
- If the navigation through the channels is transparent?
- If the text uses words that the children can comprehend?

The information collected from all the interviews will be used to improve the Kid's Page.

B. Research on Testing Children

EIA knew that evaluating the EIA Web site with grade-school students and their teachers would require a somewhat different approach to testing than EIA had used in past usability testing. The agency needed guidance for testing children, and there was some in the literature.

A few years ago, Jakob Nielson conducted usability testing with grade-school children ages 6 to 12 here in the United States and in Israel on 24 Websites designed for children and on three "mainstream" sites. Nielson argued that despite this growth in [children] users and services, very little is known about how children actually use websites or how to design sites that will be easy for them to use. "Most website designs for [children] are based on **pure folklore about how kids supposedly behave** -- or, at best, by insights gleaned when designers observe their own children, who are hardly representative of average kids, typical Internet skills, or common knowledge about the Web[8]."

One of Nielson's findings was that children had their greatest success using Websites that were intended for adults because of their "simplicity and compliance with web design conventions." He also said that children, contrary to a common folk wisdom, are not "masters of technology." They are "incapable of overcoming many usability problems." Finally, he found that "fancy wording in interfaces confused users and prevented them from understanding available choices" and that the children are "keenly aware of their age" and can sharply differentiate between material that "is appropriate for them and material for older or younger [children]." EIA's test would support these last two points.

C. Recruitment and Methodology

EIA conducted two rounds of interviews, testing a total of 37 students and 6 teachers. The first test was on Take Your Sons and Daughters to Work Day at EIA on April 24, 2003. The participants were the children and grand children, in some cases, of EIA employees. EIA tested 23 school children in grades one through eight and one high school student. The grade and middle-school children represented every grade except the first. EIA interviewed them in a makeshift lab and in offices at EIA's Washington, DC offices. About a half dozen were tested at the homes of EIA staff.

The second test was at the Washington area hotel of the June 2003 conference of the National Energy Education Development Project (NEED). EIA tested 13 students and 6 teachers. Of the students, six were in 5th through 8th grades and seven were high schools students. There were both similarities and differences in the findings from this test and those from the April test. The similarities were mainly that the June evaluation showed the same problems with the Kid's Page that the April test showed. The similarities and differences will be discussed in detail shortly. Figure 1 shows the distribution by grade of all (both April and June) the students that EIA tested.

Protocol

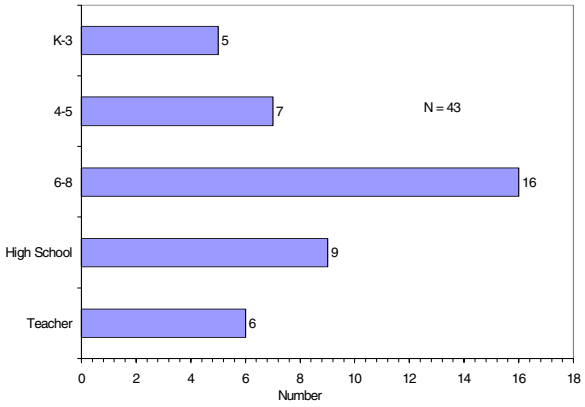
The protocol, as in previous tests, was divided into four parts[9]. There were now four recall/self-evaluation demographic questions. In addition to asking the participants how often they use the Internet and how they rate their Internet skills, EIA asked the students what grade they were in and the purpose of their Internet use.

[8] Jakob Nielson's Alertbox, Kid's Corner: Website Usability for Children, April 14, 2002. http://www.useit.com/alertbox/20020414.html

[9] EIA did not video-tape the tests as it had in previous tests.

Next, the students (and teachers in June) did the exercises. Depending on how quickly they answered the questions, the students completed between 8 and 15 exercises. Unlike previous years' tests, the students were not given a maximum time to complete the exercises and were not timed on how long they spent on each exercise.

Figure 1. Grade Distribution by Participants Tested



EIA compiled an inventory of 34 questions. From the inventory, EIA developed different sets of questions for different grade groups – i.e., K-3, Grades 4-5, Grades 6-8, High School – and for the teachers.

The stratification of the questions resulted from an exercise design problem. EIA debated whether some questions had wording that was not appropriate for a particular group, and if they would come close to finding the answer. For example, EIA constructed the question "where would you go to find significant events in energy history." EIA decided that words and phrases such as "significant," "significant events," and "energy history" were hard concepts for students from kindergarten through the 3rd grade students to comprehend. Another example is "you heard that the EIA Kid's Page had biographical material on people with distinguished careers in the energy field." EIA only asked 6th graders and above this question.

Thus, the older students and teachers were asked harder questions than the younger students. Further, all the participants were not asked the same questions. Finally, there were about half dozen questions that EIA asked in the April test that were not asked at all or only a few times in the June test.

A second issue was whether EIA should ask questions for which they knew the answers would be difficult to find because the navigation was not transparent, even if the students understood the meaning of the question. EIA decided to ask these questions.

The participants were asked five debriefing questions in the final part of the session, as opposed to three in previous tests. An example of one of these questions was "overall, what did you think of the Kid's page?" The complete session lasted between 35 and 45 minutes.

D. Test Results

Exercises

There was a clear difference in the answer patterns to the exercises between the April participants and the June participants. In the April test, a majority of the students found the correct answer for one third of the exercises and the wrong answer for another third. For the remaining third, half the students found the correct answer and half did not. Overall, the April students gave the correct answer to 55 percent of the questions asked (95 correct answers out of 187 asked).

The June participants provided the correct answer to 74 percent of the questions they were asked (106/144). Teachers provided the correct answer to 88 percent of the questions they were asked, and students 71 percent. A majority of the June participants answered only four questions incorrectly, three with which the April students also had problems. EIA thought that the participants might have problems finding the correct answers to these questions. (There was only one question that a majority of April students answered correctly and a majority of the June participants did not.) For both tests, the percent of correct answers was 65%, (201/331).

It is not surprising that the students and teachers in the June test had a higher percent

(75%) of correct answers than the April students (55%). There were, as noted, six teachers in the June test and none in the April test. Also, the June participants were older and they were attending the National Energy Education Development Project conference at that time. They were more interested in and knew more about energy than the April students.

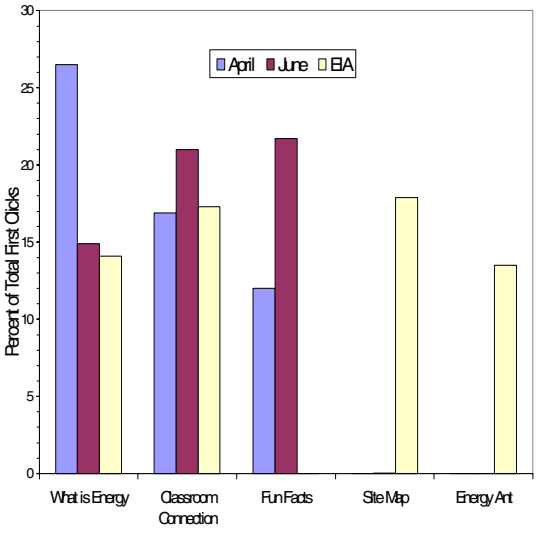
Navigation Patterns

As noted, the students and teachers said there was lots of good information on the Kid's page, but it was often difficult to find. This problem often emerged as soon as a participant started to look for information. When asked to start an exercise, a majority of participants would move the cursor up and down the channel menu before deciding which link to use to start to find the information. When asked why they did this, they invariably said that they were not sure what link to use because they did not understand the link name. The question, "What do you think you will find under the link called 'Milestones?'" which was asked of fourth graders and above, was exhibit number one of this problem. "What is Energy" is another example. Some participants said that it was a definition of energy.

Figure 2 illustrates the first click pattern of the respondents and a first click analysis by EIA using key word(s) in the question as a cue for the first click. The patterns for both the April and June tests were very similar. "What is Energy" (WE), "Classroom Connection" (CC), and "Fun Facts" (FF) were the most frequent first clicks for both groups. The only substantial difference between the two is that "What is Energy" was the most frequent first click for the April participants while "Fun Facts" was the most frequent first click for the June participants.

The Site Map, "What is Energy," "Classroom Connection," and Energy Ant were the most frequently used first clicks in EIA's key word analysis. In EIA's analysis, "Fun Facts" was seldom used to start a search, while the test participants seldom used Energy Ant and the Site Map In fact, there were only six clicks by the participants on "Energy Ant, which dominates the left side of the Kid's Page. The Ant's bubble says "Hi, I'm your host, Energy Ant. Click on me to learn more about energy." Most students saw the Ant as another decorative picture on the site and not as something that was clickable. (Even fewer (8) used the Site Map that is located right next to the Glossary link.)

Figure 2. Navigation Pattern Analysis



The use of "Fun Facts" was particularly interesting. Apparently, the word "fun," or "facts," was a cognitive clue for the students (and teachers) that there would be a wide range of great information behind the link (which there really was not). One student said she thought there was something "fun" behind the link. Another said it was the most interesting thing about the Kid's Page.

There were only about thirteen total first clicks in both tests on the Glossary, and EIA asked nearly each student and teacher to find the meaning of a term or word. The Glossary, though, was not on the main menu. It was below and to its left. Participants just did not notice the Glossary link or pay attention to it. "Focusing on the yellow," as one teacher stated when he was shown the Glossary link and asked why he did not use it.

At the end of each session during the June test, EIA showed each participant a paper mock up of a redesigned Kid's page with the Glossary and Site Map in the main menu, with some of the link names changed, and with all the link names in bigger type. The participants, with one or two exceptions, said they liked that design over the current one.

Finally, in doing the exercises, EIA thought that the students were just as patient as the adults EIA had tested. In some cases, they seemed more determined to keep looking for the answer. They also asked for more guidance from the interviewers than the adults in the previous tests.

Debriefing Questions

The participants in both the April and June tests made the same comments during the exercises and in their answers to the debriefing questions about what they saw as the strengths and weaknesses of the Kid's Page. In fact, the students were more forthcoming than the adults EIA had tested in noting strengths and weaknesses and in making suggestions to improve the Page. These are their major compliments and complaints about the Page.

Compliments:

- Good use of charts and graphs (and said there should be more)
- Liked "Fun Facts" (until they saw there was little fun or facts behind the link.)
- Liked famous people, dictionary, and timeline of events

Complaints:

- Lots of information, but difficult to find things (confusing link names)
- Too much reading (too much scrolling)
- Not enough graphics and games (and did not know that most graphics were links)
- Need a search engine

The general thrust of the comments was that EIA has a lot of interesting information on the Kid's Page, but the Page needs better organization and design. EIA needs to decrease the confusion and clutter that they saw in the Kid's Page. In short, they recommended that EIA make the Kid's Page more efficient or rational or "user friendly." Without saying so, the students suggested that EIA organize the page according to "writing for the Web" principles[10]. The following is a sample of their recommendations to do that:

- Change link names: For example, change the "Milestones" link to "History of Energy," "Energy Timeline", or "Historical Facts". Change the "Pioneers" link to "Famous People" or "Biographies" and put pictures next to the scientists and others.
- Add keyword search
- Move the Glossary and Site Map to the main navigation menu.
- Modify the content of the "Fun Facts" link (especially the "fun" part) to more accurately reflect the link name by adding graphics – e.g., pictures and cartoons about energy.
- Add more graphics.
- Add some games and add a "Games" link to the home page.
- Condense some of the text in paragraphs to more of an outline form.
- Remove redundant information that does not fit in a category (link) and already is in another category.

These findings will form the basis for the recommendations that the researchers will make to the Kid's Page Committee to make the page more useful and easier to use.

Finally, the test's findings and recommendations also support Neilson's finding that children, contrary to a common folk wisdom, are not "masters of technology." They are "incapable of overcoming many usability problems" and that "fancy wording in interfaces [confuses]

[10] Human Factors International, "Writing for Web Usability: With Focus on Content/Text," Seminar given to the Energy Information Administration (Washington, DC, January 2001).

> users and [prevents] them from understanding available choices.[11]"

IV. Conclusion

In 2002, EIA changed the focus of its usability testing. The agency moved from a macro test design to a more micro design. In tests prior to 2002, EIA had asked participants to find specific energy information. Starting in 2002, EIA also asked participants if they understood the wording on links; where they would find definitions of terms; and other specific parts of the site.

EIA also changed its recruitment process in 2002. EIA divided its participants into energy experts and non-energy experts. For 2003, EIA further focused on testing only grade school children and teachers on the EIA Kid's Page. This research design implicitly, if not explicitly, addresses the problem of how to design a site – its architecture, navigation, and content - that will serve a wide range of customer groups, such as energy experts students, and teachers who use the site more frequently than other groups.

[11] Jakob Nielson's Alertbox, Kid's Corner: Website Usability for Children, April 14, 2002.
http://www.useit.com/alertbox/20020414.html