Development and Evaluation of Web-Accessibility Tutorials

Abstract

The goals of this project were to, 1) Create prototype tutorials to help developers create accessible pages, and 2) carry out usability evaluation of these tutorials. A prototype series of tutorials were created. Five students experienced in web development and three inexperienced, used these tutorials as guides for creating accessible web pages for the usability testing. Quantitative and qualitative data were collected. An analysis of the data indicated that, in general, all participants performed adequately, though experienced participants performed notably better as anticipated. In addition, users indicated that they found the tutorials useful and easy to use, though, again, this was particularly true for the experienced users. Qualitative analyses indicated that users tended to use the tutorials as a reference, going back and forth between the tutorial pages and their assigned tasks. This tendency was most pronounced on difficult tasks and for inexperienced users.

Introduction

Web Accessibility Standards

Web Accessibility is a set of standards used to govern what acceptable web sites should look like. There are many agencies responsible for these “standards”. The World Wide Web Consortium (W3C) and the government release standards collectively to be used for web development. The U.S. government release standards that are categorized and published on the official Section508 web page, and the focus of my project involves the comprehension and implementation of these specific guidelines. The Section 508 web site explains the standards for web development as follows: “The criteria for web-based technology and information are based on access guidelines developed by the Web Accessibility Initiative of the World Wide Web Consortium. Many of these provisions ensure access for people with vision impairments who rely on various assistive products to access computer-based information, such as screen readers, which translate what’s on a computer screen into automated audible output, and refreshable Braille displays. Certain conventions, such as verbal tags or identification of graphics and format devices, like frames, are necessary so that these devices can ‘read’ them for the user in a sensible way.” The site also goes on to clarify, “The standards apply to Federal web sites but not to private sector web sites (unless a site is provided under contract to a Federal agency, in which case only that web site or portion covered by the contract would have to comply).”

Although these are “standards” the government isn’t in a position to force compliance from sites not involved with the Federal government. However, the use of these standards is what helps to make sites accessible to people with disabilities and couldn’t otherwise view your site. The National Federation for the Blind estimates that there are roughly 1.1 million blind citizens in the United States, who could be affected by a poorly designed site. There are also many people who interact with the Internet
through cell phones, PDA’s, and other mobile devices that will be able to use a standards compliant site. From a business perspective this would allow you to reach more people and potentially make more money because of it, also when you make accessible sites they become more efficient and don’t require as much memory, saving money on web services. The Section 508 web site sums this up well: “Accessible sites offer significant advantages that go beyond access. For example, those with “text-only” options provide a faster downloading alternative and can facilitate transmission of web-based data to cell phones and personal digital assistants.” So, the importance of standards compliant sites exist in the financial, feasibility, and disability realms of Internet usage.

Usability Testing

The necessity of standards compliant sites can only be undermined by the necessity of having usable web sites, because if you can’t effectively use a site it defeats the purpose of having even the most standards compliant sites. This is where the idea of usability testing comes into play. Jakob Nielsen defines usability as, “[…] a quality attribute that assesses how easy user interfaces are to use. The word ‘usability’ also refers to methods for improving ease-of-use during the design process.” So, when we add this ideal to a research context we get the idea of “usability testing”. Usability testing is the actual physical testing of sites to see if users can adequately complete their desired tasks. This is formally defined as, “Usability testing encompasses a range of methods for identifying how users actually interact with a prototype or a complete site. In a typical approach, users – one at a time or two working together – use the Web site to perform tasks, while one or more people watch listen, and take notes.” The overall goal of conducting usability tests is to find out what is working well and what isn’t. This could be determined by a number of performance measures used during testing such as, task completion, time to complete, location of target information, etc. This is very important to determining whether or not the users’ needs are being met by your interface, in this case a web site. In relation to traditional behavioral experimental research, Usability testing is done with a smaller number of participants and relies primarily on qualitative research.

Usability and usability testing relate to education because of the fact that this type of testing is an iterative process, throughout the product life-cycle. Since you constantly test many levels of prototypes and the final product, a developer is constantly learning about the product through all of the iterations of that product. From the prospective of the user, when learning in an internet-based environment, it is extremely crucial that the site is both accessible and usable in order for a user to complete their tasks. This is basically the same logic used to justify the use of accessible sites, which is basically if the site is usable it will aid in the ability to complete tasks effectively, which in this specific context means that a user would learn effectively through usable web interfaces.

Research Goals

In order to address the issues mentioned above there are two main goals to for this project:
1. Create prototype tutorials to help developers create accessible pages, and
2. carry out usability evaluation of the tutorials.

To address the specific issue of comprehension and implementation of web
standards, I created a prototype set of tutorials for the usability experiments. In
addressing the issue of usability testing, I needed to carry out usability evaluations of the
prototype set of tutorials. This evaluation of the interface was used to answer the
following questions:

1. Were the tutorials effective?
2. How did users interact with the tutorials?
3. How did tutorial effectiveness and interactions differ with respect to experts vs. novices?

These questions were addressed through implementation of traditional usability
methodology as described below.

Method

Participants

Five students participated in this experiment, five experienced web developers and
three inexperienced web developers. An experienced web developer was considered to
be a person who had developed web pages prior to this experiment, while an
inexperienced web developer was considered to be a person who had no prior experience
in developing web pages. These students were recruited from the IST web development
class and another business class within the School of Management and Information
Sciences. The experienced participants were from the web development class and the
inexperienced students were from the business class. The students, who were recruited
from the business class, were selected with the criterion that they had not developed web
pages previously. All students received extra credit for their participation.

Materials

A series of three tutorials were created. Each tutorial was a web page, which
consisted of a description of the 508 guideline and an example. One web page also
included a video, which was a dynamic capture illustrating how the example was done.
The three tutorials covered the following topics: 1) Add alternative text to an image; 2)
add column and row headers to tables; and 3) add a “skip navigation” link, which is used
to allow people with text only browsers to skip redundant links, (A screen shot of one of
the tutorial pages is shown in figure 1.)
Each tutorial page was linked to a second page that asked the user to open an .html page on the hard drive that contained errors. The tutorial contained information that should allow the user to correct these errors. An example of the correction directions page is presented in Figure 2.
Figure 2. Page that followed the tutorial page instructing user to open an .html page on the hard drive and correct it accordingly.

The testing of the participants occurred in the Laboratory for Information Technology Evaluation (LITE) lab. This gave us a private area where the participants could come in and focus on their tasks. I also used a tool called Morae produced by TechSmith. This tool was used to record both audio and visual data from each participant, edit video, and gather data such as the amount of clicks and each user’s time spent on task. This also allowed me to measure the entire time they spent on the experiment from beginning to end. (A Morae Manager screen shot is shown in figure 3.)
Figure 3. This is a screen in Morae manager that was used to create “highlight” videos of participants completing tasks.

Procedure

Experimental sessions lasted for approximately forty-five minutes participants carried out the following steps:

1. Read an introduction screen that explained the purpose of the interface.
2. Signed a consent form allowing us to record video and audio data for the purpose of the experiment.
3. Completed a pre-questionnaire explaining their background in web development.
4. Complete the three experimental tasks, which consisted of reading the tutorials and correcting web pages accordingly.
5. Complete a post-questionnaire. The questionnaire consisted of items aimed at evaluation of the interface.

Results

Quantitative Analyses
To analyze the data that resulted from the experiments I used the Morae video tool described above. This tool allowed me to not only have audio and video data, it also allowed me to record data such as the users time on task, the amount of times a web page changed, how many times a dialog box appeared, and how many times a context menu popped up. This assisted me in finding out how effective the users were in completing the required tasks. To analyze the videos, after segmenting the videos, I was able to generate data automatically based on the appropriate parameters that are necessary.

The completion rate with respect to the three tasks is displayed in Figure 3.

Pre-experiment questionnaires consisted of the following three items: 1) I am an expert in the web editor, dreamweaver; 2) I am an expert in .html; and 3) I am an expert in cascading style sheets (css). Participants responded to each items using a 9 point Likert scale that ranged from 1 (strongly disagree) to 9 (strongly disagree). Figure 5 represents mean responses for the experienced and non-experienced participants.
Web-Accessibility Tutorials

Proceedings of the 2005 University of Missouri – Rolla, Undergraduate Research Conference

Figure 5. Mean responses to pre-experimental questionnaire items for the experienced and non-experienced users. (The scale ranged from: 1 = strongly disagree and 9 = strongly agree).

The average time that a user spent on task was: 6.48 minutes on task one, 6.55 minutes on task two, 14.54 minutes on task 3, and 27.67 minutes on the whole experiment. There as no noticeable relationship between completion rate and task performance.

On the post-experimental questionnaire, participants were required to respond to items on a scale that ranged from 1 meaning strongly disagree to 9 meaning strongly agree. The statements on this questionnaire are as follows: 1) I found it easy to complete the tasks from beginning to end; 2) the tutorials helped me to perform the required tasks, and 3) the web pages that included the tutorials were easy to use. The questionnaire also included an open ended question asking for suggestions to help improve the interface, the results of which will be discussed in the qualitative section below.

Figure 6 represents the mean to items 2 and 3, which were directly related to the effectiveness of the tutorials, for the experienced and non-experienced users.

Figure 6. Mean responses to post-experimental questionnaire items for experienced vs inexperienced users (The scale ranged from: 1 = strongly disagree and 9 = strongly agree).

Qualitative Analysis
One central theme emerged from the qualitative analyses. Since the interface is set up so that users could backtrack through the interface if they needed to, all of the users did just that and very frequently. Originally I thought some users would be able to skip through the tutorials and just do the corrections without issue, especially the more experienced participants, and the inexperienced participants would do most of the back-tracking and going back and forth between the tutorials and the corrections pages. Yet, in this situation that did not occur, all of the participants went back and forth between the pages and in many instances appeared to be merely copying data from page to page. While some of the participants did this less than others, it is clearly obvious that this became an issue during testing. Ultimately, the users appeared to comprehend the material and write the code on their own, until they got stuck on a problem or impatient with the process of reading each tutorial. So, it was visible that a participant would start off well in the beginning and work on their own, and as they got further through the experiment they would become either impatient or confused and then began to copy text directly a lot more often.

Conclusions

Goal 1: Create prototype

For this project I created a prototype set of tutorials, to help evaluate the participants and how they interacted with the interface. There were some issues that came up during the creation of these web pages. First, I initially had grammar and structure issues with the interface that had to be corrected with the assistance of my advisor. There were also technical issues limiting the creation of the interface that had to be fixed before I could move on. These were basically formatting and placement errors that occurred on the page that were corrected with the assistance of my research advisor and another research student.

Goal 2a: Usability Testing: Were the Tutorials Effective?

One of the biggest questions that should be answered from my research is whether or not the tutorials were effective. In figure 4, we can see that a majority of the expert users completed at least two of the tasks, and all users completed at least one of the three tasks. Also, the post-questionnaire results indicated that the participants benefit from the tutorials and they were easy to use, based on the average Likert scale ratings (Figure 6). This leads to the idea that participants found the interface usable.

Goal 2b: Usability Testing: How Did the Users Interact with the Tutorials?

Next, I would like to examine how the users interacted with the interface. Initially, I anticipated that the experienced users would read the tutorials to get an understanding of what to do and then move to the corrections pages to correct the errors in that page. This was largely true for the first one and a half tasks, but, as the tasks increased in difficulty (which happened progressively) the experienced users began to go back and forth between the tutorials and almost appear to be copying the text at times.
The inexperienced users were more likely to do this with all three tasks, which was more consistent with my expectations.

**Goal 2c: Usability Testing: How Did Experienced and Inexperienced Users Differ?**

One clear difference between the experience and inexperience users was their performance. As illustrated in Figure 4, the majority of experienced users completed all three tasks, while the majority of inexperienced users completed only 1. This, of course, is to be expected. Referring back to figure 6, we can also see that the experienced users found the tutorials more helpful and easier to use. Based on this information I would infer that the tutorials were more effective for the experienced users, than for the inexperienced users. This could primarily be because of the fact that experienced users knew what to expect, but the inexperienced users haven’t had any training or prior knowledge in dealing with problem solving in HTML.

One similarity between the experienced and inexperienced users was the fact that all of the users tended to approach and solve the proposed problem in the same manner. They used the tutorials as a reference while writing code for the corrections pages. They differed, as mentioned previously, in the extent to which they went back and forth between the tutorial pages and the correction pages, with the inexperienced users doing this more frequently.

**Suggestions for Future Research and Development**

Based on the results of this pilot work, I am encouraged to continue work with additional tutorials for users to complete, which would further test the acceptance of the Section 508 guidelines, and allow further testing of the users’ ability to implement those guidelines.

This research also pointed to ways that future interfaces could be improved. This interface was meant to be a prototype. When writing the code for the web-based tutorials, I tended to make the corrections pages look a certain way for aesthetics instead of in a more pragmatic design. In other words, I made pages look good, but when the participants corrected the pages it drastically changed the way the pages looked, because of some coding errors on my part. These errors didn’t effect the data that was gathered, it just made the testing environment that much more difficult for participants. So, in the future I hope to code those types of applications more carefully. Also, one of the tasks required the users to check a hidden link on a page, which would only be visible when the CSS was turned off. Since, you can’t turn off CSS in the Internet Explorer browser I should have used a browser that allows you to do that so that the participants could have checked to see if they completed that task correctly. (Which was task three, and three participants did complete it correctly.) An inexperienced user suggested that there be more types of hands on examples to aid the learning process. To do this I could have just added more types of examples, with the same information shown in different ways, to
make the retention and implementation of these ideas easier. However, it’s still important to reiterate that, for the most part, the participants were positive about the effectiveness of interface as referenced in figure 6.
References


Acknowledgements

I would first like to thank my research advisor, Dr. Richard Hall for aiding me with this paper. He gave me valuable advice for the direction of this paper, usability testing, and more importantly other things not related to this research project. I would also like to thank LITE (Laboratory for Information Technology Evaluation) for allowing me to do this research and use the lab and all of the equipment that they supplied. I would also like to thank the participants in my project, despite some frustrations all stuck it out to support my project. Thank you to Carla Bates’ Web Development Class and Patricia Spence’s Management and Organizational Behavior class. Also, thanks to those instructors mentioned above for offering the extra credit to participate in the project. Last but not least, thank you to OURE for giving me the opportunity to enjoy this undergraduate research, I really appreciate the opportunity and I’m sure it will be valuable to me in the future.