December 2004

Discount Eye Tracking: The Enhanced Restricted Focus Viewer

Peter Tarasewich
Northeastern University

Stephanie Fillion
Northeastern University

Discount Eye Tracking: The Enhanced Restricted Focus Viewer

Peter Tarasewich
College of Computer and Information Science
Northeastern University
Boston, MA 02115
tarase@ccs.neu.edu

Stephanie Fillion
College of Computer and Information Science
Northeastern University
Boston, MA 02115
raven82@ccs.neu.edu

ABSTRACT

This paper discusses the design and implementation of a unique software tool, the enhanced restricted focus viewer (ERFV), for tracking the visual attention of users in hyperlinked environments such as web sites. The software collects data such as mouse clicks along with the path of the user's visual attention as they browse a site. Unlike traditional eye-tracking procedures, the ERFV requires no hardware other than a personal computer. In addition to cost and time savings, the ERFV also allows the administration of usability testing to groups of subjects simultaneously. The benefits of the ERFV are demonstrated through an experiment that evaluated the usability of two web sites that were equivalent in content but differed in terms of design.

Keywords

Usability Testing and Evaluation, Web, Restricted Focus Viewer, Eye-Tracking, Visual Attention.

INTRODUCTION

Usability testing is a crucial part of designing Web sites that not only attract users, but enable users to find what they are seeking. A well-designed site can also leave users with a satisfying experience that encourages a return visit. However, there are limitations to the methods that are currently available for such testing.

Laboratory usability testing can be very effective, but often investigates only basic metrics such as task times, error rates, and subjective satisfaction. However, Web site design and usability can be more thoroughly analyzed using additional metrics (Ivory et al. 2001, Mueller and Lockerd 2001, Tarasewich 2002). Simply recording task time does not show what people looked at or clicked on (or tried to click on) while performing that task, although this additional data can often reveal other aspects about users' behavior and the design of the Web site (Goldberg et al. 2002).

One way to more closely monitor users' behaviors during tasks is through think-aloud techniques (Card et al. 2001). This, however, requires extra user effort and the increased overhead of recording and analyzing comments. Another solution is the use of eye-tracking equipment to record users' visual attention (Collewijn 1999). Unfortunately, eye-tracking methods are relatively expensive and difficult to set up and administer, and equipment can only be used with one person at a time. These methods are also limited to the laboratory, and cannot easily replicate the user's home or work environment.

Furthermore, usability testing is often used with relatively small numbers of potential users due to time and cost constraints. While small subject populations can be satisfactory to reveal many problems with a Web site (Nielsen 2000), it can also be argued that using more subjects is preferable, especially for the purposes of statistical analysis. Questionnaires (e.g., Web surveys) can reach larger numbers of actual site users, but may not always be accurate because users may not completely understand the questions, be completely forthcoming, nor even be aware of why they did something (Schiessl et al.).

To address some of these concerns, we have developed software called the Enhanced Restricted Focus Viewer (ERFV). The software can track a user's attention as they view linked images, which could not have been done previously without eye-tracking equipment. The ERFV also tracks the time spent on each page, and the places that a user clicked (even if not on a hyperlink). This paper describes the ERFV software tool, along with an experiment that used the ERFV to track multiple users simultaneously as they searched for information on two versions of a Web site that were equivalent in content but differed in terms of design.

BACKGROUND

This section reviews 1) visual attention and its measurement, 2) recent studies that have tested Web usability using eye-tracking techniques, and 3) work that preceded the ERFV.

Visual Attention and its Measurement

Attention involves the allocation of perceptual or cognitive resources to something at the expense of not allocating them to something else (Harris and Jenkin 2001). Humans have a limited amount of resources available for allocation to different tasks, so everything cannot be attended to at once. People can attend to a modality (vision, hearing, touch, taste, smell), a color, a shape, or a location (Harris and Jenkin 2001). The decision to attend specifically to one of these over the others arises from the task at hand. However, events occurring in the unattended modalities will not go unnoticed (McCormick 1997). For example, when reading a book, a person may ignore most sounds but will respond to their name.

There are various ways to measure a person's focus of attention. One way is by using eye-tracking equipment, which can accurately measure eye movement over areas such as a computer screen. This technology has been used to perform usability testing of different user interfaces and concepts, such as menus (Byrne et al. 1999) and Web sites (e.g., Faraday 2001). Unfortunately, eye-tracking equipment is expensive to purchase, maintain, and use. Eye-tracking involves a complex and intrusive collection of specialized optical and imaging hardware and software (Collewijn 1999). Experimental sessions are often lengthy because of complicated set-up procedures and equipment recalibrations (Chen et al. 2001). Collected data can also be difficult to interpret (Blackwell et al. 2000).

Web Site Usability Testing

Various studies have used eye-tracking equipment to precisely identify what users were looking at as they browsed a Web site. Gagneux, Eglin, and Emptoz (2001) presented an approach to evaluate the quality of a Web site by analyzing the behavior of users as they explore the site. Eye-tracking equipment was used to record the scan path and fixations of users as they searched for information from home pages of online travel agencies. Overall, the study seemed to show a relationship between the distribution of fixations, the time spent on the page, and the quality and structure of the page.

Faraday (2001) used eye-tracking equipment to study how the display characteristics and organization of information on a Web page affected user scanning behavior. Subjects were exposed to different pages that were similar in content but differed in terms of variables such as position and size. Results showed that larger text was more effective at drawing a subject's initial attention than smaller text, and that an image must be larger than text to draw the initial attention of a subject. The study also found that the middle and top of any page were dominant, while the left and bottom were secondary.

Josephson and Holmes (2002) recorded the eye movements of users on three different types of Web pages – a portal page, an advertising page, and a news page – three times over the period of a week. They found that users seem to follow a habitually preferred path across the page each time, although the study was exploratory in nature.

Goldberg et al. (2002) used eye-tracking equipment to monitor user behavior as they browsed pages of a prototype Web portal that contained rectangular user-modifiable "portlets" on each page. Each subject performed a series of tasks on the portlets such as removing a link or customizing a name while their eye movements were tracked. Results showed that subjects seemed to navigate horizontally across columns when two or more columns existed on a page (rather than within a column). Subjects were also biased to first looking at the top and/or left-hand side of a given screen.

The Restricted Focus Viewer

Previous studies have looked at the correlation between a user's gaze and their cursor movements on a graphical interface, although sometimes with inconclusive or conflicting results. Work by Byrne et al. (1999) that used eye-tracking to study visual search during the use of click-down menus found that sometimes a user's gaze followed cursor movement, and sometimes it did not. Research by Chen et al. (2001) suggested a stronger relationship between gaze position and cursor position after testing using a customized Web browser. They suggested that tracking cursor movement might be an inexpensive alternative to eye-tracking systems.

Recently, the Restricted Focus Viewer (RFV) was developed as a software-based alternative to eye-tracking equipment (Blackwell et al. 2000, Jansen et al. 2003). The RFV application[1] was designed to track a participant's eye movements by

[1] The RFV software is available at http://www.csse.monash.edu.au/~tonyj/RFV/

restricting their field of vision to a small focus window, with regions outside of that area successively blurred (see Figure 1). Movement of the focus box is cursor controlled. Recording this movement provides a record of the user's attention over time, which can then be analyzed in relation to the viewed images. The images can be arbitrary combinations of text and graphics. The RFV was developed for the purpose of supporting experiments that studied how people reason with diagrams.

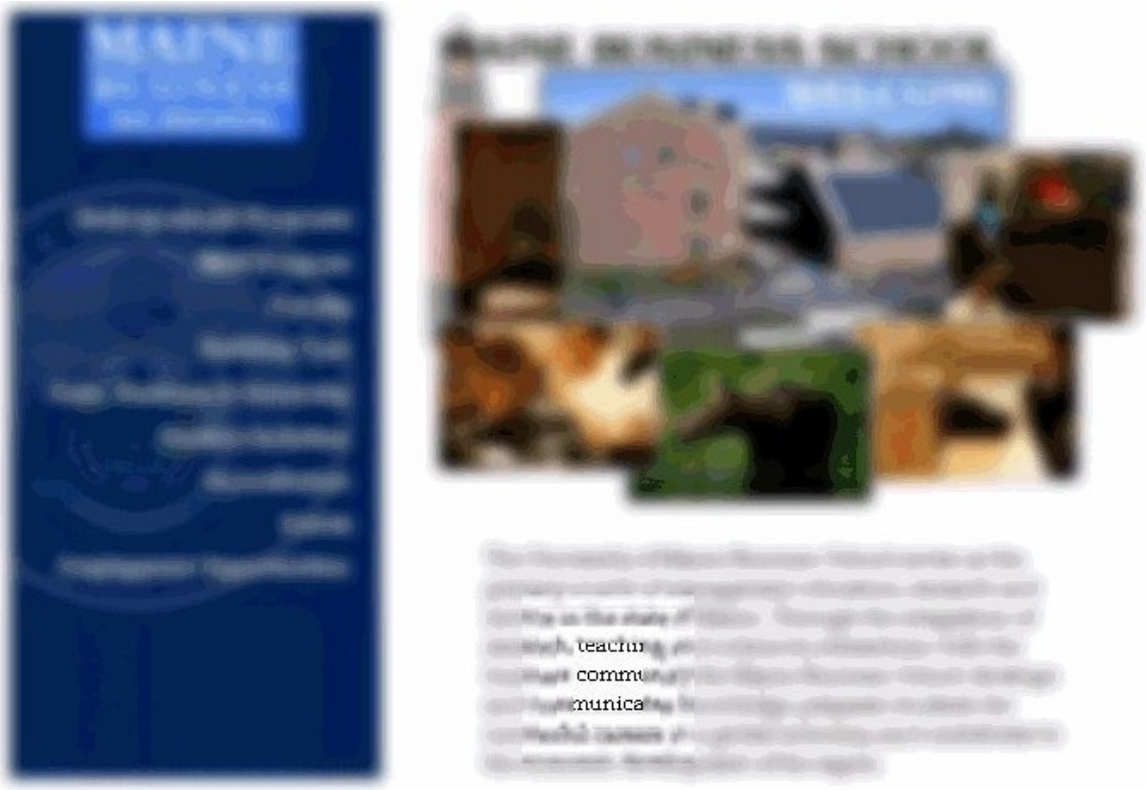

Figure 1. Original blurred home page from experiment with text window at bottom center.

The RFV application consists of an input file, a main application, and an output file. The input file provides setup information, such as focus window size, required diagrams, and subject questions for a particular experiment. Each diagram consists of five image files – four blurred images which make up the different levels of blurring used to restrict the focus of the participant, and an original unaltered image for inside the focus window. Images are normally GIF or JPEG files, and can be made with a graphics editor. The RFV application parses the input file and runs the experiment. Time limits can be placed on how long subjects can spend on each image. All images necessary to run the experiment are loaded into memory at once. Mouse movements and time stamps are recorded to an output file that can be read into a separate Replayer application. This application can play back the real-time path that the user took while viewing a diagram (see Figure 2).

Visual field restrictions such as those used in the RFV have been used in other areas of research. For example, the number of letters that can be processed in one eye fixation has been examined using visual restrictions during text reading experiments (Osaka and Oda 1994). The RFV is unique, however, in that it provides a customizable testing environment to study images using this type of mechanism.

Other researchers have successfully used the RFV to perform experiments. Futrelle and Rumshisky (2001) tracked visual attention of subjects during a study of discourse structure for documents containing both text and graphics. Romero et al. (2002) studied behavior of Java programmers who used an integrated software debugging environment that provided multiple representations of program code, functional visualizations, and output. The RFV tracked use of the different representations.

However, since the original RFV was designed to handle only single static images, the complexity of the testing it can be used for is limited. Our research goal was to extend the capabilities of the RFV to handle hyperlinked images so that it can be used in experiments such as those which analyze a user's search behavior at a Web site.

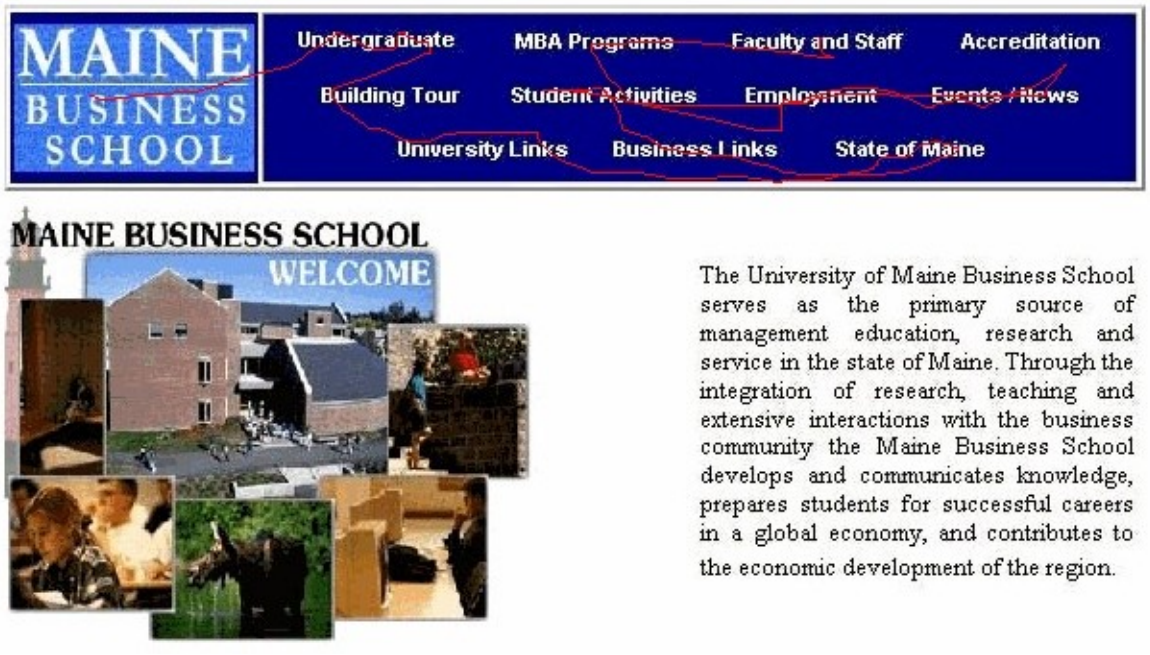

Figure 2. Redesigned home page with trace path of user's visual attention across the menu.

THE ENHANCED RESTRICTED FOCUS VIEWER

The capabilities of the RFV were extended to allow studies with hyperlinked images. With this additional capability, the software can now be used for eye-tracking studies of Web sites. We call the modified software the enhanced RFV (ERFV). Substantial changes were made to the existing RFV code. While the RFV was built using Java, it used an older graphical library (AWT) that was neither easily modifiable nor robust enough for our purposes. Instead of modifying the original code, the ERFV was built from scratch using Swing, keeping the current functionality of the RFV but adding to its capabilities.

The ERFV consists of a Link Editor, two input files, a main application, and an output file. The current version of the ERFV still works with images (rather than HTML), so static files of each Web page must be created for an experiment. To facilitate the addition of hyperlinks to a set of images, a Link Editor was created. With this application, rectangular link areas can be easily drawn and edited on a set of related diagrams. The Link Editor then generates an input file with image and link information for the ERFV.

A second input file contains basic configuration information such as frame size and image size that previously was stored inside each RFV input file. This makes it easier to change the information before running an experiment. The ERFV parses its input file like the original RFV, except that it now presents the user with a list of input files to choose from rather than requiring that the user specify a input file name before executing the program.

As stated previously, all images needed for an experiment with the RFV were loaded at once. However, the original RFV was designed to handle only several images at a time. When testing hyperlinked images, the number of potential images in an experiment can increase dramatically. Loading many pages at once into memory was problematic due to the time necessary to load the images and possible memory size restrictions. Therefore, the ERFV loads and views only one web page at a time.

During an ERFV experiment, when a user passes over a hyperlink, the cursor changes from a cross to a pointing finger. When a user clicks on a hyperlink, they are brought to the corresponding image at the location of that hyperlink.

Like the RFV, the ERFV writes all mouse movements and time stamp data to an output file that can be read by the Replayer. In addition, the ERFV also writes click data to the output file. Click data provides a record of when and where a user pressed the mouse button, even if they did not press it while over a hyperlink. The ERFV output file contains a header and a subject ID and name (all of which the original RFV lacked) to aid in reading the output file in text format if desired. Also added was a list of images used in a particular experiment.

TESTING THE ERFV

Previously, Tarasewich (2002) investigated the use of Web site design complexity and usability metrics. An informational Web site (for a business school) was redesigned with regard to a set of complexity metrics, which measure quantitative aspects of Web site design, with the goal of increasing the site's usability. An experiment was then performed to test the usability of the redesigned site against the original site. Usability was judged not only by traditional metrics such as task time and errors, but also by more Web-specific measures such as the number of links clicked. The redesigned site was perceived as more useable than the original site and allowed users to perform information retrieval tasks better. The study also illustrated how complexity metrics might be used early in the design phase to create a more usable Web site, and how the addition of more Web-specific metrics might contribute to better measurement of overall site usability.

Two sets of linked images were created for use with the ERFV. These images were created from a subset of the actual Web pages used in Tarasewich (2002). Each "Web site" (set of linked images) contained identical information, but the pages were formatted differently (i.e., information content was identical for each, but their designs differed). Two sets of search tasks, each consisting of five questions that could be answered by searching the sites, were created for testing purposes. An example question was "Who is Director of the MBA program?" Each task set was designed to be equivalent in terms of search difficulty.

The study was performed in a computer laboratory at a large university. Subjects completed a brief questionnaire asking for background data, and then performed one set (randomly selected) of five search tasks on one version of the Web site (also randomly selected). The questions were listed on a sheet of a paper. Subjects started on the home page of the site and searched until they found the answer to the question. The answer was always found on a page other than the home page, but subjects were not told this in advance. Subjects wrote the answer below the question, and went back to the home page before searching for the answer to the next question. The ERFV software tracked where the subjects searched and how long the search took. Subjects then performed a second set of five search tasks with the other version of the Web site.

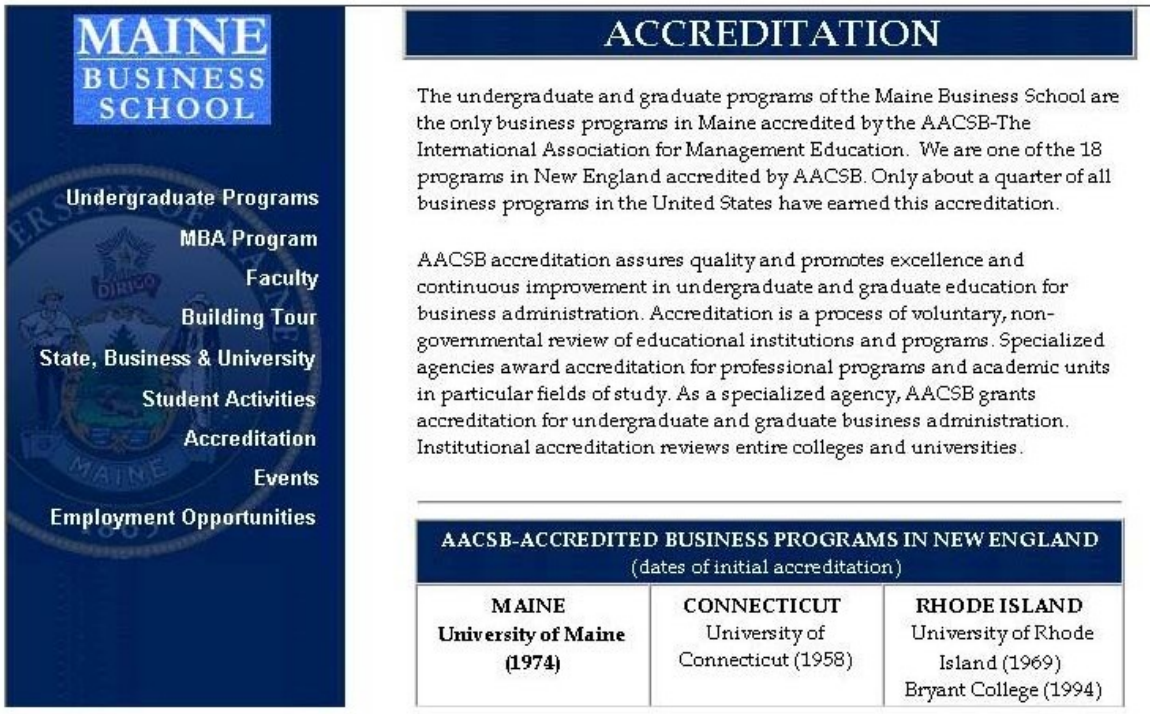

Figure 3. Original accreditation page.

Results

Nineteen subjects from an information science class participated in the experiment during a single class session. Seventeen of the subjects were male, and two were female. All but two of the subjects stated that they browsed the Web daily. The median age of subjects was 21 years. Subjects were not directly compensated for participating in the study, but their class involved designing research studies. Due to space constraints, this section only presents a subset of the results of the experiment, but they illustrate some of the benefits of the ERFV as an analysis tool.

The data collected for each subject was analyzed using the Replayer. Each task generated a set of Web page images with paths of the cursor movement (corresponding to the subjects' visual attention) superimposed (see Figure 1 as an example). These were printed using a color printer and annotated to indicate the direction of the flow of attention. The amount of time the subject spent browsing different regions of each page was also analyzed by watching a real-time playback of the page on the Replayer. Two situations that occurred while users performed their search tasks are discussed here. These are user behavior while on the home pages of the Web sites, and behavior while on the accreditation pages.

Home Page Analysis

Each home page contained a menu of links to other pages in the site, along with a paragraph of general information about the school (see Figures 1 and 2). While most subjects spent all of their time while on the home page on the menu itself, some people (7/19) did spend time reading this text (in addition to going through the menu options) for one or more of the tasks. These seven subjects did this a total of eleven times, usually during the first set of tasks. In two out of the eleven cases, subjects scanned the home page text when they began the experiment (i.e., performed their first search task). Two times subjects searched the home page text during their second set of tasks, specifically for a task that asked for accreditation information about the school. In the other cases, subjects browsed the home page text at some point during the first set of tasks (but not during the first task itself). Only one subject read through the home page text during both the first and second task sets. Reading text on the home page occurred evenly with both versions (original and redesigned) of the Web site.

ACCREDITATION

The undergraduate and graduate programs of the Maine Business School are the only business programs in Maine accredited by the AACSB-The International Association for Management Education. We are one of the 18 programs in New England accredited by AACSB. Only about a quarter of all business programs in the United States have earned this accreditation.

AACSB accreditation assures quality and promotes excellence and continuous improvement in undergraduate and graduate education for business administration. Accreditation is a process of voluntary, non-governmental review of educational institutions and programs. Specialized agencies award accreditation for professional programs and academic units in particular fields of study. As a specialized agency, AACSB grants accreditation for undergraduate and graduate business administration. Institutional accreditation reviews entire colleges and universities.

AACSB- ACCREDITED BUSINESS PROGRAMS IN NEW ENGLAND *(dates of initial accreditation)*	
MAINE	**University of Maine (1974)**
CONNECTICUT	University of Connecticut (1958)
RHODE ISLAND	University of Rhode Island (1969), Bryant College (1994)

Figure 4. Redesigned accreditation page.

Accreditation Page Analysis

One page in each site contained accreditation information about the school and other area schools. These pages (for both original and new designs) were created with summary tables at the bottom of the page (see Figures 3 and 4). There were two different questions that asked subjects to find information related to accreditation. One question asked what year the school was accredited. The other asked which school (not the featured school) was accredited in a certain year. The information needed to answer both of these questions was contained in the summary table (and nowhere else on the accreditation page or on the Web site).

Analysis of the data from the Replayer shows that twelve out of the nineteen total subjects searched areas other than the contents of the tables on one or both of the accreditation pages. This behavior was split approximately equal between the original and redesigned versions of the site. Two subjects showed this behavior on both search tasks that involved finding accreditation information. One subject showed this behavior during the second task, but not the first. The remaining nine exhibited this behavior on the first task concerning accreditation, but not the second. Of these nine, six did this while they were searching for the year the school was accredited.

DISCUSSION

The analysis of data collected from the ERFV points out potential problems and concerns with the design of the Web sites that may not otherwise be found. On the home pages, it was thought that subjects would search through the menu options first, and then go to a page that seemed relevant to what they were seeking. Overall, results showed that most subjects did this, but about a third did not. Those subjects scanned other portions of the home pages as well, specifically the descriptive text. If we had data only on task time, we would be unaware of this situation. Given the fact that some subjects were looking in places that do not seem relevant to the task at hand, we can formulate hypotheses on whether or not the design of the Web site is supporting the users' tasks. We can also try to come up with reasons for this behavior, and possible design fixes.

Most scanning outside of the menu on the home pages happened sometime during the first task set. It seems reasonable for someone to scan the home page when they began the experiment to see what information it contained. But this only accounts for two cases. The other cases may have been caused by poor wording of menu choices, resulting in confusion about the best choice for a given task. For example, one question asked what the American Marketing Association (AMA) sponsors. While the AMA is a student group, some subjects may not have realized this and may not have quickly found a menu choice that made sense. Given this, they may have looked to see if the information was available on the home page.

Since the content of the two Web sites were identical, it was also expected that subjects would learn where information was on the different pages of the sites fairly quickly. However, there were two subjects that searched the home page during the second set of tasks. The two times this happened was for a task that asked for accreditation information about the school. One explanation for this could be that people reasonably expect to find accreditation information about a school on its home page, something that is not taken into account on either of the current designs. While we do not know for certain what subjects expected to find on the home page, this is still something for consideration during a future design revision.

On the accreditation pages, it was thought that the tables would draw users' attention. They were relatively large compared to the rest of the page, had larger text than the rest of the page, and had a large proportion of white space (see Figures 3 and 4). Furthermore, tables often contain summary information. Since most subjects that searched the entire accreditation page did so for the first accreditation question, and not the second, some learning seemed to take place as to what the table contained. However, the fact that so many subjects did not find the answers in the tables without additional searching on the page could indicate a problem with the site design. Perhaps the tables need to be even larger to draw the user's attention, or placed higher up on the page.

As with any testing method or tool, the ERFV has its limitations. There are also many questions that need to be addressed in future research. Users may not actually be browsing a Web site in the same fashion with the ERFV as they would be with a browser. The focus window may also not represent exactly where the user is looking at all times. The ERFV currently cannot handle scrolling, and does not have the complete functionality of a Web browser.

The blurriness levels used in this first experiment with the ERFV were approximated from those used with the RFV. The size and shape of the focus window was also carried forward. Blurriness levels and focus window parameters need to be examined more closely to determine the most appropriate settings for this type of testing.

CONCLUSIONS AND FUTURE RESEARCH

This paper has described the design and implementation of a new software tool called the Enhanced Restricted Focus Viewer that can be used for usability testing of interfaces that contain hyperlinks. In addition, a study was performed using the ERFV to compare the usability of two Web sites in order to demonstrate some of the benefits of the ERFV. The primary benefit that the ERFV provides over many other usability testing methods is its ability to track the path of the user's visual attention. Data about visual attention can reveal information about site design and usability that would not be found by simply looking at task time and error rates. While eye-tracking methods can also provide this same information, the ERFV provides benefits over traditional eye-tracking methods because the ERFV:

- Is a platform independent program that requires no hardware other than a personal computer
- Requires no calibrations to run
- Does not require dedicated laboratory space to run, and can be run anywhere
- Requires less time to run an experiment
- Can be run with multiple subjects simultaneously

Planned improvements to the ERFV and the Replayer will address some of its current limitations. The EFRV needs additional functionalities found in browsers, such as a back button, to more closely approximate the Web environment. A longer-term goal is to create a complete Web browser with the capabilities of the ERFV. This would allow field testing of actual Web sites at multiple locations simultaneously. Another goal is improving the Replayer to increase its analysis capabilities and automate current manual analysis tasks. A direct comparison test between eye-tracking equipment and the ERFV is also required to determine how the two methods differ.

ACKNOWLEDGEMENTS

The authors would like to thank Chen Zhang, who helped modify the original RFV code, and Robert Futrelle, for his ideas and feedback over the course of this study. This work was supported in part by a grant from the Undergraduate Office at Northeastern University.

REFERENCES

1. Blackwell, A. F., Jansen, A. R., and Marriott, K. (2000) Restricted Focus Viewer: A Tool for Tracking Visual Attention, in M. Anderson, P. Cheng & V. Haarslev, (Eds.) *Theory and Application of Diagrams. Lecture Notes in Artificial Intelligence 1889,* Springer-Verlag, 162-177.
2. Byrne, M. D., Anderson, J. R., Douglass, S., and Matessa, M. (1999) Eye Tracking the Visual Search of Click-Down Menus, *Proceedings of the CHI 99*, Pittsburgh, PA.
3. Card, S. K., Pirolli, P., Van Der Wege, M., Morrison, J. B., Reeder, R. W., Schraedley, P. K. and Boshart, J. (2001) Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability, *Proceedings of the CHI 01*, Seattle, WA.
4. Chen, M., Anderson, J. R., and Sohn, M. (2001) What Can a Mouse Cursor Tell Us More? Correlation of Eye/mouse Movements on Web Browsing, *Extended Abstracts of the CHI 01,* Seattle, WA.
5. Collewijn, H. (1999) Eye Movement Recording, in R. H. S. Carpenter and J. G. Robson, (Eds.) *Vision Research: A Practical Guide to Laboratory Methods*, New York: Oxford University Press, 245-285.
6. Faraday, P. (2001) Attending to Web Pages, *Extended Abstracts of the CHI 01*, Seattle, WA.
7. Futrelle, R. P., and Rumshisky, A. (2001) Discourse Structure of Text-Graphics Documents, *Proceedings of the 1st International Symposium on Smart Graphics*.
8. Gagneux, A., Eglin, V., and Emptoz, H. (2001) Quality Approach of Web Documents by an Evaluation of Structure Relevance, *Proceedings of WDA 2001*, 11-14.
9. Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., Wichansky, A. M. (2002) Eye Tracking in Web Search Tasks: Design Implications, *Proceedings of ETRA 2002,* 51-58.
10. Harris, L. R. and Jenkin, M. (2001) Vision and Attention, in M. Jenkin and L. R. Harris, (Eds.) *Vision and Attention*, Springer-Verlag, New York, 1-17.

11. Ivory, M. Y., Sinha, R. R., and Hearst, M. A. (2001) Empirically Validated Web Page Design Metrics, *Proceedings of the CHI 01,* Seattle, WA.
12. Jansen, A. R., Blackwell, A. F., and Marriott, K. (2003) A Tool for Tracking Visual Attention: The Restricted Focus Viewer, *Behavior Research Methods, Instruments, & Computers,* 35, 1, 57-69.
13. Josephson, S. and Holmes, M. E. (2002) Visual Attention to Repeated Internet Images: Testing the Scanpath Theory on the World Wide Web, *Proceedings of ETRA 2002*, 43-49.
14. McCormick, P. A. (1997) Orienting Without Awareness, *Journal of Experimental Psychology. Human Perception and Performance,* 23, 168-180.
15. Mueller, F. and Lockerd, A. (2001) Cheese: Tracking Mouse Movement Activity on Websites, a Tool for User Modeling, *Extended Abstracts of the CHI 01*, Seattle, WA.
16. Nielsen, J. (2000) Why You Only Need to Test with 5 Users, Alertbox column (3/19/2000). Available at: www.useit.com.
17. Osaka, N. and Oda, K. (1994) Moving Window Generator for Reading Experiments, *Behavior Research Methods, Instruments, & Computers* 26, 1, 49-53.
18. Pirolli, P., Card, S. K. and Van Der Wege, M. M. (2001) Visual Information Foraging in a Focus + Context Visualization, *Proceedings of the CHI 01*, Seattle, WA.
19. Romero, P., Cox, R., du Boulay, B., and Lutz, R. (2002) Visual Attention and Representation Switching During Java Program Debugging: A Study Using the Restricted Focus Viewer, *Proceedings of Diagrams 2002,* 221-235.
20. Schiessl, M., Duda, S., Tholke, A., Fischer, R. (2002) Eye Tracking and Its Application in Usability and Media Research. Available at: www.cs.ucl.ac.uk/staff/J.McCarthy/pdf/library/eyetrack/EyeTracking-ResearchApplications.pdf
21. Tarasewich, P. (2002) An Investigation into Web Site Design Complexity and Usability Metrics. To appear in *Quarterly Journal of Electronic Commerce*.