

The Enhanced Restricted Focus Viewer¹

August 2004
Revised, January 2005
Final Revision, February 2005

Peter Tarasewich^{*2}

Marc Pomplun[#]

Stephanie Fillion^{*}

Daniel Broberg[#]

*HCI Laboratory
College of Computer & Information Science
Northeastern University
360 Huntington Avenue, 202WVH
Boston, MA 02115 USA
tarase@ccs.neu.edu, raven82@ccs.neu.edu

#Department of Computer Science
University of Massachusetts Boston
100 Morrissey Boulevard
Boston, MA 02125 USA
marc@cs.umb.edu, dbroberg@cs.umb.edu

¹ An earlier version of this paper appears in the Proceedings of AMCIS 2004.

² Corresponding author

ABSTRACT

The Enhanced Restricted Focus Viewer (ERFV) is a unique software tool 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 to operate 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. A laboratory test comparing the ERFV to a hardware-based eye-tracking system showed that the two methods compare favorably in terms of how well they track a user's visual attention. The usefulness of the ERFV as a usability testing tool was demonstrated through an experiment that evaluated two Web sites that were equivalent in content but differed in terms of design. While several open issues concerning the ERFV still remain, some of these issues are being addressed through ongoing research efforts.

INTRODUCTION

Usability testing, which determines the quality of a system with respect to ease of use, ease of learning, and user satisfaction (Rosson and Carroll, 2002), 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, Sinha, and Hearst, 2001; Mueller and Lockerd, 2001; Tarasewich, 2002). Simply recording task time does not show what people looked at or what people clicked on (or tried to click on) while performing that task, although such additional data would help to reveal other aspects about a user's behavior and the design of the Web site (Goldberg et al., 2002).

One way to more closely monitor a user's behavior during a task is through think-aloud techniques (Card et al., 2001). These techniques, however, requires extra effort on the part of the user, and the increased overhead of recording and analyzing comments. Another solution is the use of eye-tracking equipment to record a user's visual attention (Collewyn, 1999). Unfortunately, eye-tracking methods are relatively expensive and difficult to set up and administer, and equipment can only be used to test 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., 2003).

To address some of these concerns, we have developed a piece of software called the Enhanced Restricted Focus Viewer (ERFV). The software can automatically track a user's attention as they view linked images, something that previously could not have been done 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). The ERFV requires no increased effort on the part of the user for data collection, no additional cost in terms of equipment and testing overhead, and can test multiple users at the same time.

The remainder of this paper is structured as follows. A background section provides a brief review of visual attention, web site usability testing, and research leading up to the ERFV. The next section describes the development and technical specifications of the ERFV. This section is followed by the results of laboratory testing which shows that the ERFV performs very similarly to hardware-based systems in terms of tracking a user's visual attention. Next, the usefulness of the ERFV as a tool for usability testing is demonstrated through the results of an experiment which used the ERFV to compare two different designs of a Web site. Lastly, current research efforts involving the ERFV are discussed.

BACKGROUND

This section briefly reviews 1) visual attention and its measurement, 2) some recent studies that have tested Web site usability using eye-tracking techniques, and 3) work that preceded development of 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 factors over the others arises from the task at hand. However, events occurring in the unattended modalities will not go unnoticed (McCormick, 1997; Moray, 1959). For example, when reading a book, a person may ignore most sounds but will respond if their name is called.

There are two basic ways of describing how a person directs their attention. One is described by the spotlight (Posner, 1980) and zoom-lens (Erikson and St. James, 1986) models, which hypothesize that attention is directed toward a particular zone in space. Another says that a particular object or feature is specifically selected for attention (Duncan, 1984; Egly, Driver, and Rafal, 1994; Weber, Kramer, and Miller, 1997; Moore, Yantis, and Vaughn, 1998). Researchers are still not clear as to which of these models more closely matches human behavior. Delegation of attention may depend on task. Questions also remain as to how people actually select a particular object or location from all its competitors. Of everything that has been tested so far, change seems to be the most reliable stimulus for capturing someone's attention (Yantis, 1998; Yantis and Hillstrom, 1994; Yantis and Jonides, 1984), even better than motion (Hillstrom and Yantis, 1994). But additional research and testing are needed to better

understand how people allocate their attention. In order to accomplish this research, methods to track a person's attention are required.

There are various ways to measure a person's focus of attention. One way is by using eye-tracking equipment, which can accurately measure a person's 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; Hendrickson, 1989) and Web sites (Faraday, 2001; Gagneux, Eglin, and Emptoz, 2001; Goldberg, et al. 2002). 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 (Collewyn, 1999). Experimental sessions are often lengthy because of complicated set-up procedures and equipment recalibrations (Chen, Anderson, and Sohn, 2001). Collected data can also be difficult to interpret (Blackwell, Jansen, and Marriott, 2000).

Web Site Usability Testing Using Eye-Tracking Equipment

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 a set of ten home pages of online travel agencies. After the information search task, data was collected from subjects concerning their general impression of the site, along with their overall perceptions of the time they spent on each page, their comprehension of the links present, and the visibility of text (among other factors). Overall, the study seemed to show a relationship between the distribution of fixations, the time spent on the page, the quality of the page, and the 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 ten 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. If images and text were of a similar size, the study showed text dominated over the similar size images. The study also found that the middle and top of any page were dominant, while the left and bottom were secondary. Anything at the bottom of the page was not guaranteed to be seen by subjects.

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 application 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). Within a portlet, the portlet’s header bar was not usually visited before its contents, making the header an unreliable navigation aide. Subjects were also biased to first looking at the top and/or left-hand side of a given screen.

The Restricted Focus Viewer

While eye-tracking techniques have been used successfully to perform Web site usability testing, disadvantages such as high costs, time-consuming maintenance and calibrations, complicated set-up procedures, intrusiveness during user testing, and difficult interpretation of data have limited their overall use. An alternative that eliminates some or all of these disadvantages could provide a potentially useful and powerful option for the usability testing of Web sites (as well as other types of information systems).

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 his or her cursor movement, and sometimes it did not. Research by Chen, Anderson, and Sohn (2001) suggested a stronger relationship between gaze position and cursor position based on 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, Jansen, and Marriott, 2000; Jansen, Blackwell, and Marriott, 2003). The RFV application³ was designed to track a participant's eye movements by 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 controlled by the cursor. Recording this movement provides a record of the user's attention over time, which can then be analyzed in relation to the images that were viewed. 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.

[Insert Figure 1 here]

The RFV application consists of three major parts: an input file, a main application, and an output file. The input file provides setup information, such as the size of the focus window, the diagrams to be used, and questions to be posed to subjects for a particular experiment. Each diagram to be displayed 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, extracts the necessary information from it, and runs the experiment. Time limits can be placed on how long a subject can spend on each image. All images that are necessary to run the experiment are loaded into memory at once. Mouse movements and time stamps are recorded to an output file that can then be read into a separate RFV Replayer application. This application can play back the real-time path

³ The RFV software is available at <http://www.csse.monash.edu.au/~tonyj/RFV/>

that the user took while viewing a diagram (see Figure 2). The Replayer can also display multiple subjects simultaneously using different colors for each path.

[Insert Figure 2 here]

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; Rayner and Pollatsek, 1989). The RFV is unique, however, in that it provides a customizable testing environment to study images using this type of mechanism. In addition, while other research has considered the potential advantages of tracking users' attention in text-based hyperlinked environments (e.g., Seabrook and Shneiderman, 1989), the RFV tracks visual attention in a graphical environment (which can include text).

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

THE ENHANCED RESTRICTED FOCUS VIEWER

The capabilities of the RFV were extended to allow studies to be performed 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 to achieve this result.

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. Here we detail the major functional changes to the software.

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 of the desired site must be made for an experiment. To facilitate the addition of hyperlinks to a set of images, an additional Link Editor application 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 use with the ERFV. The general format for an input file is shown in Figure 3.

[Insert Figure 3 here]

A second input file contains basic application configuration information such as frame size and image size that used to be stored inside each RFV input file. This file makes it easier to change configuration information before running an experiment. The ERFV parses its input file just like the original RFV, except that it 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 did the RFV).

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. A participant may also not need to view all of the linked images during an experiment. 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. The 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 general format for an output file is shown in Figure 4. 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 the images used in a particular experiment.

[Insert Figure 4 here]

TESTING WITH THE ERFV

For the two experiments described in the remainder of this paper, two sets of linked images were created for use with the Enhanced Restricted Focus Viewer. These images were created from a subset of actual Web pages used in a previous study of Web site usability (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). An “original” version of the site reflected the Web site as it was first found by the author. A “redesigned” version of the site incorporated changes based on good Web design principles (Tarasewich, 2000). Some of the changes made included placing a navigation bar at the top of a page rather than down the left-hand side, limiting the amount of text and graphics contained on any given page, and increasing the symmetry and balance of text and graphics on a page where possible⁴. Each of the “sites” consisted of a home page and six lower level pages. Each page had navigation bar that allowed direct access to the home page or any other page. The size of the focus window used with the ERFV was equal to that used with previous RFV research (Blackwell, Jansen, and Marriott, 2000).

⁴ The redesigned version of the Web site used in the study was not meant to be the final version of the site, but simply an intermediate step that could continue to be refined through further design and testing.

The first experiment shows that the software-based ERFV functions in much the same way as hardware-based eye-tracking equipment in terms of recording a user's visual search paths. The second experiment supports the usefulness of the ERFV as a tool for performing usability testing of Web sites.

COMPARING THE ERFV TO HARDWARE-BASED EYE-TRACKING

In order to determine the validity of ERFV data, we conducted a psychophysical study in which subjects had to browse two sets of Web sites. In this study, both ERFV trajectories and gaze trajectories during free viewing (that is, browsing without the ERFV) were recorded. We compared the patterns obtained under the two different conditions to determine their similarity and thereby evaluate the validity of the ERFV method.

Eight student volunteers from one of the authors' institutions were individually tested. All participants had normal or corrected-to-normal vision. They were naïve with respect to the purpose of the study and were paid for their participation. Eye movements were recorded with the EyeLink-II system by SR Research Ltd.⁵, which operates at a sampling rate of 500 Hz and measures a participant's gaze position with an average error of 0.5 degrees of visual angle. Stimuli were presented on a 21-inch Dell Trinitron monitor with a refresh rate of 85 Hz and a screen resolution of 1024 by 768 pixels. The original and redesigned "Web sites" (sets of linked images described above) were used for the study. For each Web site, seven questions were prepared in such a way that subjects had to visit every page at least once to find all of the answers to the questions.

Each subject completed two consecutive experimental trials, each of which consisted of browsing one of the Web sites in order to answer the corresponding questions. Prior to each trial, a brief calibration procedure was performed in order to maximize the accuracy of gaze position measurement. During this procedure, subjects had to visually track a target that appeared at nine different positions on the screen. Subsequently, the trial was started and the experimenter read the first question to the subject. After the

⁵ Information on EyeLink can be found at <http://www.eyelinkinfo.com/>

subject verbally answered the question correctly, the next question was read, and so on. When all seven questions were answered, the trial terminated.

For each subject, one of the trials was performed using the ERFV and recording the movement of the focus window. The other trial was performed without any viewing restrictions (that is, without the ERFV), but the subject's eye movements were recorded using the EyeLink-II System. The order of presentation of the two Web sites, and the combination of each website with either the ERFV or with eye tracking, were counterbalanced across subjects.

Results

On average, completing the ERFV trials took subjects slightly longer (430 seconds) than completing the free viewing eye-tracking trials (378 seconds); however, this difference did not reach statistical significance ($p > 0.3$). Figure 5 shows three sample trajectories for the ERFV (left column) and the eye-tracking trials (right column) on one of the Web pages used in the study⁶. Since no subject performed both an ERFV trial and an eye-tracking trial on the same Web site, these sample trajectories were taken from different subjects. The information required to answer the question created for this particular Web page was contained in its bottom line of text. The figure illustrates a clear similarity between ERFV and gaze trajectories, but also a considerable variability of scan paths within each of the two groups. The figure also shows that the ERFV tracking is limited by the boundaries of the computer screen, while the eye-tracking equipment can show if a user looks at something other than the screen.

[Insert Figure 5 here]

In order to measure the similarity of ERFV and gaze trajectories, we compared the spatial distribution of visually scanned locations. This comparison was achieved with a technique previously used in Pomplun, Ritter, and Velichkovsky (1996). Each web page was divided into six by four rectangles, each of them subtending 133 by 150 pixels. Across all subjects, the duration of the ERFV

⁶ Note that by design, the eye-tracking system differs from the ERFV in terms of data point collection speed and accuracy.

focus window being located in each of the rectangles during the ERFV trials was summed up. Analogously, the duration of the subjects' gaze position being in each of the squares during the eye tracking trials was summed up separately. As a result, for each of the 14 Web pages used in the study, we obtained two 24-dimensional vectors, one describing the spatial distribution of locations inspected through the ERFV, and the other one indicating the spatial distribution of gaze positions during the eye tracking trials.

For each of the 14 pages, the similarity of the ERFV and gaze distributions was computed. For these calculations, the 24-dimensional ERFV and gaze distribution vectors were normalized and their inner product was calculated and taken as the similarity measure. In geometrical terms, this measure is the cosine of the angle between the two vectors in their 24-dimensional space. A value of 1 indicates identity of the two distributions, and a value of 0 indicates that the two distributions are orthogonal. It is important to notice that due to the high variability of gaze trajectories across subjects, it is impossible for any method (ERFV or other systems) to predict actual gaze trajectories perfectly, i.e. to yield a similarity value of 1. Given the current task and number of subjects, we estimated based on prior experience (e.g., Pomplun, Ritter, and Velichkovsky, 1996) that the best possible prediction could reach a maximum value of approximately 0.7.

As shown in Figure 6, the average similarity of the ERFV and gaze trajectories was computed as 0.56. To provide some baseline measure for this value, we also determined the similarity of random scan paths to the actual gaze trajectories. In other words, we asked to what extent a random pattern could predict the distribution of gaze. If the similarity of random patterns and gaze trajectories also reached the value 0.56, it would demonstrate that the ERFV scan paths were not suitable predictors of actual gaze patterns. However, the average similarity value was calculated as 0.35. A paired t-test across all 14 images revealed that the similarity between ERFV paths and gaze paths was significantly greater than the one between random paths and gaze paths, $t(13) = 3.12$, $p < 0.01$.

[Insert Figure 6 here]

All in all, based on the quantitative and visual comparison of ERFV paths and gaze trajectories, it is justified to say that the ERFV method is providing a good estimate of the areas that subjects inspect when viewing web pages. Therefore, the software-based ERFV seems to be tracking a user's visual attention in approximately the same way as the hardware-based eye-tracking system.

THE VALUE OF THE ERFV FOR WEB SITE USABILITY TESTING

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 performance time and the number of 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.

A second experiment was performed with the ERFV and the original and redesigned Web sites (set of linked images described above). Unlike the previously described experiment, which was performed to test the validity of the ERFV as tool to track a user's visual attention, the purpose of this experiment was to use the ERFV to compare the usability of the two Web sites. In this discussion, we seek to show the benefits of the using ERFV as a tool for Web site usability testing.

The study was performed in a computer laboratory at a large university. Two sets of search tasks, each consisting of five questions that could be answered by searching the sites, were created for testing purposes. For example, one question was "Who is Director of the MBA program?" Each task set was designed to be equal in terms of search difficulty. 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). Ten subjects started with the original version of the Web site, and nine started with the redesigned version. 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. Subjects were asked to work as quickly and accurately as possible, but did not have a direct incentive to enforce this behavior. The answer to each question was always found on a page other than the home page, but subjects were not told this in advance. Subjects then wrote the answer on the sheet of paper 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.

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. The experiment took approximately 45 minutes to administer. Subjects were not directly compensated for participating in the study, but their class involved designing research studies. The next section presents only a subset of the results of the experiment. It is meant to illustrate some of the benefits of the ERFV as an analysis tool, not provide a complete analysis of the two sites.

Results

The data collected with the ERFV 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 images were printed using a color inkjet 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 situations 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 Page Analysis

One page in each site contained accreditation information about the school and other schools in the area. These pages (for both original and new designs) were created with summary tables at the bottom of the page (see Figures 7 and 8). 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).

[Insert Figures 7 and 8 here]

Analysis of the data from the Replayer shows that twelve out of the nineteen 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 that they completed 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 be found by other usability methods. 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 fixes to the Web site design.

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 a subject not knowing 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.

Given that the two Web sites had identical content, 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 who searched the home page during the second set of tasks. Each time this occurred, the subject was

searching for accreditation information about the school. One possible explanation for this is that people reasonably expect to find accreditation information about a school on its home page, which was not taken into account on either of the current designs but could be considered during a future design revision.

On the accreditation pages, it was thought that the tables would draw a user's 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 7 and 8). Furthermore, tables often contain summary information. Given that most subjects who searched the entire accreditation page did so for the first accreditation question, and not the second, there did seem to be some learning that took 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 design of the site. Perhaps the tables need to be even larger to draw the user's attention, or placed higher up on the page.

Based on the results of our two experiments, the ERFV seems to work well as a tool for tracking a user's visual attention, and a tool that can be useful for the usability testing of Web sites. However, this does not mean that the ERFV is without limitations. There are many questions that still need to be addressed and investigated further. The blurriness levels used in this experiment with the ERFV were approximated from those used with the RFV. The size and shape of the focus window was also carried forward. The attributes of the focus window need to be examined more closely to determine the most appropriate settings for this type of testing. Research from studies on human vision, reading, and search behavior is being used to provide insight into the proper settings for the focus window. Given a person's normal field of view, and the way they read text, it may be better to use a focus window that is shaped like a rectangle or oval. The current ERFV levels of blurriness may not best approximate a person's vision (both focused and peripheral), but further testing is needed in these areas.

There is also further work that can be done in terms of comparing the ERFV to traditional eye-tracking equipment. An experiment might be performed where each subject browses two different Web sites, both with and without the ERFV, but using eye-tracking equipment to track visual attention in each case (synchronizing the two environments). The Web sites could be different types (e.g., search engine, e-

commerce, informational) and subjects could perform different types of tasks (e.g., purchasing a product, searching for contact information). The comparison could also be done with different sites of similar structure. With enough subjects, these types of comparison studies could provide additional insight into user behavior with the ERFV, and whether it is different than user behavior without the ERFV. Another idea involves comparing the similarity of gaze trajectories and the ERFV not to a completely random pattern but against a random pattern with more specifically defined limits (e.g., including targets such as menu items). This type of comparison may be more realistic, depending on how the limits are defined.

The ERFV still 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 that can read HTML directly. This would allow field testing of actual Web sites at multiple locations simultaneously. We are currently investigating different methods that would allow us to implement the ERFV capabilities into a Web browser platform.

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. A study was performed that showed that the ERFV seems to track visual attention in approximately the same way as a hardware-based eye-tracking system. In addition, a study using the ERFV to compare two Web site designs demonstrated some of the benefits that the ERFV can provide over other methods as a usability testing tool. The primary benefit that the ERFV provides is its ability to track the path of the user's visual attention. Data about visual attention can reveal information about the design and usability of a site 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, some of which are underway, will address its current limitations. Blurriness levels along with focus window size and shape are being reviewed and will be tested. Creation of a complete Web browser with the capabilities of the ERFV, which would allow multi-location field testing of Web sites, is in progress. Plans are also being made for a larger-scale study comparing the ERFV to traditional eye-tracking equipment. Another goal is improving the Replayer to increase its analysis capabilities and automate some of the tasks currently done by hand.

While this recent work is a significant step, there are even greater possibilities to extend the capabilities of the ERFV. The software might also be modified for use on mobile devices and multi-screen displays. Coordinating the ERFV as it runs on these two different platforms will allow a user's attention to be tracked as it shifts from the mobile device to the display. There is little that scientists understand about how people divide their attention between distractions, and even less about how to measure attention as it moves from place to place. Techniques that accurately measure divided attention will allow researchers to better test and measure the usability of mobile devices in realistic settings. Better insight into usability techniques should help to create more efficient, less frustrating, and safer mobile device and application designs.

ACKNOWLEDGEMENTS

The authors would like to thank Chen Zhang, who helped modify the original RFV code; Robert Futrelle, for his ideas and feedback over the course of this study; three anonymous referees who provided us with constructive feedback on our work; and Guest Editors Scott McCoy, Ping Zhang, and Fiona Nah for putting together this special issue. This work was supported in part by a grant from the Undergraduate Office at Northeastern University.

REFERENCES

- Blackwell, A.F., Jansen, A.R., & 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*. New York: Springer-Verlag, 162-177.
- Byrne, M.D., Anderson, J.R., Douglass, S., & Matessa, M. (1999). Eye tracking the visual search of click-down menus. *Proceedings CHI 1999*. New York: ACM Press, 402-409.
- Card, S.K., Pirolli, P., Van Der Wege, M., Morrison, J.B., Reeder, R.W., Schraedley, P.K. & Boshart, J. (2001). Information scent as a driver of web behavior graphs: Results of a protocol analysis method for web usability. *Proceedings CHI 2001*. New York: ACM Press, 498-505.
- Chen, M., Anderson, J. R., & Sohn, M. (2001). What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing. *Conference companion CHI 2001*. New York: ACM Press, 281-282.
- Collewijn, H. (1999). Eye movement recording. In R.H.S. Carpenter & J.G. Robson (Eds.) *Vision research: A practical guide to laboratory methods*. New York: Oxford University Press, 245-285.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of experimental psychology general*, 113, 501-517.
- Egly, R., Driver, J. & Rafal, R.D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of experimental psychology general*, 123, 161-177.
- Erikson, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and psychophysics*, 40, 225-240.
- Faraday, P. (2001). Attending to web pages. *Extended abstracts CHI 2001*. New York: ACM Press, 159-160.
- Futrelle, R.P., & Rumshisky, A. (2001). Discourse structure of text-graphics documents. *Proceedings Ist international symposium on smart graphics*.
- Gagneux, A., Eglin, V., & Emptoz, H. (2001). Quality approach of web documents by an evaluation of structure relevance. *Proceedings WDA 2001*, 11-14.
- Goldberg, J.H., Stimson, M.J., Lewenstein, M., Scott, N., & Wichansky, A.M. (2002). Eye tracking in web search tasks: design implications. *Proceedings ETRA 2002*, 51-58.
- Harris, L.R. & Jenkin, M. (2001). Vision and attention. In M. Jenkin & L.R. Harris (Eds.), *Vision and attention*. New York: Springer-Verlag, 1-17.
- Hendrickson, J.J. (1989). Performance, preference, and visual scan patterns on a menu-based system: implications for interface design. *Proceedings CHI 1989*. New York: ACM Press, 217-222.
- Hillstrom, A.P., & Yantis, S. (1994). Visual motion and attentional capture. *Perception and psychophysics*, 43, 346-354.
- Ivory, M.Y., Sinha, R.R., & Hearst, M.A. (2001). Empirically validated web page design metrics. *Proceedings CHI 2001*. New York: ACM Press, 53-60.
- Jansen, A.R., Blackwell, A.F., & Marriott, K. (2003). A tool for tracking visual attention: The restricted focus viewer. *Behavior research methods, instruments, & computers*, 35(1), 57-69.
- Josephson, S., & Holmes, M.E. (2002). Visual attention to repeated internet images: Testing the scanpath theory on the world wide web. *Proceedings ETRA 2002*. New York: ACM Press, 43-49.
- McCormick, P.A. (1997). Orienting without awareness. *Journal of experimental psychology. Human perception and performance*, 23, 168-180.

- Moore, C.M., Yantis, S. & Vaughn, B. (1998). Object-based visual selection: evidence from perceptual completion. *Psychological science*, 9, 104-110.
- Moray, N. (1959). Attention in a dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, 11, 56-60.
- Mueller, F., & Lockerd, A. (2001). Cheese: Tracking mouse movement activity on websites, a tool for user modeling. *Extended abstracts CHI 2001*. New York: ACM Press, 279-280.
- Nielsen, J. (2000). Why you only need to test with 5 users. Alertbox column (3/19/2000). Available at: www.useit.com.
- Osaka, N., & Oda, K. (1994). Moving window generator for reading experiments. *Behavior research methods, instruments, & computers*, 26(1), 49-53.
- Pirolli, P., Card, S.K., & Van Der Wege, M.M. (2001). Visual information foraging in a focus + context visualization. *Proceedings CHI 2001*. New York: ACM Press, 506-513.
- Pomplun, M., Velichkovsky, B.M., & Ritter, H. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception* 25(8), 931-948.
- Posner, M.I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32, 3-25.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Upper Saddle River, New Jersey: Prentice-Hall.
- Romero, P., Cox, R., du Boulay, B., & Lutz, R. (2002). Visual attention and representation switching during java program debugging: A study using the restricted focus viewer. *Proceedings diagrams 2002*, 221-235.
- Rosson, M.B., & Carroll, J.M. (2002). *Usability engineering: Scenario-based development of human-computer interaction*. San Diego, CA: Academic Press.
- Schiessl, M., Duda, S., Tholke, A., & Fischer, R. (2003). Eye tracking and its application in usability and media research. Available at: <http://www.cs.ucl.ac.uk/staff/J.McCarthy/pdf/library/eyetrack/EyeTracking-ResearchApplications.pdf>
- Seabrook, H., & Shneiderman, B. (1989). The user interface in a hypertext, multiwindow program browser. *Interacting with computers* 1(3), 299-337.
- Tarasewich, P. (2002). An investigation into web site design complexity and usability metrics. *Quarterly journal of electronic commerce*, forthcoming.
- Tarasewich, P. (2000). Designing usable web sites. *WebNet journal*, 2(2), 23-30.
- Weber, T.A., Kramer, A.F. & Miller, G.A. (1997). Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection, *Biological psychology*, 45, 159-182.
- Yantis, S. (1998). Objects, attention and perceptual experience. In R. Wright (Ed.), *Visual attention*. New York: Oxford University Press, 187-214.
- Yantis, S., & Hillstrom, A.P. (1994). Stimulus driven attentional capture: Evidence from equiluminant visual objects. *Journal of experimental psychology, human perception and performance*, 20, 95-107.
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search, *Journal of experimental psychology, human perception and performance*, 10, 601-621.

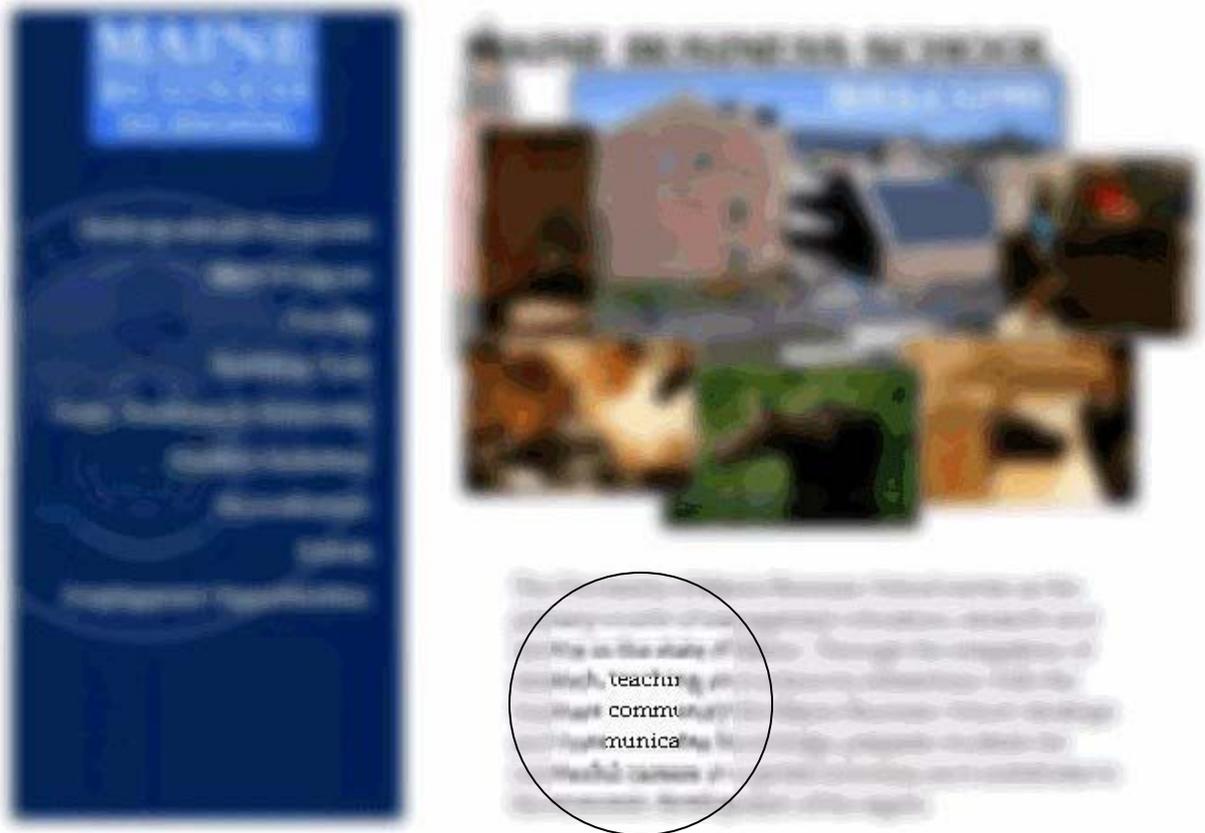
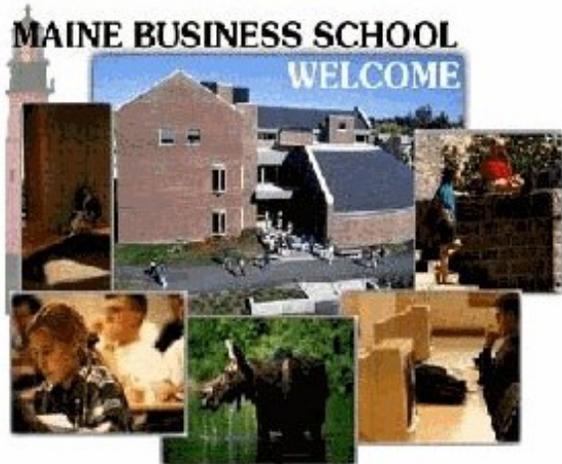


Figure 1. Original Home Page (blurred) with Text Window (Circled)



The University of Maine Business School serves as the primary source of management education, research and service in the state of Maine. Through the integration of research, teaching and extensive interactions with the business community the Maine Business School develops and communicates knowledge, prepares students for successful careers in a global economy, and contributes to the economic development of the region.

Figure 2. Redesigned Home Page with Trace Path Shown on Menu

```
Header Information
Images //five images (four blurred, and one unaltered)//
{HomePage.jpg
  HomePageA.jpg
  HomePageB.jpg
  HomePageC.jpg
  HomePageD.jpg}
Links //for specified image//
{link: student //name of link//
  location: 108, 24, 282, 112 //starting (x,y), width, height//
  dest: StudentPage //destination image//
  link: tour
  location: 89, 25, 181, 113
  dest: TourPage
  link: home
  location: 144, 108, 20, 72
  dest: HomePage}
```

Figure 3. General ERFV Input File Format

```
Subject ID: 0001
Subject Name: Subject One
Images: //set of all used in experiment//
{HomePage.jpg
StudentPage.jpg
TourPage.jpg}
Position (x, y)  Time at Position (seconds)  Clicked Dest.
<630 , 571>      <0.031>      <false>
<624 , 546>      <0.000>      <false>
<620 , 535>      <0.016>      <false>
```

Figure 4. General ERFV Output File Format



Figure 5. Visualization of three ERFV trajectories (left column) and three gaze trajectories (right column) recorded for one of the stimuli. The size of the circles indicates the duration of an ERFV or gaze fixation, respectively.

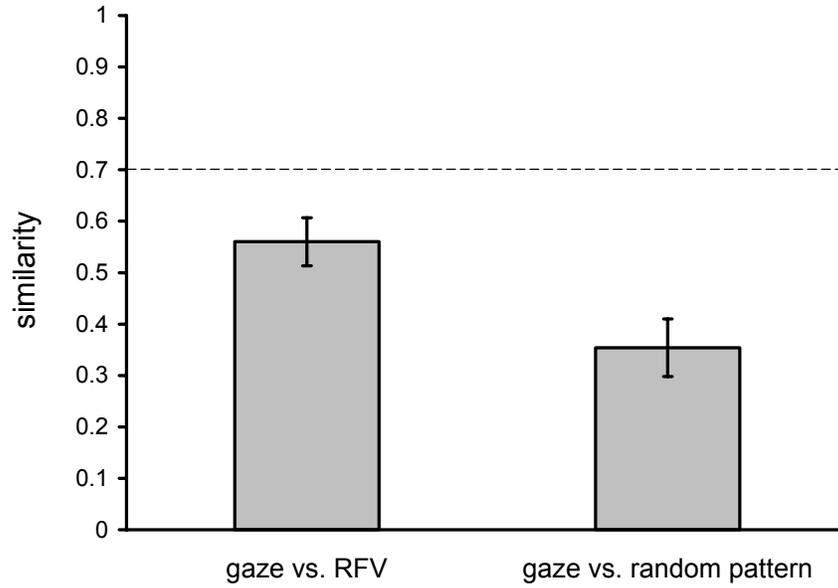


Figure 6. Similarity of the ERFV trajectories and the gaze trajectories (left column) and – as a baseline measure – similarity of the ERFV trajectories and random patterns (right column). Due to the high variability of visual scan paths across subjects, it is estimated that it is impossible in the present context to predict gaze trajectories with a higher accuracy than 0.7 (dotted line).



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 7. Original Accreditation Page



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 <i>(dates of initial accreditation)</i>	
MAINE	University of Maine (1974)
CONNECTICUT	University of Connecticut (1958)
RHODE ISLAND	University of Rhode Island (1969), Bryant College (1994)

Figure 8. Redesigned Accreditation Page