Protocol Considerations for Using Eye-Tracking in Website Usability Testing

Elisabeth Cuddihy, Zhiwei Guan, Judy Ramey*

Eye-tracking systems can enrich a Website usability study by providing an additional method for observing users' behavior. While eye-trackers can provide valuable data, the pros and cons of adding eye-tracking to a usability study need to be considered before designing the study's protocol. This paper discusses the kinds of usability questions that benefit from eye-tracking data and considerations for designing and running the study. Our findings are based on work done in the Laboratory for Usability Testing and Evaluation (LUTE) at the University of Washington, which is equipped with the Eye-gaze Response Computer Aid (ERICA) system controlled by the Gaze Tracker software, both of which were developed by Eye Response Technologies, Inc.

INTRODUCTION

Human beings perceive 80% of information through the visual channel. To gain information, human eyes consciously or unconsciously fixate on objects that might possibly be carrying essential and useful information. When people look at a web page, it takes them within microseconds to make subconscious decisions as to the importance of the various units of information. If the unit of information is perceived as more interesting or important relative to others, eye fixations on it will be considerably longer. Because the human brain processes visual information so quickly, it is very difficult to gain insight about human attention and knowledge processing by using traditional usability testing methods.

Eye-tracking data can provide cues to help us understand human visual behavior. Reading, visual search, scene perception, and visual behavior in driving, sports, and aviation are domains that have been investigated by using eye-tracking. By capturing the location and duration of the human gaze, eye-tracking studies can also enable us to assess the usability of information displays, user interfaces, and other visual designs.

CONSIDERATIONS FOR USING EYE-TRACKING IN USABILITY STUDIES

Eye-tracking can provide valuable data in addition to more traditional data from techniques such as thinking-aloud protocols during a usability study, but the pros and cons must be weighed before collecting the data in order to make the most effective use of it. The first step in deciding whether to use eye-tracking is to determine what questions you want to answer or what issues you want to explore. Is eye-tracking the most effective method for collecting data to address your questions? Also, the limitations of the specific equipment being used to capture eye movement data need to be considered. Is use of eye-tracking feasible for your study? In this paper we describe our experiences based on using the ERICA eye-tracking system from Eye Response Technologies[1] for Website usability studies.

When eye-tracking benefits Web site usability studies

The collection of eye-tracking data adds an extra layer of complexity to a usability study. It is important to first determine whether or not the data needed can be gained through other means or if the added cost of eye-tracking benefits the study. There are many factors that affect the usability of a Website [1, 2]. One aspect is how the visual design and layout of a Web page affect users' ability to quickly and correctly find needed information. While think aloud protocol (TAP) is useful for determining the cognitive thought processes that a user engages in while completing a task, users may not provide enough information during TAP to reliably determine where on the screen they are looking [3].

For usability issues related to Web page design, continuous knowledge of where a person is looking can be very important. For example, knowing if there are specific regions in a page layout that users regularly and repeatedly scan when looking for navigational links,

[1] http://www.eyeresponse.com/

orientation cues, or specific pieces of information can be used to improve information layout. It may be found, for instance, that users constantly scan a navigational menu containing links to high-traffic pages before looking at the primary navigational menus. A usability study may also want to determine why users aren't finding certain information. Eye-tracking data allows us to see whether or not users are looking at an area of interest. This allows us to determine if they are ignoring a region of the screen merely because their gaze has not fixated on any elements within that region or if they have looked at that region long enough to perceive information, but didn't think that it was relevant. It can be difficult to accurately uncover this information from TAP and observation alone. Likewise, during a stimulated recall, the users may not be able to provide information that accurately reflects their thought process during that task.

Eye-tracking data can be particularly beneficial when testing large commercial Websites that often make use of complex grid layouts with repeated elements. Unless the user accurately points with his/her mouse while looking for information on the screen, such page layouts are difficult to describe verbally. Furthermore, the time that it takes to move the eye and perceive a piece of information (100 to 200ms) is much shorter than the time that it takes for users to move and point their mouse (1 to 2 sec) or to speak [4]. Thus, if the usability study requires fine-grained knowledge about where a user is looking to find information, eye-tracking data provides a more accurate method.

Eye-tracking data can also be useful when testing Websites with long text passages. It is difficult for people to read for comprehension when speaking aloud [5]. Although one can give a comprehension test to participants after the use the Website, such a test only tells you what the users are remembering or able to infer, and not the exact passages that they read. If knowing what they are reading, rereading, scanning, or skipping is important, eye-tracking data allows one to collect this data.

Eye-tracking data can also be useful for usability studies of link name wording. For example, if the difference in meaning between two or more links does not provide enough distinction to allow users to determine the best choice, it is helpful to know which links they are considering. Likewise, if the link names do not closely match the terminology familiar to the users, the user may end up considering a set of links while trying to determine which one is the best to select. During TAP, if users provide descriptive information about what they are looking at (e.g., “Ok, so there is motorcycle endorsements and then there is licensing…”), it is easier to figure out what links they are considering. But, if users provide other kinds of information (e.g., “Ok, I'm not sure which to choose, and I'm looking around…”), it may be difficult to understand which links they are considering. In fact, even when users do a good job at providing descriptive information during TAP, they are nonetheless able to read multiple link names in a list faster than they can speak them, and thus they may fail to vocalize all of the choices that they are considering. We have observed some users hesitating for a fraction of a second just before clicking on a link; eye-tracking data revealed that they briefly looked at other possible candidate links during that moment of hesitation [3]. Eye-tracking data can be used to fill in such gaps.

Although their work is outside the realm of traditional usability studies, marketing departments have become increasingly interested in how people interact with Websites. Space on the homepages of commercial Websites is often at a premium, with many internal groups vying for prominent placement on pages. Marketing interests in Web page layout also include visibility of advertising and the effectiveness of product up-sell and cross-sell opportunities (for instance, presenting additional items for possible purchase that are related to the item selected). While click-through rates are the ultimate measure of marketing success, eye-tracking data can help marketing studies determine visibility trends of marketing information [6].

Keeping participants comfortable during the study

Not all eye-tracking systems allow participants to move freely while their eye movement is being tracked [7]. The particular equipment that we have used in the Laboratory for Usability Testing and Evaluation (LUTE) at the University of Washington does require that participants remain still during the eye-tracking portion of a study. Our lab collects eye movement data using the Eye-gaze Response Computer Aid (ERICA) hardware controlled by the Gaze Tracker software, both developed by Eye Response Technologies. The ERICA system uses an infrared light emitting diode to generate two reflections off of the surface of the eye and measures this reflection with a camera sitting underneath the participant's computer screen. These reflections appear in the camera as a bright pupil and a small glint on the iris. The distance and angle between the glint and the bright pupil is used to calibrate and then track the location of the participant's eye gaze. Once this system has been calibrated, any movement of the participant's head reduces the accuracy of the eye-tracking data. Too much movement will cause the system to lose calibrations and fail to collect data.

Despite these limitations, the eye-tracker has been successfully used at LUTE in two pilot usability studies run by graduate students in a Usability Testing course, in an early pilot study for a commercial Website,[2] in a10-person usability study for another commercial Website, and in a study that used usability techniques to investigate the accuracy and reliability of TAP data [3]. The requirement for keeping a participant still while eye-tracking data is being collected has resulted in our using the following protocol during studies:

- Tasks are kept short (under 5 minutes long), in order to reduce the possibility of the participant experiencing physical discomfort from remaining still.
- If multiple tasks use the eye-tracker, the system is recalibrated between each task, not just to maintain better accuracy, but also to allow the participant time to relax.

It is important to note that repeated recalibration of the equipment sharply breaks the participant's sense of emersion within the tasks. Such trade-offs need to be made during study design or when selecting an eye-tracking system for purchase.

There are two types of eye-tracking systems: mounted eye-tracking systems and stationed eye-tracking systems. For mounted eye-tracking systems, the eye-tracking device is mounted on the participant's head, thereby providing unrestricted head movement. Such systems include the Applied Science Laboratory's Video Eye-tracking system and the Sensomotoric MEyeTrack-Silent Vision system. Among stationed eye-tracking systems, some require that head movement is highly limited where as others are able to capture eye movement while the head moves within a larger range of space. Limited head movement eye-tracking systems include the ERICA system and the Arrington Research ViewPoint Eye Tracker. Sustained accuracy can be improved by using a chin rest with these systems. Free head movement eye-tracking systems include the Tobii monitor-integrated eye tracker. Free head movement systems can be used for longer periods of time without the participant experiencing discomfort from remaining still.

Calibrating participants who wear contact lenses or glasses

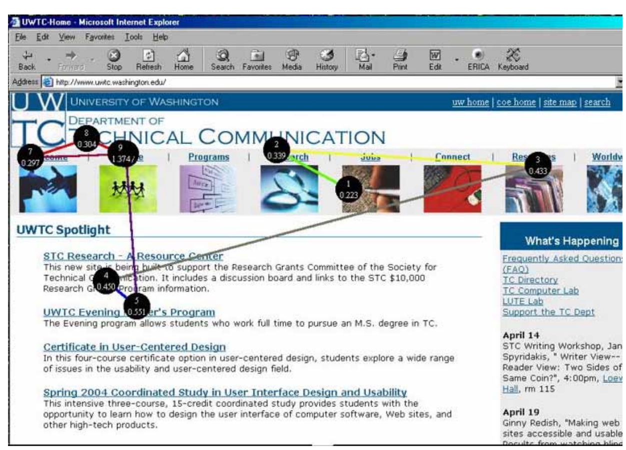
Not all systems provide accurate calibration on participants who wear contact lenses or glasses. This creates an extra concern when recruiting subjects for a usability study.

Analysis considerations during the study

Some eye-tracking systems provide a real-time data feed of the eye-location superimposed over the facilitator's view of the participant's computer screen. One example of such a system is Arrington Research's ViewPoint Eye Tracker. Other eye-tracking systems provide eye-location data only during playback, after the session has completed (the Eye Response Technologies system that we use is of this type).

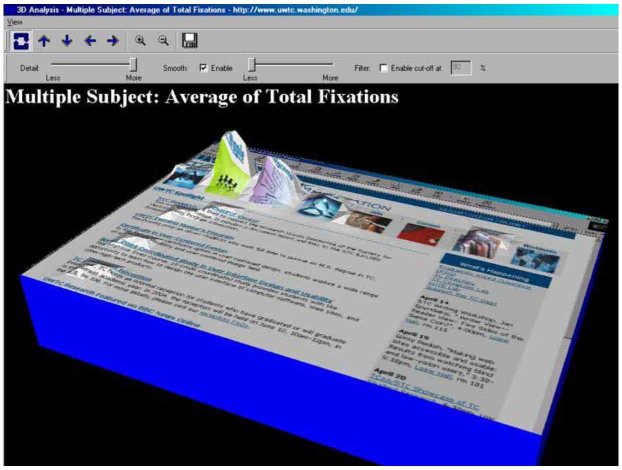
The capabilities of the eye-tracking system will affect how eye-tracking data can be incorporated into a usability study. If real-time data feeds are available, the eye-tracking data becomes another form of observation that the facilitator can leverage during the study. Knowing where the participant is looking, the facilitator can choose to ask the participant questions about his or her visual behaviors.

If the eye-tracking system provides review of the data only after the session has been captured, an additional step is added to the analysis. Appropriate time for this must be scheduled and in an iterative testing situation, extra time needs to be scheduled between participants to review the eye-tracking data. Options provided by some eye-tracking software include video playbacks of the eye position over the Website, images depicting eye-gaze trails over the pages (See Fig. 1-A), topological maps aggregating the viewed areas (See Fig. 1-B), and heat maps of repeatedly viewed areas (See Fig. 1-C). If the study question asked requires synchronizing of data from a thinking-aloud protocol with the eye-tracking data, additional steps to synchronize multiple channels of data are required.

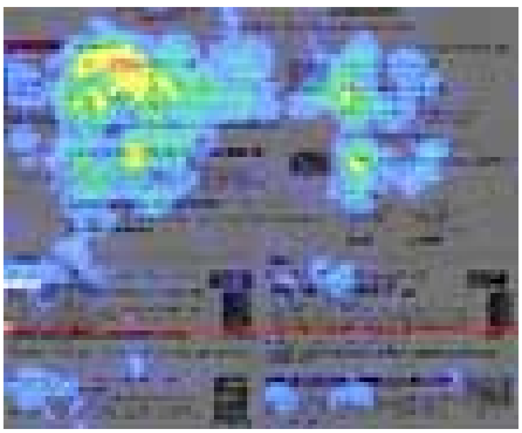


(A)

[2] Although the eye-tracking segment of the study was successful when piloted, it was cut from the actual usability study because the client eliminated some of the test questions before the full study was run.



(B)



(C)

Figure 1: Pictures for gaze-trail, topological map, and heatmap (A: A *GazeTracker* image display of a sequence of eye positions and fixation times; B: A *GazeTracker* aggregate topological map of fixations; C: An *Eyetools* heatmap showing what a group of people view on a page[3])

At LUTE, we have successfully synchronized eye-gaze playback videos with the control room videotape that captured the user's screen, a picture-in-picture of the user, and their audio. This synchronization was performed using Adobe Premiere. The first step for preparing synchronization is to identify the matching points on the videos, pictures, or audio tracks. For instance, an eye-tracking video frame that shows a newly loading Webpage can be matched with the time code for the audio track when a mouse button click is heard. Given identified matching points, lengths between adjacent matching points on audio and video can be compared. If the lengths are not equal, then, the second step is to calculate a synchronization factor to be used as a multiplier to equalize the lengths of data on video and audio channels. This synchronized data from video, audio and pictures can be merged onto a composite video. The synchronization requires the participation of a person with a high degree of technical sophistication who is handy with software tools.

[3] Approval of using this picture has been granted by Eyetools.

DESIGNING STUDY PROTOCOLS TO INCLUDE EYE-TRACKING DATA

Introducing eye-tracking into a usability lab often requires some additional steps in study protocol. This section describes the steps taken at LUTE, using the ERICA system with Gaze Tracker.

Generating Images for Analysis

Some eye-tracking systems generate illustrations and videos of participants' eye gaze data after the eye-tracking session has completed. It can be time consuming to generate these videos (e.g., 5 to 30+ minutes) and the length of time depends on the amount of data to be processed, the number of Web pages visited, and the computational speed of the computer handling the processing. Once produced, these can be viewed as standalone videos.

It is important to note that some eye-tracking software packages, such as GazeTracker, need to access the live Website when generating still images or videos of eye gaze movement. Thus, if the Web pages on the tested website are changed after the study, the eye-tracking software will be unable to correctly match the eye movements with the content on the page. Therefore, if the Website is likely to change, it is important to create all images and videos needed for analysis immediately after running the study. An alternative would be to run the study on a copy of the Website.

Task ordering and system calibration

The length of time for system calibration needs to be taken into account when designing a study protocol. For example, the ERICA system requires approximately 2 to 5 minutes to calibrate. In the calibration process, the facilitator focuses the eye-tracking camera on the participant's eye and then asks the participant to focus on a series of dots as they appear on their computer screen. The system must be recalibrated if the participant moves outside of the calibration range by shifting in their chair, slumping their posture, or moving their head. The time required to calibrate the system interrupts the flow of the study, increasing the artificiality of the study situation and increasing the chance that the participant forgets information needed understand the study tasks. If using an eye-tracking system that requires repeated calibration, the periods of time used for calibration should occur when they will have the least effect on the overall goals of the study. At LUTE, we have performed studies that collect eye-tracking data on only one or two tasks, and we have placed those tasks at the beginning of the study.

Task presentation regarding mouse and keyboard use

When using a limited head movement eye-tracking system, the head movement required to look down at the keyboard or to look at paper task booklet materials on their desk will cause the system to lose calibration. In this situation, tasks should be presented on the computer screen and should avoid keyboard use. The participant should be told to place their hand on or near the mouse before calibration begins.

CONCLUSION

Eye-tracking systems can enrich a Website usability study by providing an additional method for observing users' behavior. Eye-tracking study can help us understand aspects of human behavior and investigate some interface usability that cannot be evaluated by using traditional usability testing methods. This paper reviewed several issues that should be considered when eye-tracking is integrated with a usability study. Our findings are based on our experiences of using eye-tracking in various usability study conducted in the Laboratory for Usability Testing and Evaluation (LUTE) at the University of Washington.

REFERENCES

(1) Goldberg, J.H, Stimson, M.J., Lewenstein, M.L., Scott, N. & Wichansky, A.,(2002). "Eye tracking in web search tasks: design implications," *Proceedings of the Symposium on Eye Tracking Research & Applications*, New Orleans, Louisiana, March 2002, pp. 51-58.

(2) Jacob, R.J., and Karn, K.S. (2003). "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises," *The Minds Eye*, edited by J. Hyona, R. Radach, and H. Deubel. Elsevier Science, Oxford.

(3) Cooke, L. & Cuddihy, E. (2005). "Using Eye Tracking to Address Limitations in Think-Aloud Protocol," to appear in *Proceedings of the IEEE Professional Communication Society International Professional Communication Conference*, Limerick, Ireland, July 2005.

(4) Ware, C. (2003). *Information Visualization: Perception for Design*, Morgan Kaufman Publishers, San Francisco, CA.

(5) Nielsen, J., Clemmensen, T., & Yssing, C. (2002). "Getting access to what goes on in people's heads? Reflections on the think-aloud technique," in *Proceedings of the 2nd Nordic Conference on Human-Computer Interaction*, Aarhus, Denmark (pp. 101-110). New York: ACM Press.

(6) Duchowski, A. (2003). *Eye Tracking Methodology Theory and Practice*. Springer-Verlag, UK.

(7) DeSantis, R. & Zhou, Q. (2005). "A Comparison of Eye Tracking Tools in Usability Testing," in *STC Proceedings of the 52nd Annual Conference*, Seattle, Washington, May 2005.

Elisabeth Cuddihy
Dept. of Technical Communication
Box 352195
University of Washington
Seattle, WA 98195
ecuddihy@u.washington.edu

Elisabeth Cuddihy is a doctoral student in the Department of Technical Communication at the University of Washington. Her research focuses on methodologies supporting improved interface design, and on improving communication during the design of information systems. She is a member of the UW Laboratory for Usability Testing and Evaluation (LUTE) research group on eye tracking in usability.

Zhiwei Guan
Dept. of Technical Communication
Box 352195
University of Washington
Seattle, WA 98195
zguan@u.washington.edu

Zhiwei Guan is a PhD candidate at the Department of Technical Communication, University of Washington. Her research interests include user-centered design, usability testing, and context-aware human-computer interaction. She is a member of the UW Laboratory for Usability Testing and Evaluation (LUTE) research group on eye tracking in usability.

Dr. Judy Ramey
Dept. of Technical Communication
Box 352195
University of Washington
Seattle, WA 98195
206-543-2567
jramey@u.washington.edu

Judy Ramey is professor and chair, Department of Technical Communication, University of Washington, and Director of the UWTC Laboratory for Usability Testing and Evaluation (LUTE). She leads the research group on eye-tracking in usability research. She is a Fellow of STC and winner of STC's Jay R. Gould Award.