Comparison of Eye Movements in Searching for Easy-to-Find and Hard-to-Find Information in a Hierarchically Organized Information Structure

Yoshiko Habuchi* Muneo Kitajima† Haruhiko Takeuchi‡
National Institute of Advanced Industrial Science and Technology (AIST)

Abstract

Finding information by successively selecting hyperlinks on web pages is a typical task performed on websites. A number of web usability studies have provided important insights about how web visitors carry out a search, and have concluded that "following information scent" is the fundamental process involved in the behavior. The purpose of this paper is to explore the relationship between the strength of information scent and web visitors'eye movements. Four web page types with different usability problems were considered. In an eyetracking experiment, eleven participants were asked to find an article on a simulated encyclopedia website by first selecting a heading from among nine provided headings, then selecting the appropriate topic link under the selected heading. The number of eye fixations, the duration of the fixations, and the task completion times were analyzed. The eye-tracking study reported in this paper added further insight to the knowledge gained from traditional web usability studies, in which visitors'performance are measured by the total number of clicks and task completion times. Website visitors'performance will not exhibit any differences in the initial heading selection stage irrespective of whether or not the pages have usability problems. However, performance will deteriorate in terms of the total number of fixations in the subsequent link selection stage when the web page has any kind of usability problem.

CR Categories: H.5.2 [Information Interfaces and Presentations (e.g., HCI)]: User Interface – Evaluation/Methodology—Theory and methods User-centered design Graphical User Interfaces H.5.4 [Information Interfaces and Presentation (e.g., HCI)]: Hypertext/Hypermedia – Navigation, Architectures, Theory, User issues — [H.5.2]: Models and Principles—User/Machine Systems – Human information processing, Human factors

Keywords: information scent, LSA, latent semantic analysis, hierarchical information structure, web usability

1 Introduction

1.1 Eye Movements in Search of Hyperlinks to Go to Other Pages

In the last decade, eye-tracking studies have provided deep insights into how visitors interact with web pages in search of desired information. Nielsen [2006] demonstrated that the eye-scan patterns of website visitors are F-shaped by aggregating a large amount of eye-movement data for web pages provided by a specific search engine. Cutrell et al. [2007] and Guan et al. [2007] conducted a series of detailed studies that examined the relationships between the duration of fixation and the rank order of search results and their representations, *i.e.*, short, medium, and long descriptions. They found that adding information to the search results significantly improved performance for informational tasks but degraded performance for navigational tasks in terms of the accuracy of selecting the correct search result and task times. Using eye-tracking results, they argued that the difference in performance was due to the fact that as the length of description increased, users paid more attention to the description itself and less attention to the URL located at the bottom of the search result that would be useful for assessing the adequacy of the search result for their task.

1.2 Combining Web Usability Studies with Eye-Tracking Studies

These eye-tracking studies have provided an overall understanding of users' eye movement behavior on websites. However, they could not relate the results to the task difficulty caused by usability problems that a web page may have. Users may exhibit wandering eye movements when faced with difficult tasks. One may feel tempted to evaluate the degree of difficulty in terms of eye-movement data. This paper argues, however, that this treatment is not adequate because it would result in a circular argument.

Blackmon et al. [2002] proposed a web usability inspection method called Cognitive Walkthrough for the web (CWW) to detect a variety of usability problems that a website visitor may encounter when navigating through a website to find desired information by successively selecting labeled hyperlinks on intermediate navigational pages. The unique feature of CWW is that it uses a **L**atent **S**emantic **A**nalysis (LSA) semantic space [Landauer and Dumais 1997] for measuring information scent. CWW identifies usability problems based on an estimate of the strength of information scent that the correct hyperlinks emit for a given search target relative to the strength of information scent that the rest of the hyperlinks, *i.e.*, the wrong hyperlinks, emit. An LSA semantic space represents the meaning of words as perceived by a certain user group, *e.g.*, knowledge about words that typical university undergraduates possess, in a vector space with about 300 dimensions. Blackmon et al. [2002; 2003] reported that their experimental participants, university undergraduates, had much difficulty in selecting the correct links on pages with severe usability problems identified by CWW, and some of them could not complete the tasks within the 150-second time limit.

The purpose of this study is to investigate the relationship between task difficulty and eye movements by independently assessing task difficulty, applying a method developed for web usability studies. This paper uses Cognitive Walkthrough for the Web (CWW) [Blackmon et al. 2002; Blackmon et al. 2005; Blackmon et al. 2007] as its method for evaluating task difficulty. The tasks this paper deals with are information search tasks carried out within a hierarchically organized information structure. The task difficulty thus depends primarily on the strength of information scent each

*e-mail: habuchi.y@aist.go.jp
†e-mail: kitajima.muneo@aist.go.jp
‡e-mail: takeuchi.h@aist.go.jp

ETRA 2008, Savannah, Georgia, March 26–28, 2008.
© 2008 ACM 978-1-59593-982-1/08/0003 $5.00

navigational page presents for a given search target.

Four task difficulty types that can be diagnosed by CWW will be considered: (1) no problem, (2) weak scent correct link problem, (3) unfamiliar correct link problem, and (4) competing links nested under competing headings problem.

2 Types of Website Usability Problems Identified by CWW

2.1 Weak Scent Correct Link Problem

A weak scent correct link problem refers to the situation when the correct link is semantically unrelated to the visitor's goal (near-zero LSA cosine), with no other correct links that have a moderate or strong scent. Visitors generally ignore links that they perceive as semantically unrelated to their current goal and have no way of knowing which link is correct. Therefore, a weak scent correct link causes serious problems.

2.2 Unfamiliar Correct Link Problem

An unfamiliar correct link problem can potentially occur whenever target visitors to the website lack sufficient background knowledge to comprehend the link text and to accurately estimate its similarity to their current goals. A short LSA term vector length has been empirically proven to be the most useful CWW index of insufficient background knowledge. Low word frequency in the selected LSA semantic space is, however, another important marker, because typical users cannot recognize or comprehend technical terms and other low-frequency words.

2.3 Competing Links Nested Under Competing Headings Problem

A competing heading problem arises when any heading and its associated links is semantically very similar to the visitor's goal but its associated links do not contain a correct link that leads to accomplishing the visitor's goal. Competing headings problems are liable to be serious because they divert the visitor's attention away from the "correct" heading. Visitors often click several links under a focused-on heading before switching their attention from that heading to another semantically similar heading. Blackmon et al. [2005] found that the best measure of competing heading problems is the total number of attractive links within competing headings, called "competing links nested under competing headings." Many high-scent links increase the visitor's perception that the correct link is somewhere within a high-scent competing heading, so a visitor will probably click many links in that competing heading, even exhaustively clicking relatively unlikely links, before leaving that heading.

3 Experiment

3.1 Task

The participants' task was to find an article about an item in a simulated encyclopedia website [1] . A target word with a short description was provided to the participants. It was expected that participants would build a mental representation of the item by which information scent of navigational pages in the website would be evaluated.

[1]The original experimental website for this task can be found at http://autocww.colorado.edu/~blackmon/Expt040423Home.html. The website was translated into Japanese for this study.

Figure 1: Menus at the second levels. Upper area is a goal description area and bottom area is a menu area. All links under each heading were exposed by a mouse click of the corresponding heading, and visible for only one heading at a time.

3.2 Materials

There were 32 items, 16 Humanities items and 16 Sciences items. By applying CWW [2] , the items were classified in terms of usability problems they might have. The following eight Humanities items were selected for each usability problem type in order to examine whether participants' performance changes depending on the usability problems. **Kite** and **Perspective** had no problem. **Barak** and **Blondin** had only a weak scent correct link problem. **Boehm** and **Helvetia** had only an unfamiliar correct link problem. **Chorale** and **Wilderness Road** had only a competing links nested under competing headings problem (there were four competing links). The rest of the items, including 16 Sciences items, were pooled for further analysis, such as a effect of combined usability problems on user performance.

3.3 Participants

There were 11 participants (four males and seven females). Their ages ranged from 21 to 45 years. They had normal or corrected-to-normal vision. They were regular Internet users and were accustomed to browsing web pages with Internet Explorer.

3.4 Apparatus and Procedure

The Tobii 1750 eye-tracking system was used. Eye movements were recorded at a sampling rate of 50Hz. The experimental website was operated with Internet Explorer 6 running under Windows XP. URL visibility events and mouse movement and click events were recorded.

Participants were instructed to read the brief description about the encyclopedia article they were to find on the simulated encyclopedia website within 130 seconds. To eliminate order effects, the 32 items were divided into two blocks and two order sets.

The dependent variables were the time to complete the task, the total number of fixations, and the total duration of the fixations.

[2]A Japanese LSA semantic space was constructed for this study. It used corpora from Japanese news papers [Mainichi-Shinbunsha 1998] with 47,878 words and 258,512 contexts. The LSA space had 300 dimensions. See Takeuchi et al. [2000] for the details for making Japanese LSA spaces.

4 Results

The performance was divided into two stages: from the beginning until the first click was made, the "initial heading selection stage," during which the nine headings and the goal description were visible, and the rest of the task, the "link selection stage ", see Figure1.

4.1 Initial Heading Selection Stage

A one-way analysis of variance was conducted for the total processing time, the total number of fixations, and the total duration of the fixations. It was found, however, that there was no difference among the conditions.

4.2 Link Selection Stage

A one-way analysis of variance (ANOVA) was conducted for the processing time, the total number of fixations, and the total duration of the fixations.

4.2.1 Total Processing Time and Total Number of Fixations

The ANOVA results indicated that there is a significant difference in the total processing time $F(3, 40) = 7.26, p < .05$ and the total number of fixations $F(3, 40) = 5.73, p < .05$. Multiple comparisons revealed that participants spent more time on the pages with usability problems, *i.e.* on the pages with weak scent correct link problem, competing links nested under competing headings problem, and unfamiliar correct link problem, than on the no problem pages.

4.2.2 Total Duration of the Fixations

Figures 2 depict the total duration of the fixations. The bar on the left is for the initial heading selection stage, and the bar on the right is for the link selection stage.

The total duration of the fixations indicated a significant difference, $F(3, 40) = 5.36, p < .05$. Multiple comparisons indicated significant differences between the no problem pages and the weak scent correct link problem pages, and between the no problem pages and the unfamiliar correct link problem pages.

Figure 3 illustrates another difference between selecting strong-information-scent links versus weak-information-scent links. Each figure aggregates the fixation durations of five participants on second-level pages with each condition, representing them as heat maps. The two maps at the top are from a page with no problem and a page with competing links nested under competing headings. On these pages, information scent of the links is strong, and thus the participants immediately selected the link. On the competing links nested under competing headings page, the participants were likely to select the wrong link with strong information scent, requiring them to re-examine the links. This would increase the number of fixations. However, there is no need in this case to re-build the representation of the goal. In contrast, the two maps at the bottom, a page with weak scent correct link problem and a page with unfamiliar correct link problem reveal repetitive eye movements between the goal description area and menu area. On these pages, the participants could not find an appealing link and thus would try to rebuild the representations of the task goal, the heading and the links, and re-examine the goal description on the page, resulting in longer duration of fixations on those pages with weak information scent than those with strong information scent.

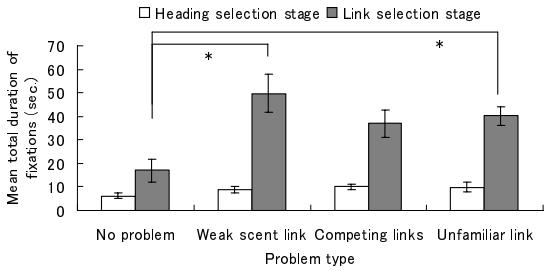


Figure 2: Mean total duration of fixations. Vertical lines depict standard errors of the means.

5 Discussion

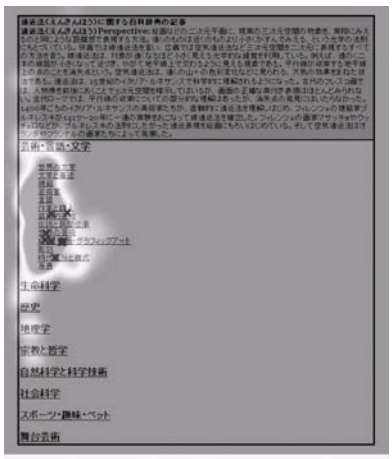
The results indicated no difference in the three performance measures, *i.e.*, the total processing time, the total number of fixations, and the total duration of the fixations, among the usability problem conditions in the initial heading selection stage. However, there were differences between the performance on pages with no problem and problematic pages. It seems that pages with a weak scent problem or unfamiliar problem are more problematic than pages with the other usability problems because they required participants to examine links and goal description *more frequently and longer* than pages with no problems.

The result of the eye-tracking study suggests that the characteristics of visitors' eye movements to select a topic link will vary depending on the nature of the usability problem. Pages with a weak scent problem or an unfamiliar problem seemed to require more time to make a decision. This is presumably because it was necessary for the participants to determine the meaning of the correct link by carefully examining the link label. This situation is different from that of selecting competing links nested under competing headings, for which they could immediately select a competing link because it was a good match with the goal and required less effort to comprehend the meaning of the link.

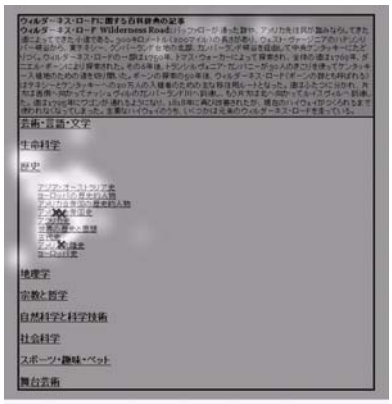
6 Conclusions

This paper sought to determine how website visitors looked at web pages when those pages had usability problems. Three types of website usability problems were considered, the weak scent correct link problem, the unfamiliar correct link problem, and the competing links nested under competing headings problem. The eye-tracking study that this paper conducted adds further insight to the knowledge from traditional web usability studies, in which visitors' performance has been measured by the total number of clicks and by task completion time. Website visitors' performance does not exhibit any differences in the initial heading selection stage irrespective of whether the pages have usability problems or not. However, the performance deteriorates in terms of the total number of fixations in the subsequent link selection stage when the web page has any kind of usability problem. It is worst on pages with a weak scent problem or an unfamiliar problem, which will cause longer fixations due to careful examination of the correct link before selecting it.

This study has attempted to clarify the relationship between visitors' performance and website usability problems. However, the types of usability problems considered in this paper were limited. It is therefore desirable to advance this line of study by now considering pages with combinations of usability problems.



A) Searching for *Perspective* on a no problem page



B) Searching for *Wilderness Road* on a page with competing links nested under competing headings problem

C) Searching for *Barak* on a page with weak scent correct link problem

D) Searching for *Boehm* on a page with unfamiliar correct link problem

Figure 3: Comparison of duration of fixations at the second level page as a function of website usability problems

References

BLACKMON, M. H., POLSON, P. G., KITAJIMA, M., AND LEWIS, C. 2002. Cognitive walkthrough for the web. In *Proceedings of the conference on human factors in computing systems (CHI'2002)*, 463–470.

BLACKMON, M. H., KITAJIMA, M., AND POLSON, P. G. 2003. Repairing usability problems identified by the cognitive walkthrough for the web. In *Proceedings of the conference on human factors in computing systems (CHI'2003)*, 497–504.

BLACKMON, M. H., KITAJIMA, M., AND POLSON, P. G. 2005. Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *Proceedings of the conference on human factors in computing systems (CHI'2005)*, 31–40.

BLACKMON, M. H., MANDALIA, D. R., POLSON, P. G., AND KITAJIMA, M. 2007. Automating usability evaluation: Cognitive walkthrough for the web puts LSA to work on real-world HCI design problems. In *Handbook of Latent Semantic Analysis*, T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, Eds. Lawrence Erlbaum Associates, 345–375.

CUTRELL, E., AND GUAN, Z. 2007. What are you looking for? an eye-tracking study of information usage in web search. In *Proceedings of the conference on human factors in computing systems (CHI'2007)*, 407–416.

GUAN, Z., AND CUTRELL, E. 2007. An eye tracking study of the effect of target rank on web search. In *Proceedings of the conference on human factors in computing systems (CHI'2007)*, 417–420.

LANDAUER, T. K., AND DUMAIS, S. T. 1997. A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review 104*, 211–240.

MAINICHI-SHINBUNSHA. 1998. *CD Mainichi Shinbun '97*.

NIELSEN, J., 2006. F-shaped pattern for reading web content. Jakob Nielsen's Alertbox, http://www.useit.com/alertbox/reading_pattern.html, April 17.

TAKEUCHI, H., KITAJIMA, M., AND AKAMATSU, M. 2000. Extracting knowledge from texts: On the effect of context length. In *Proceedings of the Fourth Asian Fuzzy Systems Symposium (AFSS 2000)*, 482–486.