How Task Types and User Experiences Affect Information-Seeking Behavior on the Web: Using Eye-tracking and Client-side Search Logs

Hitomi SAITO
Aichi University of Education
1 Hirosawa, Igaya-cho,
Kariya-shi, Aichi, Japan
hsaito@auecc.aichi-edu.ac.jp

Hitoshi TERAI
Tokyo Denki University
2-1200 Muzai Gakuendai,
Inzai-shi, Chiba, Japan
terai@sie.dendai.ac.jp

Yuka EGUSA
National Institute for Educational Policy Research
3-2-2 Kasumigaseki,
Chiyoda-ku, Tokyo, Japan
yuka@nier.go.jp

Masao TAKAKU
National Institute for Materials Science
1-2-1 Sengen, Tsukuba–shi,
Ibaraki, Japan
TAKAKU.Masao@nims.go.jp

Makiko MIWA
The Open University of Japan
2-11 Wakaba, Mihama, Chiba, Japan
miwamaki@code.u-air.ac.jp

Noriko KANDO
National Institute of Informatics
2-1-2 Hitotsubashi,
Chiyoda-ku, Tokyo, Japan
kando@nii.ac.jp

ABSTRACT

We investigated what influence task type and user experience had on information-seeking behaviors on the Web by using screen-capture logs and eye-movement data. Five graduate students in library and information science and eleven undergraduate students with other majors performed two different Web searches, a report-writing and a trip-planning task, and their think-aloud protocols, behaviors, and eye movements were recorded. Analyses of the screen-capture logs and eye-movement data revealed that the task type and user experience affected the participants' information-seeking behaviors.

1. INTRODUCTION

Originally developed as a means of searching the Web for information, search engines have become fairly routine and increasingly important in our everyday lives [6]. Considerable research has been done using a variety of methodologies, e.g., analysis of search-engine logs, user experiments, questionnaires, and interviews, to determine how ordinary people use search engines. Because searching for information on the Web is a process of browsing through individual Web pages that are offered by a search engine in response to a query, the ability to support exploratory searches is crucial [4]. This motivated us to pursue quantitative user trials and experiments with the goal of clarifying the exploratory search process by collecting various data from a pre-test questionnaire, client-side search logs, think-aloud protocols, eye-tracking, and post-experiment interviews [7].

One of the main objectives of this study is to deepen our understanding of the relationship between search behavior and the characteristics of different tasks. A number of studies have examined differences in search behavior in dealing with different tasks [3, 8]. In this study, we compare a report-writing task with a trip planning task. These tasks correspond, respectively, to informational and transactional in Broder's taxonomy [1].

We also studied how different levels of knowledge and experience affected the search behaviors of participants conducting exploratory searches. We compared the search behaviors of undergraduate students of various majors with those of graduate students of library and information science. There have been many studies examining the effects of experience on search behaviors [5]. Yet, very few of these studies have analyzed the kind of information that users are searching for. We used eye-tracking data to analyze what students were looking at on the screen, and we then determined whether these viewing tendencies were correlated with differences in experience. The following sections detail our experimental methodology and analytical findings.

2. METHODOLOGY

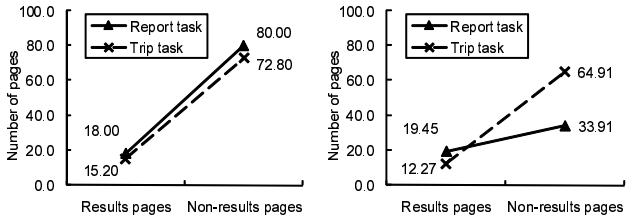
2.1 Participants and Tasks

The participants were 11 undergraduate (ages: between 19 and 21; male: 5, female: 6) and 5 graduate students (ages: between 23 and 28; male: 4, female: 1). The undergraduate students' academic majors included economics, literature, electronics engineering, Spanish, psychology, chemistry, and civil engineering, and the graduate students' were in library and information science.

Two groups differed in terms of web browsers and search engines that they used. Most of the undergraduate students used Internet Explorer 6 (IE6: 10, Firefox: 1). In contrast, almost all of the graduate students used tab browsers (Sleipnir: 2, Firefox: 1, Opera: 1, Others: 1). Almost half of the undergraduate students used Yahoo!Japan as their search engines; the graduate students all used Google.

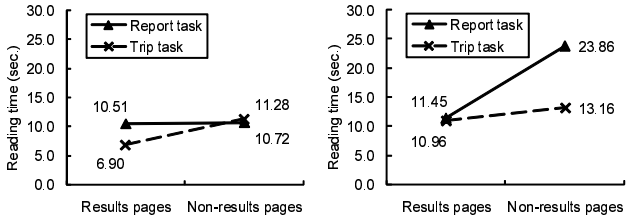
The participants were requested to conduct two different Web searches: a report-writing (report task) and a trip-planning task (trip task). They selected a particular topic for each task based on their own interests because we wanted their search to be exploratory in nature.

SIGIR'09, July 19-23, 2009, Boston, USA.



(a) Graduate students (b) Undergraduate students

Figure 1: Average number of results pages and non-results pages viewed



(a) Graduate students (b) Undergraduate students

Figure 2: Average viewing time for results pages and non-results pages

2.2 Procedure and methods

The participants answered questions in a pre-test questionnaire about their information-seeking experience with Web-search engines. They were instructed to use their favorite search engine in the experiment. They were given a five-minute period to conduct a Web search and practice the "think-aloud" method, in which they orally described their thought processes. Two experimental search tasks (report and trip tasks) were then conducted for fifteen minutes. The order of the searches was counterbalanced between participants. Their eye-movements during the experiments were recorded with an eye-tracking system (EMR-AT VOXER, NAC Image Technology Inc.). They were required to think aloud, and the log data were recorded.

After each search, the participants completed a questionnaire about the degree of difficulty and satisfaction with their searches. We subsequently interviewed them about their information-seeking process while watching screen-capture video of their PC use together with eye movements to facilitate episodic memory retrieval.

3. RESULTS AND DISCUSSION

We next report the results of analysis based on the browser logs, screen-capture video, and the eye-movement data.

3.1 Behavioral Data Analysis

Analysis of Number of Pages and How Long They Were Viewed

We analyzed the number of pages participants viewed and how long they were viewed for two types of tasks and two groups. The pages were classified into two types: results and non-results pages. The results pages were of results or hits that were presented by the search engine in response to queries, and the non-results pages were Web pages other than these.

Table 1: Number of search actions per task

Categories of Action	Report-writing task				Trip-planning task			
	Graduates		Undergraduates		Graduates		Undergraduates	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Search	9.20	3.35	8.00	4.63	7.80	5.89	6.36	5.18
Link	25.80	14.13	19.18	6.66	29.00	14.47	35.73	9.12
Next	0.80	0.84	0.36	0.81	0.20	0.45	0.91	1.14
Back	10.40	9.07	17.18	7.97	10.80	8.04	22.00	14.61
Jump	2.20	1.92	2.36	1.69	3.20	2.68	2.45	1.75
Browse	0.80	1.30	1.82	2.36	0.60	1.34	0.18	0.60
Submit	7.60	12.62	1.27	2.97	4.40	4.34	3.00	3.03
Bookmark	8.00	1.41	4.55	2.16	8.00	6.44	4.55	2.42
Change	43.60	26.38	2.45	5.63	28.40	19.96	3.36	3.11
Close	4.20	3.96	0.36	0.67	6.00	9.82	2.36	1.86

Figure 1 shows the average number of pages viewed for each task by the graduate and the undergraduate students, and Figure 2 shows the average viewing time per page for each task. First, we found that the graduates looked at significantly more non-results pages than results pages for both tasks in terms of the number of pages ($F(2,16) = 73.86, p < .01$). The undergraduates also looked at significantly more non-results pages than results pages for both tasks ($F(1,43) = 6.39, p < .05$; $F(1,43) = 107.82, p < .01$). The undergraduates looked at a significantly greater number of non-results pages particularly for the trip task ($F(1,43) = 43.39, p < .01$).

We found that the graduates showed no significant task-specific differences in the number of pages viewed during the search time. However, in the report task, the undergraduates spent significantly longer browsing non-results pages compared with the results pages ($F(1,43) = 7.60, p < .01$).

Analysis of Web-search Categories

We analyzed the number of search-related actions for the two tasks and two groups. We defined 10 categories of action to analyze user behavior on the Web. Table 1 lists the averages and standard deviations for the number of actions carried out for each task by the graduate and undergraduate students. The 2-factor analysis results revealed significant differences between the two groups for the Search, Next, Jump, and Browse actions. The undergraduates were significantly more likely to click links during the trip task than during the report task. The undergraduates tended to return to previous pages more often than the graduates, but the graduates tended to submit more forms than the undergraduates. The graduates bookmarked significantly more pages and switched to different tabs or windows significantly more often than the undergraduates. We also found that the graduates switched to different tabs or windows significantly more often for the report task than for the trip. The graduates also tended to close windows or tabs more often than the undergraduates. We noticed that both groups tended to close more windows and tabs in doing the trip task than during the report.

Summary of Behavioral Data Analysis

First, we consider task-specific differences in search behavior. We found that there were no task-specific differences commonly observed both groups. However, the two groups did share certain characteristics in the number of results pages they looked at and in their actions of Search, Next, Jump, and Browse. This suggests that both groups pur-

Figure 3: Location of blocks in *lookzone*

sued similar search procedures, particularly with respect to results pages, regardless of the type of task or level of experience.

After that, we considered the differences in search behavior that could be attributed to different levels of experience. In contrast to the graduates who tended to look at about the same number of pages for the same length of time in both tasks, we found that the undergraduates examined non-results pages for longer periods when doing the report task. Moreover, the search action data revealed that the graduates tended to change between windows and tabs and close them more frequently than the undergraduates. This reflected a tendency on the part of the graduates to search in parallel by frequently switching back and forth between a number of pages that were open at the same time. By contrast, the undergraduates were more likely to search sequentially by using the Link and Back functions to go back and forth between links.

3.2 Analysis of Eye-Movement Data

This subsection explains our analysis of the eye-movement data. Because Web searches involved dynamic changes in screen (scrolling and page transitions), no thorough assessment of search behavior could be based sorely up on quantitative analysis using stationary point coordinates. Tagging was also needed to determine exactly what the participants were looking at on the screen. We therefore employed a results page with a relatively simple structure in our investigations.

Definition of Lookzone

We defined 22 *lookzone* blocks on the page to classify exactly where participants were looking on the page. Figure 3 shows the 22 lookzone blocks superimposed on the Google-search results page. These same lookzone block items were applied to the search-engine pages used by the participants in this study.

Next, we captured images from the eye-tracking data of the participants at 0.5-second intervals, beginning as soon as the results pages were presented to them. We then manually tagged where the eye-gaze points in the extracted images fell within the lookzone. On the basis of this tagged data, we analyzed the number of eye-gaze points per block, and the eye-gaze points and number of clicks per search-result ranking.

Table 2: Average number of eye-gaze points for each Lookzone block

Lookzone	Report-writing task		Trip-planning task	
	Graduates	Undergraduates	Graduates	Undergraduates
1 Title bar	0.40	3.78	0.80	1.00
2 Menu	1.80	0.22	0.00	0.11
3 Bookmark	0.00	3.78	0.20	0.00
4 Tool bar	0.40	1.78	0.40	1.22
5 URL bar	0.40	0.78	0.00	0.11
6 Search bar	4.00	0.00	4.00	0.00
7 Search bar button	0.20	0.00	0.20	0.00
8 Tab	10.20	8.11	6.00	9.22
9 Link for services	2.40	17.67	2.20	5.00
10 Query box	5.40	36.89	3.00	12.56
11 Search button	0.00	0.89	0.20	0.67
12 Scroll bar	0.60	0.11	0.00	0.00
13 Number of hits	0.00	0.44	0.60	0.00
14 Sponsor link	0.00	6.67	11.40	12.11
15 Spell check	0.00	0.00	0.20	0.00
16 Title	38.80	60.67	39.20	42.11
17 Snippet	70.00	91.11	28.40	37.00
18 URL	16.60	40.89	12.40	15.44
19 Related search	1.20	3.00	1.20	2.56
20 Link for next page	1.00	0.78	1.00	0.78
21 Find in a page	0.00	0.00	0.00	0.00
22 Status bar	0.00	1.78	0.00	0.00
Out of lookzone	19.60	52.89	17.00	18.22
Lack of eye position	12.00	83.44	7.20	70.78

Analysis of Eye Gaze Points for Each Lookzone Block

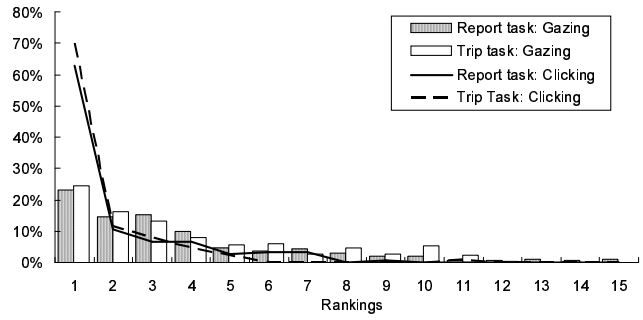
Table 2 shows the average number of eye gaze points for each Lookzone block broken down by task for the two groups. The category "Out of lookzone" is the number of eye-gaze points elsewhere on the page besides the 22 lookzone blocks, and the category "Lack of eye position" is the number of images in which the eye-gaze points could not be determined. As we can see from the table, most of the eye-gaze points on the results pages were focused on information pertaining to the hit pages (titles, snippets, and URLs).

The 2-factor analysis of variance results revealed clear differences between the two groups of students for a number of lookzone blocks. The undergraduates exhibited significantly more eye-gaze points on the tool bar ($F(1, 12) = 12.40$, $p < .01$). They also tended to focus more attention on the query box ($F(1, 12) = 3.87$, $p < .10$) and the search button ($F(1, 12) = 4.72$, $p < .10$). The graduates were significantly more prone to look at the search bar ($F(1, 10) = 6.02$, $p < .05$).

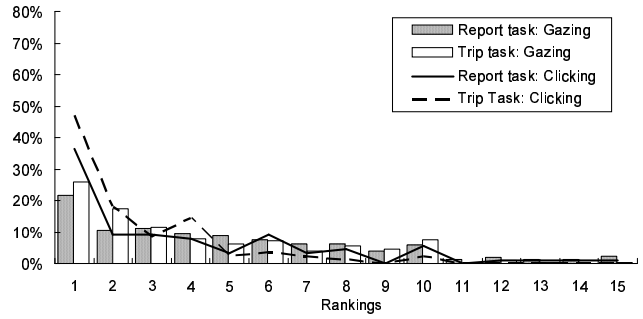
We also found differences between the tasks for a number of lookzone blocks. There were significantly more eye-gaze points on the scroll bars ($F(1, 12) = 4.77$, $p < .05$) and snippets ($F(1, 12) = 8.89$, $p < .05$) for the report task as opposed to the trip task. By contrast, in the trip task, students were more inclined to look at the sponsor's information ($F(1, 12) = 5.95$, $p < .05$).

Analysis of Eye-Gaze Points and Clicks for Each Ranking

As previously noted, there was a clear tendency for students to focus on the titles, snippets, and URLs of the hits displayed on the results pages. We consequently grouped the eye-gaze points on titles, snippets, and URLs and assigned rankings, then analyzed which rankings attracted the most views. We next extracted actual click-ranking data from the search-log data and investigated the relationship between eye-gaze points and clicks.



(a) Graduate students



(b) Undergraduate students

Figure 4: Percentages of clicks and eye-gaze points for each ranking

Figure 4 plots the percentages of clicks and eye fixations for all rankings for the two groups of students. One can see that the percentages are greatest for Rank 1 for both graduate and undergraduate students. These results are similar to those reported in previous studies [2]. After Carrying out a 2-factor analysis of variance on all ranks as a function of clicks, we found that the number differed significantly depending on the level of experience and type of task for Ranks 1, 6, 8, and 10. First, we found that the graduates tended to select rank 1 much more frequently than the undergraduates in doing the report task ($F(1,12) = 3.18, p < .10$). The graduates also selected rank 1 more often when doing the report task as opposed to the trip task ($F(1,12) = 5.68, p < .05$). Moreover, we found that both groups of students tended to select rank 6 more often for report tasks than for trip tasks($F(1,12) = 3.85, p < .10$). The undergraduates chose ranks 8 and 10 more often than the graduates did for both tasks (rank8:$F(1,12) = 5.36, p < .05$, rank10:$F(1,12) = 4.20, p < .10$). This reveals that the graduates tended to favorably assess and choose higher ranking pages from the search results, while the undergraduates tended to choose pages ranked 5 and below.

Next, we did a 2-factor analysis of variance on eye-gaze points for all rankings and found that the main effect of the task was quite significant in ranks 4 and 7 (rank4: $F(1,12) = 5.10, p < .05$, rank7: $F(1,12) = 6.12, p < .05$). This demonstrated that both graduates and undergraduates tended to examine lower ranking pages when conducting report tasks.

Summary of Analysis of Eye-Movement Data

We investigated to see if the eye-gaze points for each look-zone block in the eye-movement data, the eye-gaze points for each rank, and the number of clicks were correlated in any way with the different tasks and levels of experience. Analysis of the eye-movement data did reveal any task-specific differences. We found that the students in the report task perused from higher to lower ranking pages and scrutinized snippets revealing the content of the pages. By contrast, the participants had much less inclination to look at lower ranking pages for the trip task and focused more attention on the sponsors' information. This means that the type of task clearly did affect the information that was regarded as important and how students viewed the rankings. We found a clear tendency in graduates to look at the search bar at the top of the browser and to select more rank 1 pages. In contrast, the undergraduates tended to look more at the query boxes and search buttons at the top and bottom of the results page. Moreover, they were more likely to choose lower ranking pages. These characteristics observed in two groups suggests that the level of experience was clearly reflected in different search strategies and in the criteria for selecting ranked pages.

4. CONCLUSIONS

We studied how different tasks and levels of experience affect the behavior of students searching for information on the Web. Based on our analysis of search behaviors and eye-movement data, we found that the type of task and level of experience did indeed affect their search behaviors.

However, there were too few participants to allow reliable conclusions. In the future work, we will conduct more large-scale experiments to verify our findings.

5. REFERENCES

[1] A. Broder. A taxonomy of web search. *SIGIR Forum*, 36(2):3–10, 2002.

[2] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. G. ay. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 154–161, New York, NY, USA, 2005. ACM.

[3] M. Kellar, C. Watters, and M. Shepherd. A field study characterizing web-based information-seeking tasks. *Journal of the American Society for Information Science and Technology*, 58(7):999–1018, 2007.

[4] G. Marchionini. Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.

[5] J. L. Moore, S. Erdelez, and W. He. The search experience variable in information behavior research. *Journal of the American Society for Information Science and Technology*, 58(10):1529–1546, 2007.

[6] A. Spink and B. J. Jansen. *Web Search: Public Searching of the Web*. Kluwer Academic Publishers, Dordrecht, the Netherlands, 2004.

[7] H. Terai, H. Saito, M. Takaku, Y. Egusa, M. Miwa, and N. Kando. Differences between informational and transactional tasks in information seeking on the web. In *Proceedings of the Second Symposium on Information Interaction in Context (IIiX 2008)*, pages 152–159, 2008.

[8] A. Thatcher. Web search strategies: The influence of web experience and task type. *Information Processing & Management*, 44(3):1308–1329, 2008.