
A method to capture information encountering embedded in exploratory Web searches

[Makiko Miwa](#)

The Open University of Japan: 2-11 Wakaba, Mihama-ku, Chiba, Japan

[Yuka Egusa](#)

National Institute for Educational Policy Research: 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, Japan

[Hitomi Saito](#)

Aichi University of Education: 1 Hirosawa, Igaya-cho, Kariya-shi, Aichi, Japan

[Masao Takaku](#)

National Institute for Materials Science: 1-2-1 Sengen, Tsukuba, Ibaraki, Japan

[Hitoshi Terai](#)

Nagoya University: Furo-cho, Chikusa-ku, Nagoya, Japan

[Noriko Kando](#)

National Institute of Informatics: 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

Abstract

Introduction. This paper reports on a method for closely observing, eliciting and visualizing exploratory search processes with embedded information encountering in context.

Method. Exploratory Web search processes of sixteen participants for report task and trip task with thinking aloud in Japanese were captured in video, together with their eye-movements. In post-search interviews, participants' thoughts, feelings and intentions for particular actions at each point of the Web search processes were elicited while showing video-recorded search processes with and without eye-movement.

Analysis. Think-aloud and interview data were recorded together with the type of pages viewed, twenty-two eye-gaze points on search results pages, ten search-related actions and the link depth into the annotation tool using time-stamps for sequencing. A qualitative content analysis of the integrated datasets identified thirty instances of information encountering.

Results. Information encountering sometimes led to a temporary deviation from the initial task, but may have resulted in a change of the topic of the paper and/or destination of the trip.

Conclusions. Results lead to a reconsideration of the definition of information encountering, which may include learning by searching. Showing eye-movement data during the post-search interviews helped participants recall their thoughts at each moment during the search, including information encountering in the context.

Introduction

Erdelez (1997), Williamson (1998) and Toms (2000) brought attention to the concept of opportunistic discovery of information as a common information behaviour in information-rich environments. Erdelez used interviews and surveys of critical incident techniques to capture participants' *'memorable experience of unexpected discovery of useful or interesting information'* (Erdelez 2005: 179). Erdelez defined information encountering as *'an instance of accidental discovery of information during an active search for some other information'* (Erdelez 2005: 180). She identified the Web as an information-rich environment for information encountering, particularly for those who had infrequent experiences of information encountering (Erdelez 2000).

Williamson (1998) used in-depth interviews with elderly women to capture incidental information acquisition, based on the assumption that background information needs transform to foreground information needs only when information is discovered. Based on the study, she identified the process of the phenomenon of incidental information acquisition, which is more likely to occur in intimate personal networks.

Toms (2000) conducted an experimental study of browsing by comparing two interfaces (a list of suggested items in a hierarchical menu and a keyword search function) for newspaper articles, using observation, questionnaire and retrospective verbal protocols. She also compared two different tasks: one with a specified goal of answering predefined questions and the other with no predefined goal. She found that participants without a predefined goal, who used the interface with a suggested list of items, tended to encounter information through browsing by recognizing affordances of texts.

Several methods have been applied to capture information encountering, including critical incident interview or questionnaire by Erdelez (1997) and experiments under controlled environments by Toms (2000). Rubin *et al.* (2010) collected blog messages relevant to serendipity using Google Blog. The content of collected messages was analysed for the acquisition of accounts of serendipitous experiences. They reported the strengths of the method as being the non-obtrusive nature and free availability of useful data. However, the above-mentioned methods rely on retrospective recollections of information-encountering experiences, which may not provide searchers' micro-level accounts of search contexts, such as external behaviour captured by client-side search logs and internal processes of thoughts and feelings before, during and after the incidence of information encountering.

This paper reports on a method for closely observing, eliciting and visualizing real-time micro-level accounts of exploratory search processes with embedded information encountering in the context. The analysis reported here forms part of a larger study that aims to understand better exploratory search processes through the close observation of Web search processes for ill-defined search goals, conducted in a laboratory setting.

Background

Many researchers have analysed Web search logs to identify characteristics of Web search processes (Silverstein *et al.* 1999; Spink and Jansen (2004). These logs give useful insights on a variety of externally observable Web search actions such as the sequence of queries, query duplication, correlation of log entries, interaction of terms within queries and lengths of queries and search sessions. However, we may not be able to obtain a better understanding of Web search behaviour without capturing searchers' internal processes, such as intentions, thoughts and feelings and connecting them with the search logs.

Recently, many researchers have collected empirical and descriptive evidence of Web search processes by employing various techniques, such as survey, interview, questionnaire, video capturing, observation and think-aloud, to collect qualitative and quantitative data (Martzoukou 2005; Choi 2010). However, some searchers may not articulate their thoughts and feelings if they are not accustomed to thinking aloud. They may not be able to recall their thoughts and feelings during Web search processes at the time of the post-search interviews. Thus, it is not easy to capture internal processes that occur concomitantly and uninterruptedly with externally observable behaviour or search logs.

With this in mind, we have been looking for ways of connecting searchers' internal processes with their external behaviour (Miwa and Kando 2007; Egusa *et al.* 2010). Through these studies, we came up with the idea that showing video-recorded search process screens with superimposed eye-movement data may help

searchers retrieve their episodic memory, the long-term memory of temporally unique events ([Buckner and Barch 1999](#)).

We asked the following research questions as guides to the analysis of this paper:

1. How are information-encountering incidents embedded in exploratory search processes?
2. Does watching recorded eye movements during the exploratory search process facilitate episodic memory retrieval of information encountering?

Method

Participants

The participants were eleven undergraduate students of various majors (aged 19–21; 5 male, 6 female) and five graduate students of library and information science (aged 23–28; 4 male, 1 female). When we recruited them, we associated students' search expertise with their background and assumed undergraduate students to be novice and graduate students to be expert searchers to compare the differences in their search processes ([Moore et al. 2007](#)).

Tasks

The participants were asked to conduct two different types of Web searches: gathering information for writing a research paper (report task), which we interpreted as a type of informational task and planning a trip (trip task) as a type of navigational search ([Broder 2002](#)). In the report task, the participants were asked to gather information from multiple Web pages on any topic of world history, a requisite subject for every high-school student in Japan. In the trip task, participants were required to gather information from multiple Web pages to plan a domestic trip with family members or friends. Participants selected a particular topic for the paper in the report task and the destination, season and accompanying people in the trip task, based on their own interests. We intended to encourage natural exploratory search behaviour even in a laboratory setting.

The report task was presented as follows.

Assume that you are an undergraduate student who is required to submit a term paper on your own topic of interest for a world history class. Your topic is _____ (participant decides his/her own topic.). Please collect related information for the term paper through the Web. You have 15 minutes to accomplish this. Start your search for useful sites. Please add pages to the browser's bookmarks if you find useful information.

The trip task was presented as follows.

You plan to take a trip with _____ (type of people). The time for the trip will be in _____ (season of a year) and you will go to _____ (location). Assume that you need to give your trip partners useful information, such as transit points, accommodation, locations, events, etc. Please collect related information through the Web to prepare for the trip. You have 15 minutes to accomplish this. Start your search for useful sites. Please add pages to the browser's bookmarks if you find useful information.

Data collection

Figure 1 shows the sequence of the experiment. After signing a consent form, participants were first asked to fill in a questionnaire concerning their Web search experience. Before starting the two experimental Web search tasks, participants were asked to conduct a Web search for five minutes to practice a 'think-aloud' technique in which they orally described their thoughts and feelings while searching for information. We used a 19-inch liquid crystal display with 1024 x 786 resolution and the Firefox Web browser on Windows XP. For eye-movement analysis, we gave instructions to use the full-window mode and to use the New Tab function rather than the New Window function on the Web browser. The Web browser was set up so that the default page was blank and the bookmarks included Yahoo! Japan and Google as common search engines. Participants were instructed to use their favourite search engine.

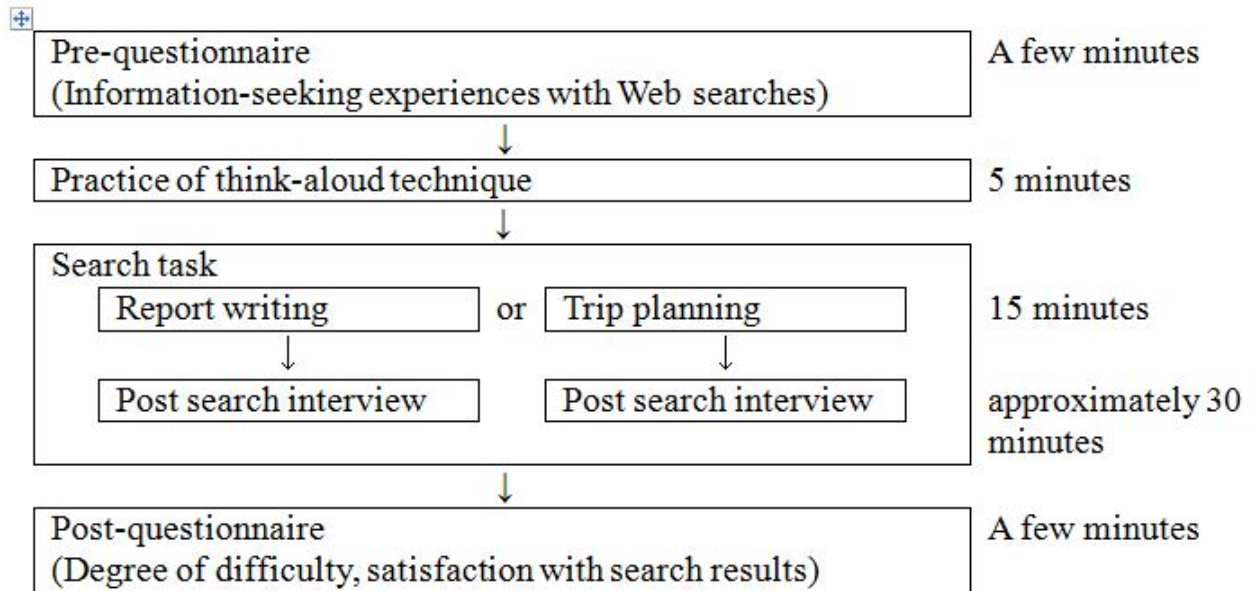


Figure 1: Sequence of experiment

The participants were asked to conduct either the report or the trip task for fifteen minutes. The sequence of the two tasks was counterbalanced between participants. In each task, participants were requested to think aloud while searching. Their utterances were captured with a voice recorder, eye movements were captured with an eye-tracking system (EMR-AT VOXER) and screen shifts of Web search processes were recorded with screen capture software (HyperCam 2.14.02 Japanese version). After each task, the participants answered a questionnaire about the degree of difficulty and the level of satisfaction with the search results and completed a post-search interview. In the interview, participants were asked to articulate their thoughts, feelings and intentions for particular actions at each point of the Web search processes in response to the interviewer while watching a video of their search process/behaviour. Either the first or last half (counter-balanced between participants) of the video of their Web search process was shown with their recorded eye movements to facilitate episodic memory retrieval.

Data analysis

Think-aloud and interview data were transcribed in detail. These data were recorded, together with several variables including type of pages viewed (either search results page or content pages [non-search results pages]), ten search-related actions (Table 1) and the eye-gaze points of search results pages viewed, categorized into twenty-one areas as shown in Figure 1, were recorded manually into the Cognitive Process Annotation Tool by Terai (COPATT), using time-stamps for sequencing.

Action	Description
Search	Searching using search engine
Link	Clicking on page links
Next	Going to next page
Back	Going back to previous page
Jump	Going to bookmarked or history page
Browse	Browsing new search results
Submit	Clicking submit button
Bookmark	Adding a bookmark
Change	Changing from one tab to another
Close	Closing tabs or windows

Table 1: Web action categories

Graduates tended to use 'Change' and 'Close' more frequently than undergraduates, indicating that graduates tend to open multiple pages simultaneously and switch between pages for both tasks. Details of the analysis are reported elsewhere ([Saito et al. 2009](#)); [Egusa et al. 2010](#)).

In analysing their Web-search activities, we classified the Web pages browsed by the participants into two categories: search results pages (ranked lists of the search results) and non-search results pages (content pages). Twenty-two areas of the eye-gaze points or *look zones* in the search results pages were identified manually and recorded into the Cognitive Process Annotation Tool, as shown in Table 2 and Figure 2.

As shown in Table 2, participants looked at information (snippet, title and URL) on retrieved pages more frequently than other parts of the search results pages. Again, details of the analysis are reported elsewhere ([Saito et al. 2009](#)); [Egusa et al. 2010](#)).

Look zone	Report	Trip
Title Bar	34	9
Menu	2	1
Bookmark	34	0
Tool bar	16	22
URL bar	7	2
Search bar	0	0
Search bar button	0	0
Tab	73	83
Link for services	159	45
Query box	332	113
Search button	8	6
Scroll bar	1	0
Number of hits	4	0
Link for ads	60	109
Spell check	0	0
Title	546	379
Snippet	820	333
URL	368	139
Related search	27	23
Link for next page	7	7
Find a page	0	0
Status bar	16	0
Out of look zone	476	164
Lack of eye position	751	637

Table 2: Look zones (0.5-second intervals)

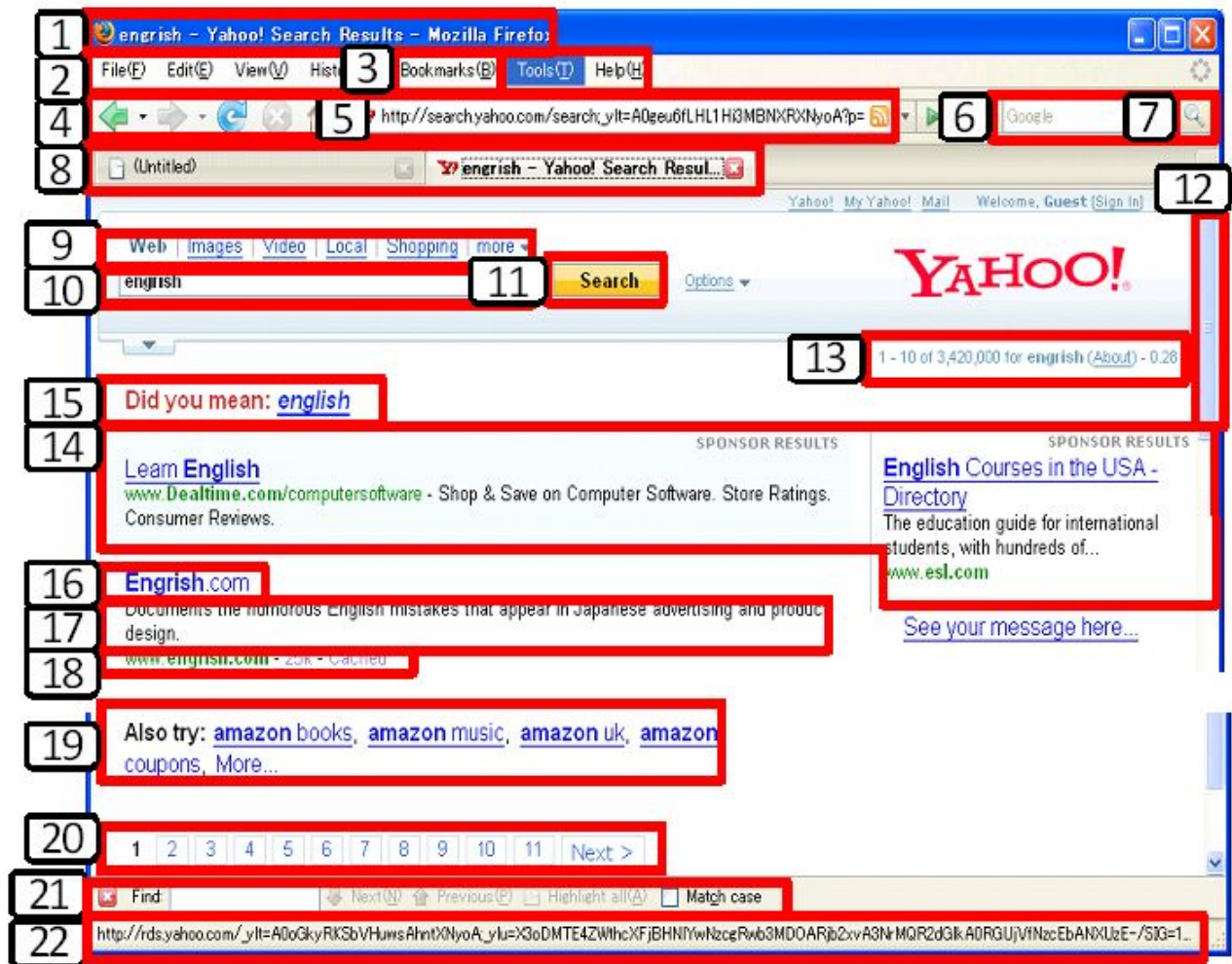


Figure 2: Areas of look zones

We defined link depth as an index for indicating the extent to which a searcher moved deeper into the Web, flowing links from search results pages and visualized it (Egusa et al. 2010).

The transcribed think-aloud and interview data, together with the types of pages viewed, search-related actions, gaze points of result pages and link depth, organized in a time line, were content analysed employing a top-down strategy with a constant comparative technique (Strauss and Corbin 1990), using the *taxonomy of knowledge modification patterns* (Table 3) and the *taxonomy of knowledge utilization patterns* (Table 4) developed in previous studies (Miwa and Kando 2007; Miwa and Takahashi 2008).

Type		Definition
Knowledge acquisition	adding	Acquire novel information to increase knowledge
	limiting	Narrow down the scope of the concept
Modification of existing knowledge	specifying	Increase the depth of meaning of the concept by increasing specificity
	correcting	Clear up misunderstanding
	recalling	Recalling existing knowledge
	verifying	Confirm or validate existing knowledge
Combination of existing and acquired knowledge	relating	Relate a concept to another concept
	transforming	Understand a concept in a different framework
	clarifying	

		Ambiguous knowledge becomes clearer
	interpreting	Combine existing knowledge and acquired knowledge for heuristics

Table 3: Taxonomy of knowledge modification patterns

Type		Definition
start	start searching	Begin searching for a topic
	start browsing	Begin browsing withing a site
stop	stop browsing	Stop browsing within a site
	stop retrieval	stop retrieval within a site
	stop searching	Stop searching for a topic
change	change keyword	Change or modify keyword(s)
	change logic	Change or modify search logic
	change strategy	Change the entire strategy of searching for a task
further searching		Continue searching to explore the topic
plan post-search activity		Plan what to do after finishing Web searching

Table 4: Taxonomy of knowledge utilization patterns

These two coding taxonomies, accompanied by coding rules, were given to a second analyst (a co-author of this paper) who coded the same subset of think-aloud and interview data of two participants' Web search processes, annotated with the Web action categories, the type of page viewed, the look zones and the link depth. After a trial coding of one participant's data, the coding results were compared to obtain the level of reliability in coding. The final coding by the second analyst was compared with the initial coding done by the first analyst (the first author of this paper). We used " κ " (Kappa) (Cohen 1960), as a coefficient of inter-judge agreement for nominal scales and obtained 0.72 for taxonomy of knowledge modification patterns and 0.82 for taxonomy of knowledge utilization patterns, both of which we considered acceptable. In addition, we developed a *taxonomy of knowledge type* (Table 5) using a bottom-up strategy of a constant-comparative technique. Again, the coding results of the same subset were compared between the analysts using " κ " and obtained 0.87, which we considered acceptable.

Type	Definition
topic	Knowledge of the topic of a paper and/or what to do during the trip
system	Knowledge of the functionality, interface and/or usability of the search system
source	Knowledge of the existence and/or usage of information sources
scheme	Knowledge of how to approach Web search for a particular task

Table 5: Taxonomy of knowledge type

Results

In this section, we report on the results of the qualitative analysis of think-aloud and post-search interview data regarding the incidence of information encountering. We translated the quotations from the think-aloud and/or interview data from Japanese into English. In these quotations, supplemental information was added in square brackets, irrelevant information was omitted using ellipses, and each participant's identification and category are provided in parentheses at the end.

Identification of information encountering

We initially defined information encountering as '*an instance of accidental discovery of information during an active search for some other information*' ([Erdelez 2005](#): 180). However, we also included those incidents in which participants encountered novel information which caused them surprise or amazement because they did not previously know about it. After identifying data indicating information encountering in think-aloud and interview data, we chose only those information-encountering incidents that were confirmed by both think-aloud and interview data.

Table 6 presents the thirty incidents of information encountering we identified through the analysis. In this table, each identification label consists of a number and a capital letter indicating the participant category of undergraduate students (U) and graduate students (G). Twelve out of 16 participants indicated that they experienced information encountering in one or both tasks, while four (S01, S03, S98, S09) did not.

Identification label	Task	Duration (seconds)	Post action	Page	Look zone	Link depth	Consequence
02U	report	5.5	link	content	NA	1	temporary deviation
04U	report	30.7	link	content	NA	1	temporary deviation
04U	report	25.4	bookmark	content	NA	1	temporary deviation
05U	report	37.4	search	SR	URLbar	NA	temporary deviation
05U	report	69.8	search	SR	snippet	NA	temporary deviation
05U	report	17.9	search	SR	title	NA	change
05U	report	126.9	return	content	NA	1	change
06U	report	7	return	content	NA	1	temporary deviation
06U	trip	97.6	link	content	NA	2	change
06U	trip	84.2	link	content	NA	2	change
07U	report	64.7	link	content	NA	3	change
07U	trip	14.1	link	content	NA	3	change
10U	report	44.9	bookmark	content	NA	2	change
10U	trip	47.7	link	content	NA	2	change
11U	report	50	link	content	NA	3	change
11U	trip	76.5	return	content	NA	3	change
20G	trip	36.4	link	content	NA	7	change
21G	report	33.6	change	content	NA	1	change
21G	report	64.2	search	content	NA	1	change
21G	trip	62.1	link	content	NA	5	change
21G	trip	29.2	link	content	NA	1	change
22G	trip	16.2	change	content	NA	1	temporary deviation
23G	report	22.9	link	content	NA	1	change
23G	report	46.4	link	content	NA	1	change
23G	report	11.3	change	content	NA	3	change
23G	report	20.0	jump	content	NA	1	change
23G	trip	9.6	link	SR	snippet	NA	change
23G	trip	43.2	link	content	NA	3	

							temporary division
23G	trip	10.1	link	content	NA	3	temporary deviation
24G	report	25.4	return	content	NA	4	temporary deviation
'SR' in 'Page' column refers to search result page; 'content' refers to content page. 'NA' means 'not applicable'							

Table 6: Incidence of information encountering

The 'Duration' of the information encountering (in seconds) was calculated from the time of information encountering to the time the participant moved to the next indicated action not directly connected to the information encountering, as shown in the 'post action' column. The 'page' shows whether the information encountering occurred in the search results page (SR) or the content page. The look zone was specified if it occurred in the search results page. The link depth indicates how far the participant was from the search results page at the moment of information encountering (Egusa *et al.* 2010). The *consequence* is the result of the information encountering: some participants changed the topic of the paper or the destination of their trip as a result of the information encountering, while others temporarily deviated from the main search processes but got back to the initial path after a while.

We superimposed the information-encountering incidents upon the graphic representation of the link depths for exploratory search processes of all participants. Figure 3 shows examples of visualized link depth for report task and trip task of a single participant (07U). The first diagram is the visualization of information encountering on link depth for the report task and the second diagram shows the trip task. The x-axis represents link depth and the y-axis shows the time (in seconds) that passed during execution of the search task. SE on the x-axis represents the search engine page and SR represents the search results page. The green dots in the diagram represent the *search*, a triangle indicates a *bookmark*, and the purple dots represent *submit*. The Japanese expressions listed on the left side of the diagram are query terms used for 'search' actions. The red line represents link depth: as the line extends towards the right, the link depth increases. The part of the link depth line overlaid with a thick black line indicates an incidence of information encountering. From these visualization results, we can see at a glance the occurrences of information encountering and the outline of search and page viewing behaviour.

In the report task, the participant chose the topic 'Causes of World War II' and input 'Wikipedia' as a query. When he entered the top page of Wikipedia, he submitted 'World War II' and followed the link to examine the topic while trying to develop an outline of the paper. Because the topic he chose was too broad, he almost gave up the search just before he saw a link to the 'Kristallnacht (Night of Broken Glass)', a series of attacks against Jews throughout Nazi Germany and parts of Austria in November 1933.

During the post-search interview, he explained:

This is the part of the report task when I was absorbed in reading... This may lead to a good paper and the article was relatively short... I clicked the link without any preliminary knowledge on this event, but it describes the causes of the Nazi attack on the Jews, but the name of the event does not imply what it is all about... After reading this article, I thought I could write a term paper on this topic... (07U)

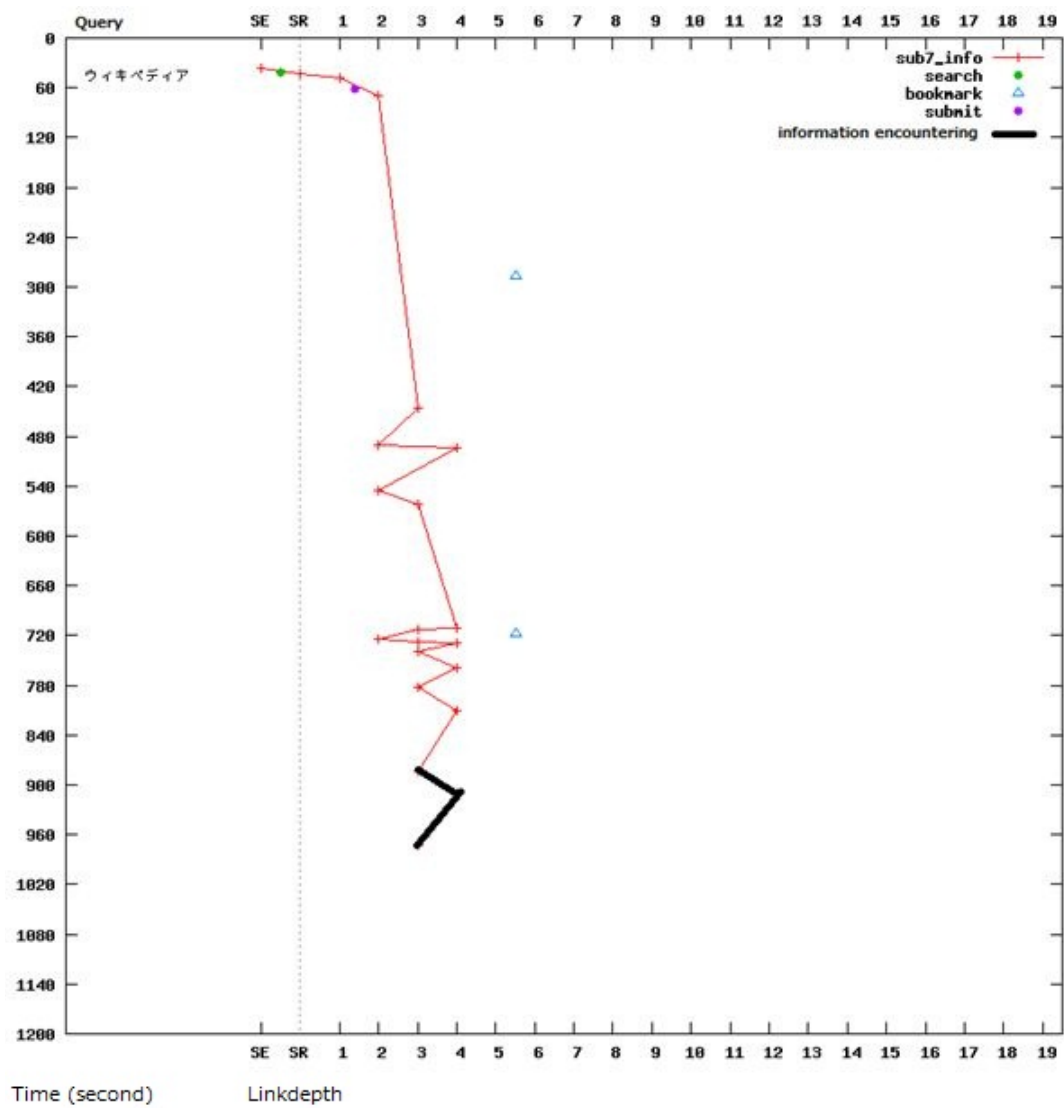


Figure 3: Visualization of report task

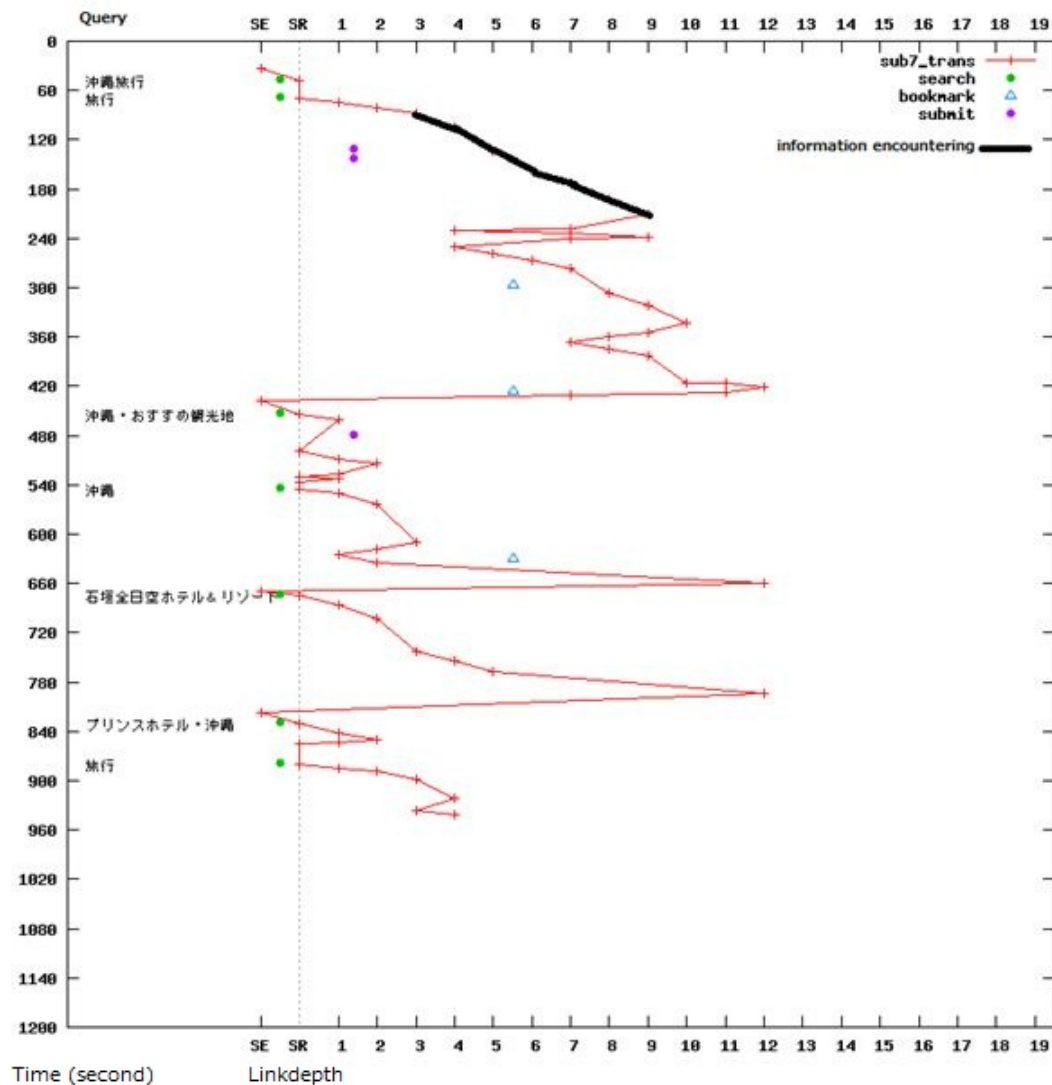


Figure 4: Visualization of trip task

Contexts of information encountering

As indicated in Table 6, almost all information encountering (26/30 or 87%) occurred when participants were looking at content pages. When participants encountered unknown but seemingly interesting information not directly related to their initial search task, the consequence of information encountering resulted in two distinctive patterns as indicated in the 'consequence' column of Table 6: (1) a temporary deviation from the task goal, with an ultimate return to the initial task, as described by Erdelez (1997) and (2) a deviation from the initial task goal, which led to a change in the topic of a paper and/or trip destination.

An example of articulating the first pattern in the post-search interview is:

When I was looking for information on Lincoln, I found the name of George Washington, who I did not know. I thought 'who is he?' and clicked the link to identify the person. (05U)

An example of articulating the second pattern is:

When I was examining a tourist site, I found there was a tree called a Malayan banyan, which I knew from the lyrics of my favourite song, which I was listening to this morning. I immediately felt that I wanted to see it... (21G)

Thus, information encountering in an exploratory search may not only lead to a temporary deviation from the task goal, but may also give a new direction to the search.

Wikipedia and information encountering

As mentioned earlier, all but one participant used Wikipedia in their Web search for the report task. However, the mode of use differed somewhat between undergraduates and library and information science graduates.

Undergraduates tended to use Wikipedia to clarify their topics and obtain chunks of information, which they thought could be used as part of the body of their papers. For example, one undergraduate explained in the interview:

[In Wikipedia], I first thought about whether I could use the information in the page as part of my paper... I found that the information was structured by sub-entries and clicking one of them led me to seemingly good information, which I might use in writing the paper... the whole content of the Wikipedia page... (03U)

Graduates also used Wikipedia to clarify their topics, but they used it to obtain reliable sources of information to be cited in their papers from references and external links. For example, a library and information science graduate explained in the interview:

The last portion of the Wikipedia entries consists of references and external links that lead to reliable information sources. This is more convenient than a Google search. Since my topic is 'History of Hawaii', I thought I could obtain useful information sources from the entry of 'Hawaii State'. (23G)

One of the reasons for such a difference might be the amount of experience in using Wikipedia and in writing research papers. Graduates may have learned the structure of Wikipedia entries through experience in using the site. They may also have developed a shared practice of not using Wikipedia to structure part of the body of their paper, but of writing original research papers based on reliable books and journal articles and providing appropriate citations.

Five out of 12 (41%) information-encountering incidents experienced by undergraduates occurred when they were on Wikipedia, while only one graduate experienced an information-encountering incident on Wikipedia. It seems as if Wikipedia is a site for information encountering predominantly for undergraduates.

Amazon.com and information encountering

One undergraduate (6% of undergraduate participants) and three library and information science graduates (60% of graduate participants) used Amazon.com in the report task. They used this site not only for locating relevant books but also for verifying the reliability and usefulness of books they found in earlier part of their search by checking customer reviews.

A graduate who encountered information on Amazon.com explained in the interview:

I knew of the existence of a Hawaiian phrase such as 'Aloha', but it was not in my mind while searching until I saw a Hawaiian-Japanese dictionary in 'frequently bought together' on Amazon.com... (23G).

One of the reasons more graduates than undergraduates used Amazon.com may be that they had developed a report scheme for identifying reliable books for writing research papers using customer reviews.

Online public access catalogue and information encountering

Three library and information science graduate students used the online public access catalogue for the report task, while none of the undergraduates used it. The three graduates used the online public access catalogue to identify the locations of books they had found in previous part of their search. One graduate used the interlibrary loan service incorporated in the online public access catalogue to obtain a book he could not locate in the library. Based on these observations, we assume that the use of the online public access catalogue for report task may be unique to library and information science graduate students. Their report scheme may include acquiring useful or reliable books in preparing for writing papers.

One graduate student who had focused on a topic from the beginning of the report task encountered a book relevant to his dissertation topic, which led him to deviate from the well-planned path for preparing a world history report. He explained in the interview:

This book is quite relevant to my dissertation research. Very much related. Therefore, even though this book is not relevant to this report task, I could not help looking at the contents of the book. This kind of unexpected finding happens quite often during Web searches. (24G)

Above all, participants' information encountering seemed to have occurred when they were searching for information on their familiar sites. Thus, undergraduates tended to encounter information on Wikipedia.

Effects of showing eye-movement data

During the post-search interviews, participants watched a video with and without their own eye movement recorded in either the first or latter half of their search processes. At the end of the post-search interviews, participants were asked to report on perceived differences in recalling their search process while watching the video with and without their recorded eye movements. A majority of participants reported that the eye-movement data on the screen helped them recall their thoughts at that moment during the search.

Some of their reports are as follows:

[When there was eye-movement data] I could identify whether I was looking at a picture or a part of the sentence, which helped me recall what I was thinking. (22G)

[In the report task] my search pattern was relatively concrete because I was following the initial plan. Therefore, the availability of eye-movement data did not affect my recall. On the other hand, there were more uncertain elements in the trip task... the eye-movement data helped me recall the process. (24G)

Thus, watching one's own eye-movement data should help the participant recall the instance of information encountering and their thoughts at the time. It may help them retrieve episodic memory of Web search processes.

Discussion

As stated above, we initially defined information encountering as '*an instance of accidental discovery of information during an active search for some other information*' (Erdelez 2005: 180). During the data analysis, however, we also captured those incidents in which participants encountered novel information that caused them surprise or amazement because they did not previously know about it; we consider that this should be included in information encountering. The former type of information encountering shifts participants' search goals from a foreground task to a background task, but eventually returns to the initial task (Erdelez 2005). The latter type of information encountering modifies participants' search goals in a way that is more focused on the limited goal than on the initial goal.

Table 7 contrasts some participants' task goals before and after the incidence of information encountering

Identification label	Task	Before information encountering	After information encountering	Comment
02U	Report	East India Company	British East India Company	Participant did not know there were so many East India Companies from different countries
07U	Report	Causes of World War II	Kristallnacht	Participant did not know about Kristallnacht
21G	Report			

		Establishment of Buddhism in India	Establishment of Buddhism in Nepal	Participant did not know the Buddha was born in Nepal
23G	Report	History of Hawaii	History of Hawaiian language	Discovery of a Hawaiian-Japanese dictionary lead the participant to focus on the topic
24G*	Report	Weimar Republic and Nazi	Weimar republic and Nazi	Discovery of a book relevant to the participant's dissertation led him to a temporary deviation from the initial goal
11U	Trip	Hokkaido and floating ice	Hokkaido and hot air ballooning	Encountering hot air ballooning prompted him to recollect his experience as a youth of missing an opportunity to go up in a balloon
20G	Trip	Hot spring	museum	Participant discovered a museum featuring a famous photographer and shifted his interests
21G*	Trip	Cedar tree in Yakushima	Cedar tree in Yakushima	Participant discovered that there is a tree she recalled from the lyrics of a song
*The participant's task or destination did not change.				

Table 7: Modification of task goals by information encountering

In Table 7, two incidents of information encountering, marked with an asterisk, did not change the participants' topic or destination. Both were found in graduate students' (24G Report task; 21G Trip task) search processes in which the initial task goals were well defined and focused. Others had ill-defined, open-ended task goals.

The finding implies that in the early stages of exploratory search in which searchers do not have much knowledge about the topic of the paper or the destination of the trip, they tend to follow links to novel information that surprises them or arouses their curiosity. This may be because such novel information helps them to focus on defining their topic of the paper or the destination of the trip. In an exploratory search in which the searcher does not have much knowledge on task goals, they need to learn to define their goals as the anomalous states of knowledge hypothesis suggested by Belkin *et al* (1982). It may be that incidents of information encountering in exploratory searches provide searchers with the opportunity to learn about the goals of search tasks so that they can reduce uncertainty (Marchionini 2006).

The above argument may lead us to a definitional issue of information encountering as well as the larger concept of opportunistic discovery of information. Do information encountering and opportunistic discovery of information occur only during the look-up search for well-defined tasks? Alternatively, do they occur in the exploratory search in which searchers learn to fill their knowledge gaps by opportunistic discovery of information? If we include the latter type of information discovery as a kind of information encountering, we may need to define information encountering more broadly to allow many opportunities of learning by searching.

Conclusion

A qualitative content analysis of think-aloud protocols and post-search interviews of Web search processes, combined with searchers' actions, look zones and link depth for report and trip tasks performed by sixteen participants was conducted using the constant comparative technique. Through analysis, we uncovered information encountering in a rich context of searchers' external behaviour and internal (cognitive and affective) processes and visualized them.

We found that encountering unexpected but interesting information in well-defined search tasks may lead to temporary deviation from the search goals, while discovery of novel information induces searchers' surprise, which may cause some modification of search goals or help searchers to develop the focus of their search goals.

We found that information encountering tended to occur in exploratory searches in which searchers intended to learn by searching for information. We argued that inclusion of this type of information discovery may lead to reconsideration of the definition of information encountering.

We also found that watching a video of their eye movements recorded during the Web search process helped searchers recall their thoughts and feelings from that moment in the search, including those at information encountering.

Acknowledgements

Part of this research was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science and a National Institute of Informatics Joint Research Grant. We are grateful to anonymous reviewers who helped us organize the structure and focus of this paper.

About the authors

Makiko Miwa is a Professor at the Center of ICT and Distance Education at the Open University of Japan. She received her Bachelor's degree in History from Japan Women's University, her Master of Library science from the University of Pittsburgh, U.S.A. and her Ph.D. from Syracuse University, U.S.A. She can be contacted at miwamaki@ouj.ac.jp.

Yuka Egusa is a researcher at the Educational Resources Research Center in the National Institute for Educational Policy Research of Japan. She received her Bachelor's degree in Library and Information Science from University of Library and Information Science, Japan, her Master of Information Science from University of Library and Information Science, Japan and her Ph.D. in Information Science from University of Tsukuba, Japan. She can be contacted at yuka@nier.go.jp.

Hitomi Saito is a Associate Professor and member of the Faculty of Education at Aichi University of Education. She received her Bachelor's degree in Library and Information Science, her Master of Information Science from University of Library and Information Science, Japan and her Ph.D. from Nagoya University, Japan. She can be contacted at hsaito@aecc.aichi-edu.ac.jp.

Masao Takaku is a Senior Engineer at National Institute for Materials Science, Japan. He received his Ph.D. in Information Science in 2004 from University of Tsukuba. His research interests include digital libraries, information retrieval and information seeking behaviour. He can be contacted at takaku.Masao@nims.go.jp.

Hitoshi Terai received a Ph.D. in Information Science from Nagoya University, Japan, in 2006. Currently he is a contract associate professor of the Graduate School of Information Science, Nagoya University. His research interests include problem solving, scientific discovery, insight and information seeking behaviour. He can be contacted at terai@is.nagoya-u.ac.jp.

Noriko Kando received a Ph.D. in Library and Information Science from Keio University, Japan, in 1994. Currently she is a professor of the National Institute for Informatics. Her research interests include information access technique for information seeking and learning, evaluation of information access technique, cognitive research for exploratory searching, text mining and knowledge-based text compilation. She can be contacted at nii@ac.jp.

References

- Belkin, N.J., Oddy, R.N. & Brooks, H.M. (1982). ASK for information retrieval. Parts I and II. *Journal of Documentation*, **38**(2), 61-71; (3). 145-164.
- Broder, A. (2002). A taxonomy of Web search. *SIGIR Forum*, **36**(2), 3–10.
- Buckner, R.L. & Barch, D. (1999). [Cognition: memory, 1: episodic memory retrieval](#). *American Journal of Psychiatry*, **156**(9), 1311. Retrieved 13 August, 2011 from <http://ajp.psychiatryonline.org/cgi/reprint/156/9/1311.pdf> (Archived by WebCite® at <http://www.webcitation.org/60ut4Y5Pp>)
- Choi, Y. (2010). Effects of contextual factors on image searching on the Web. *Journal of the American Society for Information Science and Technology*, **61**(10), 2011–2028.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**(1), 37–46.
- Egusa, Y., Saito, H., Takaku, M., Terai, H., Miwa, M. & Kando, N. (2010). Link depth: measuring how far searchers explore Web. *Proceedings of the 43rd Hawaii International Conference on System Sciences*, (pp. 1–8), Washington, DC: IEEE Computing Society.
- Erdelez, S. (1997). Information encountering: a conceptual framework for accidental information discovery. In P. Vakkari, R. Savolainen & B. Dervin, (Eds.), *Information seeking in context: Proceedings of the International Conference on Research in Information Needs, Seeking and Use in Different Contexts*, (pp. 412–421). London: Taylor Graham.
- Erdelez, S. (2000). Towards understanding information encountering on the Web. *Proceedings of the 63rd Annual Meeting of American Society for Information Science*, **37**, 363-371.
- Erdelez, S. (2004). Investigation of information encountering in the controlled search environment. *Information Processing and Management*, **40**(6), 1013–1025.
- Erdelez, S. (2005). Information encountering. In Karen E. Fisher, Sanda Erdelez & Lynne (E.F.) McKechnie (Eds.). *Theories of information behavior*, (pp.179-184). Medford, NJ: Information Today.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, **49**(4), 41–46.
- Martzoukou, K.(2005). [A review of Web information seeking research: considerations of methods and foci of interest](#). *Information Research*, **10**(2). Retrieved 19 July, 2011 from <http://informationr.net/ir/10-2/paper215.html> (Archived by WebCite® at <http://www.webcitation.org/60urmSebF>)
- Miwa, M. & Takahashi, H. (2008). [Knowledge acquisition and modification in students' exploratory Web search processes while career planning](#). *Information Research*, **13**(4). Retrieved 19 July, 2011 from <http://informationr.net/ir/13-4/paper376.html> (Archived by WebCite® at <http://www.webcitation.org/60urqX3nF>)
- Miwa, M. & Kando, N. (2007). [A naïve ontology for concepts of time and space for searching and learning](#). *Information Research*, **11**(2). Retrieved 19 July, 2011 from <http://informationr.net/ir/12-2/paper296.html> (Archived by WebCite® at <http://www.webcitation.org/60urw4QdO>)
- Moore, J.L., Erdelez, S. & He, W. (2007). The search experience variable in information behaviour research. *Journal of the American Society for Information Science and Technology*, **58**(10), 1529–1546.
- Rubin, V.L., Burkel, J. & Quan-Haase, A. (2010). Everyday serendipity as described in social media. *Proceedings of the American Society for Information Science and Technology*, **47**, Article 152.
- Saito, H., Terai, H., Egusa, Y., Takaku, M., Miwa, M. & Kando, N. (2009). [How task types and user experiences affect information-seeking behavior on the Web: using eye-tracking and client-side search logs](#). In Nicholas J. Belkin, Ralf Bierig, Georg Buscher, Ludger van Elst, Jacek Gwizdzka, Joemon Jose & Jaime Teevan, (Eds.). *Proceedings of the Workshop on Understanding the User - Logging and Interpreting User Interactions in Information Search and Retrieval (UIIR-2009)*, Boston, MA, USA, July 23, 2009, (pp. 19–22). Retrieved 19 July, 2011 <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-512/paper05.pdf> (Archived by WebCite® at <http://www.webcitation.org/60usGGWPp>)
- Silverstein, C., Henzinger, M., Marais, H. & Moricz, M. (1999). [Analysis of a very large Web search engine query log](#). *SIGIR Forum*, **33**(1), Retrieved 19 July, 2011 from <http://www.acm.org/sigir/forum/F99/Silverstein.pdf> (Archived by WebCite® at <http://www.webcitation.org/60usLXs6G>)

- Strauss, A. & Corbin, J. (1990). Basics of qualitative research: grounded theory procedures and techniques. Newbury Park, CA: Sage Publications.
- Toms, E.G. (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human Computer Studies*, **52**(3), 423–452.
- Williamson, K. (1998). Discovered by chance: the role of incidental information acquisition in an ecological model of information use. *Library and Information Science Research*, **20**(1), 23–40.

How to cite this paper

Miwa, M., Egusa, Y., Saito, H., Takaku, M., Terai, H. & Kando, N. (2011). "A method to capture information encountering embedded in exploratory Web searches" *Information Research*, **16**(3) paper 487. [Available at <http://InformationR.net/ir/16-3/paper487.html>]

Find other papers on this subject

[Scholar Search](#)

[Google Search](#)

[Bing](#)

Check for citations [using Google Scholar](#)

■ [Bookmark This Page](#)

1877

© the authors, 2011.
Last updated: 13 August, 2011



[Contents](#) | [Author index](#) | [Subject index](#) | [Search](#) | [Home](#)