Right Info, Wrong Answers: Eyeing the Search for Why with Tracking Technology

Feng-Ru Sheu
University Libraries, Kent State University

Abstract

This poster reports an exploratory study about the process or the quality of the search online information, including how the information was found and "read" (or "seen"). Inspired by the eye-tracking technology and the idea of "eye-mind hypothesis," the study explores students' online information searching/browsing behaviors and what eye movement data can provide to better understand Web searching/browsing behaviors, and therefore, to inform practitioners (i.e., instructional designers, librarians, and information specialists etc.) for better design.
Keywords: information seeking behavior; scan path; eye tracking

Introduction

In the information age, students constantly search for information on the internet. Sometimes a search for information to specific inquiries is successful; sometimes it is not. Even with success, scant information is available about the process or the quality of the search, including how the information was "read" (or "seen"), not to mention whether the reader engaged with the information that was seen or read. Understanding how users perform online information searching is a pervasive interest of librarians, information specialists, and instructional designers. For example, researchers have been focused on how visual design or presentation/display format was associated with learning performance, reading comprehension, and decision making. (Agostinelli et al., 2012; Bettman & Kakkar, 1997; Biehal & Chakravarti, 1982; Tomita, 2017; Zikmund-Fisher et al., 2016). While conventional user research methods are effective in obtaining users feedback on usability and their experiences, they have several key limitations. For example, traditionally one can verify whether the users actually "look at" the information on the present screen (or Web pages) by asking the participants follow-up interview questions and retrospective analyses of their online behaviors or to engage in self-report/think-aloud activities as they search. However, the latter approach may interrupt participants as they engage in resource intensive search behaviors, while the former methods rely on limited human memory. With eye tracking data (i.e., scan paths, fixation counts, and fixation duration), the researcher can explore a user's searching and 'reading' behaviors by accessing the visual path track. Therefore, this study investigates online information browsing behavior with deeper exploration by incorporating eye-tracking data into conventional user research methods.

Background

The research detailed here was driven by a previous usability study, which the instructional designer accidently found that some users responded with the 'wrong' answer to a common library inquiry when they all arrived at the 'target' pages where information was located. A pilot study (Sun, Sheu, & Tsai, 2018), as part of a large project, was meant to be a test round for the protocol and the tasks were serve as warm-up tasks. Therefore, the tasks were designed to be simple and easy to complete. The tasks were to find answers from a selected website for two simple questions derived from the commonly asked questions at the library settings (e.g., library hours, directions, and eligibility of 'membership' etc.). The two questions were: (1) what's the library hours for Tuesday? and (2) Can a 15-year old high school student apply for a library card?

In the pilot study with eight participants, all participants completed both tasks/questions within 90 seconds, meaning that all participants were efficiently locate the information on the website and turned in the task sheet with their answers. The total number of Web pages visited was less than six for both tasks, excluding the home page/starting page. It was another indicator for efficient performance. All participants fulfilled the first inquiry correctly, but surprisingly three out of the eight respondents got the second question wrong even though they all successfully 'locate' the information. In other words, the participants have 'arrived' the target page that contains the necessary information to answer the question correctly but failed to choose the right answer.

Based on the data from the task sheets, desktop screen recording, and direct observations, we knew that all participants efficiently "arrive" on the target page and successfully locate the necessary information. They also felt confident about their search and answers. From a usability standpoint, it was a success. However, we do not know why some of participants get the second answer wrong when they have the information in front of them. It raises questions: what are the possible explanations and how can this be improved by design if applicable. The motivation to find answers has led to the present study.

Purpose of the study

The primary purpose of the present study was to explore possible explanations of why some participants respond with the “wrong” answer when they have “right” information in front of them by investigating user's information search process, including visual browsing/reading patterns, presentation style of the information used, and general performance (e.g. time spent, number of pages visited, completion rate etc.)

The study was an extended study from the previous pilot study, using the same research design with a larger number of participant data set. Descriptive results regarding the search process and performance were reported. However, instead of focusing on how the users ‘arrive’ or ‘found’ information on the website, the question becomes how the users ‘read’ or ‘look at’ the information.

Methods

As mentioned earlier, the present study was a study extended from previous pilot study aiming to seek possible explanation with a larger data set. User research methods were adopted, which involved asking participants to perform a given task in a lab setting and directly observing what they do. A well-designed library website was applied to test out two search tasks commonly seen in libraries, which are ‘searching for library hours’ and ‘eligibility of applying for a library card.’ All participants were performing these two tasks after the study was introduced and the consent of participation was obtained. Computer screen was recorded during the process. In addition, an eye-tracking device (Tobii 4C eye-tracker) was used to detect the visual pathway.

Participant

The target audience of the library website is general adults. Therefore, the general selection criteria were adults aged 20 years old or older with no severe visual impairments. Recruitment information was posted on Facebook and PTT, a social media tool used by students in Taiwan. Due to the geographic convenience, all participants were students from a public university where the lab was located. A total of 37 valid data sets were collected from the eligible participants. They are 27 females and 10 males with average age 26 years old.

Data Analysis

Study included both qualitative and quantitative data. In order to get an overall picture of searching and browsing process, descriptive data analysis was employed, describing demographic data, task completion, time spent, and the average number of web pages browsed. In addition, screen recording with scan path video was further used to analyze the user’s visual behavior patterns. Qualitative content analysis was applied to analyze videos, which inquired about visual patterns and type of message display. Video were reviewed and coded into emerged categories. The majority of respondents were under the age of 30, and 4 of the 37 people aged 30 or older.

Results & Discussion

The efficiency and effectiveness of online information searching performance were evaluated via various indicators, such as completion rate, time spent on tasks, number of pages/clicks count, accuracy etc. (Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002). In this study, we reported completion rate, number of pages browsed, time spent on tasks, and accuracy of responses. To illustrate the qualitative differences as well as triangulate possible explanations, results also include the type of information display (e.g., table vs. paragraph) and visual patterns that emerged from the scan path video.

Completion rate

Same as defined in the pilot study, the definition of technical completion is to ‘arrive at the “target web page,” which refers to where the necessary information for completing tasks is located. The soft definition for completion was when the participant indicated they were finished (writing down answers and return the task sheet to the facilitator). All participants completed the tasks. Overall the participants have no problem finding information in responding to the given tasks. All participants completed both tasks within four minutes.

Browsed page count

As mentioned earlier, the browse page count is one of indicators associated with search efficiency (Goldberg et al., 2002.) Results showed that all participants complete both tasks within eight pages, excluding the home page/starting page. Most of the participants completed the two tasks within two pages respectively. Twenty-nine out of 37 participants (78.38%) finished task one within two pages, and 32 out of 37 respondents (86.49%) completed task two within two pages. In other words, most of the participants completed each task with 2 clicks, which is considered rather efficient (see Table 1).

Table 1. Browsed page count per task (n=37)

Number of pages visited	Task 1	Task 2
	N	N
1-page	20	24
2-page	9	8
3-page	4	3
4-page	4	1
7-page	0	1
Total	37	37

Time spent

As shown on Table 2, the average time spent on task one and task two were 63.59 seconds, 51.26 seconds, respectively. The average time to complete task one and two was 114.85 seconds. We further split time spent into two parts: one was the time spent on searching (from the departure point to the arrival of the target page); the other was the time spent on the target page (where the necessary information was located). For task one, the average *searching time* was 33.14 seconds, and the average time spent on the result page (target page) was 30.45 seconds. For task two, the average time that participants spent on searching targeted page was 22.03 seconds, and the average time spent on the result page was 29.23 seconds. Comparing the time participants spent on the two different tasks, results showed spent longer time on the first task than those on the second task, both searching the target page and stayed on the target page.

Table 2. Time spent on search, reading and completion

Time (Seconds)	Task 1	Task 2
Average search time (starting point to target page)	33.14	22.29
Average time stayed at the result page	30.45	28.80
Average total time to complete task	63.59	51.26

Accuracy

All participants got the first question right and 11 out of the 37 respondents (29.7%) answered the second question incorrectly. The number of incorrect responses for the second question was “relatively” high while considering the ease of the task. During the observation process, all participants filled out answers confidently, and most importantly, they all had the right information in front of them. This aligns with the findings from the pilot study. The answer to the second question was “yes,” a high school student can apply for a library card. In general, only 16 years old or older can apply for a library card. However, full-time students under 16 years old can also apply for a card. Both pieces of information were listed on the target pages. Unlike task one, which was straightforward, task two requires some thinking to process both ‘conditions. We know that participants have “arrived” the “target” pages that contain the necessary information for the question/inquiry and possible “look at” the information before answering it. However, we do not know if the participants actually “read” the information and how they read it. This is one of the areas where eye tracking data can provide additional information for better understanding without interrupting task process. We turned to desktop recording with eye tracking data for better insights and possible explanation.

Display/layout of information

How information is structured and display often plays an important role for reading (or browsing). The tasks were live search and participants were free to go to any page and use any strategies to find answers for the given inquires except using ‘search tool.’ Based on the data we collected, there were three ‘target pages’ where the participants “found” or “landed” for the inquiry #2 (task 2). One page presents the information in table format, whereas one was all text in paragraph format. The third page was a pdf document regarding library policy. The information was displayed in text form in paragraph as well. Therefore, we coded videos into two categories based on the format of display: table vs. paragraph.

As shown in Table 3, 23 participants landed on table format and 14 participants landed on paragraph format. Among the participants in table group, 18 (78.26%) got the answer right while eight (57.14%) participants in paragraph group got the task 2 right. One possible explanation was the table format of information display helped participants read the information.

Table 3. Task 2 responses by display format (n=37)

Response	Table		Paragraph	
	N	%	N	%
Correct	18	78.26	8	57.14
Incorrect	5	21.74	6	42.86
Total	23	100	14	100

Moreover, for task 1, participants spent more time on ‘searching’ than ‘reading,’ whereas for task 2, participants took longer time on ‘reading’ than ‘searching.’ While reading time between the two tasks was not relatively meaningful, it is interesting to see the time difference on reading/finding information between two display styles (table vs. paragraph) regarding the task 2 responses.

In general, participants who get the answer wrong spent more time on target page too. It could be that most participants who get the wrong answer landed on paragraph format, which naturally take more time to read and process information. As mentioned earlier, 23 out of 37 participants “landed” on the target page with table display and 14 on paragraph display. The average time for “table” group of participants to spend on the target page before writing down answer was 23.46 seconds while the average time spent on the target page for “paragraph” group 38.72 seconds. This could mean that it took longer for participants to process information in the paragraph format than table format. This finding aligned with common design principles for organizing information. Table helps to visualize information with simplify form. The paragraph format also contains more text. However, this is not to suggest that all information needs to be organized in table format. It will depend on the context that is designed for, such as the content, the purpose, needs, and the audience etc.

Table 4. Time spent on task 2 (n=37)

Time (Seconds)	Display format		Response	
	Table (n=23)	Paragraph (n=14)	Correct (n=27)	Incorrect (n=11)
Average search time	18.71	27.47	21.4	23.6
Average time stayed on result page	23.46	38.72	26.5	35.7
Average total time to complete task	42.17	66.19	35.7	59.3

Visual Pattern

After reviewing all scan path video, two general patterns were identified. One was "focus," which most of visual movement was more congregate or assembled. They also fall into the area of interest, the area that contain answers or necessary information to make a judgement (see Figure 1). The other type of visual pattern was "spread" or "scattered," which by contrast, majority of visual attention fall outside of the area of interest or the visual path was (more) spread out (see Figure 2).

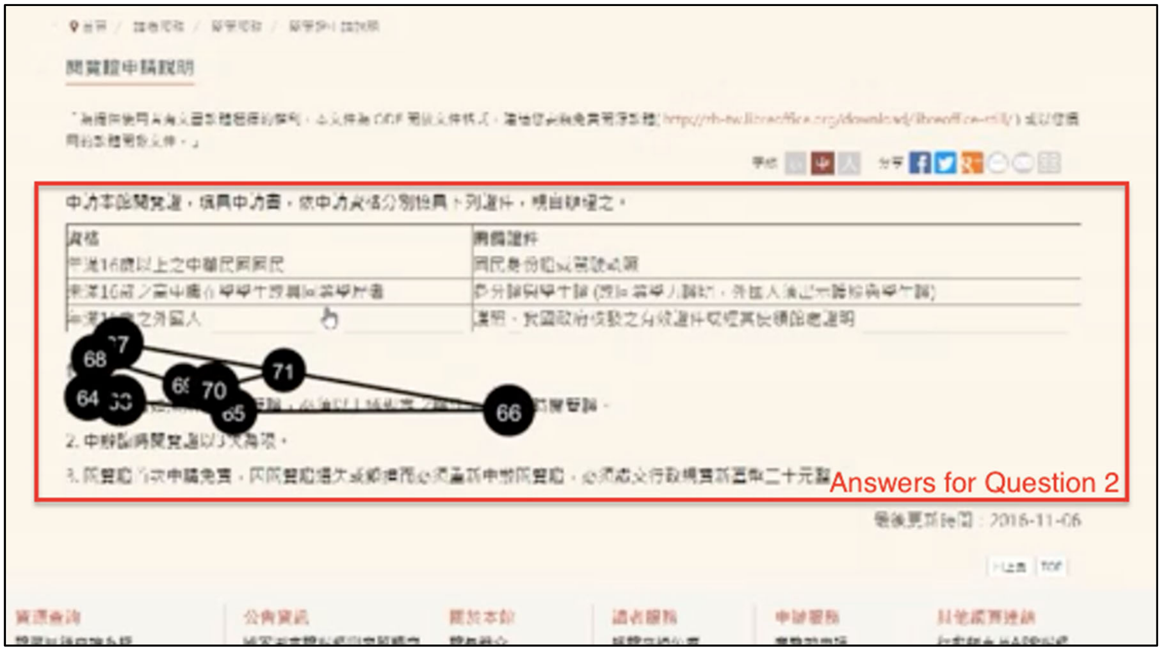

Figure 1. Example of "focus" visual pattern

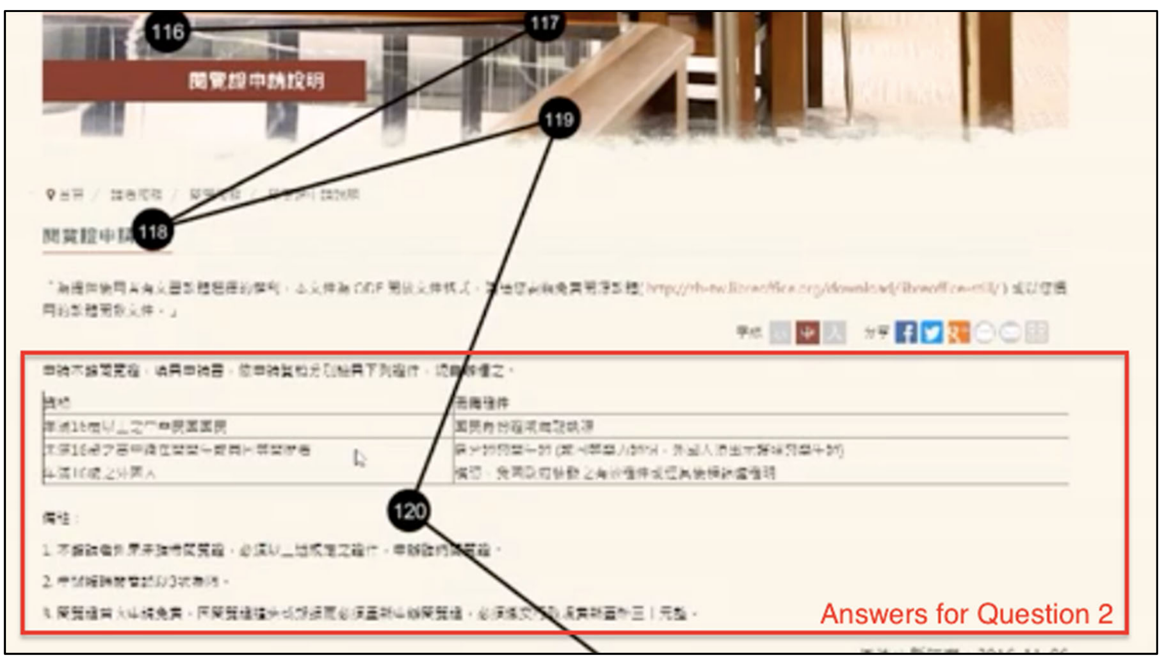

Figure 2. Example of "spread" visual pattern

As shown in Table 5, the participants whose visual pattern was more focused or strategic have higher percentage getting the answer correct (96%) while the participants in spread or scattered visual pattern have a higher percentage of people get the answer wrong. For a few cases, it was challenging to determine which category when there was a mix of patterns with high speed of visual movements. For example, the visual attention might focus (fixation) on one spot or one focused area but quickly move around for the rest of time and in some cases come back to the same spot. When it happened, the video was reviewed by the third person and discussed with all coders (total of three) until reach consensus. With that in mind, some cases are in the grey area. Another interpretation can be that the visual pattern of the participants who answered correctly was more 'focused' or strategic searching/browsing pattern, whereas the participants who did not respond with the correct answer showed no patterns or spread out pattern regarding the visual path when conducting the task. That could mean the participants 'skimming' through a lot of text quickly without 'reading'.

Table 5. Task 2 responses by visual pattern (n=37)

Response	Focus		Spread	
	N	%	N	%
Correct (n=26)	23	96	3	23.07
Incorrect (n=11)	1	4	10	76.93
Total	24	100	13	100

Among 23 participants who landed on the page with table format, 18 participants' visual pattern was "focus." And only one out of 18 got the wrong answer. While among 14 participants who landed on the paragraph format, there were only six participants' visual pattern was "focus" and all six responded with correct answer (see Table 6). This finding could indicate that table format helps allocate visual attention and ultimate help the "reading."

Table 6. Task 2 responses correction, type of display used, and users' scan path patterns (n=37)

Type of display Used	N	Scan path pattern	N	Task two response	N
Table	23	Focus	18	Correct	17
				Incorrect	1
		Scatter	5	Correct	1
				Incorrect	4
Paragraph	14	Focus	6	Correct	6
				Incorrect	0
		Spread	8	Correct	2
				Incorrect	6
Total	37		37		37

Conclusion

With a larger data set (larger number of participants) and eye-tracking data, we were able to get more insights on how participants search/browsing the web content for inquiry-based tasks. We find that the findings regarding general search/browsing behaviors from the pilot still hold true. The website is well designed from a usability standpoint. All participants efficiently completed the tasks by "conventional definition," which users arrived at the designated locations and spent approximately the same amount of time to search information and completed the tasks (whether they got the answer wrong). They have the same patterns to use both top navigation menu and the sitemap on the bottom of each page to get around the site.

Eye-tracking data provide additional information that conventional methods cannot obtain, which helps to confirm or triangulate a specific phenomenon, in this case, their visual browsing/reading patterns on the target information. It took longer for participants to process information in paragraph format than table format, which is normal. However, more participants who landed on table format have the correct answer for task 2 than percentage in paragraph format.

In general, we can conclude that table format works better than paragraph format for this type of information or inquiry. The table format in general guides visual attention to the target information if designed well. We may be able to design in a way to guide visual attention for better performance beyond just usability. However, we cannot guarantee users to engage with the information they see even if they read it. It might require a different type of research for user engagement; specifically, the type of engagement with the content. More future research is needed.

Based on the results, we can conclude that the characteristics of visual pattern for most participants in 'right answer' group was more focused while the pattern was more spread for the other group. We can conclude that for people in the right answer group seem to have more focused visual patterns while the other group have more spread or scatter visual patterns. In addition, the information in the table format seem to help users in 'reading' the information.

Finally, individuals may have seen the information but not 'read' it carefully or the reading was not 'registered' in the mind. Since it requires to put two conditions together for task 2, somehow the information that has been 'seen' or 'read' need to be engaged with the mind as well. However, we are not able to know for sure whether the information that has been 'seen' or even 'read' was 'processed'. Further research is needed in exploring higher level information processing.

References

Agostinelli, A., Specchia, M. L., Liguori, G., Parlato, A., Siliquini, R., Nante, N., ... & La Torre, G. (2012). Data display format and hospital ward reports: effects of different presentations on data interpretation. *The European Journal of Public Health, 23*(1), 82-86.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research, 3*(4), 233-240.

Biehal, G., & Chakravarti, D. (1982). Information-presentation format and learning goals as determinants of consumers' memory retrieval and choice processes. *Journal of Consumer Research, 8*(4), 431-441.

Goldberg, J. H., Stimson, M. J., Lewensteln, M., Scott, N., Wichansky, A. M. (2002) Eye Tracking in Web Search Tasks: Design Implications. In: Proceedings of the symposium on Eye tracking research & applications, pp. 51-58. ACM, New Orleans Louisiana, USA (2002). doi:1-58113-447-3/02/0

Sun, W.-T., Sheu, F.-R., & Tsai, M.-J. (2018, August). Understanding inquiry-based searching behaviors using scan path analysis: a pilot study. In: Wu TT., Huang YM., Shadiev R., Lin L., Starčič A. (eds) Innovative Technologies and Learning. ICITL 2018. Lecture Notes in Computer Science, vol 11003. Springer, Cham. DOI: 10.1007/978-3-319-99737-7_17

Tomita, K. (2017). Visual Design Tips to Develop an Inviting Poster for Poster Presentations. *TechTrends, 61(*4), 313-315.

Zikmund-Fisher, B. J., Scherer, A. M., Witteman, H. O., Solomon, J. B., Exe, N. L., Tarini, B. A., & Fagerlin, A. (2016). Graphics help patients distinguish between urgent and non-urgent deviations in laboratory test results. *Journal of the American Medical Informatics Association, 24*(3), 520-528.