Understanding Health-related Information Searching Behavior Through Eye Tracking

Chongyan Chen, Kexin Yu, Zihan Yang
School of Information, The University of Texas at Austin, Austin, Tx
chongyanchen_hci@utexas.edu, kyu@utexas.edu, zihan95@utexas.edu

ABSTRACT

Nowadays the Internet has become a favored source to seek health-related information across different cultures and backgrounds. Worldwide, about 4.5% of all Internet searches are for health-related information and most users are looking for information about specific health conditions. However, the quality of online health information is not always as good as we expect. To understand how individuals explore the search result page of health-related information and provide insights for reliable content providers, we designed an eye tracking experiment. During the experiment, participants' eye movements will be tracked using Tobii TX300 screen-based eye-tracker and iMotions software. A post interview will be conducted where participants can view their recorded eye movements. Ultimately we find out that the ranking position of health-related search results affects the fixation duration and participants are more likely to click the search result which they pay more attention to. In addition, Date AOI also has influence on participants' viewing behavior. However, for this demographic group we don't have significant evidence to support that viewing behavior is different on symptom and treatment search result page.

1. INTRODUCTION

As there are no time and space constraints, the Internet has become a favored source to find health information. According to a research conducted by the Pew Internet & American Life Project, in total, approximately 80% of American Internet users have searched for health related information. Compared to a study released in 2001 by the Washington research firm, this number has up from 62%. Jeremy M Asch (2019) reported that the number of online
searches for health-related information increases from 6% to 15% of all search queries right before a patient visits Emergency Department.

In China, Baidu provides most of the online medical search portals. 60 million search queries in Baidu search box are related to health. About 25% of search queries are about diseases and 5% of them are about hospitals. After the scandal involving Wei Zexi, a 21-year old Chinese college student who died after receiving misleading treatment
information from a promoted result on Baidu, general public began to realize that some search results on Baidu are sponsored by certain hospitals and these results are not distinguished clearly from organic results.

Worldwide, about 4.5% of all Internet searches are for health-related information (Morahan-Martin, 2004), which makes online health information a hot market. Most users of

online health information are looking for information about specific health conditions because they or someone they know was diagnosed with a medical condition (Morahan-Martin, 2004). However, the quality of online health information is not always good and reliable. Since the mid-1990s, researchers and clinicians show interest in exploring how individuals seek health-related information, what resource they go to get such information, what kind of information they trust, and how such information is applied (Lambert 2007).

Therefore, our research tries to understand how individuals explore the search result page(SERP) of health-related information and explore what affects their decision on clicking the link to access to the content and browsing the page. Since Asch, J. M.'s research also points out that 56% of patients who searched health-related information online searched for symptoms, while 23% of them searched about the treatment or management of a certain disease. We also designed experiment to see if individual's viewing behavior is different on SERP about symptom and treatment. Overall, this research can drive to a better understanding of health-related information searching behavior and provide insights for trustful health-related content providers as well.

2. RELATED WORK

2.1 Web Viewing Behavior in General

Some prior studies show that the relationship between the performance of an information retrieval system and the "success" achievable by human searchers is weak (Smith and Kantor, 2008). Users pay more attention towards the top organic search results and they rarely pay more attention to sponsored results. Also, the quality of the search result and their ranks, the type of search task and individual differences can affect search behavior (Buscher, Georg, Susan T. Dumais, and Edward Cutrell, 2010). Culture and language skills cause difference in information viewing behavior (Marcos, M. C., Garcia-Gavilanes, R. O. G. G., Bataineh, E., & Pasarin, L., 2014).

2.2 Viewing behavior on Health-related information

Users of online health information usually stop browsing on the first SERP. They don't trust sites with too much commercial advertising and don't pay attention to the indicator of credibility (Morahan-Martin, 2004). Another factor affects health information searching is the epistemic belief. Users are more

inclined to select the web with an epistemic belief that contains correct knowledge. Also, epistemic beliefs moderated the effects of the search interface. Users with strong beliefs that the Web contains correct knowledge showed a more focused information selection and better search outcomes (Kammerer & Gerjets, 2012).

2.3 Research with Eye Tracking technique on SERPs

In the previous research, researchers conducted studies with eye tracking technique on the viewing behavior on SERPs and indicates that rank matters largely on viewing behavior while only three to five results were viewed on average (Lorigo, Haridasan, Hrönn Brynjarsdóttir, Xia, Joachims, & Gay, et al. 2010). About the attributes of search results, 43% of total fixation is on snippet while title and URL get only a little less (Granka, L., Feusner, M., & Lorigo, L., 2008).

However, only a few work applies Eye Tracking to analyze 1) viewing behavior with different query intents on health-related information; 2) the influence of different attributes of search results on individual's viewing behavior and decision on clicking the results, which provides the motivation for our work. Pian, Khoo, & Chi, J (2017) use Logistic Regression to build health information searching model. Their data are a combination of participants' demographic information, Eye-Tracker data, and Mouse-Click data which inspired our experiment design.

3. RESEARCH QUESTIONS & HYPOTHESES

From previous work conducted by Cristina et al.(2011), we know that ranking position of the search results and the areas of interest(title, snippet, url and image) influence individual's browsing behavior. Based on that finding we infer that users who search for health-related information may also be affected by rank of search results and the areas of interest and proposed our first and second research questions:

RQ1: What is the relationship between the elements viewed and the links clicked on the health-related information search result page?
RQ2. Does the rank of search results affect above behaviors?

Before experiment design we defined four areas of interest(AOI): 1) title, 2) url, 3) date and 4) snippet (see Figure 1). Title, url and snippet are three common defined AOIs in previous work about SERPs, so we also include them in our research. Since we excluded sponsored results and only kept organic results, all the SERPs doesn't include any image. In addition, there is no former work defines Date as an AOI that could affect individual's decision on clicking search result, but we think that when seeking medical information, professional online users may pay attention to the date of the information to make sure they get the latest health-related information. Research conducted by Morahan-Martin, J. M. (2004) also mentioned that checking the date information is posted is one of the methods that used to evaluate online information. In that case we came up with the following hypotheses:

H1. Rank of the search result affects the time spent on the health-related information search results.
H2. Time spent of each search result affects whether the results are clicked.
H3. Search results with date can capture more attention than those without date.

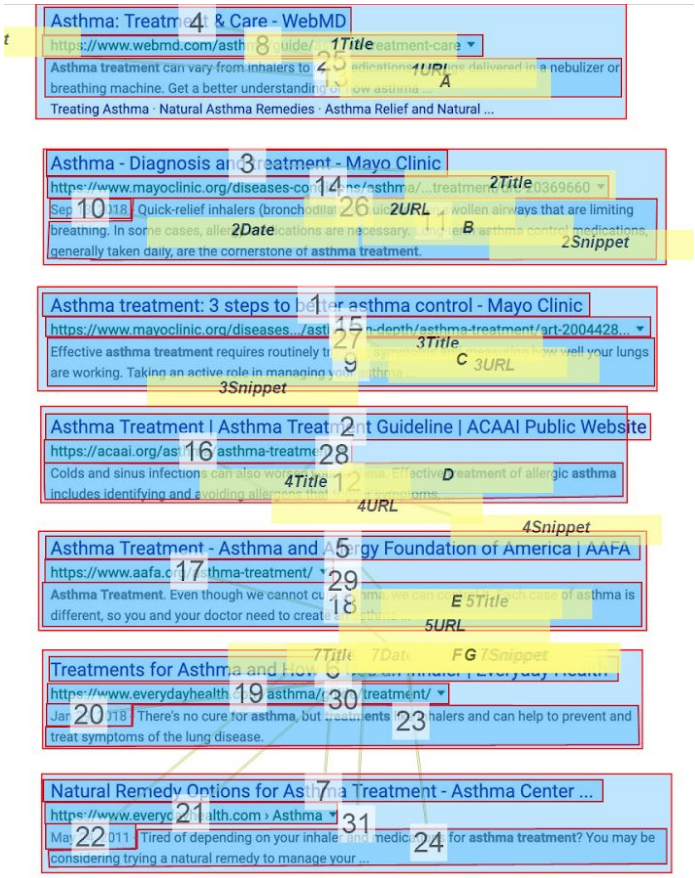


Figure 1 Example of AOI definition

Cristina et al.(2011) classified three types of intent of the query: informational, navigational and transactional. What's more, Asch, J. M. et cl. (2019) in their research pointed out that 56% of patients who searched health-related information online searched for symptoms and 23% of them searched about the treatment or management of a certain disease. We would like to know if individual's viewing behavior is different on the search result page of different query intents: symptom and treatment. In that case, we crafted our third research question and forth hypothesis:

RQ3: How users explore the search result page of different query intents: symptom and treatment?
H4. Users attend to different AOIs of the search results for queries about symptoms and treatments.

4. METHOD

This experiment analyzes ocular behavior on health-related SERPs with eye tracking. During the experiment, each participant viewed 10 Google search result pages. They could spend up to 2 minutes on each page and click one search result which would direct them to the next page.

4.1 Participants

Participants are students from The University of Texas at Austin, who at least have experience in using Google Search. In all, 15 participants completed the experiment in the IX Lab at the School of Information. Individuals who need glasses to see the computer screen were excluded because the glasses provide potentially disturbing reflections.

4.2 Stimuli

We selected 5 common diseases: Flu, Asthma, Migraine, Hypertension and Gastritis. All these diseases are curable and non-fatal. Then we screenshotted original 10 Google search result pages on the above-mentioned 5 diseases' symptom and treatment. Supported results and graphics were removed to keep only organic results on the page (see Figure 2). All the pages include seven units of search result.

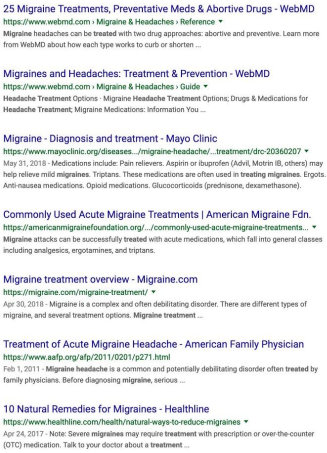
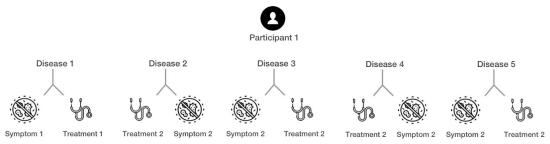


Figure 2 Example of SERP in the experiment

4.3 Tasks

Before we take the formal test all the participants should read and sign up the consent letter. During the experiment each participant is asked to view the 10 Google search result pages and click one search result. We will randomize the order of 5 diseases and the order of symptom and treatment of the same disease (see Figure 3) after each participant complete all the tasks to reduce the bias of order.



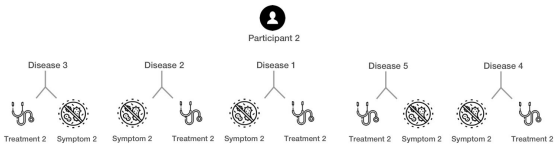


Figure 3 Randomize order of search result pages

4.3.1 Task Script

Study Introduction:

Thank you for your participation in our study. As a reminder, this is a study in which we will be using eye-tracking technology to study health-related information search behavior. In this study, you will be asked to complete 10 specific tasks using google. The whole experiment will take 20-30 minutes.

It is important that you know this study is completely voluntary and you have the right to pause or completely end this study at any time. It is also important that you know, we are not evaluating your performance for these tasks, we are simply here to observe how you interact with the system. Any issues or difficulties you encounter when using the system is not a reflection of your performance.

Before we get started, I would like to ask you to fill out this consent form. [Present the consent form to the participant]

[Receive consent]

Now, we are going to get started with our tasks. You are going to experience 10 different scenarios. In each scenario, you will view one search result page of disease symptom/treatment. The task will end when you click on the link that you would like to click or when you have looked at the page for 2 minutes.

Do you have any questions for me?

4.3.2 Scenarios

1. Treatment

- Flu
  Your friend Elsa didn't go to school today because she got flu yesterday. And now you are searching for the flu treatment for her.
- Asthma
  Your cousin Austin has asthma problems since his childhood. Recently it gets worse. Now, you are searching for Asthma treatment for her.
- Migraine
  Your friend Kavin got Migraine, so he doesn't want to use any electric device now but he wants to know the Migraine treatment. Could you help him to search for the Migraine treatment?
- Hypertension
  Your friend's father has Hypertension for years. Your friend is worried about his father health and you want

to search the Hypertension treatment to help your friend.

- Gastritis
  You accompanied your friend Lisa to the doctor. After diagnose, Lisa told you she got gastritis. You are worried about her so you would like to search for the Gastritis treatment for Lisa.

2. Symptom

- Flu
  As winter approaches to Austin, the temperature decreases rapidly. Your nephew Elsa thinks she might get the flu. Please search the symptom of flu for her.
- Asthma
  The first Tuesday of May is World Asthma Day. To enhance people's knowledge of Asthma and thus prevent it, please search for the symptom of asthma.
- Migraine
  Your friend Kavin thinks he might get the migraine. Please search the symptom of migraine for him.
- Hypertension
  Your grandma is calling you, saying that she forgot where the sphygmomanometer is and she thinks she might get hypertension. Please search for the symptom of hypertension for her.
- Gastritis
  Your friend Jay always skips breakfast and now she thinks she might get gastritis and asks you to search the symptom of gastritis for her.

4.4 Procedure

After recruiting participants qualified for our experiment, we scheduled a 20 min eye-tracking test with each participant. The test was taken in the IX Lab of School of Information. During the experiment, participants' eye movements were tracked using Tobii TX300 screen-based eye-tracker and iMotions software. A post interview was conducted where participants could view their recorded eye movements.

Participants are expected to 1) explain their choice of certain search result; 2) explain and rank the elements that affect their decision; 3) provide basic demographic information including web search skills and knowledge on the 5 diseases.

Post Test Questionnaire

1. How often do you use a search engine to search health-related information?

a. Many times everyday
b. About once a day
c. A few times a week
d. Once a week
e. A couple of times a month
F. Never

2. In what degree do you know about these disease symptoms and treatments?

Note. Judgments were made on 5-point Likert scales (1 = not applicable at all; 2=rarely applicable; 3=moderately applicable; 4=largely applicable; 5=totally applicable).

No.	Item	Symptoms	Treatments
1	I know a lot about flu		
2	I know a lot about asthma		
3	I know a lot about migraine		
4	I know a lot about hypertension		
5	I know a lot about gastritis		

3. Except our experiment have you ever searched any of these five diseases online?

□Flu
□Asthma
□Migraine
□Hypertension
□Gastritis

4. What affects your decision of clicking on this link? You could see the videos recorded to help recall.

5. Could you rank the following elements according to how much attention you pay to when you are searching these five diseases?

□Result title
□Snippet
□Date
□Rank of the search result
□Link URL

6. Follow up question: Why did you spend a long time reading xxx?

7. Our research is about health-related information search behavior. Do you have any suggestions or comments for this study? We appreciate your responses.

Screener:

1. Do you use corrective eyewear (Select all that apply):
   a. None
   b. Glasses
   c. Contacts
   d. Other (please specify): _____
   e. If yes, when do you use them:
      i. Always
      ii. Never
      iii. For seeing objects/screens close-up
      iv. For seeing objects/screens from far away
      v. Other (please specify): _____

4.5 Measures

A. Eye tracking metrics

We use Tobii TX300 screen-based eye-tracker and iMotions software to collect data.

Time to first fixation, fixation point, fixation duration, and time spent are the four metrics that help us explore the determinants of visual behavior on health-related information search result pages. The Time to First Fixation (TTFF) indicates the amount of time that it takes a participant to look at a specific AOI. Fixation points show what the eyes are looking at. Fixation duration refers to the time visual gaze on a single location, around 200-300 milliseconds. Time spent quantifies the amount of time that participants have spent looking at a specific AOI. During the experiment, fixations in each AOI and fixation durations(ms) were recorded.

B. Retrospective Verbal Protocol

During the experiment, participants' eye movements will be tracked using Tobii TX300 screen-based eye-tracker and iMotions software. A post questionnaire will be delivered to collect demographic information and inquiry cognitive process.

C. Hypothesis Test

We applied Kruskal-Wallis test to test H1 because valid fixation follows non normal distribution (see Figure 4 and Figure 5). H3 and H4 were tested by Mann-Whitney U Test since it works fine with unequal sample sizes and applies to non normal distribution sample.

5. RESULTS & DISCUSSION

5.1 Data Analysis for Hypothesis 1

This section is to see whether the rank of the search result affects the time spent on the health-related information search results. For each page, we have seven search results. So the input has seven samples and their sample sizes (valid AOI fixation number) are different. From Figure 4 and Figure 5 we can know that the second rank AOI receives the most attention (time spent-fixation (ms)). The possible reason is that the second results are shown on the center of the screen in our experiment environment and thus could receive more valid fixations. What's more, from Figure 5 we know that generally, the higher the rank is, the more time participants would spend on it.

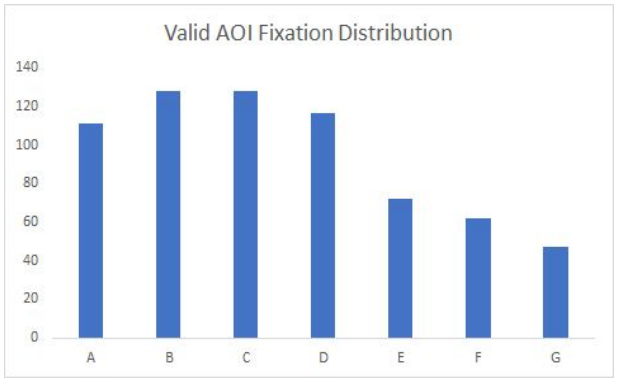


Figure 4. Valid AOI Fixation Count Distribution.

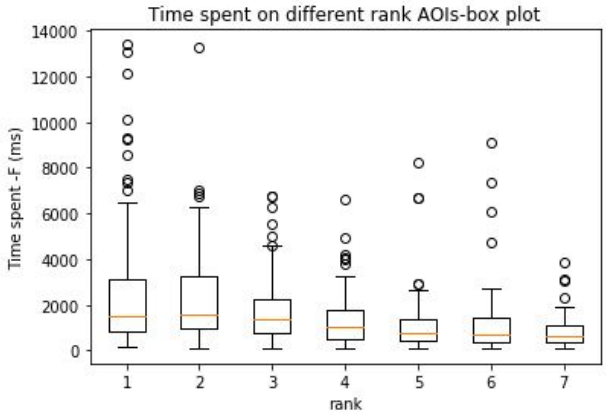


Figure 5. Time spent on different rank AOIs-box plot.

We then use Kruskal-Wallis test by rank because these seven samples have different sample sizes. This non-parametric method is suitable since we could see whether samples derived from same distribution, and know whether the rank would affect time spent. We also tried one-way anova to test H1. The independent variable is the rank while the dependent variable is the time spent.

One-way Anova Results are $F(1,7) = 59.18$, p value<0.001. Kruskal-Wallis test by rank results are Statistic=77, p value <0.001. Both Kruskal-Wallis test by rank and one way anova show that there is significant difference between the 7 big AOI search results, which means rank affects time spent.

Figure 6. Asthma Treatments Heatmap

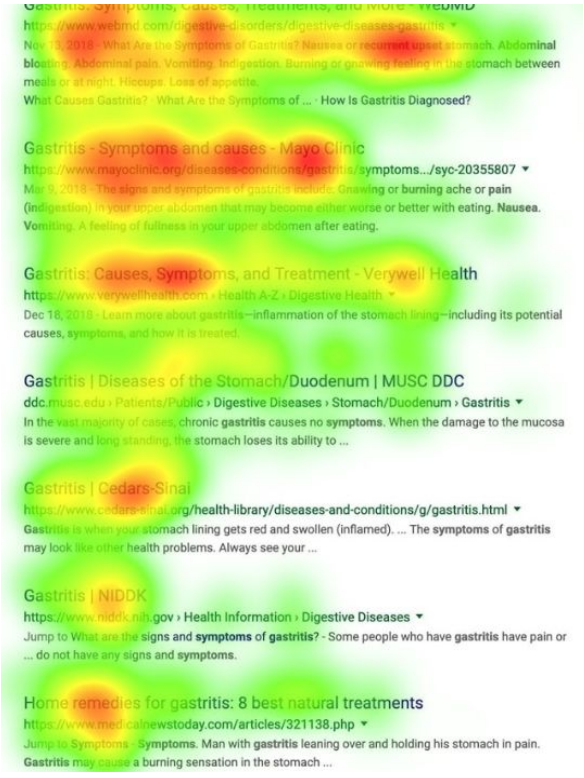


Figure 7. Gastritis Symptoms Heatmap

5.2 Data Analysis for Hypothesis 2

This section explores the relationship between whether participants click on the results or not (click behavior) and time spent.

Pearson correlation coefficient between the Time Spent-F and click behaviors is 0.301, which is a median correlation. We use one-way ANOVA T-test to see whether there is a significant difference in Time Spent-F between the results clicked and not clicked. The independent variable is whether click or not while the dependent variable is the Time Spent-F (ms).

The result of one way ANOVA T-test is $F(1,2) = 66$ and p value<0.01, which indicates that there is a significant difference in Time Spent-F between the results the clicked results and not clicked results.

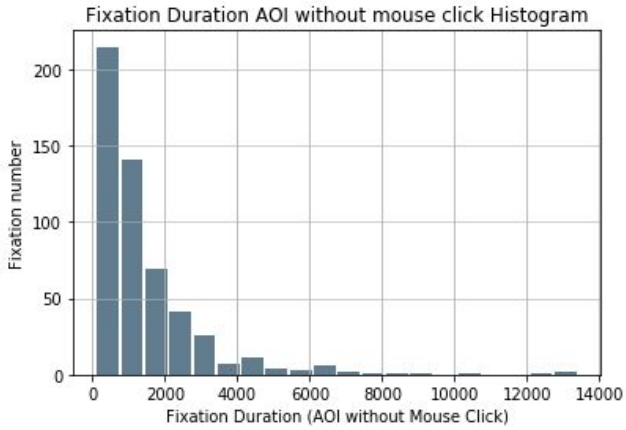


Figure 8. Fixation Duration AOI without mouse click

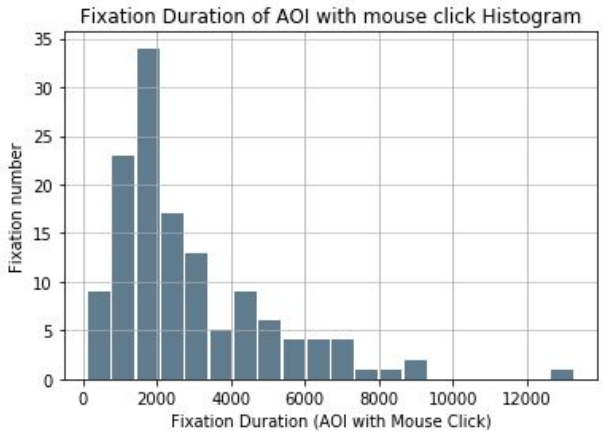


Figure 9. Fixation Duration of AOI with mouse click

Fig. 8 and Fig. 9 show that participants are more likely to click the search result which they pay more attention to.

5.3 Data Analysis for Hypothesis 3

This section is to see whether the attention captured by AOI with date and that captured by AOI without date have same distribution.

We use Mann-Whitney U test because it is non-parameters test . Our samples follow non normal distribution and have unequal sample sizes: the number of valid fixation points with date is 62 while the number of valid fixation points with date is 603.

Table 1. Mann Whitney U Result

	TTFF	Fixation Count	Time spent-F (ms)
statistic	18468.5	7436.5	7315.5
pvalue	0.438	< 0.001	< 0.001

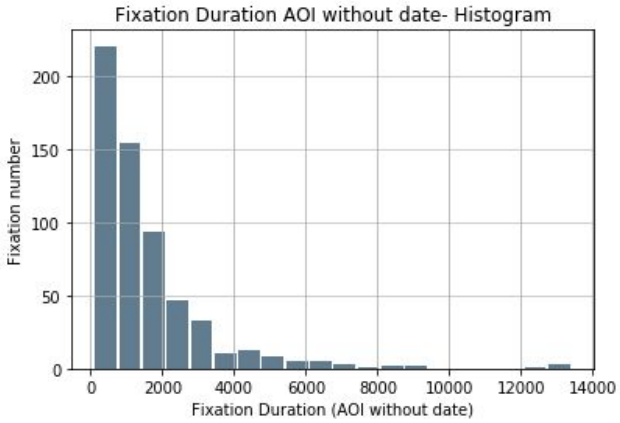


Figure 10. Fixation Duration AOI without Date

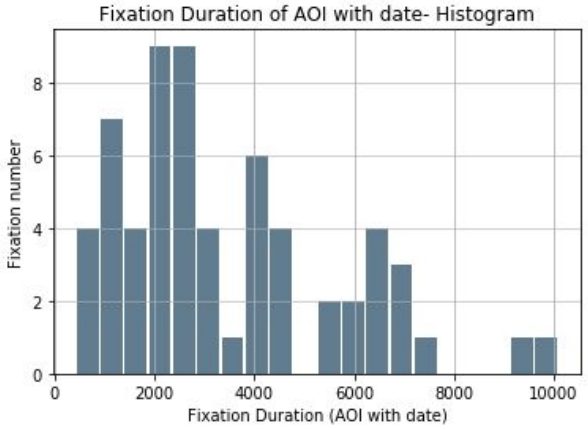


Figure 11. Fixation Duration of AOI with Date

The results of Mann Whitney U Test indicate that there is no significant difference between AOIs with and without Date at TTFF, which makes sense because Date AOI is small and thus cannot catch the immediate attention. However, the time spent and fixation count indicate that overall, Date has a positive influence on participants' viewing behavior.

5.4 Data Analysis for Hypothesis 4

In this section we use Mann-Whitney test to compare the TTFF, Fixation Count, and Time Spent on Title, URL, Date, Snippet when searching symptoms and treatments to see if their distribution is the same.

Table 2. TTFF - Mann Whitney U Result

	Title	Url	Date	Snippet
statistic	25043.0	20558.5	425	13977.5
pvalue	0.43	0.46	0.22	0.19

Table 3. Fixation Count - Mann Whitney U Result

	Title	Url	Date	Snippet
statistic	24240	20061	348.5	13827.5
pvalue	0.22	0.29	0.01	0.14

Table 4. Time Spent -F (ms) - Mann Whitney U Result

	Title	Url	Date	Snippet
statistic	23940	20279	347.5	14284
pvalue	0.16	0.36	0.03	0.29

Only Fixation and Time Spent in Date AOI show significant difference on symptom and treatment, but that maybe caused by few data on Date AOI.

6. CONCLUSION

In this paper, we studied users' viewing behavior on health-related SERPs. We found that the rank of search results affects the fixation duration. Fixation points on the second and third search results are more than those on other search results. This is because when participants were presented with the stimuli, they spent a third of a second looking at the middle of the interfaces before doing any acting (Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W., 2008). From Figure 8 and Figure 9, we note that participants are more likely to click on the search results which they pay more attention. About AOIs, although there is no significant difference between TTFF on AOIs (search results) with and without Date AOI, AOIs with Date has higher Fixation Count and Time Spent-F. In our research, we asked participants to view the information with different intents (looking for symptom or treatment). The result shows that there is no significant difference in participants' viewing behavior between the two different query intents. This result needs further research because 1) the demographic of participants are too simple, 2) the small amount of the dataset

7. LIMITATIONS & FUTURE WORK

Our research offers the very first attempt to understand how rank, attributes, query intent in search result influent viewing behavior on health-related SERPs. In our study, we only focused on the 10 static SERPs in the format of results on the Google

search engine which were made of 7 textual organic results each of which includes 3 or 4 AOIs. One next step for our research is to conduct more experiments on SERPs with more variations, e.g., including promotional ads, search results with images or sub-links, or results without our current AOIs. And we can explore whether viewing behavior is different on scrolling page and static page by add scrolling page to stimuli.

Another limitation we have is that all of our participants are from 20 to 30 years old and we asked them to search and view search results with only two different query intents. In addition, more tests on viewing behavior on different query intents can enrich our discussion on the H4. For example, we can ask participants to browse the page to find health-related information for self-diagnosing, giving advice, academic reference, etc. Finally, we would like to include diverse age/ level of education subjects and collect more data to decrease demographic bias in this research.

REFERENCES

1. Lorigo, L. , Haridasan, M. , Hrönn Brynjarsdóttir, Xia, L. , Joachims, T. , & Gay, G. , et al. (2010). Eye tracking and online search: lessons learned and challenges ahead. *Journal of the Association for Information Science & Technology, 59*(7), 1041-1052.
2. Pian, W. , Khoo, C. S. , & Chi, J. . (2017). Automatic classification of users' health information need context: logistic regression analysis of mouse-click and eye-tracker data. Journal of Medical Internet Research, 19(12), e424.
3. Granka, L. A., Joachims, T., & Gay, G. (2004, July). Eye-tracking analysis of user behavior in WWW search. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 478-479). ACM.
4. Yvonne Kammerer, , Peter Gerjets, (2012), Chapter 10 How Search Engine Users Evaluate and Select Web Search Results: The Impact of the Search Engine Interface on Credibility Assessments, in Dirk Lewandowski (ed.) *Web Search Engine Research (Library and Information Science, Volume 4)* Emerald Group Publishing Limited, pp.251 - 279
5. Yvonne Kammerer & Peter Gerjets (2012) Effects of search interface and Internet-specific epistemic beliefs on source evaluations during Web search for medical information: an eye-tracking study, Behaviour & Information Technology, 31:1, 83-97, DOI: 10.1080/0144929X.2011.599040
6. Morahan-Martin, J. M. (2004). How internet users find, evaluate, and use online health information: a cross-cultural review. CyberPsychology & Behavior, 7(5), 497-510.
7. Graham, Dan J., and Robert W. Jeffery. "Location, location, location: eye-tracking evidence that consumers preferentially view prominently positioned nutrition information." *Journal of the American Dietetic Association* 111.11 (2011): 1704-1711.
8. Shaffer, Victoria A., Justin Owens, and Brian J. Zikmund-Fisher. "The effect of patient narratives on information search in a web-based breast cancer decision aid: an eye-tracking study." *Journal of medical Internet research* 15.12 (2013).
9. Buscher, Georg, Susan T. Dumais, and Edward Cutrell. "The good, the bad, and the random: an eye-tracking study of ad quality in web search." Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. ACM, 2010.
10. Smith, Catherine L., and Paul B. Kantor. "User adaptation: good results from poor systems." *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval.* ACM, 2008.
11. González-Caro, Cristina, and Mari-Carmen Marcos. "Different users and intents: An eye-tracking analysis of web search." *Proc. WSDM* (2011): 9-12.
12. Asch, J. M., Asch, D. A., Klinger, E. V., Marks, J., Sadek, N., & Merchant, R. M. (2019). Google search histories of patients presenting to an emergency department: an observational study. BMJ open, 9(2), e024791.
13. Lambert, S. D., & Loiselle, C. G. (2007). Health information—seeking behavior. Qualitative health research, 17(8), 1006-1019.
14. Marcos, M. C., Garcia-Gavilanes, R. O. G. G., Bataineh, E., & Pasarin, L. (2014, April). Using eye tracking to identify cultural differences in information seeking behavior. ACM SIGCHI.
15. Morahan-Martin, J. M. (2004). How internet users find, evaluate, and use online health information: a cross-cultural review. CyberPsychology & Behavior, 7(5), 497-510.
16. Granka, L., Feusner, M., & Lorigo, L. (2008). Eyetracking in online search. *Passive eye monitoring*, 283-304.
17. Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: A Bayesian framework for saliency using natural statistics. *Journal of vision*, *8*(7), 32-32.