The Impact of Ad Quality and Position on Mobile SERPs

Afrah Olayan Alanazi*
La Trobe University
AAlanazi@ltu.edu.au

Mark Sanderson
RMIT University
mark.sanderson@rmit.edu.au

Zhifeng Bao
RMIT University
zhifeng.bao@rmit.edu.au

Jaewon Kim†
Twitter
jaewonk@twitter.com

ABSTRACT

In this paper we aim to explore the effects of advertisement (ad) quality and position in search engine result pages (SERPs) on a mobile device. We conducted a lab-based eye-tracking study to investigate search time, behavior, and user satisfaction with ads of *good* or *bad* qualities positioned at the *top* or *middle* of organic results. Our findings suggest that users pay attention to ads regardless of their quality or position. However, they tend to pay different amounts of attention to organic results and SERPs because of ad quality. We also found that user satisfaction and the chance of clicking on an ad vary according to ad quality and position.

ACM Reference Format:
Afrah Olayan Alanazi, Mark Sanderson, Zhifeng Bao, and Jaewon Kim. 2020. The Impact of Ad Quality and Position on Mobile SERPs. In *2020 Conference on Human Information Interaction and Retrieval (CHIIR '20), March 14–18, 2020, Vancouver, BC, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3343413.3377990

1 INTRODUCTION AND RELATED WORK

Search engines, e.g., Google and Bing, often display advertisements (ads) at the top/middle/bottom of search engine result pages (SERPs). The ads look similar to the organic results on SERPs except for some visual clue such as an ad icon. The ads underpin the business model for many search engine companies; for marketers, they are an important method for attracting users to websites [7, 23].

In desktop web search, the ads in SERPs can be useful for searches for a specific product and service [4]. They are known to attract greater user attention if located on the top of a SERP [10]. Search behaviors with ads can differ because of their presentation styles [21]. However, people often ignore ads while searching [10, 12], and search engines sometimes provide bad quality ads, i.e., irrelevant to the information the user is looking for, which may cause lower search efficiency [4].

Mobile devices are commonly used for web search. Around 53% of paid-search clicks in the US come from mobile devices[1]. The cumulative value of such clicks was estimated to reach $28.25 Billion in 2019[2].

According to past work, e.g., [14, 15, 20, 24], mobile web search behavior is different from desktop search behavior because of the smaller screen sizes and the touch screen keyboards. Some studies [6, 7] support that there is no "ad blindness" on mobile devices, unlike on desktops; i.e., users will look at ads if they are provided by mobile search engines, which is a different behavior on desktop screens [10, 12]. However, other survey research shows that two-thirds of respondents skip ads while searching on mobile devices [23]. Although there have been a few studies related to how users interact with ads on mobile SERPs, we wondered how users interact with ads on mobile devices according to different ad qualities and positions in SERPs.

Considering the previous work on the effects of ads, we conducted a lab-based eye-tracking user study to investigate the following research questions:

- **RQ1. How does *ad quality* (good or bad) affect mobile web search behavior?** In desktop web searches, users pay more visual attention to good quality ads [4]. We ask if this observation is the same for mobile web search. To the best of our knowledge, this question has not been asked before.
- **RQ2. How do users interact with ads in different SERP *positions* (top or middle)?** Users may show different behavior due to different ad positions. This aspect of ad placement has not been studied before on mobile devices.

Eye–tracking provides information on user cognition [11, 25], and is commonly adopted to explore search behavior, e.g., [6, 9, 10, 15, 16, 18, 20]. To investigate the research questions, we employed an eye-tracker to measure users' attention to ads and organic results. In addition, we measured the search behavior, e.g., click and scrolling, and determined user satisfaction. The main difference between the previous eye-tracking user studies on mobile devices is that we measured the user attention on the areas beyond the page fold (i.e., scrolling from the initial viewports), which provides the comparison between user attention on ads and organic results.

2 USER STUDY

2.1 Subjects

We recruited 27 students with varying backgrounds, e.g., computer science, mathematics, nursing, and physics from a local university

*Work completed while studying at RMIT University.
†Work completed while employed by RMIT University.

[1]https://searchengineland.com/mobile-desktop-search-traffic-split-may-have-stabilized-at-roughly-60-40-317091
[2]https://www.invespcro.com/blog/mobile-search-engine-advertising/1

CHIIR '20, March 14–18, 2020, Vancouver, BC, Canada

ACM ISBN 978-1-4503-6892-6/20/03…$15.00
https://doi.org/10.1145/3343413.3377990

Table 1: Examples of task descriptions and queries and good and bad quality ads for each task.

Task description (given task query)	Good quality ad	Bad quality ad
• Assuming you get married soon and you are preparing a wedding ring for your partner. Find an online shop to purchase a wedding ring. (wedding ring online)	• Title: Buy Wedding Rings Online Now • Snippet: Wedding Rings From Only 350! Shop Online or In-store Today. Master Jewellers. Lifetime Guarantee. 12 Locations. Superior Quality.	• Title: Wedding Party Band | Melbournes Best Wedding Band • Snippet: No backing tracks - real musicians & the perfect atmosphere for your wedding.
• You want to transfer money to your mother's account, but the issue is that your mother is living in India. Find ways to transfer money online to another account overseas. (send money to india)	• Title: Get 30 Paise Better Rate on | Money Transfer to India Safely | icicibank.com • Snippet: With ICICI Bank Money2India. Send 10000 AUD in one Go. Completely online. Use Promo Code EASYN20. Official Website. Secure Transfer. Safe & Fast. Services: Phone Bill Pay...	• Title: Heading Overseas? | NAB Smart Travel Tools • Snippet: Learn About How Our Smart Travel Tools Can Make Your Overseas Travel Easier.Overseas Card Protection. Order Temporary Card. Made For Travel. What To Do If Stolen Card. Block and Unblock
• You have a MacBook and you want to install Microsoft office 2016 for the device. Check the price of a copy. (Microsoft Office 2016 for MAC)	• Title: 2016 microsoft office for mac on eBay | Fantastic prices on Top Items • Snippet: Free Shipping on eBay. Buy It Now Available. Featured Collections. Buyer Protection Program. Free Shipping Available. Huge Selection.	• Title: How to activate your Office 2016 license - unlicense.shop • Snippet: Get the best experience here. 24/7-Support. Online Chat-Support. Secure Payment Methods.

campus. We excluded three sessions due to low eye–gaze tracker calibration accuracy, leaving us with 24 subjects aged 23–39 (mean: 29.4, standard deviation (SD): 4.0). Using 7-point Likert-scale questions (1: completely unfamiliar, 7: familiar), subjects were asked how familiar they were with using search engines and mobile devices. The subjects marked high scores for both questions (mean: 5.42, SD: 1.44 and mean: 5.67, SD: 1.25, respectively). That is, most subjects considered themselves familiar with search engines and good at using mobile devices.

2.2 Tasks

All tasks were informational, in which the subjects were required to find a particular piece of information [3, 22]. Considering the results from previous studies–revealing no significant difference in subject attention on ads according to task type [4, 21]–we excluded navigational tasks, which are easier to complete [9, 22]. Each participant completed eight tasks (see Table 1 for examples of task descriptions and given queries).

We focused on the effects of ad quality (RQ1) and position (RQ2). Each SERP contains ten organic search results with two ads. Inspired by the selections of ad quality from Buscher et al. [4], we manually selected good quality ads in response to the given initial task queries. The bad quality ads were selected from SERPs which were generated by queries from a subset of terms in the initial queries. All organic results and ads were extracted from the Google mobile search engine. As shown in Table 1, a good quality ad included relevant information (i.e., the answer) for the task, whereas a bad quality ad did not contain the answer in the search result.

We varied ad positions, i.e., *top* and *middle*. SERPs with 'top' ads consisted of ⟨A1, A2, O1, O2, O3, … O10⟩, and the 'middle' position SERPs contained ⟨O1, O2, A1, A2, O3, O4, … O10⟩, where "A" and "O" denote ads and organic results, respectively. That is, in the top setting, the initial viewport displayed two ads in the top position, which occupied most of the mobile device's initial viewport and required using the scroll function to see the organic results. With the middle setting, two organic results were displayed at the top.

2.3 Design and procedure

We adopted a within-subject design to investigate search time, behavior, user attention and satisfaction due to ad quality (2) × position (2). Each participant completed two task sets for each ad quality with each set containing tasks with two top and two middle ad positions. To minimize any carry-over effect [13], we randomized task order within each task set, but the orders for ad quality and position were counter-balanced, and every task was evenly shown with different qualities and positions of ads across the subjects.

Once the subjects agreed to the study and gave consent[3], they were given instructions regarding the procedure and tasks. We showed them two sample tasks with each position and quality until the subjects were familiar with the experimental set up. We then calibrated subject gaze recording using a nine-point procedure. Here, the first task description and query were shown on screen. After this, the subjects proceeded to the first SERP. Once the subjects found the relevant information to the task description and wanted to move to next task, we considered the task to be completed. After each task, the subjects were asked about their overall satisfaction using a seven-point Likert scale, i.e., 1: completely dissatisfied, 7: completely satisfied. The process was repeated to the last task.

At the end of the experiment, the subjects were asked to fill out a post-experiment questionnaire, which included their demographic information, their experience, e.g., familiarity with search engines and using mobile devices, and their preference of ad quality and position. The subjects spent about 30–40 minutes in our eye-tracking lab, from reading the consent form to filling out the questionnaire. At the end of the session, the subjects were compensated with a $20 voucher for their participation.

2.4 Apparatus

We used an iPhone 8 plus for the experiment, which has a popular screen size (5.5 inches with a 1080×1920 pixel resolution). To collect the gaze data, we used a Tobii X2-60[4]. The mobile phone was placed

[3]The ethics approval number: SEHAPP 52-18
[4]https://www.tobiipro.com/product-listing/tobii-pro-x2-60/

Table 2: Search time, behavior, and satisfaction for each position, broken down by their quality.

		Good		Bad		p-value		
		Top	Middle	Top	Middle	Quality	Position	Interaction
Search time	Time to first click [s]	29.39	35.77	48.04	35.52	0.16	0.53	0.94
Fixation duration	on SERPs [s]	10.10	10.29	12.85	16.34	***	0.16	0.21
	on ads [s]	2.78	2.78	2.70	3.06	0.93	0.68	0.74
	on organic results [s]	7.32	7.51	10.14	13.28	***	0.08	0.58
Search behavior	Scroll rate	0.71	0.79	0.75	0.73	0.56	0.84	0.31
	Click rate on ads	0.35	0.17	0.08	0.02	***	*	0.94
User satisfaction	7-point Likert scale	6.35	6.10	5.52	5.90	**	0.56	*

*Significant at 0.05 level. ** Significant at 0.01 level. *** Significant at 0.001 level.

on a Tobii mobile stand device[5] and connected to the eye-tracking system as a secondary monitor using the Twomon software[6]. With this setup, we collected the gaze data directly from the mobile device rather than using a scene video camera, e.g., [6, 7, 20]. This allowed us to assign Areas Of Interest (AOIs) to each organic result or ad for fixation data, which is useful in understanding how long subjects spent time reading in a particular place [8, 15]. This setup differs from previous work on mobile search behavior with ads, e.g., [6, 7].

3 RESULTS AND DISCUSSION

We obtained data from 192 tasks, which contained 96 tasks for each good and bad ad quality, and 48 tasks for each position of ads within each ad quality. We tested the statistical power of our design [5] with the significant level $\alpha = 0.05$, and we confirmed that our data sets would maintain the power, $1 - \beta \geq 0.95$ for all comparison in this paper. We focused on the effects of both by ad quality, i.e., good or bad, and position, i.e., top or middle.

Several analysis techniques were used in this study. First, we employed the analysis of variance (ANOVA) for continuous data such as time to first click and Fixation Duration data with a log-transformation $log(x + 1)$ to maintain the normality assumption, if necessary. Second, a generalized linear mixed model (GLMM) [2] with a binomial distribution and a logit link function was used for binary data. Third, we used a linear mixed model (LMM) [27] for the score data from the 7-point Likert scale. To consider individual differences, we used a block structure (subject) for ANOVA, and adopted a GLMM and LMM instead of a generalized linear model (GLM) and a linear model (LM), because the observed random effects between subjects (σ_s^2) were greater than standard errors (SEs) in all variables. All analyses were conducted using the GenStat statistical package [26].

3.1 Search time and user attention

We first looked at how much time the subjects spent and how much attention they paid to the initial SERPs. For the search time, we defined the *time to first click* as the elapsed time to the first click on a SERP, because the initial SERPs were evenly shown to the subjects and users' paths might diverge after their first click. The mean of time to first click as shown in Table 2 (min: 29 s and max: 48 s) may imply that there are some effects caused by both research variables: ad quality or position. However, neither research variable had any significant effect on time to first click.

As mentioned earlier, the Fixation duration is a useful indicator of how long a subject spends obtaining information [4, 9, 15, 17, 20]. In term of the average fixation duration on each SERP, we found significant effects on the Fixation duration on the SERPs ($F_{(1,165)} = 11.45$, $p < 0.001$) due to ad quality. This result indicates that the subjects spent more time reading in the presence of bad quality ads (mean: 14.59s, 12.85s and 16.34s for the top and middle, respectively) than good ads (mean: 10.20s, 10.10s and 10.29s for the top and middle, respectively).

To investigate which element between the ads and organic search results leads to the result, we assigned two Areas Of Interest (AOIs) to ads and organic results to compare the subject's attention. There was a significant effect on the organic results ($F_{(1,165)} = 11.89$, $p < 0.001$) due to ad quality, whereas no effect was observed on the Fixation duration of the ads. This implies that the effect on the average fixation duration on each SERP appears to come from the difference in the subject attention on organic results (7.41s and 11.7s with good and bad ad qualities, respectively). In other words, if search engines provide ads on mobile SERPs, users look at the ads regardless of the ad quality and position, but they spend more time reading organic search results when bad ads are presented. Although an additional qualitative study such as survey and interview approaches is required, it seems that our subjects lost their trust on ads due to the bad quality, so they tried to find an answer from the organic results with the longer reading time.

The result of Fixation duration on ads was consistent with some past work concerning ad effects in mobile web search [6, 7], i.e., subjects will look at ads if they are provided by mobile search engines. However, our result does not support the survey result from other previous work [23], i.e., users skip ads on mobile devices. Our subjects spent some time reading the ads regardless of their quality and position (2.70–3.06s). This was different from past results with a desktop [4], i.e., good quality ads received more attention from users than bad quality ads, the difference seems to be caused by the smaller screen size on a mobile device, which requires using the scroll function.

[5]https://www.tobiipro.com/product-listing/mobile-device-stand/
[6]http://www.easynlight.com/en/twomonusb/

3.2 Search behavior

We describe the results of two metrics: scroll and click. First, the different conditions had no significant effect on the scroll rate. Unlike the results from previous studies regarding mobile web search with no ads [15, 17], which suggested scroll rates of less than 50%, our subjects with ads recorded relatively higher chances of using the scroll function than past results: about 71–79% as can be seen in Table 2. This suggests that ads on mobile web search tend to increase the scroll rate, although we need to observe the frequency and depth of scrolling to look into the effect in detail.

Second, the chance of clicking on ads exhibited a significant difference due to ad quality and position (σ_s^2 = 1.326, X^2 = 17.99, df = 1, p < 0.001, and X^2 = 3.90, df = 1, p < 0.05, respectively). As can be seen in Figure 1, subjects who saw good quality ads tended to click on them more (mean: 26% and 5% for good and bad ad qualities, respectively). Across both ad qualities, the chances of clicks were higher (mean: 22%) when the ads were located in top ranks compared to middle ranks (mean: 9%). This may indicate that the ad quality was not only important, but the top ranks of ads were also attractive in receiving clicks. In addition, the lower chance of clicking on bad quality ads appears to result in a longer fixation duration on organic results with such ads. We leave a correlation analysis for the further analysis.

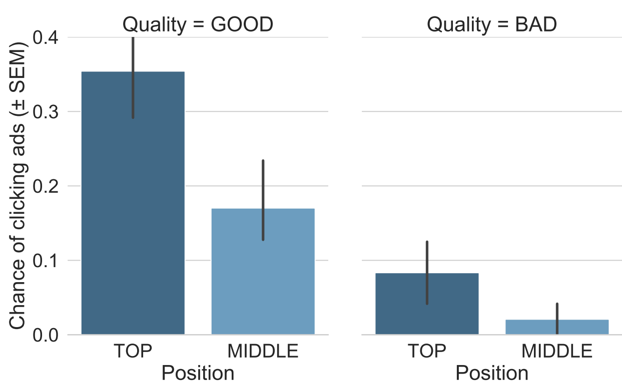

Figure 1: Chance of clicking ads for each position, broken down by their quality.

3.3 User preference and post-experiment questionnaire

We measured the overall user satisfaction after each task. Table 2 shows that an interaction between ad quality and position had significant effects on user satisfaction (σ_s^2 = 0.295, X^2 = 10.14, df = 1, p < 0.01, and X^2 = 4.02, df = 1, p < 0.05, respectively). Our subjects preferred SERPs with good quality ads over bad quality ads. However, according to a post-hoc test using standard errors of difference (SEDs) for the investigation of the interaction between two variables, subjects expressed different patterns of satisfaction when the bad quality ads were in the middle positions. As can be seen in Figure 2, when bad quality ads were located in top ranks, the subjects expressed lower satisfaction (5.52) compared to the satisfaction with good quality ads (6.23). However, when the bad ads were in the middle condition, the subjects recorded similar scores (5.90) to the satisfaction recorded with good quality ads.

At the end of the experiment, we asked subjects about their preferences for ad quality and position. Sixteen of the twenty four subjects replied that they preferred the SERPs with good quality ads, the remaining eight answered that the ad quality did not matter for their search. This may suggest that some subjects considered that good quality ads were helpful for their search. Regarding the ad position, half of the subjects replied that the ad position did not matter in mobile web search, whereas nine of the subjects preferred the top rank ads.

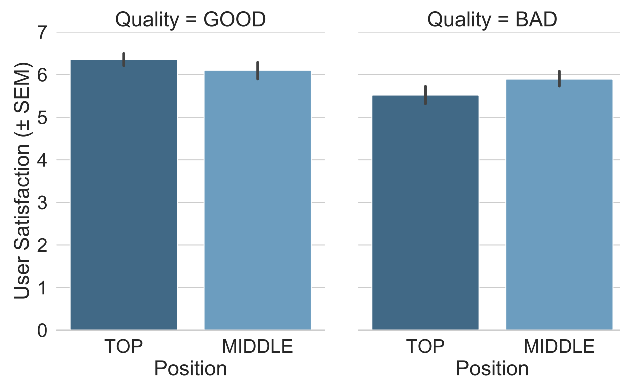

Figure 2: User satisfaction (7-point Likert scale) for each position, broken down by their quality.

3.4 Limitations

Although we carefully designed the user study, there are limitations. The subjects used were university students. Consequently, we acknowledge that these results may not represent the search behavior of the general public, and the result may vary according to the screen size of mobile devices [17] and different subject groups.

4 CONCLUSIONS

We explored how users interacted with ads in mobile web search. We investigated the effects of ad quality (good or bad) and position (top or middle) in SERPs. For RQ1–ad quality–the subjects presented with bad quality ads tended to pay more attention to the organic results and the overall SERP. They were less likely to click ads and expressed lower user satisfaction. For RQ2–ad position–subjects clicked more on ads found in top ranks compared to ads in middle ranks. Subjects expressed the least satisfaction when top ranked ads were bad quality.

As a preliminary analysis, this paper provides basic information on the effects of ad quality and position on mobile web search. We plan to conduct further analyses to investigate the relations among the measurements, e.g., between the chance of clicking on ads and user attention and between the user satisfaction and user attention, and also to explore the effects of the scanpath [15, 17, 22] on the user's attention, and search strategy [1, 8, 19].

Acknowledgment. This work was partially supported by the Australian Research Council's Discovery Project Scheme (DP170102726).

REFERENCES

[1] Anne Aula, Päivi Majaranta, and Kari-Jouko Räihä. 2005. Eye-tracking reveals the personal styles for search result evaluation. In *IFIP Conference on Human-Computer Interaction*. Springer, 1058–1061.

[2] Norman E Breslow and David G Clayton. 1993. Approximate inference in generalized linear mixed models. *J. Amer. Statist. Assoc.* 88, 421 (1993), 9–25.

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

[4] Georg Buscher, Susan T Dumais, and Edward Cutrell. 2010. The good, the bad, and the random: an eye-tracking study of ad quality in web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 42–49.

[5] Shein-Chung Chow, Jun Shao, Hansheng Wang, and Yuliya Lokhnygina. 2017. *Sample size calculations in clinical research*. Chapman and Hall/CRC.

[6] Soussan Djamasbi, Adrienne Hall-Phillips, and Ruijiao Rachel Yang. 2013. SERPs and ads on mobile devices: An eye tracking study for generation Y. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 259–268.

[7] Alexa Domachowski, Joachim Griesbaum, and Ben Heuwing. 2016. Perception and effectiveness of search advertising on smartphones. In *Proceedings of the 79th ASIS&T Annual Meeting: Creating Knowledge, Enhancing Lives through Information & Technology*. American Society for Information Science, 74.

[8] S.T. Dumais, G. Buscher, and E. Cutrell. 2010. Individual differences in gaze patterns for web search. In *Proceedings of the third symposium on Information interaction in context*. ACM, 185–194.

[9] Zhiwei Guan and Edward Cutrell. 2007. An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 417–420.

[10] Adrienne Hall-Phillips, Ruijiao Rachel Yang, and Soussan Djamasbi. 2013. Do ads matter? An exploration of web search behavior, visual hierarchy, and search engine results pages. *Computer Society Press* (2013).

[11] Robert JK Jacob and Keith S Karn. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The mind's eye*. Elsevier, 573–605.

[12] Bernard J Jansen and Marc Resnick. 2006. An examination of searcher's perceptions of nonsponsored and sponsored links during ecommerce Web searching. *Journal of the American Society for information Science and Technology* 57, 14 (2006), 1949–1961.

[13] Diane Kelly. 2009. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval* 3, 1–2 (2009), 1–224.

[14] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, and Tom Gedeon. 2012. Comparing scanning behaviour in web search on small and large screens. In *Proceedings of the Seventeenth Australasian Document Computing Symposium*. ACM, 25–30.

[15] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, Tom Gedeon, and Hwan-Jin Yoon. 2015. Eye-tracking analysis of user behavior and performance in web search on large and small screens. *Journal of the Association for Information Science and Technology* 66, 3 (2015), 526–544.

[16] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, Tom Gedeon, and Hwan-Jin Yoon. 2016. Pagination versus scrolling in mobile web search. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 751–760.

[17] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, Tom Gedeon, and Hwan-Jin Yoon. 2016. Understanding eye movements on mobile devices for better presentation of search results. *Journal of the Association for Information Science and Technology* 67, 11 (2016), 2607–2619.

[18] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, Tom Gedeon, and Hwan-Jin Yoon. 2017. What snippet size is needed in mobile web search?. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*. ACM, 97–106.

[19] K. Klöckner, N. Wirschum, and A. Jameson. 2004. Depth-and breadth-first processing of search result lists. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1539–1539.

[20] Dmitry Lagun, Chih-Hung Hsieh, Dale Webster, and Vidhya Navalpakkam. 2014. Towards better measurement of attention and satisfaction in mobile search. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. ACM, 113–122.

[21] Zeyang Liu, Yiqun Liu, Min Zhang, and Shaoping Ma. 2014. How Do Sponsored Search Results Affect User Behavior in Web Search?. In *Asia Information Retrieval Symposium*. Springer, 73–85.

[22] Lori Lorigo, Bing Pan, Helene Hembrooke, Thorsten Joachims, Laura Granka, and Geri Gay. 2006. The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management* 42, 4 (2006), 1123–1131.

[23] Enrique Murillo. 2017. Attitudes toward mobile search ads: a study among Mexican millennials. *Journal of Research in Interactive Marketing* 11, 1 (2017), 91–108.

[24] Kevin Ong, Kalervo Järvelin, Mark Sanderson, and Falk Scholer. 2017. Using information scent to understand mobile and desktop web search behavior. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 295–304.

[25] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372.

[26] VSN International. 2017. GenStat for Windows 19th Edition. VSN International, Hemel Hempstead, UK. Web page: www.vsni.co.uk.

[27] Brady T West, Kathleen B Welch, and Andrzej T Galecki. 2014. *Linear mixed models: A practical guide using statistical software*. CRC Press.