

How users' knowledge of advertisements influences their viewing and selection behavior in search engines

Sebastian Schultheiß  | Dirk Lewandowski 

Department of Information, Hamburg
University of Applied Sciences, Hamburg,
Germany

Correspondence

Sebastian Schultheiß, Hamburg
University of Applied Sciences,
Department of Information, Finkenau
35, 22081 Hamburg (Germany).
Email: sebastian.schultheiss@haw-
hamburg.de

Abstract

According to recent studies, search engine users have little knowledge of Google's business model. In addition, users cannot sufficiently distinguish organic results from advertisements, resulting in result selections under false assumptions. Following on from that, this study examines how users' understanding of search-based advertising influences their viewing and selection behavior on desktop computer and smartphone. To investigate this, we used a mixed methods approach ($n = 100$) consisting of a pre-study interview, an eye-tracking experiment, and a post-study questionnaire. We show that participants with a low level of knowledge on search advertising are more likely to click on ads than subjects with a high level of knowledge. Moreover, subjects with little knowledge show less willingness to scroll down to organic results. Regarding the device, there are significant differences in viewing behavior. These can be attributed to the influence of the direct visibility of search results on both devices tested: Ads that were ranked on top received significantly more visual attention on the small screen than the top ranked ads on the large screen. The results call for a clearer labeling of advertisements and for the promotion of users' information literacy. Future studies should investigate the motivations of searchers when clicking on ads.

Search engines like Google predominantly make money through search-based advertising, that is, advertisements shown in response to users' queries. These "sponsored results" are usually shown on the search engine results pages (SERPs) at the top and before the not paid for, so-called organic results. The question arises in how far users are aware of this distinction between paid-for and not paid-for results, and whether this knowledge influences their choice of results on the SERPs.

Distinguishing between organic results and advertisements has probably become more difficult over the years, as the snippets on the SERPs for these two result types look very similar (cf. Lewandowski, Kerkmann, Rümmele, &

Sünkler, 2018). In addition, search engines seem to further blur the lines between organic results and advertisements through changes in labeling.¹ As Ginny Marvin writing in industry newsletter Search Engine Land put it, "text ads have never looked more native." (Search Engine Land, 28.5.2019).

Prior research had found that the information literacy of search engine users is rather low, for example, when it comes to formulating precise questions (Stark, Magin, & Jürgens, 2014) or solving complex tasks (Singer, Norbistrath, & Lewandowski, 2012). This low level of information literacy was also observed regarding advertisements. Users are hardly able to distinguish between ads and organic results on search

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Journal of the Association for Information Science and Technology* published by Wiley Periodicals LLC on behalf of Association for Information Science and Technology.

engine result pages (Lewandowski et al., 2018) and are not well informed about Google's business model (Lewandowski et al., 2018). In an experiment, users who were not able to distinguish ads from organic results clicked on ads about twice as often as users who were able to differentiate both result types (Lewandowski, 2017).

This paper, which describes an eye-tracking experiment in combination with pre-study interviews and post-study questionnaires, makes several contributions to the literature. First, we investigate the extent to which users' actual viewing and clicking behavior on search engine ads correlates with their understanding of search engine advertising. Second, we wanted to find out whether there are differences in behavior patterns when results are presented on a desktop computer (large screen) versus a smartphone (small screen). We consider search on mobile devices in addition to desktop search, as this now exceeds desktop search in terms of search volume (Sterling, 2016). Third, with $n = 100$ subjects, the study differs from the majority of comparable eye-tracking studies in terms of the number of participants. Finally, we provide the software code and the questionnaire so that they can be used for further research.

1 | LITERATURE REVIEW

1.1 | Search engine results pages

Search engine results pages (SERPs) consist of various result types: *organic results*, *ads*, *universal search results*, and *knowledge-graph results*. *Organic results* are generated by algorithms from the search engine's index of the web and ranked according to equal criteria. *Text ads* are context-based ads that match a query and closely resemble organic results. *Shopping ads* differ from text ads in that they contain a product photo, the price and the name of the retailer, or other information. *Universal search results* are results from other, so-called vertical collections (e.g., images, videos). *Knowledge-graph results* give factual information directly on the SERP in answer to various questions, such as questions about famous personalities (Lewandowski et al., 2018, p. 421).

Google regularly revises the design of SERPs, for example, by adding new features such as "infinite scroll" on mobile devices (Schwartz, 2018). Furthermore, elements of result snippets vary between results, that is, some having additional information to the usual elements title, URL, and description.

1.2 | Search engine user behavior

Users' selection behavior on the SERPs is heavily influenced by the position, visibility, and design of a

search result. Thus, users prefer the first ranked results (Granka, Joachims, & Gay, 2004; Petrescu, 2014), results that are in the so-called "visible area" of the SERP, that is, results that can be seen without scrolling down the page (Cutrell & Guan, 2007; Sachse, 2019), and are encouraged to click due to the size and graphic design of a result (Liu, Liu, Zhou, Zhang, & Ma, 2015).

Users heavily trust the rankings generated by search engines. This has been shown in a representative questionnaire-based study (Purcell, Brenner, & Rainie, 2012), where 73% of US users said they feel confident that most or all of the information they find through search engines is accurate and trustworthy, and 66% said search engines are a fair and unbiased source of information.

In eye-tracking studies (Kammerer & Gerjets, 2014; Pan et al., 2007), it has been shown that there is a strong trust in Google's list interface to rank the most relevant results at the top of the SERP. The participants selected the first results even if they were irrelevant (by experimental manipulation) or were less trustworthy sources. The rank thus had the strongest influence on the selection behavior of the subjects. In a replication study of Pan et al. (2007), Schultheiß, Sünkler, and Lewandowski (2018) confirmed the huge influence of the result order on users' fixations and clicking behavior found in the original study. They found, however, that the crucial factor for a result to be clicked was the relevance and not solely its position on the SERP. However, eye-tracking studies in general have limitations regarding laboratory environment and sample sizes. The external validity of the results is threatened by the laboratory environment with often pre-defined SERPs and the very small samples of about 30 subjects on average (see Lewandowski & Kammerer, 2020; Strzelecki, 2020).

Using search engine transaction log data, a study by Keane, O'Brien, and Smyth (2008) showed results similar to the eye-tracking study by Pan et al. (2007). On reversed SERPs, most clicks were still on the first results. Only a few subjects seemed to have searched for the originally first ranked results and clicked on them even though they were at the bottom of the list. This search for a satisfactory result is supported by another study by O'Brien and Keane (2006) which found that a low-ranked result was more likely to be clicked if no similarly relevant results that could have satisfied the user were listed before it.

1.3 | Google's market share and business model

The Google search engine has a Europe-wide market share of over 90% (European Commission, 2017). In the

United States, the share is slightly lower at about 88% (StatCounter, 2019). In 2018, Google generated a profit of 30.7 billion dollars on a turnover of 136.8 billion dollars. About 83% of the turnover was generated by advertising (Alphabet Inc., 2019). However, a representative study of German internet users showed that a large proportion of users do not understand Google's business model. When asked how Google generates its revenue, 40% mentioned wrong sources of revenue or said they did not know the answer (Lewandowski, 2017).

1.4 | Ads labeling

In Google's desktop search, ads are currently labeled with a green ad label within a green frame. In mobile search, the previously identically designed label was replaced by a non-framed, black "ads" label. The labeling is regularly changed by Google (Marvin, 2020), with a trend toward more subtle labeling, which in turn leads to an increase in the number of clicks on ads (Edelman, 2014). For shopping ads on desktop computers and mobile devices, as well as for mobile text ads, an additional info button is displayed, which provides information on how the ads are generated (Google.com, 2019).

However, the differentiability of the ads from the organic results is insufficient. According to a representative study in Germany, the majority of search engine users cannot distinguish advertising from organic results (Lewandowski et al., 2018). Users who were unable to make that distinction clicked on the first ad twice as often as users who were able to do so (Lewandowski, Sünkler, & Kerkmann, 2017). On the one hand, the lack of differentiability is considered to be caused by the high similarity of text ads and organic results, as already mentioned. On the other hand, the labeling of ads may be inadequate.

The lack of ability to differentiate ads from organic content leads to the problem that users select ads under the false assumption that they are organic results (Lewandowski et al., 2018, p. 24). It follows that the trust users have in search engines' results extends to ads when users are unable to distinguish these two types of results. Hence, since ad content is produced by the advertisers and ads are ranked based on the payment, not on relevance, users' trust is to be questioned (Lewandowski, 2017, p. 22).

1.5 | Search engine advertising and its influence on users

Eye-tracking studies have shown that the visual attention on ads on desktop computers is higher when they are at the top of the SERP (Buscher, Dumais, & Cutrell, 2010;

González-Caro & Marcos, 2011). If the top-listed ads are relevant, the organic results directly below receive significantly less attention (Buscher et al., 2010). The top placed ads also attract the users' visual attention on mobile devices, where viewing behavior is independent of the presence of ads. In both cases (presence or absence of ads), the results are reviewed from top to bottom (Djamasbi, Hall-Phillips, & Yang, 2013). Even if they are of low quality, ads receive high visual attention on mobile devices (Alanazi, Sanderson, Bao, & Kim, 2020). Viewing behavior also depends on the task type. Ads get higher visual attention if the task is transactional, as the study of González-Caro and Marcos (2011) shows. The design of the ads also affects the viewing behavior when searching on mobile devices, as Lagun, McMahon, and Navalpakkam (2016) show. In their study, shopping ads received more attention than text ads in the same position. Using a sample of 20,297 client accounts advertising on Google in 2017, an industry study by the company WordStream showed that ads are clicked more often on mobile devices than on desktop computers (Donnelly, 2019). Loading time of the SERP also plays a role in the users' interaction with ads, as Bai and Cambazoglu (2019) determined in their large-scale transaction log analysis. If the response latency of the SERP increased, the willingness of the users to click on ads decreased, as did the search engines' revenues. Through surveys, further studies showed that, in addition to the ads' position, the presence of ad avoidance (ignoring paid results on a SERP) plays a crucial role in a user's attitude toward advertising and thus his or her selection behavior (Li, 2019; Yu & Marakas, 2019).

In summary, ads attract much visual attention on desktop computers and mobile devices, as well. In particular, this holds true for the top ranked search results. Furthermore, ads are clicked more frequently on mobile devices.

2 | RESEARCH QUESTIONS AND HYPOTHESES

Since we wanted to investigate how users' selection and viewing behavior regarding ads correlates with their understanding of ads on devices with large vs. small screens, our research questions are as follows:

RQ1: Is there a correlation between search engine users' knowledge of ads and their viewing and clicking behavior on ads?

RQ2: Does the viewing and clicking behavior on ads in desktop search differ from that in mobile search?

As we measure user understanding of advertisements on the SERPs using a scale and measure viewing and clicking behavior in a lab study, the hypotheses based on RQ1 all relate to the correlation between results from the questionnaire measuring user knowledge on advertisements and user viewing and clicking behavior in the lab study.

The hypotheses relating to RQ1 are based on the results of Lewandowski et al. (2017, 2018) and are as follows:

H1a: The score from the survey on understanding the ads correlates negatively with the viewing frequency of the ads (the better the understanding of the ads, the less often ads are viewed).

H1b: The score from the survey on understanding the ads correlates negatively with the click frequency of the ads (the better the understanding of the ads, the less often ads are clicked).

H2a: Users who have a poor understanding of ads pay the first listed ad the highest visual attention and the following results decreasing attention. The first organic search result therefore receives less visual attention than the ads.

H2b: Users with a high understanding of ads pay little visual attention to the ads. The first organic search result, on the other hand, receives the highest visual attention and the following results decreasing attention.

The second research question relates to differences between desktop and mobile user behavior. We expect a more significant effect of ads on the smaller mobile screen, as only ads are visible “above the fold”, that is, without scrolling down (e.g., Kim, Thomas, Sankaranarayana, Gedeon, & Yoon, 2016; Sachse, 2019). We test the following hypotheses:

H3a: Users pay more visual attention to an ad when it is displayed on a mobile device than when it is displayed on desktop computers.

H3b: Users click an ad more often when it is displayed on a mobile device than when it is displayed on desktop computers.

3 | METHODS

Since it was necessary to investigate both selection and gaze behavior in order to answer our research questions, we conducted an eye-tracking experiment with

100 participants. Additional data were collected through an interview at the beginning of the study and a questionnaire asking users about their knowledge of advertising on search engine result pages at the end of the study. From this questionnaire, we derived a scale measuring user knowledge of search-based advertisements.

We aimed for a diverse sample of participants, as it is well known that only using students as test participants, as is common practice, may lead to biases in the data (e.g., Basil, 1996; Bello, Leung, Radebaugh, Tung, & Van Witteloostuijn, 2009; Falk, Meier, & Zehnder, 2013). While we are aware that even a diverse sample in the size range feasible for a lab study is by no means representative, we are confident that it provides a much better data set than using a rather homogeneous sample.

As we wanted to generate a diverse sample of participants, we recruited individuals from two different groups. On the one hand, we invited German-speaking students from the Hamburg University of Applied Sciences (HAW), Germany. The students were enrolled in different academic disciplines (e.g., library and information science, communication design, and aircraft engineering). Students were contacted through an internal mailing list targeting all the university's students. The only inclusion criterion here was participants' enrolment at the HAW at the time the study was conducted.

On the other hand, we invited German-speaking non-students. There were no further inclusion criteria, apart from the exclusion of students enrolled during the conduct of the study. Non-student participants were reached via Ebay classified ads ($n = 30$) as well as flyers in the neighborhood of the university ($n = 20$).

The study was conducted in two parts: The student group took part in September 2018, and the non-students took part from February to April 2019. The study took place at the usability lab at the Hamburg University of Applied Sciences.

Before taking part in the study, each participant had to sign a declaration of consent and a privacy agreement (see Supplemental Material S1 and S2). For their participation, the test subjects received a compensation of 10 Euros each.

We intended and achieved a sample size of 100 participants, 50 of whom were students and 50 were non-students. From the interviews and questionnaires, the data of all 100 subjects could be used. This does not apply to the eye-tracking data. Due to technical problems, such as incomplete data sets due to contact losses of the eye-tracker, some data could not be used. About 960 of the 1,000 search tasks carried out on the desktop computer (96%), and 850 of the 1,000 tasks carried out on the smartphone (85%) were suitable for analysis.

3.1 | Data collection

Data were collected through an interview, through an eye-tracking experiment, and through a questionnaire after the eye-tracking experiment was completed.

In the interview, we asked participants for demographic data and about their use of search engines. The test supervisor read the questions to the participants and took down their answers.

The eye-tracking experiment constitutes the core component of the present study. The design is a one-factor within-subjects design with device as the independent variable. The conditions of the independent variable are desktop computer and smartphone. All participants used both devices (in random order).

Each participant had to complete a total of 20 tasks on desktop computer and smartphone (10 each). We created two task blocks with 10 search tasks each. This was necessary because when we switched from one condition to the other (i.e., desktop computer to smartphone, and vice versa) we had to change the hardware and software. Thus, on a technical level, the eye-tracking-experiment consisted of two stand-alone eye-tracking-tests.

The task types of the two blocks are distributed according to Broder (2002). In a log analysis with 400 queries, Broder (2002, p. 8) found a distribution of 48% informational, 30% transactional, and 20% navigational queries. Each task block of the present study thus contains five informational, three transactional, and two navigational tasks. In the following, we list exemplary tasks of each task type:

“Imagine you want to build a desktop computer yourself. A Google search returned the following results. Please click on a result.” (informational)

“Suppose you want to buy a refrigerator. A Google search returned the following results. Please click on a result.” (transactional)

“Suppose you want to visit the Apple website. A Google search returned the following results. Please click on a result.” (navigational)

A complete list of tasks and queries (in German and English), as well as their SERP elements, can be found in Supplemental Material S3. A time limit of 1 min was set for each search task to ensure the comparability of the performed tasks.

The SERPs of the tasks were presented as clickable screenshots. We used screenshots from Google, which we modified as follows: All SERPs were limited to the result

types *organic result*, *text ad*, and *shopping ad*, which can be seen in Figure 1.

Other vertical search results, such as maps, were removed to allow results to be evaluated without interfering elements. SERPs with vertical search results sometimes do not contain the customary 10 organic results, as shown on regular SERPs. For these SERPs, vertical results were first removed, and the list of organic results was stocked up to 10 in order to obtain a realistic SERP. These added results were taken from the second SERP, which was not displayed to the subjects in the study.

SERP screenshots in desktop computer and smartphone layouts were created for each search query. The latter were adapted to the desktop computer SERPs of the same search queries using an image processing program.

Two eye trackers were used for the experiment: Tobii T60,² which has a 17-in. screen, was used as the stationary eye tracker (desktop condition). Tobii X2-60³ and the Tobii Mobile Device Stand⁴ were used to test the SERPs on the smartphone. We used the smartphone P8 lite⁵ from Huawei.

The eye-tracking software with which the study was conducted is iMotions.⁶ In iMotions, we defined 100 ms as the minimum fixation duration. Fixations are moments when the eyes are relatively stationary to receive or encode information (Poole & Ball, 2006). We chose 100 ms, as this is a common value for eye-tracking studies on SERPs (see Buscher et al., 2010; Cutrell & Guan, 2007) and also the default setting in the iMotions software. For evaluation purposes, all organic and paid results of all SERPs were defined as Areas of Interest (AOIs) in iMotions. We analyzed the number of fixations on each AOI as an indicator of its importance for the subject (Poole & Ball, 2006).

Also, various browser extensions for window resizing, user-agent switching, and screen capturing⁷ were used to create desktop computer- and smartphone-like screenshots. Adobe Photoshop Elements⁸ version 8 was used for image editing. To make the SERP screenshots clickable for the test subjects, image maps were created with the Online Image Map Editor⁹ for all SERP screenshots. With a tool specifically developed for this study, the experimental conditions were made available via URL.¹⁰ When one of the four URLs was accessed, the 10 search tasks of the appropriate condition appeared in randomized order, with the task text and the associated clickable SERP following one another.

After the eye-tracking experiment, each test subject completed a questionnaire, which served to determine the participants' level of knowledge about search engine advertisements. It consisted of two sections.

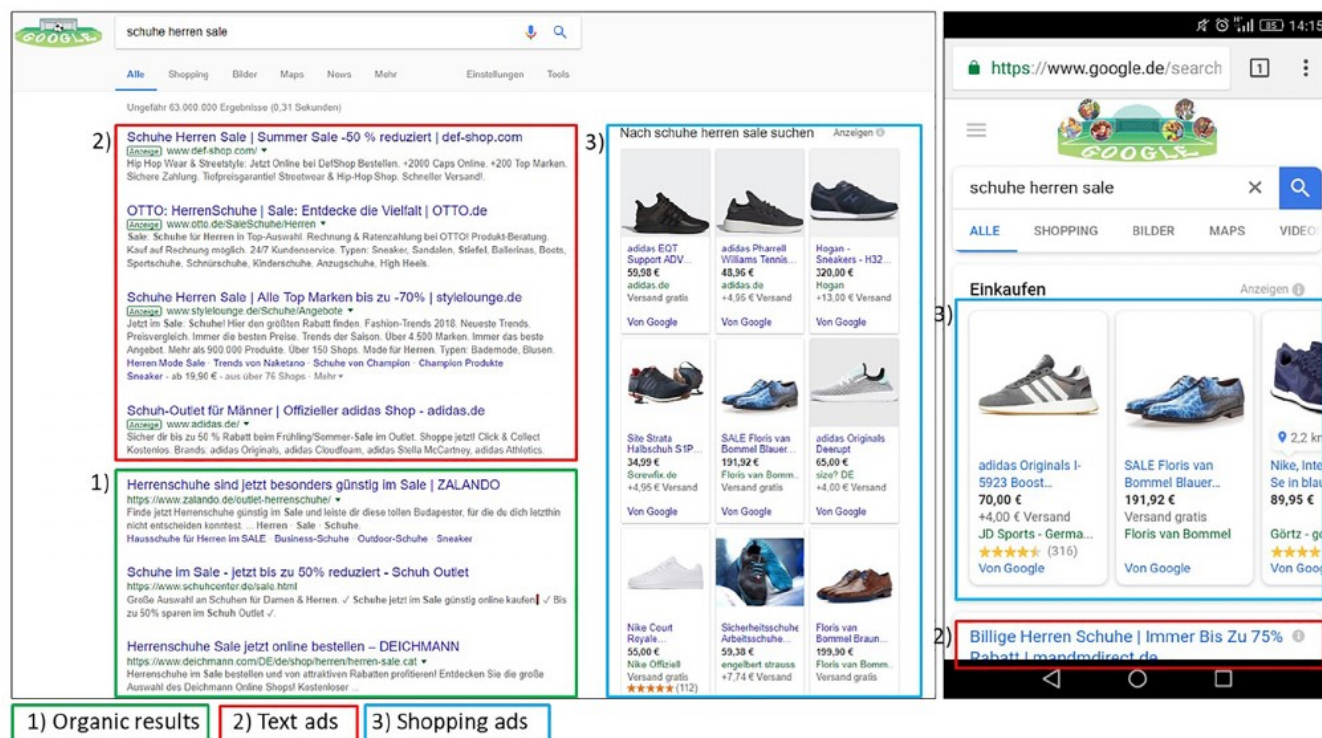


FIGURE 1 Results types of the experiment [Color figure can be viewed at wileyonlinelibrary.com]

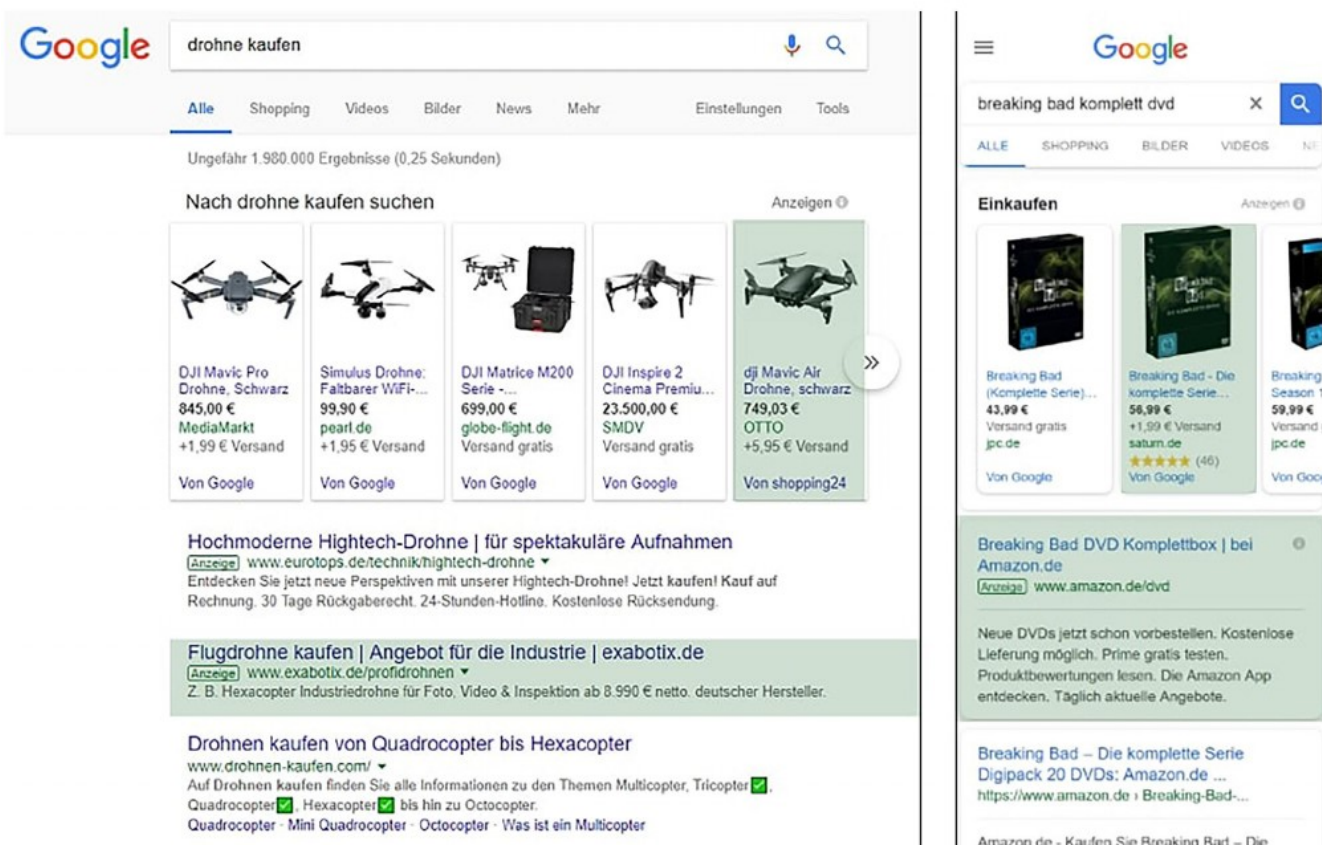


FIGURE 2 Exemplary markings of results for questionnaire tasks 4 (desktop computer) and 6 (smartphone) [Color figure can be viewed at wileyonlinelibrary.com]

In the first section, participants were asked questions about their understanding of Google's business model. The second section (questions 3–6) examined the extent to which the respondent was able to differentiate between organic results and ads. On four SERPs, the participants had to mark either ads or organic results, as shown in Figure 2. When clicking on a search result, it was highlighted in green and thus marked as ad or organic result according to the particular question. By clicking on the same result again, the marking was removed.

Questions used, their justifications and the scores assigned to each question can be found in Table 1. The weighting of the questions is based on the error rates of the subjects from the study by Lewandowski et al. (2018), whose questions served as the basis for this questionnaire. As this previous study used a sample representative of the German online user population, we can use these error rates to derive the average difficulty of the different questions/tasks. We normalized the scores to get a scale ranging from 0 to 100. For this purpose, we summed up all error rates by the subjects in (Lewandowski et al., 2018) and calculated what share the error rate of a

certain question (e.g., question 1:19%) has in the sum of all error rates (456%). In the example of question 1, this share is 4.2%, which consequently is the score of the question.

We formed groups of subjects with a low and high understanding of ads using the questionnaire scores. These scores were divided into four blocks, each containing about 25% of the data, using quartiles. The first block thus forms the group with a low understanding (10.7 [lowest score of all subjects] – 47.1 points), the fourth block the group with a high understanding of ads (92.4–100 points). Both groups consist of 25 subjects each and will be used later (Table 2).

3.2 | Procedure

Each lab session was scheduled for 1 hr. At the beginning, the test procedure was briefly outlined. We did not reveal the actual aim of the study in order to reduce demand characteristics, that is, the adaptation of the subjects' behavior to the perceived requirements of the

TABLE 1 Questionnaire

No.	Question	Justification	Score weightings related to error rates by subjects in (Lewandowski et al., 2018)
1	How does Google generate its revenues?	Self-assessment of users' knowledge about Google's revenue model	4.2
2	Is it possible to pay Google for preferentially listing one's company on the search results pages, as an answer to a search query?		5.8
2.1	[if "yes" on the previous question]:Is it possible to distinguish between paid advertisements and unpaid results on Google's search engine results pages?		9.2
2.1.1	[if "yes" on the previous question]:How do paid advertisements differ from unpaid results?		2.4
3	Task to select organic results on desktop computer SERP	Click-based test to identify problems in distinguishing between organic results and ads;tasks on desktop computer and smartphone with SERPs containing text and shopping ads, similar to eye-tracking tasks	19.6
4	Task to select ads on desktop computer SERP		19.6
5	Task to select organic results on smartphone SERP		19.6
6	Task to select ads on smartphone SERP		19.6
Sum			100

experiment (Orne, 1962). Each participant was instructed to complete the search tasks in the same way as he or she would in a private search situation.

The introduction was followed by handing over the payment. Then, signatures on the receipt, the privacy agreement, and the declaration of consent were collected. Drinks and snacks were offered to create a comfortable atmosphere for the subjects.

The study itself began with a brief interview of the participant's demographic data and his or her search engine usage. This was followed by the eye-tracking experiment consisting of two stages. Each subject completed one task block on the desktop computer and a second task block on the smartphone with random sequences of devices, task blocks, and individual tasks within the blocks. At the end of the study participants filled in the questionnaire. In the questionnaire, participants were asked about Google's business model and tested regarding their ability to differentiate between organic results and ads. Figure 3 shows the flow chart of the study.

4 | RESULTS

4.1 | Data analysis

Regarding research question RQ1 and hypotheses H1a and H1b, we performed Spearman's rho analyses, as this

is an appropriate method for variables that are not normally distributed, which is the case with our data. For testing hypotheses H2a and H2b, we conducted Mann–Whitney *U* tests as the data did not meet the requirements of t-tests in terms of normal distribution. For answering RQ2, chi-square tests of independence were performed to examine the relation between fixation rates on ads and device for each ad (H3a, H3b).

4.2 | Characteristics of the participants

Of the 100 participants, 64 were female and 36 were male. The mean age was 34.1 years ($SD = 14.2$; range between 18 and 75). Regarding educational level, 17 participants had a university degree, 59 a higher education entrance qualification, 17 a secondary education (high school diploma), and seven a lower secondary education level.

Based on our recruitment strategy, half of the participants (50) were students, 26 of which were at the department of information and 24 at other departments. The rest of the sample consisted of people with various occupational backgrounds.

The information provided by the test subjects about their search engine use gives a uniform picture—96 persons named Google as their most frequently used search engine. Three participants named Ecosia and one DuckDuckGo. The majority did not use any other search engine than Google ($n = 55$). If they did, it was Bing ($n = 17$), Ecosia ($n = 9$), Yahoo ($n = 7$), DuckDuckGo ($n = 6$), Metager ($n = 2$) or StartPage ($n = 1$). Three participants named the Safari or Tor browser as a search engine they used in addition to Google.

Almost all test persons used search engines via desktop computer or laptop ($n = 94$) as well as via smartphone ($n = 95$), while some used search engines on one of these device types only. Twenty-eight respondents stated that they also used tablets for Web search.

Concerning the question of how the respondents assess their own competency to search with Web search

TABLE 2 Quartiles of questionnaire scores

<i>N</i>	
Valid	100
Missing	0
Minimum	10.7
Maximum	100.0
Percentiles	
25	47.1
50	81.2
75	92.4

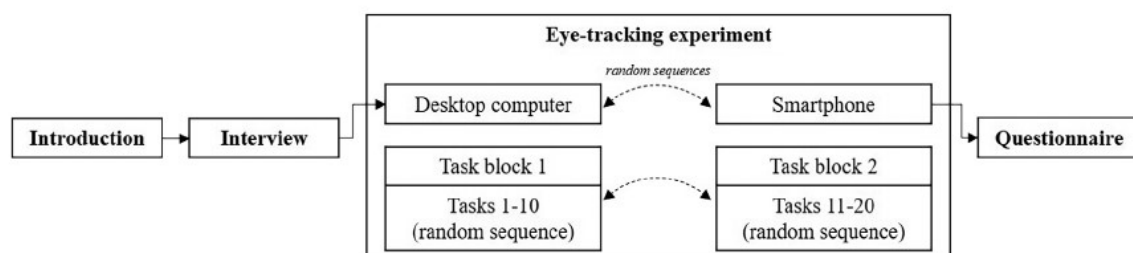


FIGURE 3 Flow chart of the study

engines such as Google, the majority ($n = 65$) rated themselves as “good” (grade 2). In the German grading system that we used, 1 is the best and 6 the worst grade. The other participants rated their competency as “satisfactory” (grade 3, $n = 21$), “very good” (grade 1, $n = 11$), and “sufficient” (grade 4, $n = 3$). No respondent rated his search skills as grade 5 or 6, resulting in an average grade of 2.16 ($SD = 0.64$).

4.3 | Clicks and fixations on results types in the desktop and the mobile condition

In the following analyses, we distinguish between four types of results:

1. Text ads top: Text ads shown above the organic results at the beginning of the SERPs (on all SERPs tested)

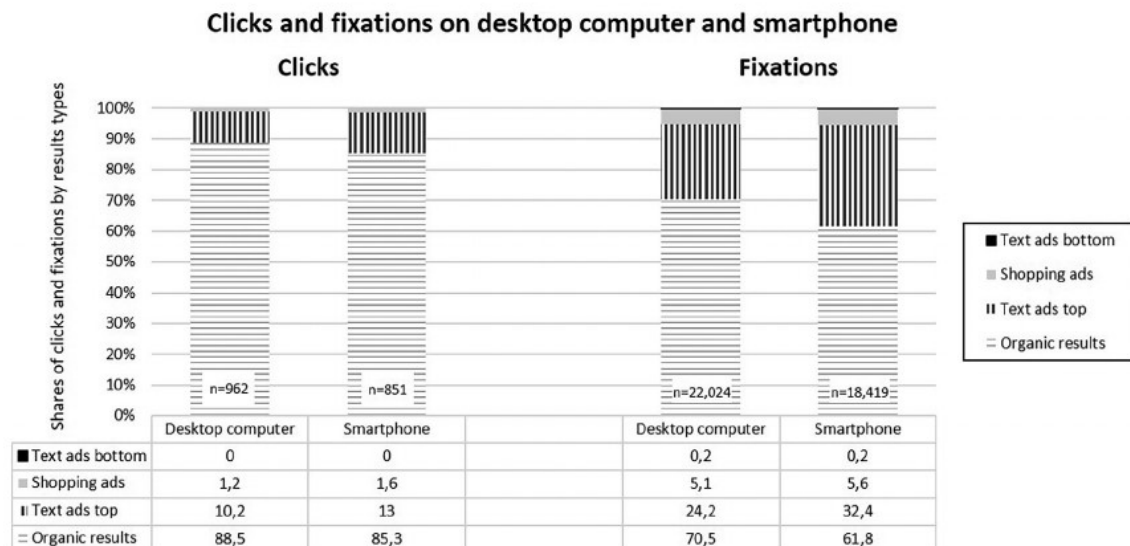


FIGURE 4 Clicks and fixations on results types

TABLE 3 Clicks on ads by task type

Task type	Device	Tasks	Clicks on ads	Click rate on ads (%)
Navigational	Desktop computer	191	48	25
	Smartphone	161	47	29
Transactional	Desktop computer	289	22	8
	Smartphone	259	22	8
Informational	Desktop computer	480	15	3
	Smartphone	430	17	4

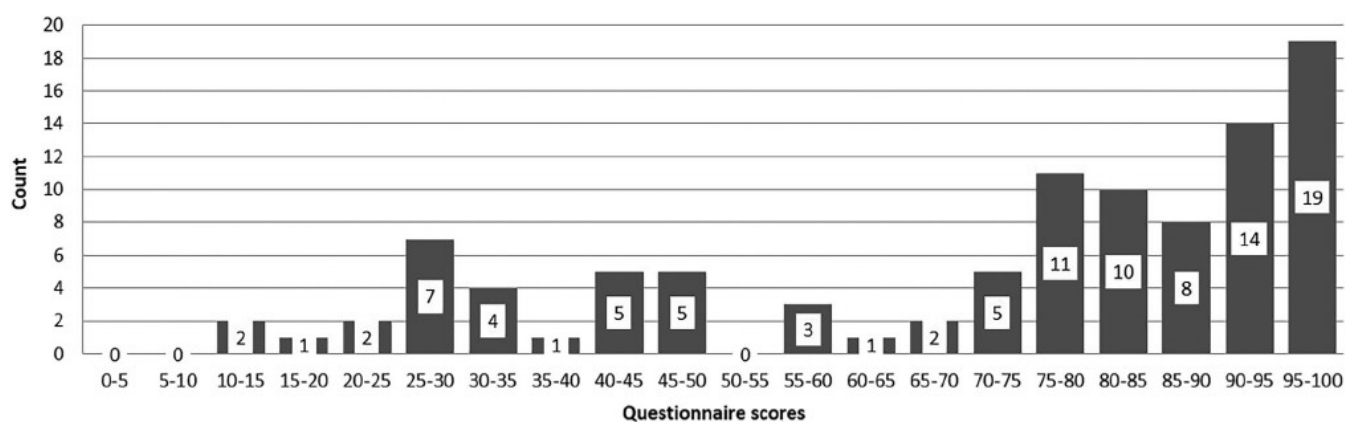


FIGURE 5 Questionnaire scores

2. Text ads bottom: Text ads shown below the organic results at the end of the SERP (on 3 of the 10 SERPs tested)
3. Shopping ads: Ads with product image (etc.), shown for transactional queries (on 6 of the 10 SERPs tested)
4. Organic results: List of 10 organic results (on all SERPs tested)

Figure 4 indicates that the text ads at the top of the SERP (“text ads top”) were selected and fixated on more frequently when searching with the smartphone than the same ads displayed on the desktop computer. About 13% of the clicks on the smartphone were made on text ads top, whereas the rate on desktop computer is 10.2%. The differences are even more apparent when looking at the fixations. 32.4% of all fixations made on the smartphone were done on the text ads top, on desktop computer the rate is 24.2%. No clicks were made on text ads at the end of a SERP (“text ads bottom”).

Participants clicked on ads mainly in navigational tasks, as shown in Table 3.

4.4 | Correlation between users' understanding of ads and their viewing and clicking behavior

As Figure 5 shows, the subjects mostly achieved high values in the questionnaire. On average, they achieved 71.8 (SD = 25.9) points.

The results of Spearman's rho analysis showed no significant correlation between the variables "questionnaire score" and "fixations on ads" ($r = .093, p = .179$). Therefore, hypothesis H1a cannot be confirmed.

The results of Spearman's rho analysis showed a significant negative correlation between the variables “questionnaire score” and “number of clicks on ads” ($r = -.196, p = .025$). Therefore, low questionnaire values are accompanied by high click counts on ads. Hypothesis H1b can be confirmed.

4.5 | Differences in viewing behavior between subjects with low and high understanding of ads

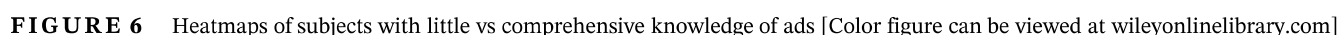
In the following, we will examine Hypotheses H2a and H2b, first considering the gaze data from the desktop computer tasks. This was realized in three steps: First, two groups of participants were formed with low and high questionnaire scores, as described above. For these groups, the second step was to check whether the differences mentioned in the hypotheses were identifiable

TABLE 4 Fixations of participants with low and high questionnaire scores for tasks on desktop computer

Average number of fixations per task																								
	Text ads, top of SERP								Organic results								Text ads, bottom of SERP							
Pos.	01	02	03	04	01	02	03	04	05	06	07	08	09	10	01	02	03							
Low scores (mean)	4.0	4.1	3.5	2.6	4.0	3.2	1.9	0.9	0.7	0.5	0.4	0.3	0.2	0.1	0.0	0.0	0.0							
Low scores (SD)	5.7	5.4	4.8	3.3	4.8	5.3	4.0	2.5	2.4	1.6	2.0	1.1	0.9	0.5	0.0	0.0	0.0							
High scores (mean)	3.6	2.9	5.3	2.7	5.6	4.4	2.8	2.1	1.4	0.9	0.9	0.4	0.6	0.2	0.1	0.3	0.1							
High scores (SD)	4.4	5.3	6.2	2.8	5.6	5.5	4.2	3.9	3.7	2.6	3.2	1.6	2.8	0.8	0.6	1.2	0.3							
p Values (Mann–Whitney U tests)	.844	.255	.133	.702	.000*	.002*	.002*	.000*	.002*	.020*	.020*	.165	.046*	.115	.327	.327	.407							

Mann-Whitney U tests were carried out to check whether the fixation counts of the two groups differed. Table 4 shows the average fixation counts of both groups on the different result ranks. The tasks Q6, Q8, Q17, and Q18 were not considered. These SERPs contained shopping ads, whereas the other SERPs did not. Since it can

For text ads on desktop computer, none of the Mann–Whitney U tests produced significant results. Thus, the visual attention on ads was similar. In contrast,



eight of the 10 organic results were fixated significantly more often by the subjects with a high level of ads knowledge. Accordingly, this group considered the complete SERP more intensively than the group with low knowledge of ads, which tended to linger on the ads in their viewing behavior.

To visualize the results shown before, heatmaps were created for all desktop computer SERPs.¹¹ The heatmaps show the gaze data of the previously formed groups (low/high scores). Due to the random assignment of task blocks and devices, the subjects saw either the first or the second task block on the desktop computer. Thus, of the 25 subjects in the group with high values, nine participants saw the SERPs of the first and 16 the SERPs of the second task block. In the case of the subjects with low scores, the distribution is 14 (block 1) to 11 (block 2).

However, in order to make the heatmaps comparable, we had to adjust the size of the groups. For block 1, this means that we compare the nine subjects with the highest scores with the worst performing nine subjects. For block 2, we used the same procedure and compared 11 with 11 subjects.

In the following, we will look at the heatmaps for task Q14 in more detail (Figure 6), which we use as an illustrative example. The heatmaps visualize the previously shown results from the Mann–Whitney *U* tests for the desktop computer tasks. It should be noted that the subjects with little knowledge of ads scrolled into the area “below the fold” less often.

Next, we will examine Hypotheses H2a and H2b using the gaze data from the smartphone tasks, using the same procedure as for the previously described desktop computer tasks. Since only the first of the bottom text ads received fixations from both groups, we did not compare the other text ads at the end of the SERP. The results of the Mann–Whitney *U* tests are shown in Table 5.

The results are very similar to those of the desktop computer tasks. The Mann–Whitney *U* tests show that four organic results were more frequently fixated by the subjects with high scores. In addition, as a rather surprising result, we found a difference in the third text ad at the top of the SERP, which was fixated more often by the subjects with high scores.

We can confirm hypothesis H2a since users who have a poor understanding of ads paid less attention to the organic results than users with high understanding of ads.

Hypothesis H2b cannot be confirmed. Although the subjects with comprehensive knowledge of ads considered the complete SERP, including the organic results, their visual attention on ads was not considerably different from the subjects with little knowledge; one of the text ads on the smartphone was fixated even more frequently by the subjects with high scores.

4.6 | Differences between the desktop computer and smartphone condition

The following analyses serve to answer RQ2. Figure 7 shows the fixation rates on ads for both devices. For “Text ad top 1”, $n = 960$ is the sum of all text ads in the first position that the subjects saw on the desktop computer. Each of the 20 desktop computer SERPs of the experiment contained a “Text ad top 1”, whereby each SERP was viewed by 50 subjects ($20 \times 50 = 1,000$). The difference between 1,000 and 960 results from missing gaze data for 40 tasks.

The first-placed text ads at the top and the bottom, as well as the first shopping ad were fixated more frequently when searching on the smartphone. The fixation rates of the other ads were higher in the desktop computer

TABLE 5 Fixations of participants with low and high questionnaire scores for tasks on smartphone

	Average number of fixations per task														
											Text ad, bottom of SERP				
	Text ads, top of SERP				Organic results										
Pos.	01	02	03	04	01	02	03	04	05	06	07	08	09	10	01
Low scores (mean)	7.5	2.2	2.1	2.7	4.1	2.3	1.6	1.0	0.5	0.2	0.4	0.3	0.3	0.1	0.1
Low scores (SD)	9.3	2.8	3.3	3.3	4.9	4.0	4.0	3.2	1.7	1.2	2.0	1.1	1.9	0.3	0.3
High scores (mean)	6.9	2.6	3.3	3.1	4.5	3.4	2.1	1.3	0.9	0.6	0.5	0.6	0.4	0.4	0.5
High scores (SD)	5.4	2.9	2.4	2.6	5.2	4.6	3.7	2.9	2.4	1.6	1.9	2.4	1.5	1.7	1.4
<i>p</i> Values (Mann–Whitney <i>U</i> tests)	.163	.357	.031*	.358	.431	.001*	.053	.029*	.020*	.005*	.179	.330	.945	.449	.501

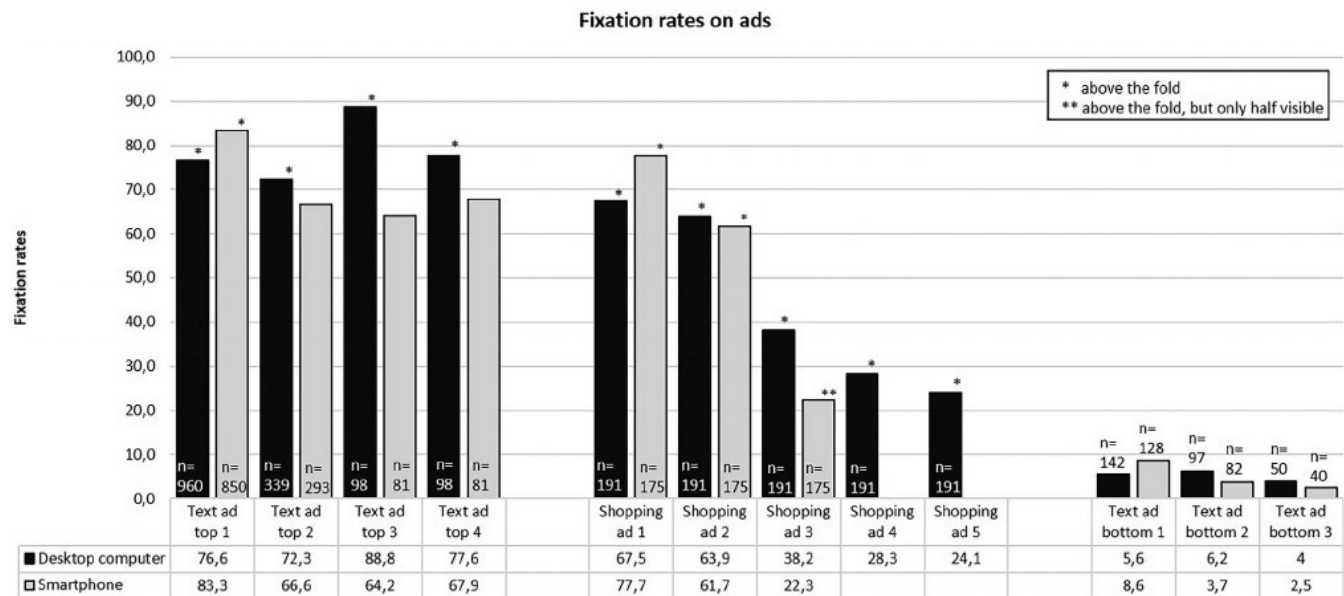


FIGURE 7 Fixation rates on ads

condition. It also turns out that the text ads at the bottom of the SERP were viewed in only 2.5–8.6% of all cases.

The relation between these variables was significant for Text ad top 1, $\chi^2(1) = 12.638$, $p = .000$ and for Shopping ad 1, $\chi^2(1) = 4.732$, $p = .030$. In both cases, the fixation rates on ads were significantly higher on the smartphone than on desktop computer. The relation between the variables was also significant for Text ad top 3, $\chi^2(1) = 15.438$, $p = .000$ and for Shopping ad 3, $\chi^2(1) = 10.918$, $p = .001$. In these cases, the fixation rates on ads were significantly higher on desktop computer than on the smartphone.

The results show that the hypothesis H3a can be confirmed when considering the results placed above the fold on both devices. Thus, the first text ad on the top of the SERP and the first shopping ad received more visual attention on the smartphone. On the other hand, ads that require scrolling to be visible on both devices do not differ in fixation rates.

Chi-square tests of independence were performed to examine the relation between click rates on ads and device for each ad. None of the tests was significant. Therefore, H3b cannot be confirmed.

5 | DISCUSSION

Our study shows that users with a low level of knowledge on search advertising are more likely to click on ads than subjects with a high level of knowledge. Additionally, subjects with little knowledge show less willingness to scroll down to the organic results. This might be caused

by the fact that the distinction of both lists (ads, organic) is not understood. The subjects with little knowledge consider all results to be the same type and therefore have no reason to scroll to further results. Regarding the device, there are large differences in viewing behavior. These can be attributed to the influence of the direct visibility of a search result on both devices tested.

Our study confirms findings by Lewandowski (2017) who found that knowledge about ads influences selection behavior on search engine result pages. He found that users unable to distinguish between ads and organic results selected the first ad about twice as often as users who were able to make that distinction. Our study confirms this, as users with less knowledge of ads clicked on ads more often. This does not seem to be, however, a result of these users fixating on ads more often.

Users do not seem to avoid looking at advertisements on the search engine results pages. Users more knowledgeable of ads do not fixate on ads less often. While, therefore, H1a had to be rejected, we found a significant negative correlation between users' understanding of ads and their clicking behavior (H1b), meaning that users less knowledgeable of ads click on them more often. These users also pay less visual attention to organic results (H2a), as more of their fixations are distributed to the ads. The inverse effect, however, cannot be confirmed for knowledgeable users (H2b rejected). We can therefore assume that users who have little understanding of ads consider paid and organic results to be one coherent list. Consequently, these users do not feel the need to consider the organic list separately, resulting in a high concentration of visual attention on the ads. We have

obtained a surprising result, which is that the third text ad was fixated more often by the subjects with a high level of knowledge on ads. This result may have been caused by the rather small size of the analyzed AOIs. Thus, it might have been more appropriate to analyze the results as blocks (e.g., text ads top, organic results).

Regarding the device the ads are displayed on, we found that users pay more visual attention to ads on the smartphone than on the desktop screen (H3a). This result is not surprising, since on the smartphone only ads were visible without scrolling down. The higher visual attention on top results of small screens has already been shown by Kim et al. (2016). This shows how important it is for results to be placed in positions “above the fold”, that is, results that can be seen without scrolling down. This does not, however, lead to more clicks on the ads (H3b rejected).

Our results lead to the question of why users more knowledgeable of ads still fixate ads. Firstly, ads may be relevant to the users' queries. Especially in the context of e-commerce related queries, it has been found that ads may be of similar relevance as organic results (Jansen & Resnick, 2006). Secondly, the design of search engine result pages may have a strong influence on users' viewing behavior toward ads. As ads are not shown only for a part of all queries, the position where the list of organic results begins changes per query. Thus, up to four text ads can be shown before the organic results and, in the case of product searches, additional shopping ads. This may lead to users fixating the ads. When ads are shown on SERPs on the smartphone, organic results can only be viewed when a user scrolls down. This may explain why users knowledgeable of ads still fixate them (even when the ads are not relevant). Further to viewing behavior, knowledgeable users did not fixate the first organic result most often. This can be explained by the factors regarding relevance and SERP design, as well.

Results regarding clicks differ from those for viewing behavior. Ad clicks in the smartphone condition were not higher than in the desktop condition. Contrary to our assumption formulated in H3b, most users selected an organic result regardless of the device used. One should note, however, that we cannot rule out that the laboratory situation influenced user behavior here. In consideration of the demand characteristics (Orne, 1962), it can be assumed that the subjects have dealt with the tasks more intensively due to the laboratory situation than they would have done in a real situation.

A striking result is that ad clicks differ considerably between different query intents (informational, navigational, and transactional). We found that users most often click on ads in navigational tasks. In these cases, ad clicks may be well-informed and rational decisions, as it does

not make a difference to the user whether he selects the same URL as an organic result or as an ad. It makes a difference to the advertiser, however, given that every ad click has to be paid for (Jansen, 2011).

Our study has some limitations. Firstly, as with most lab-based studies, the sample size is rather small. However, we aimed for a larger than usual sample size and a more diverse sample, as well. We are therefore confident that our findings are generalizable at least to a certain degree. A further limitation is the unnatural search situation due to the laboratory setting. The experimenter was present while users were working on the tasks and the queries, as well as the SERPs, were pre-defined. Furthermore, when searching on the smartphone, it was not possible for the subjects to hold the smartphone in their hands because of the laboratory devices (“mobile device stand” for the eye-tracking experiment on the smartphone).

A limitation regarding the questionnaire instrument is that while this specific questionnaire has been used in the past already, its reliability is unknown. In future research, it may be worthwhile to systematically test the questionnaire and develop it further into a standard tool of measurement.

Some suggestions for further research are to look more closely on users' motivations when selecting ads. As our study showed, this motivation seems to depend on the query intent. Future studies should distinguish between different motivations for selecting ads, for example, cases where clicks on ads are rational decisions (as in the case of navigational queries where the same URL is shown as an ad and as an organic result), versus cases where users click on ads seemingly unaware that the selected result is an ad. In this context, the influence of the quality and relevance of an ad on the selection behavior of the users should also be investigated.

As search result presentation has changed over time (and is continuing to do so), future research should also focus on more result types that are shown in the SERPs. Most obviously, vertical results (e.g., video, shopping) should be considered here. Some small-scale studies (Liu et al., 2015) and results from more extensive industry studies (Lewandowski & Sünkler, 2013) indicate a huge influence of vertical results on users' selection behavior. However, vertical results are sometimes organic (as in the case of news) and sometimes paid advertisements (as in the case of Google's shopping results). It would be interesting to find out more about users' understanding, viewing behavior and clicking behavior regarding these results.

Our results have implications for search engine providers and regulation, for advertisers, and information literacy training.

Search engine providers should take measures to label ads in a way that users can easily distinguish them from

organic results, as required by regulation (e.g., Lewandowski et al., 2018; Sullivan, 2013a, 2013b). This may, however, interfere with current business practices aiming to maximize revenue generated through ads. To this end, search engine providers may be tempted to blur the lines between paid advertisements and organic results (Edelman & Gilchrist, 2012; Lewandowski et al., 2018).

Advertisers would also profit from a clearer labeling of ads. As our study showed, users often select ads in response to a navigational query. In these cases, the same result is present in the first position of the list of advertisements, and on the first position of the list of organic results, as well. While from the user's perspective, selecting the ads is a rational decision, for the advertiser, this means paying money for a click on a result that is also present at the top position of the regular results. If ads were labeled more clearly (or even the correct result for the navigational query would be shown above the advertisements), companies would not have to spend money on searches for their own company name or the name of one of their products.

Regarding information literacy training, we deem it imperative to help users understand that search engines do not necessarily act in their best interest, but search engine providers have interests of their own. Therefore, results ranking and presentation may not strictly focus on relevance but other, predominantly business-related criteria, as well. The results of our study are in line with previous research on the low information literacy of search engine users. They also highlight the need for information literacy training, which can guide users to select results under correct (e.g., paid vs. unpaid result) assumptions.

6 | CONCLUSION

In this article, we presented a study investigating the influence of users' understanding of search-based advertising on their search behavior on desktop computer and smartphone. For this purpose, we conducted a mixed-method study consisting of an interview, eye-tracking experiment, and questionnaire with $n = 100$ subjects.

We showed that participants with a low level of knowledge of search-based advertisements were more likely to click on ads than subjects with a high level of knowledge. Also, subjects with little knowledge showed less willingness to scroll down to the organic results, which is especially important when searching on mobile devices where organic results can only be seen after scrolling down. Large differences in visual behavior were found concerning the device. These differences can be attributed to the influence of the direct visibility of a search result on both devices tested.

RESEARCH DATA

See (Schultheiß & Lewandowski, 2019) for the eye-tracking and click data of the experiment.

ACKNOWLEDGMENTS

The authors would like to thank Sebastian Sünkler for his support of the study by programming the tool for the eye-tracking experiment and the questionnaire. Besides, the authors would like to thank Daniela Sygulla for her help in collecting the data. Many thanks also to SUMA-EV for the financial support of a part of the study through a student scholarship. Open access funding enabled and organized by Projekt DEAL.

ORCID

Sebastian Schultheiß  <https://orcid.org/0000-0003-2704-7207>

Dirk Lewandowski  <https://orcid.org/0000-0002-2674-9509>

ENDNOTES

- ¹ An overview of Google's ads labeling over time can be found in Marvin (2020).
- ² <https://www.tobiiipro.com/product-listing/tobii-t60-and-t120/>.
- ³ <https://www.tobiiipro.com/product-listing/tobii-pro-x2-30/>.
- ⁴ <https://www.tobiiipro.com/product-listing/mobile-device-stand/>.
- ⁵ <https://consumer.huawei.com/ch/support/phones/p8-lite/>.
- ⁶ <https://imotions.com>.
- ⁷ <https://chrome.google.com/webstore/detail/window-resizer/kkelicaakdanhinjdeammilcgefongfh> (Window Resizer), <https://chrome.google.com/webstore/detail/user-agent-switcher-for-c/djflhoibgkdhkhcedjklpkjnoahfmg> (User-Agent Switcher), and <https://chrome.google.com/webstore/detail/full-page-screen-capture/fdpohaocaechifimbbbbbknolclacl> (Screen Capture).
- ⁸ https://help.adobe.com/archive/de_DE/photoshopelements/8/photoshopelements_8_help.pdf.
- ⁹ <http://maschek.hu/imagemap/imgmap/>.
- ¹⁰ For the code of the tool see <https://dx.doi.org/10.5281/zenodo.3978382>.
- ¹¹ For all desktop computer, SERP heatmaps see <https://dx.doi.org/10.5281/zenodo.3978382>.

REFERENCES

- Alanazi, A. O., Sanderson, M., Bao, Z., & Kim, J. (2020). The impact of ad quality and position on Mobile SERPs. In: *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, pp. 318–322. <https://doi.org/10.1145/3343413.3377990>.
- Alphabet Inc. (2019). *Alphabet Announces Fourth Quarter and Fiscal Year 2018 Results*. Retrieved from https://abc.xyz/investor/static/pdf/2018Q4_alphabet_earnings_release.pdf.
- Bai, X., & Cambazoglu, B. B. (2019). Impact of response latency on sponsored search. *Information Processing & Management*, 56(1), 110–129. <https://doi.org/10.1016/j.ipm.2018.10.005>

- Basil, M. D. (1996). The use of student samples in communication research. *Journal of Broadcasting and Electronic Media*, 40(3), 431–440. <https://doi.org/10.1080/08838159609364364>
- Bello, D., Leung, K., Radebaugh, L., Tung, R. L., & van Witteloostuijn, A. (2009). From the editors: Student samples in international business research. *Journal of International Business Studies*, 40(3), 361–364. <https://doi.org/10.1057/jibs.2008.101>
- Broder, A. (2002). A taxonomy of web search. *ACM SIGIR Forum*, 36(2), 3–10. <https://doi.org/10.1145/792550.792552>
- Buscher, G., Dumais, S. T., & Cutrell, E. (2010). The good, the bad, and the random. In: *Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '10*, 42. <https://doi.org/10.1145/1835449.1835459>
- Cutrell, E., & Guan, Z. (2007). What are you looking for? In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, pp. 407–416. Retrieved from <https://doi.org/10.1145/1240624.1240690>.
- Djamasbi, S., Hall-Phillips, A., & Yang, R. (2013). An examination of ads and viewing behavior: An eye tracking study on desktop and mobile devices. In: *19th Americas Conference on Information Systems*, AMCIS, 2013, pp. 350–355. <https://digitalcommons.wpi.edu/uxdmrl-pubs/36>.
- Donnelly, G. (2019). Google Ads Mobile Benchmarks for YOUR Industry. WordStream. Retrieved from <https://www.wordstream.com/blog/ws/2018/08/13/google-ads-mobile-benchmarks>.
- Edelman, B. (2014). Google's advertisement labeling in 2014. Ben-Edelman Org. Retrieved from <http://www.benedelman.org/adlabeling/google-colors-oct2014.html>.
- Edelman, B., & Gilchrist, D. S. (2012). Advertising disclosures: Measuring labeling alternatives in internet search engines. *Information Economics and Policy*, 24(1), 75–89. <https://doi.org/10.1016/j.infoecopol.2012.01.003>
- European Commission. (2017). *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service—Factsheet*. Retrieved from http://g8fip1kplyr33r3krz5b97d1.wpengine.netdna-cdn.com/wp-content/uploads/2017/06/Google-MEMO-17-1785_EN.pdf.
- Falk, A., Meier, S., & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association*, 11(4), 839–852. <https://doi.org/10.1111/jeea.12019>
- González-Caro, C., & Marcos, M.-C. (2011). Different users and intents: An eye-tracking analysis of web search. *Web Search and Data Mining*. Retrieved from <http://ciio.unab.edu.co/gti/images/documentos/umwa-8.pdf>.
- Google.com. (2019). *Why you're seeing an ad - Ads Help*. Retrieved October 4, 2019, from <https://support.google.com/ads/answer/1634057#info>.
- Granka, L. A., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. In: *Proceedings of the 27th Annual International Conference on Research and Development in Information Retrieval - SIGIR '04*, pp. 478–479. <https://doi.org/10.1145/1008992.1009079>.
- Jansen, B. J., & Resnick, M. (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), 1949–1961. <https://doi.org/10.1002/asi.20425>
- Jansen, J. (2011). *Understanding sponsored search*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511997686>.
- Kammerer, Y., & Gerjets, P. (2014). The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid Interface. *International Journal of Human-Computer Interaction*, 30(3), 177–191. <https://doi.org/10.1080/10447318.2013.846790>
- Keane, M. T., O'Brien, M., & Smyth, B. (2008). Are people biased in their use of search engines? *Communications of the ACM*, 51(2), 49–52. <https://doi.org/10.1145/1314215.1314224>
- Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., & Yoon, H.-J. (2016). Understanding eye movements on mobile devices for better presentation of search results. *Journal of the Association for Information Science and Technology*, 67(11), 2607–2619. <https://doi.org/10.1002/asi.23628>
- Lagun, D., McMahon, D., & Navalpakkam, V. (2016). Understanding Mobile Searcher Attention with Rich Ad Formats. In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management—CIKM '16*, pp. 599–608. <https://doi.org/10.1145/2983323.2983853>.
- Lewandowski, D. (2017). Users' understanding of search engine advertisements. *Journal of Information Science Theory and Practice*, 5(4), 6–25. <https://doi.org/10.1633/JISTaP.2017.5.4.1>
- Lewandowski, D., & Kammerer, Y. (2020). Factors influencing viewing behavior on search engine results pages: A review of eye-tracking research. *Behavior & Information Technology*, 0(0), 1–31. <https://doi.org/10.1080/0144929X.2020.1761450>
- Lewandowski, D., Kerkmann, F., Rümmele, S., & Sünkler, S. (2018). An empirical investigation on search engine ad disclosure. *Journal of the Association for Information Science and Technology*, 69(3), 420–437. <https://doi.org/10.1002/asi.23963>
- Lewandowski, D., & Sünkler, S. (2013). *Representative online study to evaluate the commitments proposed by Google as part of EU competition investigation AT. 39740-Google: Report for Germany*. Retrieved from http://searchstudies.org/wp-content/uploads/2015/10/Google_Online_Survey_DE.pdf.
- Lewandowski, D., Sünkler, S., & Kerkmann, F. (2017). Are ads on Google search engine results pages labeled clearly enough? The influence of knowledge on search ads on users' selection behavior. In M. Gäde, V. Trkulja, & V. Petras (Eds.), *Proceedings of the 15th International Symposium of Information Science (ISI 2017)* (pp. 62–74). Glückstadt: Verlag Werner Hülsbusch. <https://doi.org/10.18452/1450>
- Li, Y. (2019). User perception affects search engine advertising avoidance: Moderating role of user characteristics. *Social Behavior and Personality: An International Journal*, 47(4), 1–12. <https://doi.org/10.2224/sbp.7855>
- Liu, Z., Liu, Y., Zhou, K., Zhang, M., & Ma, S. (2015). Influence of vertical result in web search examination. In R. A. Baeza-Yates, M. Lalmas, A. Moffat, & B. A. Ribeiro-Neto (Eds.), *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in information retrieval—SIGIR '15* (pp. 193–202). New York, NY, USA: ACM Press. <https://doi.org/10.1145/2766462.2767714>
- Marvin, G. (2020). *A visual history of Google ad labeling in search results*. Search Engine Land. Retrieved from <https://searchengineland.com/search-ad-labeling-history-google-bing-254332>.

- O'Brien, M., & Keane, M. T. (2006). Modeling result-list searching in the world wide web: The role of relevance topologies and trust bias. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pp. 1881–1886.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776–783. <https://doi.org/10.1037/h0043424>
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823. <https://doi.org/10.1111/j.1083-6101.2007.00351.x>
- Petrescu, P. (2014). Google organic click-through rates in 2014. Moz. Retrieved from <https://moz.com/blog/google-organic-click-through-rates-in-2014>.
- Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-59140-562-7.ch034>
- Purcell, K., Brenner, J., & Rainie, L. (2012). Search engine use 2012. Retrieved from https://www.eff.org/files/pew_2012_0.pdf.
- Sachse, J. (2019). The influence of snippet length on user behavior in mobile web search. *Aslib Journal of Information Management*, 71(3), 325–343. <https://doi.org/10.1108/AJIM-07-2018-0182>
- Schultheiß, S., & Lewandowski, D. (2019). *How users' knowledge of advertisements influences their viewing and selection behavior in search engines (dataset)*. Zenodo. <https://doi.org/10.5281/zenodo.3407551>.
- Schultheiß, S., Sünkler, S., & Lewandowski, D. (2018). We still trust in Google, but less than 10 years ago: An eye-tracking study. *Information Research*, 23(3) paper 799. <http://www.informationr.net/ir/23-3/paper799.html>
- Schwartz, B. (2018). Google expands 'more results' button to remove paginated search results. *Search Engine Roundtable*. Retrieved from <https://www.seroundtable.com/google-expands-more-results-button-test-25542.html>.
- Singer, G., Norbistrath, U., & Lewandowski, D. (2012). Ordinary search engine users assessing difficulty, effort, and outcome for simple and complex search tasks. In: *Proceedings of the Fourth Information Interaction in Context Symposium*, pp. 110–119. <https://doi.org/10.1145/2362724.2362746>.
- Stark, B., Magin, M., & Jürgens, P. (2014). Navigieren im Netz: Befunde einer qualitativen und quantitativen Nutzerbefragung. In B. Stark, D. Dörr, & S. Aufenanger (Eds.), *Die Googleisierung der Informationssuche* (pp. 20–74). Berlin, Boston: De Gruyter.
- StatCounter. (2019). *Search Engine Market Share United States of America, StatCounter Global Stats*. Retrieved July 2, 2019, from <http://gs.statcounter.com/search-engine-market-share/all/united-states-of-america>.
- Sterling, G. (2016). Report: Nearly 60 percent of searches now from mobile devices. Search Engine Land. Retrieved from <https://searchengineland.com/report-nearly-60-percent-searches-now-mobile-devices-255025>.
- Strzelecki, A. (2020). Eye-tracking studies of web search engines: A systematic literature review. *Information*, 11, 300. <https://doi.org/10.3390/info11060300>
- Sullivan, D. (2013a). FTC updates search engine ad disclosure guidelines after “decline in compliance”. Search Engine Land. Retrieved from <https://searchengineland.com/ftc-search-engine-disclosure-164722>.
- Sullivan, D. (2013b). Revealed: The 17 other search engines the FTC warned over paid ad disclosures. Search Engine Land. Retrieved from <https://searchengineland.com/revealed-17-other-search-engines-ftc-warned-168603>.
- Yu, T., & Marakas, G. (2019). How people perceive sponsored search ads: Toward a unified research model. *25th Americas Conference on Information Systems, AMCIS 2019*, pp. 1–5. Retrieved from https://aisel.aisnet.org/amcis2019/human_computer_interact/human_computer_interact/2/.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Schultheiß S, Lewandowski D. How users' knowledge of advertisements influences their viewing and selection behavior in search engines. *J Assoc Inf Sci Technol*. 2021;72:285–301. <https://doi.org/10.1002/asi.24410>