MDPI

Eye-Tracking Studies of Web Search Engines: A Systematic Literature Review

Artur Strzelecki[ID]

Department of Informatics, University of Economics in Katowice, 40-287 Katowice, Poland;
artur.strzelecki@ue.katowice.pl

check for updates

Abstract: This paper analyzes peer-reviewed empirical eye-tracking studies of behavior in web search engines. A framework is created to examine the effectiveness of eye-tracking by drawing on the results of, and discussions concerning previous experiments. Based on a review of 56 papers on eye-tracking for search engines from 2004 to 2019, a 12-element matrix for coding procedure is proposed. Content analysis shows that this matrix contains 12 common parts: search engine; apparatus; participants; interface; results; measures; scenario; tasks; language; presentation, research questions; and findings. The literature review covers results, the contexts of web searches, a description of participants in eye-tracking studies, and the types of studies performed on the search engines. The paper examines the state of current research on the topic and points out gaps in the existing literature. The review indicates that behavior on search engines has changed over the years. Search engines' interfaces have been improved by adding many new functions and users have moved from desktop searches to mobile searches. The findings of this review provide avenues for further studies as well as for the design of search engines.

**Keywords:** eye-tracking; search engines; literature review

1. Introduction

Since the 1990s, when the first search engines were created, how results are displayed has changed many times. The Google search engine had a significant impact in terms of how results were displayed. The same applies to the Microsoft search engine, which, in addition to changes in the layout of results, underwent significant changes owing to the fusion of MSN Search, Live, and Bing with Yahoo. Just as search engines have evolved, the use of search results has also changed. The present paper reviews the eye-tracking studies available in the literature, which show how the consumption of search results has changed over the years. The research review allows an analysis of the way search results over the last 16 years have been perceived. During this period (2004–2019), the devices used for searching, screen sizes, how search queries are entered, and many other elements that determine how the search results are perceived have all changed.

Eye-tracking is defined as the recording and study of eye movements when following a moving object, lines of text, or other visual stimuli; it is used as a means of evaluating and improving the visual presentation of information. Eye-tracking research spans the domains of psychology, ergonomics, quality, marketing, and information technology. Eye-tracking research in the area of computer science in the literature usually concerns software [1] or websites [2]. Currently, there are two types of eye-tracking devices: stationary, which looks like a computer monitor; and mobile, which should be put on the head [3]. The latter are used to design outdoor advertising and product packaging [4]. These eye-tracking studies can be carried out with the help of specialist equipment, e.g., Tobii, SMI, or EyeLink [5]. However, some researchers have noted that it is expensive and not everyone has access to it [6]. It is possible to use cheaper webcams [7] or a JavaScript library that uses a built-in camera [8].

Eye-tracking is not a new research method; the first eye-tracking study was carried out in 1879 [9]. However, due to the high costs of equipment used in such research, as well as some difficulties in interpreting the results, it is not widely used [5]. Eye-tracking research conducted in different age groups can give different results [10]. For example, results for older people will differ from those for younger people who have more experience with new technologies; differences will also result from eye defects that are more common in older people [11].

The interest in eye-tracking studies is also reflected in an academic context: the number of papers published on eye-tracking is growing. Figure 1 provides an overview of the increase in research on the topic, revealing that the appearance of the term "eye-tracking" in paper titles has been increasing. This suggests that eye-tracking is becoming a more popular subject for academic inquiry. Figure 1 does not have details of the database size as graphs of this nature depends on many factors; its goal is simply to illustrate the increasing amount of eye-tracking studies.
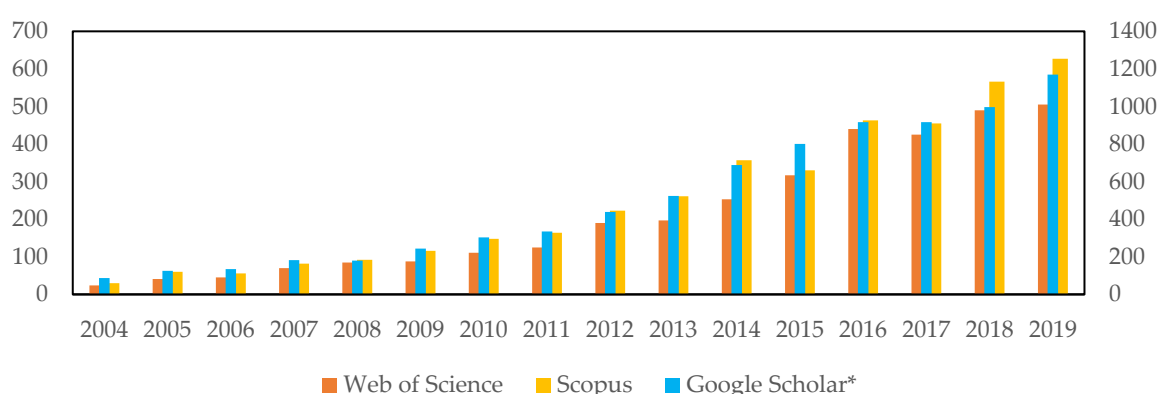
Figure 1. Search results for "*eye-tracking*" in paper titles. Note: * Uses the right axis.

There is, however, no published review of eye-tracking studies for search engines. Published reviews of eye-tracking studies generally tend to focus on medical, aviation, sports, tourism, fashion, product design, psychology, and computer science areas. Eye-tracking reviews of studies that have contributed to framework development belong to the areas of computer science and social science and tend to come from more specific areas like education [12–16], marketing [4], organizational research [17], information science [18], and software [19].

Eye-tracking allows the study of the movements of a participant's eyes during a range of activities. This can reveal aspects of human behavior such as learning patterns and social-interaction methods. It can be used in many different environments and settings and adds value to other biometric data. In the context of web searches, eye-tracking provides unbiased, objective, and quantifiable data on where participants have looked at search results and for how long.

Despite the large amount of hits on the topic, there is little coherent understanding of what kinds of studies have been conducted under the term eye-tracking on search engines, which methods they have used, what kinds of results they have yielded, and under which circumstances. Understanding whether eye-tracking for search engines is effective and provides an understanding of how users are looking at search results is also a pertinent practical issue. Search engines are continuously testing and improving their layout and how they present results to improve users' experience. This paper contributes to literature reviews in the field of eye-tracking for search engines by reviewing the existing body of empirical research on the topic.

This paper presents a literature review of 56 papers relevant to the uses of eye-tracking technology for search engines from 2004 up to 2019. Conducting this literature review is beneficial because it brings together all the studies that have been performed in the past and could help researchers to avoid misusing eye-tracking technology in search-engine research. This review provides an overview of all the different eye trackers, metrics, presentation forms, scenarios, and tasks used in previous eye-tracking

studies. It also discusses the limitations associated with eye-tracking technology. Therefore, it can be a starting point for researchers who are interested in performing eye-tracking studies, helping them to become acquainted with this technology and its limitations, to find related works, and to decide whether or not to use this modern technology [19].

In summary, the contributions of this review are the following:

1.　To provide descriptive statistics and overviews on the uses of eye-tracking technology for search engines;
2.　To examine and analyze the papers and discuss procedural and methodological issues when using eye trackers in search engines;
3.　To propose avenues for further research for researchers conducting and reporting eye-tracking studies.

The annotated bibliography presents information about the selected studies in a structured way. It allows researchers to compare studies with one another concerning the selection of search engines, the selection of participants, and the study scenarios and tasks.

The paper is organized as follows: Section 2 provides the necessary background on eye-tracking technology. Section 3 discusses the process of selecting papers for the literature review, poses research questions, and proposes a coding procedure based on the research questions. Answers to research questions and key findings from each study are presented in Section 4. Section 5 provides the conclusions, discusses the limitations, both of individual studies and the validity of this review as a whole, and details avenues for future research.

2. The Concept of Eye-tracking Studies for Search Engines

Eye-tracking studies of search results from search engines allow a better understanding of how users browse through specific parts of the text and how they select search results. To recognize patterns of user interaction with search results, numerous types of visual behaviors are observed using an eye-tracking camera. Behaviors distinguished by the camera that observes the user are: fixations; saccades; pupil dilation; and scanpaths [20]. Eye fixations are defined as a spatially stable gaze lasting for approximately 200–300 milliseconds, during which visual attention is directed to a specific area of the visual display [5]. Saccades are the continuous and rapid movements of eye gazes between fixation points. They are extremely rapid, often only 40–50 milliseconds, and can have velocities approaching 500 degrees per second [5]. Pupil dilation is a measure that is typically used to indicate an individual's arousal or interest in the viewed content matter, with a larger diameter reflecting greater arousal [5]. A scanpath encompasses the entire sequence of fixations and saccades, which can present the pattern of eye movement across the visual scene [5].

2.1. Eye-Movement Measures

Lai et al. (2013) [13] proposed a framework of eye-movement measures (temporal, spatial, and count) that can be identified and applied in reviews of eye-tracking studies. First, the temporal scale measures eye movement in a time dimension, e.g., durations of time spent on particular areas of interest: fixation duration; total reading time; time to first fixation; and gaze duration. Second, the spatial scale measures eye movement in a space dimension. It concerns locations, distances, directions, sequences, transactions, spatial arrangement, or relationships of fixations or saccades. Indices such as fixation position, fixation sequence, saccade length, and scanpath patterns belong to this scale. Third, the count scale measures eye movements on a count or frequency basis. For example, fixation count, revisited fixation count, and probability of fixation count belong to this category. This framework is adopted in the present review study [16]. A similar framework was also proposed by Sharafi et al. (2015), in which eye-tracking measures are number of fixations (count), duration of fixation (temporal), and scanpaths (spatial) [19].

Temporal measures may answer the "when" and "how long" questions about cognitive processing, and are often used to imply the occurrence of reading problems [21]. Spatial measures may answer the "where" and "how" questions about the cognitive process. Saccadic eye movements and scanning behaviors are important in that they reveal the control of selective processes in visual perceptions, including visual searching and reading [22]. Count measures are usually used to reveal the importance of visual materials. Sometimes, fixation counts are strongly correlated with measures such as total fixation duration. This suggests that measurements in different categories might reflect the same cognitive process.

2.2. Eye-Tracking Results Presentation

The eye-tracking device enables the presentation of research results in several ways:

1. Heat map: in the image being examined there are spots in the colors from red to green, which represent the length of the user's eye concentration in a given area.
2. Fixation map: presented using points defining the areas of concentration of the line of sight. Points are numbered and connected with lines.
3. Table: the elements on which the eyes were concentrated, together with the duration of the concentration and the order of their observation, are presented in the rows.
4. Charts: presents the time of fixation and clickability, together with the position of each result from search engine results pages.

These presentation forms show the eye-tracking results and how participants have interacted with an environment or responded to a task.

2.3. Eye-Tracking Participants

Participants are invited to the study and receive prepared tasks. Each participant receives an identical set of tasks, while each should be calibrated with a device that tracks eye movements. The greater the number of participants, the more reliable the results, although there are also studies where there have been only a few participants. To guarantee the reliability of results, the literature recommends that the study group should have more than 30 respondents to ensure that the group is internally consistent [23]. When preparing the test, it is necessary to consider data losses resulting from poor calibration or other unexpected factors.

3. Literature Review

The method used in this systematic literature review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [24].

3.1. General Database Search

The literature source for this review was the Web of Science and Scopus. The period was set from 2004 to 2019 and the document type was limited to journal articles and conference proceedings. The procedures implemented to identify the research papers of this study can be classified into four stages. In the first stage, two sets of keywords were organized for searches using the Boolean operator AND, including first: "eye track*" AND "search engine", second "eye track* AND web search". The word "eyetracking" word was not used for the search since it significantly reduced the number of results. More commonly used words are either "eye-tracking" or "eye tracking". The search terms were used for all fields (including title, abstract, keywords, and full text). Searches revealed 71 papers in Web of Science and 200 papers in Scopus. Scopus returned more articles since content from IEEE and the ACM Digital Library are listed in the Scopus database. After removing duplicates between the two queries and 57 overlapping results between these two databases, 214 papers were left.

3.2. Focused Searches

After narrowing down the results to peer-reviewed studies, the following criteria were implemented to further refine the results. In the second stage, the article titles and abstracts were manually and systematically screened to confirm that the selected articles were studies on search engines, used eye-tracking devices, and provided empirical evidence or evaluation. Most papers were published in various computer science/HCI conference proceedings such as SIGIR or SIGCHI. Additionally, a few papers had been published in management information systems journals, such as *Journal of the Association for Information Science and Technology*, *Behaviour and Information Technology*, *Information Processing and Management*, and *Aslib Journal of Information Management*.

3.3. Additional Searches Through References

The references of the initially found papers and the references made to those papers were further investigated. This method comes from snowball sampling [19]. Using this method, nine papers not covered in the databases, yet highly relevant for the literature review, were discovered.

The papers in the literature search that did not satisfy the set criteria mostly fell into the following three categories:

1. Papers on click rate and log analysis, which were only compared to eye-tracking;
2. Papers on eye-tracking for search engines but for which the scientific rigor was poor and presented obvious results;
3. Eye-tracking was mentioned in the text but the actual substance was not eye-tracking-related.

3.4. Analysis

After performing these steps for the literature search, 56 peer-reviewed, empirical research papers on eye-tracking for search engines were identified for the review. A total of 43 papers were published conference proceedings and 13 papers were published in journals. Webster and Watson (2002) presented the concept of a matrix for the systematic literature review [25]. The modified matrix, based on the matrix for systematic literature review provided by Webster and Watson (2002) [25], contained 56 peer-reviewed papers on eye-tracking for search engines. The literature searches and selection processes were completed by the author (Figure 2). The full list of the chosen papers can be found in the Appendix A.

### 3.5. Research Questions and Coding Procedure

The present review focused on examining 10 research questions, which were the basis for the coding procedure:

1. Which search engine(s) were tested?
2. What type of apparatus was used for the study?
3. What kind of participants took part in the studies? (e.g., age, gender, education)
4. What type of interface was tested?
5. Which types of results or parts of the results were tested?
6. Which eye-movement measures were used?
7. What kind of scenarios were prepared for participants?
8. What types of tasks were performed using the search engines?
9. Which language was used when queries were provided by participants?
10. How were the results of the eye-tracking study presented?
11. What research question are addressed by each study?
12. What key findings are presented in each study?

The coding procedure was created based on the content analysis of selected papers. It was performed in preliminary form and common parts of each paper were selected:

1.　Search engine. Each paper described an eye-tracking study on at least one search engine. Some studies presented research on two different search engines as a comparison study. Although some papers did not mention the chosen search engine, it was possible to obtain this information from the screenshots provided.

2.　Apparatus. Most papers had a detailed section about the device(s) used and basic settings like size of screen in inches, screen resolution, web browser used, and distance from the screen.

3.　Participants. Each study described how many participants took part in tests and almost all of them described the basic characteristics of participant groups, e.g., age, gender, and education.

4.　Interface. Either desktop or mobile.

5.　Search engine results. Generally, search results can be either organic or organic and sponsored. Organic results are produced and sorted by ranking using the search engine's algorithm [26]. Sponsored search results come from marketing platforms that usually belong to the search engines. On these platforms, advertisers create sponsored results and pay to display it and when it is clicked. Search engines display sponsored results in a way to distinguish them from organic results [27]. Some studies were focused on other areas of search engines' results pages, e.g., knowledge graphs.

6.　Eye-movement measures. This measure is based on a temporal–spatial–count framework [13].

7.　Scenario. Search results can be displayed using the normal output format of the search engine or modified by researchers and displayed in a changed form. The most common changes involve switching the first two results to see if a participant can recognize this change or reversing the order of all results.

8.　Tasks. Tasks are grouped into navigational, informational, or transactional. Navigational tasks involve finding a certain website. Informational tasks involve finding information on any website. Transactional tasks involve finding information about a product or price.

9.　Language. Participants always belong to one language group and use this language in the search engine.

10.　Presentation of results. Researchers can illustrate the results of eye-tracking studies in several ways, e.g., heat maps, fixation maps, gaze plots, charts, and tables.

11.　Research questions. Summary the sort of research interests and questions on user search behavior utilizing eye tracking and how effective eye tracking methodology was for providing insights on these specific questions.

12.　Key findings. Summary of key finding presented in each study.
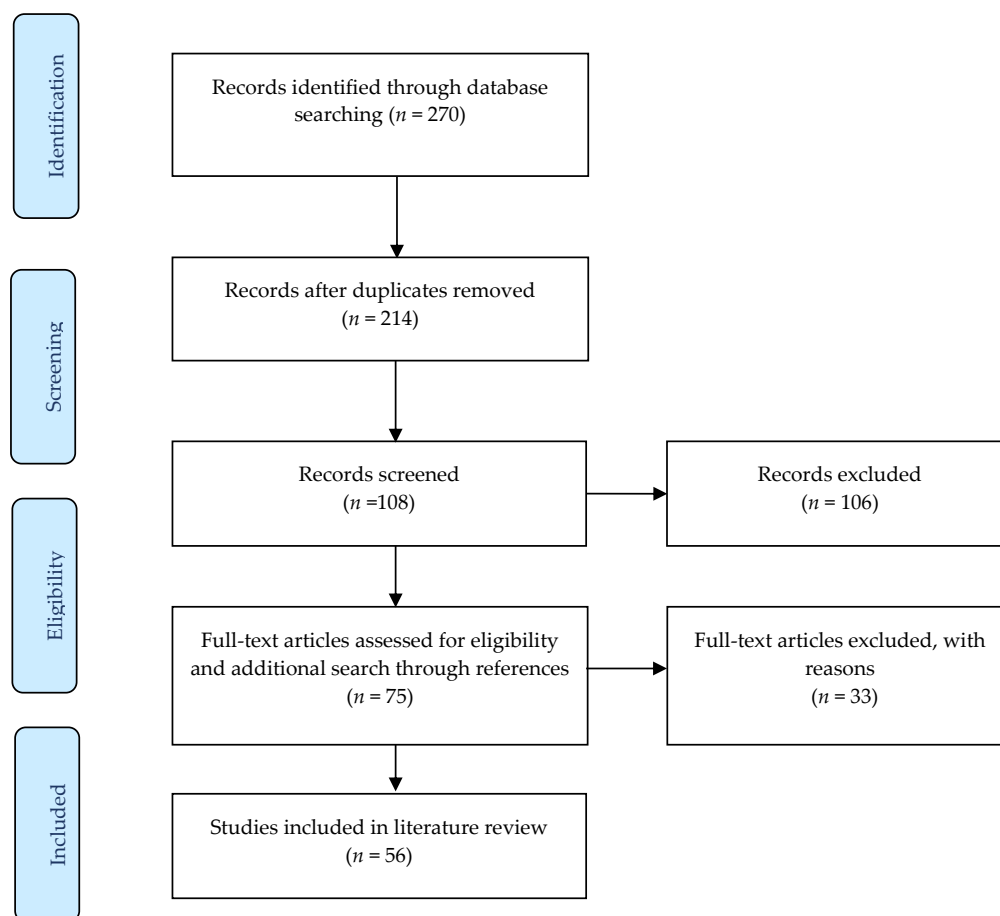
Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) chart of search strategy.

4. Results

From the content analysis, it was found that the 56 papers reviewed here discussed the topic of web-search behavior based on informational and navigational queries, recognition of organic and sponsored listings, detailed studies about different parts of presented results (title, description, URL) and studies involving the search engine's results page being modified by reversing results, changing the position of results, switching to horizontal navigation, and dividing results into a grid. The initial analysis suggested that when the eye-movement method was applied to studies related to search engines, the focus of discussion was largely on the acquisition of data regarding how results from search engines are perceived. Despite each study having many common parts, such as those highlighted in the coding procedure, there was one main category of question that they were all trying to answer: Each study presented results for a set of particular elements described in the coding procedure, but the main question remains the same: How do you search?

4.1. Search Engine

The selected studies focused on eight different search engines: 40 studies on Google; nine studies provided on MSN, Live, and Bing (which are the same search engine but with different names in different years); five studies on Yahoo; three studies on Baidu; two studies on Sogou; and one study on Blindeh-Kuh (a German search engine for children). Some papers provided results for more than one search engine or let participants choose one of two proposed search engines. In some studies, researchers created a controlled interface for the search engine, but results were still downloaded from the commercial search engine.

The focus of eye-tracking studies is limited to commercial search engines. Currently, there is a limited number of search engines operating on the Internet, with Google having the largest market share. However, this review revealed that eye-tracking studies can be performed on other search engines. Results from these studies can provide valuable contributions to search-engine studies overall.

4.2. Apparatus

Researchers mostly used professional apparatus for their tests. In 35 studies, Tobii devices were used. Usually, these were 1750, x50, T120, or T60 models. In 10 studies, researchers used an ASL 504. In four studies, researchers used Facelab 5. In six studies, researchers used common web cameras, either built into the laptop or mounted on a monitor. The usage of the apparatus was strictly aligned with the research center from which the paper was published. Different researchers, therefore, commonly used the same apparatus. For example, Cornell University [28–35] used ASL 504, Microsoft Research Center [11,36–41] used Tobii x50, Worcester Polytechnic Institute [42–45] used Tobii X120, Tsinghua University [6,46–48] used Tobii X2-30, and The Australian National University [49–52] used Facelab5.

Sharafi et al. (2015) described the criteria that should be considered when researchers are choosing eye-tracking devices for their studies [19]. There is an extreme variability in the costs of eye trackers. Thus, when considering the use of eye trackers, researchers must consider a tradeoff between the cost and the quality of the eye trackers. They provided a list of factors that should be considered while comparing eye trackers: accuracy; sampling rate; customer support; time needed for setting up the study; and software features of the eye-trackers' driver and analysis software systems.

4.3. Participants

Probably the most important element of each test for eye-tracking studies is the participants. Most studies briefly described metrics about participants and how they were recruited for the study. In most studies (36 papers), recruited participants were students from universities or colleges where researchers worked. For taking part in tests, students were often rewarded with additional points, grades, gifts, vouchers, or payments. These groups were usually between 18 to 24 years of age, almost equally divided into males and females, and most of them declared that they had advanced or expert levels in using the Internet and search engines. Almost all of them were always familiar with the tested search engine. In 15 papers, participants taking part in tests were more diverse, usually in the range of 18 to 60 years of age, both male and female. In three papers, participants were children [10,53,54]. One study was lacking information about the participants. One paper only asked female college students to take part in the study [55].

Often, when recruitment was advertised, the e-mails or leaflets contained information that participants should not have problems with vision. Before participants are allowed to do the test, the eye-tracking device must be calibrated for each individual. For several studies, there were problems with calibration with some participants, so the overall number of participants taking part in the study was lower than the number invited for tests. Problems usually occurred with participants with corrected vision or who were older.

As mentioned before, 30 participants for eye-tracking studies is the recommended minimum number. In the present review, the number of participants for each reviewed paper is given after subject losses due to issues with calibration. The rounded mean value is 30 participants in the whole set of papers. The median was 29, with a standard deviation of 12.32. A total of 28 papers had less than 30 participants and 27 papers involved 30 or more participants (see Table 1 for details).

Some studies revealed that there were different searching behaviors even if participants were from internally consistent groups. If the scenario of the study contained modified results and more than one type of task, participants were divided into smaller groups, so that results could be compared between these subgroups.

Table 1. Descriptive statistics of participants in eye-tracking studies.

	Mean	**Median**	**SD**	**Min.**	**Max.**
Participants	30.19	30	12.21	5	58

4.4. Interface

In the early days of eye-tracking studies, the tested interface was always presented on desktop computers, with a monitor and fixed resolution, with the eye-tracking device being connected to the computer and monitor. A total of 45 of the chosen studies were performed using the desktop version of search engines. More recently, however, mobile phones and smartphones have gained popularity. The first study on a mobile-sized screen was performed by Kim et al. (2012) [49]. In seven papers mobile versions or mobile devices were studied [49,51,52,56–59]. Kim et al. (2015) tested the mobile version, but not on a mobile device or smartphone [50]. In three papers performed a comparison test between the desktop and mobile versions of the search engine [43,44,50].

The tested interface depends on the eye-tracking device. Eye-tracking devices were often used for desktop computers. This explains why 45 of the 55 chosen studies were performed using desktop computers and why the study by Kim et al. (2015) only simulated the mobile version on a computer monitor [50]. Given the limited number of studies on mobile devices, technological advances appear to be underutilized at present. A similar limitation was observed by Meissner and Oll (2019) [17].

4.5. Search Engine Results

In the early days, search engines always presented organic search results. These are created based on search engine crawler data, indexed, and displayed by the search engine. Later, search engines often launched platforms for sponsored results to provide funds to maintain their infrastructure. Researchers mainly tested organic results (41 papers), organic and sponsored results (6 papers; [8,11,39,40,60,61]), or only sponsored results (7 papers; [41,43–46,58,59]. Three papers tested granular parts of organic results like title, description, and URL separately [60–62]. In one paper, knowledge graphs were tested [57] and in one, image results were tested [63].

Most academic research centers performed studies on organic results. In several studies, the scenario contained filtered results from the search engine, and additional elements displayed on the screen like ads, maps, and knowledge graphs were removed; thus, participants were not distracted by other elements. Most of the reviewed papers focused on behavior with organic results. Only a few researchers, mainly representing search engines, performed studies on sponsored search results; these studies were designed to check how well sponsored ads were performing [41,58].

4.6. Eye-Movement Measures

As far as eye movement measures are concerned, temporal measures were the most frequently employed overall (54 times), followed by frequency measures (48 times), and spatial measures (30 times). Temporal measures usually reveal how long participants looked at a certain place on the page with results. Count measures usually show how often participants looked at the tested area. Spatial measures are presented in papers in the form of heat maps or gaze plots and show areas where participants looked.

4.7. Scenario

For each eye-tracking study, the researchers prepared scenarios and tasks. A scenario is a general environment in which the test was provided. It covers the selected search engine and whether the results were taken as provided by each engine or were modified. Usually, modification covered downloading results from search engines and displaying them from the cache, so for each tested participant, the results would be the same. Another type of modification involved changes to the user search interface, e.g., results being displayed in tabs [64], in reversed order [29,34,41,63,65], or with

additional words and descriptive categories [66]. In 26 papers, participants used regular results from the search engine, unmodified by the researchers. In 30 papers, results were somehow modified and prepared for the tests.

The scenario also covers the time-frame window. In some studies, times for accomplishing each task were set: 20 minutes for query sessions [66]; 10 minutes for each task [53]; and four minutes for each task [64]. In other studies, researchers only reported the time needed to accomplish a study session: 20 minutes to complete the session [56]; 25 to 30 minutes to complete a session in a laboratory [52]; and from 45 minutes up to 90 minutes [10,67].

4.8. Tasks

Detailed tasks contain a set of queries to search for. The literature describes informational, navigational, and transactional queries [68]. The most common task for participants was informational searching (54 papers). In several papers, researchers used the same informational queries, e.g., "Where is the tallest mountain in New York located?" [29–32,34] or "best snack in Boston" [42–45]. In some papers, informational tasks were divided into closed or open ones. Closed tasks relate to direct informational queries. Open tasks relate to descriptive information on what to search, with the actual search queries chosen by participants. The second most common type of task was navigational searches (26 papers). In several papers, researchers used the same navigational queries, e.g., "Find the homepage of Michael Jordan, the statistician" [29–32,34]. The least common task was transactional searches [48,58–61,69]. One study also utilized multimedia queries [60].

4.9. Language

The most common language for participants was English (36 studies), followed by German (nine studies), Chinese (five papers), Spanish (four papers), Japanese (one study; [70]) and Finnish (one study; [10]). Some studies were originally not provided in English, with results translated into English for publication purposes. Search engines can recognize in what language the query is written. Recent technological advancements also allow search engines to recognize spoken language in questions asked via voice search [71]. Users are willing and able to ask queries in their native language.

4.10. Presentation of Results

The results of the studies were presented in different ways. Heat maps, where colors illustrate the intensity of participants' eye fixation, proved to be the most interesting and valuable way to present results, although this type of presentation is the least precise. More precise presentations like charts or tables did not produce results of equal interest value. In many papers, results were presented in more than one presentation form. A total of 17 papers presented results via heat maps, 17 papers presented results via charts, 14 papers presented results via tables, 5 papers via area-of-interest (AOI), two papers illustrated results via gaze plots, and one did not have any form of presentation [72].

Heat maps represent all the participants. In some studies, each task was presented on a different heat map [60], while in other studies all participants were included on one map [38,57]. Showing averages on a heat map, especially if the study has less than 30 participants, could provide different results. Pernice and Nielsen (2009) reported that, in their study, when 60 participants were divided into six groups (each group of 10 participants), heat-map results were different for each group [23]. The same happened when three groups were tested, each group containing 20 participants; results presented on heat maps were different for each group.

4.11. Research Questions

Research questions addressed in each paper in this review are summarized. Almost all of the papers have stressed the research question. In some of them, research questions were repeated by the same researchers conducting several studies or repeated by other researchers, who tried to repeat the same study and see the differences in the results. If there was no clearly stated research question,

it was possible to draw it from the title or achieved results. Research questions in reviewed papers can be grouped into seven following topics: basic search behavior (BSB), complex search behavior (CSB), ranking presentation (RP), clicking and cursor moving (CCM), ads recognition (AR), different group age behavior (AG) and mobile vs. desktop (MD).

Research questions on the topic of basic search behavior are concentrated on delivering answers to questions like: Does the user read from top to bottom? How many abstracts are read before clicking? Does the user read the descriptions? How long does the user spend on the search engine result page? Does the user look on its query? Research questions on the topic of complex search behavior are concentrated on complex tasks to solve like: How to find the answer to a medical problem using search engines? How do users' search behaviors change over time in a search session? Research questions on the topic of ranking presentation are concentrated on delivering answers to questions like: Is tabular presentation better than scrolling? How does a user behave when the position of the results is changed? What are the visualization techniques for the user behavior using search engine results pages?

Research questions in the topic of clicking and cursor moving are concentrated on delivering answers to questions like: To what extent do eye movements correlate with cursor behavior? What does cursor behavior reveal about search engine users' result examination strategies? How to estimate the credibility of user clicks with mouse movement and eye-tracking Information on SERPs (Search Engine Results Page)? Research questions in the topic of ads recognition are concentrated on delivering answers to questions like: How does the user distribute visual attention on results page, especially on ads and sponsored links? Does the presence of ads affect the attention to search results? Research questions on the topic of different group age behavior are concentrated on delivering answers to questions like: Do children and adults have different search strategies? how children examine search results? Research questions on the topic of mobile vs. desktop are concentrated on delivering answers to questions like: What is the difference between user search behavior on large and small screens? How does the diversification of screen sizes on hand-held devices affect how users search? Is horizontal swiping better than vertical scrolling? In Table 2 is a list of every research question stated in the reviewed papers. Research questions are grouped into topics.

4.12. Key Findings

Over 16 years, eye-tracking research for search engines has changed. In the early years, researchers were interested how users were using search engines, i.e., where and how often they looked at results. In later years, other aspects of using search engines were studied, like different screens, different result types, modified or regular results, and other behaviors of users. Table 3 presents the key findings from each study.

These 56 studies (in which the researchers tested either all or specific parts of the search engine results pages) show how search engine results pages are perceived. The review of the studies leads to two key conclusions. First, the research has evolved over time. Initially, it addressed simple questions, e.g., how long users look at the results and which results are observed. Later studies addressed questions such as what the differences are between sponsored search results and organic search results or what the differences are between desktop search results and mobile search results. The most recent research has addressed more detailed questions, such as which elements of snippets are observed: title; URL; description; or additional features beyond 10 blue links.

Second, these studies have resulted in significant findings that could not have been obtained by other means. Studies comparing the attention paid to the search results page by children and adults show that these two groups have different behaviors in terms of reading results from the search results page. Children read carefully, whereas adults read quickly and do not read every result.

Table 2. Research questions in the reviewed eye-tracking studies.

Study Reference.	Research Questions	Topic
Granka et al. (2004) [28]	How do users interact with the list of ranked results of WWW search engines? Do they read the abstracts sequentially from top to bottom, or do they skip links? How many of the results do users evaluate before clicking on a link or reformulating the search?	BSB
Klöckner et al. (2004) [72]	In what order do users look at the entries in a search result list?	BSB
Pan et al. (2004) [33]	Are standard ocular metrics such as mean fixation duration, gazing time, and saccade rate, and also scan path differences determined by individual differences, different types of web pages, the order of a web page being viewed, or different tasks at hand?	BSB
Aula et al. (2005) [73]	What are the personal styles for search result evaluation?	BSB
Joachims et al. (2005) [29]	Do users scan the results from top to bottom? How many abstracts do they read before clicking? How does their behavior change, if we artificially manipulate Google's ranking?	BSB
Radlinski and Joachims (2005) [35]	How query chains can be used to extract useful information from a search engine log file?	RP
Rele and Duchowski (2005) [62]	Will tabular interface increase efficiency and accuracy in scanning search results which will be achieved through a spatial grouping of results into distinct element/category columns?	RP
Lorigo et al. (2006) [31]	How does the effort spent reading abstracts compare with selection behaviors, and how does the effort vary concerning user and task? Or, can we safely assume that when a user clicks on the nth abstract, that they are making an informed decision based on n-1 abstracts preceding it?	BSB
Guan and Cutrell (2007) [37]	How users' search behaviors vary when target results were displayed at various positions for informational and navigational tasks?	RP
Joachims et al. (2007) [30]	How users behave on Google's results page? Do users scan the results from top to bottom? How many abstracts do they read before clicking? How does their behavior change, if we artificially manipulate Google's ranking?	BSB
Pan et al. (2007) [34]	What is the relevance through human judgments of the abstracts returned by Google (abstract relevance), as well as the pages associated with those abstracts (Web page relevance)?	BSB
Cutrell and Guan (2007) [38]	Do users read the descriptions? Are the URLs and other metadata used by anyone other than expert searchers? Does the context of the search or the type of task being supported matter?	BSB
Lorigo et al. (2008) [32]	How do users view the ranked results on a SERP? What is the relationship between the search result abstracts viewed and those clicked on, and whether gender, search task, or search engine influence these behaviors?	BSB
Egusa et al. (2008) [70]	What are the visualization techniques for the user behavior using search engine results pages?	RP
Xu et al. (2008) [7]	How to prepare a personalized online content recommendation according to the user's previous reading, browsing, and video watching behaviors?	RP

Table 2. *Cont.*

Study Reference.	Research Questions	Topic
Buscher et al. (2009) [36]	What is the relationship between segment-level display time and segment-level feedback from an eye tracker in the context of information retrieval tasks? How do they compare to each other and a pseudo relevance feedback baseline on the segment level? How much can they improve the ranking of a large commercial Web search engine through re-ranking and query expansion?	BSB
Kammerer et al. (2009) [66]	How do authority-related epistemological beliefs and salience of source information influence the evaluation of web search results?	CSB
Buscher et al. (2010) [41]	How do users distribute their visual attention on different components on a SERP during web search tasks, especially on sponsored links?	AR
Dumais et al. (2010) [11]	How is the visual attention devoted to organic results influenced by page elements, in particular ads and related searches?	AR
Marcos and González-Caro (2010) [60]	What is the relationship between the users' intention and their behavior when they browse the SERPs?	BSB
Gerjets et al. (2011) [74]	What type of spontaneous evaluation processes occur when users solve a complex search task related to a medical problem by using a standard search-engine design?	CSB
González-Caro and Marcos (2011) [61]	Whether user browsing behavior in SERPs is different for queries with an informational, navigational, and transactional intent?	BSB
Huang et al. (2011) [40]	To what extent does gaze correlate with cursor behavior on SERPs and non-SERPs? What does cursor behavior reveal about search engine users' result examination strategies, and how does this relate to search result clicks and prior eye-tracking research? Can we demonstrate useful applications of large-scale cursor data?	CCM
Balatsoukas and Ruthven (2012) [75]	What is the relationship between relevance criteria use and visual behavior in the context of predictive relevance judgments using the Google search engine?	CSB
Huang et al. (2012) [39]	When is the cursor position is a good proxy for gaze position and what is the effect of various factors such as time, user, cursor behavior patterns, and search task on the gaze-cursor alignment?	CCM
Kammerer and Gerjets (2012) [64]	How both the search interface (resource variable) and Internet-specific epistemic beliefs (individual variable) influenced medical novices' spontaneous source evaluations on SERPs during Web search for a complex, unknown medical problem?	CSB
Kim et al. (2012) [49]	On small devices, what are users' scanning behavior when they search for information on the web?	BSB
Marcos et al. (2012) [76]	What is the user behavior during a query session, that is, a sequence of user queries, results page views and content page views, to find a specific piece of information?	MD
Nettleton and González-Caro (2012) [77]	How learning to interpret user search behavior would allow systems to adapt to different kinds of users, needs, and search settings?	BSB

Table 2. *Cont.*

Study Reference.	Research Questions	Topic
Djamasbi et al. (2013a) [42]	Can competition for attention theory help predict users' viewing behavior on SERPs?	BSB
Djamasbi et al. (2013b) [43]	Would the presence of ads affect the attention to search results on a mobile phone? Do users spend more time looking at advertisements than at entries? Can advertisements distract users' attention from the search results?	AR
Djamasbi et al. (2013c) [44]	Is there the phenomenon of banner blindness on SERPS on computers and mobile phones?	AR
Hall-Phillips et al. (2013) [45]	Would the presence of ads affect the attention to the returned search results? Do users spend more time looking at ads than at entries?	AR
Bataineh and Al-Bataineh (2014) [55]	How do female college students view the web search results?	AG
Dickerhoof and Smith (2014) [78]	Do people fixate on their query terms on a search engine results page?	BSB
Gossen et al. (2014) [53]	Are children more successful with Blinde-Kuh than with Google? Do children prefer BK over Google? Do children and adults have different perception of web search interfaces? Do children's and adults' employ different search strategies?	AG
Hofmann et al. (2014) [67]	How many suggestions do users consider while formulating a query? Does position bias affect query selection, possibly resulting in suboptimal queries? How does the quality of query auto-completion affect user interactions? Can observed behavior be used to infer query auto-completion quality?	BSB
Jiang et al. (2014) [69]	How do users' search behaviors, especially browsing and clicking behaviors, vary in complex tasks of different types? How do users' search behaviors change over time in a search session?	CSB
Lagun et al. (2014) [57]	Does tracking the browser viewport (the visible portion of a web page) on mobile phones could enable accurate measurement of user attention at scale, and provide a good measurement of search satisfaction in the absence of clicks?	CCM
Y. Liu et al. (2014) [6]	What is the true examination sequence of the user in SERPs?	BSB
Z. Liu et al. (2014) [46]	How do sponsored search results affect user behavior in web search?	AR
Lu and Jia (2014) [63]	How users view the results returned by image search engines. Do they scan the results from left to right and top to bottom, i.e., following the Western reading habit? How many results do users view before clicking on one?	BSB
Mao et al. (2014) [48]	How to estimate the credibility of user clicks with mouse movement and eye-tracking Information on SERPs?	CCM
Kim et al. (2015) [50]	Is there any difference between user search behavior on large and small screens?	MD
Z. Liu et al. (2015) [47]	What is the influence of vertical results on search examination behaviors?	MD
Bilal and Gwizdka (2016) [54]	What online reading behavior patterns do children between ages 11 and 13, demonstrate in reading Google SERPs? What eye fixation patterns and dwell times do children in ages 11 and 13, exhibit in reading Google SERPs?	AG

Table 2. *Cont.*

Study Reference.	Research Questions	Topic
Domachowski et al. (2016) [59]	How is attention distributed on the first screen of a mobile SERP and which actions do users take? How effective are ads on mobile SERPs regarding the clicks generated as an indicator of relevance? Are users subjectively aware of ads on mobile SERPs and what is their estimation of ads?	AR
Kim et al. (2016a) [56]	How the diversification of screen sizes on hand-held devices affects how users search?	MD
Kim et al. (2016b) [51]	Is horizontal swiping for mobile web search better than vertical scrolling?	MD
Lagun et al. (2016) [58]	How do we evaluate rich, answer-like results, what is their effect on users' gaze and how do they impact search satisfaction for queries with commercial intent?	AR
Kim et al. (2017) [52]	What snippet size is needed in mobile web search?	MD
Papoutsaki et al. (2017) [8]	Can SearchGazer be useful for search behavior studies?	BSB
Bhattacharya and Gwizdka (2018) [79]	Are the changes in verbal knowledge, from before to after a search task, observable in eye-tracking measures?	BSB
Hautala et al. (2018) [10]	Are sixth-grade students able to utilize information provided by search result components (i.e., title, URL, and snippet) as reflected in the selection rates? What information sources do the students pay attention to and which evaluation strategies do they use during their selection? Does the early positioning of correct search results on the search list decrease the need to inspect other search results? Are there differences between students in how they read and evaluate Internet search results?	AG
Schultheiss et al. (2018) [65]	Can the results by Pan et al. (2007) be replicated, despite temporal and geographical differences?	BSB
Sachse, (2019) [80]	What snippet size to use in mobile web search?	BSB

Table 3. Key findings in the reviewed eye-tracking studies.

Study Reference	Key Findings
Granka et al. (2004) [28]	The amount of time spent viewing the presented abstracts, the total number of abstracts viewed, as well as measures of how thoroughly searchers evaluate their results set.
Klöckner et al. (2004) [72]	Defining depth- and breadth-first processing of search result lists.
Pan et al. (2004) [33]	Gender of subjects, the viewing order of a web page, and the interaction effects of page order and site type on online ocular behavior.
Aula et al. (2005) [73]	Based on evaluation styles, the users were divided into economic and exhaustive evaluators. Economic evaluators made their decision about the next action faster and based on less information than exhaustive evaluators.
Joachims et al. (2005) [29]	Users' clicking decisions were influenced by the relevance of the results, but that they were biased by the trust they had in the retrieval function, and by the overall quality of the result set.
Radlinski and Joachims (2005) [35]	Presented a novel approach for using clickthrough data to learn ranked retrieval functions for web search results.
Rele and Duchowski (2005) [62]	The eye-movement analysis provided some insights into the importance of search result's abstract elements such as title, summary, and the URL of the interface while searching.
Lorigo et al. (2006) [31]	The query result abstracts were viewed in the order of their ranking in only about one fifth of the cases, and only an average of about three abstracts per result page were viewed at all.
Guan and Cutrell (2007) [37]	Users spent more time on tasks and were less successful in finding target results when targets were displayed at lower positions in the list.
Joachims et al. (2007) [30]	Relative preferences were accurate not only between results from an individual query, but across multiple sets of results within chains of query reformulations.
Pan et al. (2007) [34]	College student users had substantial trust in Google's ability to rank results by their true relevance to the query.
Cutrell and Guan (2007) [38]	Adding information to the contextual snippet significantly improved performance for informational tasks but degraded performance for navigational tasks.
Lorigo et al. (2008) [32]	Strong similarities between behaviors on Google and Yahoo!
Egusa et al. (2008) [70]	New visualization techniques for user behaviors when using search engine results pages.
Xu et al. (2008) [7]	New recommendation for an algorithm for online documents, images, and videos, which is personalized.
Buscher et al. (2009) [36]	Feedback based on display time on the segment level was much coarser than feedback from eye-tracking for search personalization.
Kammerer et al. (2009) [66]	Authority-related epistemological beliefs affected the evaluation of web search results presented by a search engine.
Buscher et al. (2010) [41]	The amount of visual attention that people devoted to ads depended on their quality, but not the type of task.
Dumais et al. (2010) [11]	Provided insights about searchers' interactions with the whole page, and not just individual components.

Table 3. *Cont.*

Study Reference	Key Findings
Marcos and González-Caro (2010) [60]	A relationship existed between the users' intention and their behavior when they browsed the results page.
Gerjets et al. (2011) [74]	Measured spontaneous and instructed evaluation processes during a web search.
González-Caro and Marcos (2011) [61]	Organic results were the main focus of attention for all intentions; apart from for transactional queries, the users did not spend much time exploring sponsored results.
Huang et al. (2011) [40]	The cursor position was closely related to eye gaze, especially on SERPs.
Balatsoukas and Ruthven (2012) [75]	Novel stepwise methodological framework for the analysis of relevance judgments and eye movements on the web.
Huang et al. (2012) [39]	Gaze and cursor alignment in web search.
Kammerer and Gerjets (2012) [64]	Students using the tabular interface paid less visual attention to commercial search results and selected objective search results more often and commercial ones less often than students using the list interface.
Kim et al. (2012) [49]	On a small screen, users needed relatively more time to conduct a search than they did on a large screen, despite tending to look less far ahead beyond the link that they eventually selected.
Marcos et al. (2012) [76]	Proposed a model of user search behavior which consists of five possible navigation patterns.
Nettleton and González-Caro (2012) [77]	Successful users formulated fewer queries per session and visited a smaller number of documents than unsuccessful users.
Djamasbi et al. (2013a) [42]	Viewing behavior can potentially have an impact on effective search and thus user experience of SERPs.
Djamasbi et al. (2013b) [43]	The presence of advertisements and location on the screen can have an impact on user experience and search.
Djamasbi et al. (2013c) [44]	Top ads may be more effective in attracting users' attention on mobile phones compared to desktop computers.
Hall-Phillips et al. (2013) [45]	Findings provided support for the competition for attention theory in that users were looking at advertisements and entries when evaluating SERPs.
Bataineh and Al-Bataineh (2014) [55]	Making searches in not native language required more time for the scanning and reading of results.
Dickerhoof and Smith (2014) [78]	Users fixated on some of the displayed query terms; however, they fixated on other words and parts of the page more frequently.
Gossen et al. (2014) [53]	Children used a breadth-first-like search strategy, examining the whole result list, while adults only examined the first top results and reformulated the query.
Hofmann et al. (2014) [67]	Focus on top suggestions (query auto completion) is due to examination bias, not ranking quality.
Jiang et al. (2014) [69]	User behavior in the four types of tasks differed in various aspects, including search activeness, browsing style, clicking strategy, and query reformulation.
Lagun et al. (2014) [57]	Identified increased scrolling past direct answer and increased time below direct answer as clear, measurable signals of user dissatisfaction with direct answers.

Table 3. *Cont.*

Study Reference	Key Findings
Y. Liu et al. (2014) [6]	Proposed a two-stage examination model: skimming to the reading stage; and reading to clicking stage.
Z. Liu et al. (2014) [46]	Different presentation styles among sponsored links might lead to different behavior biases, not only for the sponsored search results but also for the organic ones.
Lu and Jia (2014) [63]	Image search results at certain locations, e.g., the top-center area in a grid layout, were more attractive than others.
Mao et al. (2014) [48]	Credible user behaviors could be separated from non-credible ones with many interaction behavior features.
Kim et al. (2015) [50]	Users had more difficulty extracting information from search results pages on the smaller screens, although they exhibited less eye movement as a result of an infrequent use of the scroll function.
Z. Liu et al. (2015) [47]	Influence of vertical results in web search examination.
Bilal and Gwizdka (2016) [54]	Grade level or age had a more significant effect on reading behaviors, fixations, first result rank, and interactions with SERPs than task type.
Domachowski et al. (2016) [59]	There was no ad blindness on mobile searches but, similar to desktop searches, users also tended to avoid search advertising on smartphones.
Kim et al. (2016a) [56]	Behavior on three different small screen sizes (early smartphones, recent smartphones, and phablets) revealed no significant difference concerning the efficiency of carrying out tasks.
Kim et al. (2016b) [51]	Users using pagination were: more likely to find relevant documents, especially those on different pages; spent more time attending to relevant results; and were faster to click while spending less time on the search result pages overall.
Lagun et al. (2016) [58]	Showing rich ad formats improved search experience by drawing more attention to the information-rich ad and allowing users to interact to view more offers.
Kim et al. (2017) [52]	Users with long snippets on mobile devices exhibited longer search times with no better search accuracy for informational tasks.
Papoutsaki et al. (2017) [8]	Introduced SearchGazer, a web-based eye tracker for remote web search studies using common webcams already present in laptops and some desktop computers.
Bhattacharya and Gwizdka (2018) [79]	Users with higher change in knowledge differed significantly in terms of their total reading-sequence length, reading-sequence duration, and number of reading fixations, when compared to participants with lower knowledge change.
Hautala et al. (2018) [10]	Students generally made a flexible use both of eliminative and confirmatory evaluation strategies when reading Internet search results.
Schultheiss et al. (2018) [65]	Although the viewing behavior was influenced more by the position than by the relevance of a snippet, the crucial factor for a result to be clicked was the relevance and not its position on the results page.
Sachse, (2019) [80]	Short snippets provide too little information about the result. Long snippets of five lines lead to better performance than medium snippets for navigational queries, but worse performance for informational queries.

5. Discussion

The scope of this review was to identify the use of eye-tracking technology as a research method in the study of search engines. The purpose of the review was also to identify gaps in how eye-tracking technology is used for search engines and the possible use of eye-tracking in future research. The review was guided following the PRISMA recommendations and, after applying the selection criteria described in Section 2, 56 papers from the Web of Science and Scopus databases were found relevant for this review.

In this paper, the current efforts in empirical eye-tracking studies for search engines have been divided into components to analyze the results and state of the research. A coding procedure for eye-tracking studies on search engines based upon the (1) search engine, (2) apparatus, (3) participants, (4) interface, (5) results, (6) measures, (7) scenario, (8) tasks, (9) language, and (10) presentation was provided and the studies were categorized based on the coding procedure.

As these results show, in these eye-tracking studies, most attention was received by the Google search engine. The most frequently used device for measurement was Tobii. Participants were usually students recruited from the university campus. The most tested results were organic. Eye-tracking apparatus was mainly calibrated with desktop computers; thus, the desktop interface displayed search-engine results. As far as eye-movement measures are concerned, the temporal forms were the most frequently used. Researchers used both modified and regular scenarios, setting informational and navigational tasks on search engines. English was the main language of most studies and results were most often presented on heat maps.

In answering the main question behind every study ("How do you search?"), the literature review suggests that users search for answers in different ways. There is no single path for searching. Users use different strategies like breadth-first strategy (user scans several results and then revisits and opens the most promising ones), depth-first strategy (user examines each entry in the list, starting from the top, and decides immediately whether to open the page), "only top 3 results" strategy (user opens only the first three (top) results), or other different observed strategies. Search behavior also depends on the age of users. In some studies children participated in tests, revealing that children examine search results more deeply and in more detail than adults [10,53,54].

In the next section, methodological limitations in the reviewed studies are discussed, followed by avenues for future research and the limitations of this literature review.

5.1. Methodological Limitations in The Reviewed Studies

Several limitations were identified during the literature review. The major issue was that people who were invited for the tests but could not be calibrated with the eye-tracking apparatus were excluded from the study. This limitation shows that not everyone can be a participant of such a study, even though these excluded participants certainly use search engines and data collected from them could yield more in-depth results. The small sample size seems to be an additional source of concern. About half of the studies do not seem to reach the recommended minimum of 30 participants.

The second major limitation was that mainly students from university campuses were invited for the studies. This is a shortcoming because it narrows participants to internally consistent groups in which everyone is similar in terms of age and education. The study could only be representative of this kind of group, not for other users with different age and education. The same conclusion was drawn by Alemdag and Cagiltay (2018) [12].

The reviewed studies mainly focused on regular search-engine results—mostly organic, with some sponsored. Only two studies were designed to test other elements like knowledge graphs [57] or image searches [63]. There is a lack of studies for other known elements of search engines like news searches or video searches. In addition, in the past, search engines had extensions like instant search [81] and real-time search [82], which have not been tested in eye-tracking studies.

The reviewed studies were only in six different languages (English, German, Spanish, Chinese, Japanese, and Finnish). All of them are left-to-right written, since the results in the search engines

were displayed from left to right. There is no study where right-to-left written language was used, e.g., Arabic or Hebrew.

Finally, few studies have replicated results from other studies. Egusa et al. (2008) [70] repeated the study of Lorigo et al. (2008) [32] and proposed a different presentation form. Papoutsaki et al. (2017) [8] repeated the studies of Cutrell and Guan (2007) and Busher et al. (2010) [38,41] to see if the proposed software could replace eye-tracking devices. Schultheiss et al. (2018) [65] repeated the study of Pan et al. (2007) [34] to see how trust in Google had changed after 10 years.

5.2. Avenues for Future Research

Several gaps in eye-tracking research on search engines have been determined in this review. Regarding the methodologies of the reviewed studies, the majority were conducted with college students and mainly on Google. It is important to replicate existing studies with different types of participants and using other search engines. More research studies could be conducted with children and high-school students within the age range of 13 to 18 years and older. One of the reasons for this in an analysis of the readability and level of word complexity of results snippets and associated pages [83].

There are still some search engines on which eye-tracking studies have yet to be conducted: the Chinese Shenma and Haosou; the Russian Yandex and Mail.ru; the Czech Seznam; the Vietnamese CocCoc; the Korean Naver; and the US DuckDuckGo. It is recommended that more studies be performed on a wider range of search-engine languages as users are increasingly using their language.

This review revealed that eye-tracking studies on search engines are being undertaken only in few academic research centers: Cornell University (eight studies); Microsoft Research Center (six studies); Worcester Polytechnic Institute (four studies); Tsinghua University (four studies); The Australian National University (four studies); Knowledge Media Research Center (three studies); and Pompeu Fabra University (three studies). More research centers and universities could start eye-tracking studies on search engines. With this comes another suggestion: the eye-tracking technology needs to be cheaper than it is now to be used more widely [14,15,18,19].

Considering the limited number of eye-tracking studies on search engines, this review covered only 56 papers in the period of 16 years. Therefore, it is critical for researchers from other countries to contribute to this research area. They can use the proposed scenarios and tasks in their countries and their search engines to discuss findings over different cultures and provide strong empirical findings to be used in future analyses [12]. With only 13 studies identified in journals, the application of eye-tracking in journals is not only surprisingly rare but also restricted to a very limited number of outlets. They are not even selected from leading journals but come from across scientific databases.

The reviewed studies have used only languages for which the displayed results are written left-to-right. Studies on languages written right-to-left could reveal additional behaviors of participants. Common search engines are also operating on markets and in countries where the language is written right-to-left, e.g., Arabic or Hebrew language. Researchers from this particular area could run eye-tracking studies on search engines.

Another possibility to extend research in eye-tracking studies on search engines is to test displayed results from voice searches. Every task in the reviewed studies was based on typing queries on the keyboard. Nowadays, users are increasingly using voice searches to find information [71]; however, as well as hearing the results provided by search engines, users also see them displayed on screen, either desktop or mobile.

The rapid development of mobile technologies and devices like mobile phones or smartphones has resulted in a change in how users search for information using search engines [18]. This development raises some new challenges in how to study users' behavior on search engines and their interaction with the interface, either keyboard or voice search, using mobile devices. Mobile technology has also driven the development of responsive user interfaces where the user interface changes according to the screen size of the mobile device. Eye-tracking could be used as a means of studying the usability

challenges of mobile devices and thereby also the information behavior of users. Only in 11 papers were studies on mobile devices, and five of these were performed in only one research center [49–52,56].

Eye-tracking devices are adopted in these search-related studies, mostly because people either want to study the attention distribution on search interfaces or study the examination behavior during search processes. The progresses made by existing works on desktop versions of the search interface have been explored. We know a lot about how users behave using the desktop version. What is still unknown is the use of search engines on a variety of mobile devices, especially those with large screens and full touch screens.

Most eye-tracking studies suffer from the large cost of devices and therefore cannot involve many participants. The experiments are also usually performed in a controlled environment instead of practical ones. Investigating possible errors that are caused by these common settings is another interesting topic. However, this avenue can be explored by every eye-tracking study, not only in web search engines.

Modern search engine systems go far beyond 10 blue links in the presentation of results. However, most existing studies made many simplifications in their experimental settings. The future studies may look into these settings and try to find out which factors (or combination of factors) are still not investigated.

5.3. Limitations of This Literature Review

Two reviewed papers were published in the form of an extended abstract or preliminary study. Applying the filter of excluding not fully presented work, these two publications should have been excluded from this review. However, these works by Granka et al. (2004) and Klöckner et al. (2004) [28,72] received a large amount of attention in the literature by being cited in several other works; they were the very first published pieces of research in eye-tracking studies on search engines and both are covered in Scopus. Granka et al. (2004) [28] has been cited 391 times and is cited in 19 of the 56 reviewed papers, while Klöckner et al. (2004) [72] has been cited 49 times and is cited in 8 of the 56 reviewed papers. In this review, therefore, both have been treated as methodological foundations in the area of eye-tracking studies on search engines.

This review only focused on eye-tracking studies. Some parts of the papers also studied click-tracking or cursor-tracking and the possible correlation between gaze position and click-tracking or cursor position [39,40]. Despite these possible correlations, they were not considered in this review. There are research papers that focus only on eye–mouse coordination and cursor position, and this could be one direction for future literature-review studies.

There is also a limitation that applies to all literature reviews, i.e., the question of whether the major papers have been found adequately or not. The search was restricted to computer science and social science to find relevant papers based on the research query. However, the Web of Science and Scopus use the most trusted and well-recognized literature repositories, including ACM Digital Library and IEEE. Because no previous literature review exists with regards to the usage of eye-tracking techniques for search engines, the quality of the search string used for finding papers cannot be evaluated. Although, to reduce the possibility of missing a relevant paper, reference analysis was performed and snow-balling was applied to detect missing papers, some published papers may still have been missed in national journals and conferences. Thus, the results must be qualified as considering only papers that have been published in major international journals, conferences, and workshops in the areas of computer science and social science.

During the revision rounds of this review, a review of eye-tracking studies by Lewandowski and Kammerer (2020) was recently published with a focus on factors influencing viewing behavior on search engine results pages [84].

Appendix A

Study Reference	N	Search Engine	Participants	Device	Interface	Results	Measure	Scenario	Tasks	Language	Presentation
Granka et al. (2004) [28]	26	Google	Students	ASL 504	Desktop	Organic	Temporal, count	Regular	Informational	English	Table
Klöckner et al. (2004) [72]	41	Google	Random	ASL 504	Desktop	Organic	Temporal	Modified	Informational	German	n/a
Pan et al. (2004) [33]	30	Google, Yahoo	Students	ASL 504	Desktop	Organic	Temporal	Regular	At will	English	Table
Aula et al. (2005) [73]	28	Google	Students	ASL 504	Desktop	Organic	Temporal	Regular	Informational, navigational	English	AOI
Joachims et al. (2005) [29]	29	Google	Students	ASL 504	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Table
Radlinski and Joachims (2005) [35]	36	Google	Students	ASL 504	Desktop	Organic	Temporal	Regular	At will	English	Table
Rele and Duchowski (2005) [62]	16	Google	Random	Tobii 1750	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Heat map
Lorigo et al. (2006) [31]	23	Google	Students	ASL 504	Desktop	Organic	Temporal, spatial, count	Regular	Informational, navigational	English	Gaze plot
Guan and Cutrell (2007) [37]	18	MSN	Random	Tobii x50	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Table
Joachims et al. (2007) [30]	29	Google	Students	ASL 504	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Chart
Pan et al. (2007) [34]	22	Google	Students	ASL 504	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Table
Cutrell and Guan (2007) [38]	18	MSN	Random	Tobii x50	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Heat map
Lorigo et al. (2008) [32]	40	Google, Yahoo	Students	ASL 504	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	English	Chart
Egusa et al. (2008) [70]	16	Yahoo	Students	n/a	Desktop	Organic	Temporal, count	Regular	Informational	Japanese	Gaze plot
Xu et al. (2008) [7]	5	Google	n/a	web camera/ OpenGazer	Desktop	Organic	Spatial	Modified	Informational	English	Chart
Buscher et al. (2009) [36]	32	Live	Students	Tobii 1750	Desktop	Organic	Temporal, count	Modified	Informational	German	Chart
Kammerer et al. (2009) [66]	30	Google	Students	Tobii 1750	Desktop	Organic	Temporal, count	Modified	Informational	German	Chart
Buscher et al. (2010) [41]	38	Bing	Random	Tobii x50	Desktop	Paid	Temporal, count	Modified	Informational, navigational	English	Chart
Dumais et al. (2010) [11]	38	Bing	Random	Tobii x50	Desktop	Organic, paid	Temporal, spatial, count	Modified	Informational, navigational	English	Heat map
Marcos and González-Caro (2010) [60]	58	Google, Yahoo	Random	Tobii T120	Desktop	Organic, paid	Temporal, spatial, count	Regular	Informational, navigational, transactional, multimedia	Spanish	Heat map

Study Reference	N	Search Engine	Participants	Device	Interface	Results	Measure	Scenario	Tasks	Language	Presentation
Gerjets et al. (2011) [74]	30	Google	Students	Tobii 1750	Desktop	Organic	Temporal, count	Modified	Informational	German	Table
González-Caro and Marcos (2011) [61]	58	Google, Yahoo	Random	Tobii T120	Desktop	Organic, paid	Temporal, spatial, count	Regular	Informational, navigational, transactional	Spanish	Heat map
Huang et al. (2011) [40]	36	Bing	Random	Tobii x50	Desktop	Organic, paid	Temporal, spatial, count	Regular	Informational, navigational	English	Heat map
Balatsoukas and Ruthven (2012) [75]	24	Google	Students	Tobii T60	Desktop	Organic	Temporal, count	Regular	Informational	English	Table
Huang et al. (2012) [39]	36	Bing	Random	Tobii x50	Desktop	Organic, paid	Temporal, count	Regular	Informational, navigational	English	Chart
Kammerer and Gerjets (2012) [64]	58	Google	Students	Tobii 1750	Desktop	Organic	Temporal	Modified	Informational	German	Chart
Kim et al. (2012) [49]	32	Google	Students	Facelab 5	Mobile	Organic	Temporal	Modified	Informational, navigational	English	AOI
Marcos et al. (2012) [76]	57	Google	Students	Tobii 1750	Desktop	Organic	Temporal, count	Regular	Informational	Spanish	Table
Nettleton and González-Caro (2012) [77]	57	Google	Students	Tobii 1750	Desktop	Organic	Temporal, spatial, count	Regular	Informational	Spanish	Heat map
Djamasbi et al. (2013a) [42]	11	Google	Students	Tobii X120	Desktop	Organic	Temporal, spatial, count	Regular	Informational	English	AOI
Djamasbi et al. (2013b) [43]	16	Google	Students	Tobii X120	Desktop/ mobile	Paid	Temporal, spatial, count	Regular	Informational	English	Heat map
Djamasbi et al. (2013c) [44]	34	Google	Students	Tobii X120	Desktop/ mobile	Paid	Temporal, spatial, count	Regular	Informational	English	Heat map
Hall-Phillips et al. (2013) [45]	18	Google	Students	Tobii X120	Desktop	Paid	Temporal, spatial, count	Regular	Informational	English	AOI
Bataineh and Al-Bataineh (2014) [55]	25	Google	Students	Tobii T120	Desktop	Organic	Temporal, spatial, count	Modified	Informational	English	Heat map
Dickerhoof and Smith (2014) [78]	18	Google	Students	Tobii T60	Desktop	Organic	Temporal, spatial, count	Regular	Informational	English	Table
Gossen et al. (2014) [53]	31	Google and Blinde-Kuh	Children, adults	Tobii T60	Desktop	Organic	Temporal, spatial, count	Regular	Informational, navigational	German	Heat map
Hofmann et al. (2014) [67]	25	Bing	Random	Tobii TX-300	Desktop	Organic	Temporal, spatial, count	Regular	Informational, navigational	English	Heat map
Jiang et al. (2014) [69]	20	Google	Students	Tobii 1750	Desktop	Organic	Temporal, count	Modified	Informational, transactional	English	Chart
Lagun et al. (2014) [57]	24	Google	Random	Tobii X60	Mobile	Knowledge graph	Temporal, spatial	Regular	Informational	English	Heat map
Y. Liu et al. (2014) [6]	37	Sogou	Students	Tobii X2-30	Desktop	Organic	Temporal, count	Regular	Informational, navigational	Chinese	Chart
Z. Liu et al. (2014) [46]	30	Baidu	Students	Tobii X2-30	Desktop	Paid	Temporal, count	Regular	Informational, navigational	Chinese	Chart
Lu and Jia (2014) [63]	58	Baidu	Students	Tobii T120	Desktop	Image	Temporal, spatial, count	Regular	Informational	Chinese	Chart

Study Reference	N	Search Engine	Participants	Device	Interface	Results	Measure	Scenario	Tasks	Language	Presentation
Mao et al. (2014) [48]	31	Baidu	Students	Tobii X2-30	Desktop	Organic	Temporal, count	Regular	Informational, navigational, transactional	Chinese	Table
Kim et al. (2015) [50]	35	Google	Students	Facelab 5	Desktop/ mobile	Organic	Temporal, count	Modified	Informational, navigational	English	Table
Z. Liu et al. (2015) [47]	32	Sogou	Students	Tobii X2-30	Desktop	Organic	Temporal, spatial, count	Modified	Informational	Chinese	Heat map
Bilal and Gwizdka (2016) [54]	21	Google	Children	Tobii X2-60	Desktop	Organic	Temporal, count	Modified	Informational	English	Table
Domachowski et al. (2016) [59]	20	Google	Random	Tobii X2-60	Mobile	Paid	Temporal, spatial, count	Modified	Informational, transactional	German	Heat map
Kim et al. (2016a) [56]	18	Google	Students	Eye Tribe	Mobile	Organic	Temporal, spatial, count	Modified	Informational	English	AOI
Kim et al. (2016b) [51]	24	Google	Students	Facelab 5	Mobile	Organic	Temporal, spatial, count	Modified	Informational	English	Chart
Lagun et al. (2016) [58]	24	Google	Random	SMI Glasses	Mobile	Paid	Temporal, spatial, count	Modified	Transactional	English	Heat map
Kim et al. (2017) [52]	24	Google	Students	Facelab 5	Mobile	Organic	Temporal, count	Modified	Informational, navigational	English	Chart
Papoutsaki et al. (2017) [8]	36	Bing and Google	Random	SearchGazer	Desktop	Organic, paid	Temporal, count	Modified	Informational, navigational	English	Heat map
Bhattacharya and Gwizdka (2018) [79]	26	Google	Students	Tobii TX300	Desktop	Organic	Temporal, count	Regular	Informational	English	Table
Hautala et al. (2018) [10]	36	Google	Children	EyeLink 1000	Desktop	Organic	Temporal, count	Modified	Informational	Finnish	Chart
Schultheiss et al. (2018) [65]	25	Google	Students	Tobii T60	Desktop	Organic	Temporal, spatial, count	Modified	Informational, navigational	German	Chart
Sachse (2019) [80]	31	Google	Random	Pupil ET	Mobile	Organic	Count	Modified	Informational, navigational	German	Chart

References

1. Goldberg, J.H.; Wichansky, A.M. Eye tracking in usability evaluation: A practitioner's guide. In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*; Hyönä, J., Radach, R., Deubelk, H., Eds.; Elsevier: New York, NY, USA, 2003; pp. 493–516. ISBN 9780444510204.

2. Goldberg, J.H.; Stimson, M.J.; Lewenstein, M.; Scott, N.; Wichansky, A.M. Eye tracking in web search tasks. In *Proceedings of the Symposium on Eye Tracking Research & Applications—ETRA '02*; ACM: New York, NY, USA, 2002; pp. 51–58.

3. Duchowski, A.T. *Eye Tracking Methodology*; Springer: Champangne, IL, USA, 2017; ISBN 978-3-319-57881-1.

4. Wedel, M.; Pieters, R. A review of eye-tracking research in marketing. In *Review of Marketing Research*; Malhotra, N.K., Ed.; Emerald Group Publishing Limited: Bingley, UK, 2008; pp. 123–147.

5. Granka, L.; Feusner, M.; Lorigo, L. Eye monitoring in online search. In *Passive Eye Monitoring: Signals and Communication Technologies*; Hammoud, R., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 347–372.

6. Liu, Y.; Wang, C.; Zhou, K.; Nie, J.; Zhang, M.; Ma, S. From skimming to reading. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management—CIKM '14*; ACM: New York, NY, USA, 2014; pp. 849–858.

7. Xu, S.; Lau, F.C.M.M.; Jiang, H.; Lau, F.C.M.M. Personalized online document, image and video recommendation via commodity eye-tracking categories and subject descriptors. In *Proceedings of the 2008 ACM Conference on Recommender Systems*; ACM: New York, NY, USA, 2008; pp. 83–90.

8.   Papoutsaki, A.; Laskey, J.; Huang, J. SearchGazer: Webcam eye tracking for remote studies of web search. In *Proceedings of the 2nd ACM SIGIR Conference on Information Interaction and Retrieval, CHIIR 2017*; ACM: New York, NY, USA, 2017; pp. 17–26.

9.   Wade, N.J.; Tatler, B.W. Did Javal measure eye movements during reading. *J. Eye Mov. Res.* **2009**, *2*, 1–7. [CrossRef]

10.   Hautala, J.; Kiili, C.; Kammerer, Y.; Loberg, O.; Hokkanen, S.; Leppänen, P.H.T. Sixth graders' evaluation strategies when reading Internet search results: An eye-tracking study. *Behav. Inf. Technol.* **2018**, *37*, 761–773. [CrossRef]

11.   Dumais, S.; Buscher, G.; Cutrell, E. Individual differences in gaze patterns for web search. In *Proceedings of the Third Symposium on Information Interaction in Context—IIiX '10*; ACM: New York, NY, USA, 2010; pp. 185–194.

12.   Alemdag, E.; Cagiltay, K. A systematic review of eye tracking research on multimedia learning. *Comput. Educ.* **2018**, *125*, 413–428. [CrossRef]

13.   Lai, M.-L.; Tsai, M.-J.; Yang, F.-Y.; Hsu, C.-Y.; Liu, T.-C.; Lee, S.W.-Y.; Lee, M.-H.; Chiou, G.-L.; Liang, J.-C.; Tsai, C.-C. A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educ. Res. Rev.* **2013**, *10*, 90–115. [CrossRef]

14.   Leggette, H.; Rice, A.; Carraway, C.; Baker, M.; Conner, N. Applying eye-tracking research in education and communication to agricultural education and communication: A review of literature. *J. Agric. Educ.* **2018**, *59*, 79–108. [CrossRef]

15.   Rosch, J.L.; Vogel-Walcutt, J.J. A review of eye-tracking applications as tools for training. *Cogn. Technol. Work* **2013**, *15*, 313–327. [CrossRef]

16.   Yang, F.-Y.; Tsai, M.-J.; Chiou, G.-L.; Lee, S.W.-Y.; Chang, C.-C.; Chen, L.L. Instructional suggestions supporting science learning in digital environments. *J. Educ. Technol. Soc.* **2018**, *21*, 22–45.

17.   Meissner, M.; Oll, J. The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues. *Organ. Res. Methods* **2019**, *22*, 590–617. [CrossRef]

18.   Lund, H. Eye tracking in library and information science: A literature review. *Libr. Hi Tech* **2016**, *34*, 585–614. [CrossRef]

19.   Sharafi, Z.; Soh, Z.; Guéhéneuc, Y.-G. A systematic literature review on the usage of eye-tracking in software engineering. *Inf. Softw. Technol.* **2015**, *67*, 79–107. [CrossRef]

20.   Rayner, K. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol. Bull.* **1998**, *124*, 372–422. [CrossRef] [PubMed]

21.   Liversedge, S.P.; Paterson, K.B.; Pickering, M.J. Eye Movements and Measures of Reading Time. *Eye Guid. Read. Scene Percept.* **1998**, 55–75. [CrossRef]

22.   Liversedge, S.P.; Findlay, J.M. Saccadic eye movements and cognition. *Trends Cogn. Sci.* **2000**, *4*, 6–14. [CrossRef]

23.   Pernice, K.; Nielsen, J. *How to Conduct and Evaluate Usability Studies Ssing Eyetracking*; Nielsen Norman Group: Fremont, CA, USA, 2009.

24.   Moher, D. Preferred reporting Items for systematic reviews and meta-analyses: The PRISMA statement. *Ann. Intern. Med.* **2009**, *151*, 264–267. [CrossRef] [PubMed]

25.   Webster, J.; Watson, R.T. Analyzing the past to prepare for the future: Writing a literature review. *MIS Q.* **2002**, *26*, xiii.

26.   Lewandowski, D. The retrieval effectiveness of web search engines: Considering results descriptions. *J. Doc.* **2008**, *64*, 915–937. [CrossRef]

27.   Jansen, B.J.; Mullen, T. Sponsored search: An overview of the concept, history, and technology. *Int. J. Electron. Bus.* **2008**, *6*, 114–171. [CrossRef]

28.   Granka, L.A.; Joachims, T.; Gay, G. 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*; ACM: New York, NY, USA, 2004; pp. 478–479.

29.   Joachims, T.; Granka, L.; Pan, B.; Hembrooke, H.; Gay, G. Accurately interpreting clickthrough data as implicit feedback. *ACM SIGIR Forum.* **2005**, *51*, 154–161.

30.   Joachims, T.; Granka, L.; Pan, B.; Hembrooke, H.; Radlinski, F.; Gay, G. Evaluating the accuracy of implicit feedback from clicks and query reformulations in Web search. *ACM Trans. Inf. Syst.* **2007**, *25*, 7. [CrossRef]

31.   Lorigo, L.; Pan, B.; Hembrooke, H.; Joachims, T.; Granka, L.; Gay, G. The influence of task and gender on search and evaluation behavior using Google. *Inf. Process. Manag.* **2006**, *42*, 1123–1131. [CrossRef]

32. Lorigo, L.; Haridasan, M.; Brynjarsdóttir, H.; Xia, L.; Joachims, T.; Gay, G.; Granka, L.; Pellacini, F.; Pan, B. Eye tracking and online search: Lessons learned and challenges ahead. *J. Am. Soc. Inf. Sci. Technol.* **2008**, *59*, 1041–1052. [CrossRef]

33. Pan, B.; Hembrooke, H.A.; Gay, G.K.; Granka, L.A.; Feusner, M.K.; Newman, J.K. The determinants of web page viewing behavior. In *Proceedings of the Eye Tracking Research & Applications—ETRA'2004*; ACM: New York, NY, USA, 2004; Volume 1, pp. 147–154.

34. Pan, B.; Hembrooke, H.; Joachims, T.; Lorigo, L.; Gay, G.; Granka, L. In Google we trust: Users' decisions on rank, position, and relevance. *J. Comput. Commun.* **2007**, *12*, 801–823. [CrossRef]

35. Radlinski, F.; Joachims, T. Query chains: Learning to rank from implicit feedback. In *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*; ACM: New York, NY, USA, 2005; pp. 239–248.

36. Buscher, G.; van Elst, L.; Dengel, A. Segment-level display time as implicit feedback. In *Proceedings of the 32nd international ACM SIGIR Conference on Research and Development in Information Retrieval—SIGIR '09*; ACM: New York, NY, USA, 2009; pp. 67–74.

37. Guan, Z.; Cutrell, E. 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—CHI '07*; ACM: New York, NY, USA, 2007; pp. 417–420.

38. Cutrell, E.; Guan, Z. What are you looking for? An eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '07*; ACM: New York, NY, USA, 2007; pp. 407–416.

39. Huang, J.; White, R.W.; Buscher, G. User see, user point. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems—CHI '12*; ACM: New York, NY, USA, 2012; pp. 1341–1350.

40. Huang, J.; White, R.W.; Dumais, S. No clicks, no problem: Using cursor movements to understand and improve search. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI '11*; ACM: New York, NY, USA, 2011; pp. 1225–1234.

41. Buscher, G.; Dumais, S.; Cutrell, E. The good, the bad, and the random: An eye-tracking study of ad quality in web search. In *Proceedings of 33rd International ACM SIGIR*; ACM: New York, NY, USA, 2010; pp. 42–49.

42. Djamasbi, S.; Hall-Phillips, A.; Yang, R. Search results pages and competition for attention theory: An exploratory eye-tracking study. In *Human Interface and the Management of Information. Information and Interaction Design*; Yamamoto, S., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 576–583.

43. Djamasbi, S.; Hall-Phillips, A.; Yang, R. SERPs and ads on mobile devices: An eye tracking study for generation Y. In *Universal Access in Human-Computer Interaction. User and Context Diversity*; Stephanidis, C., Antona, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 259–268.

44. Djamasbi, S.; Hall-Phillips, A.; Yang, R. An examination of ads and viewing behavior: An eye tracking study on desktop and mobile devices. In Proceedings of the AMCIS 2013, Chicago, IL, USA, 15–17 August 2013.

45. Hall-Phillips, A.; Yang, R.; Djamasbi, S. Do ads matter? An exploration of web search behavior, visual hierarchy, and search engine results pages. In *Proceedings of the 2013 46th Hawaii International Conference on System Sciences*; IEEE: Piscatway, NJ, USA, 2013; pp. 1563–1568.

46. Liu, Z.; Liu, Y.; Zhang, M.; Ma, S. How do sponsored search. In *Information Retrieval Technology*; Jaafar, A., Ali, N.M., Noah, S.A.M., Smeaton, A.F., Bruza, P., Bakar, Z.A., Jamil, N., Sembok, T.M.T., Eds.; Springer: Champagne, IL, USA, 2014; pp. 73–85.

47. Liu, Z.; Liu, Y.; Zhou, K.; Zhang, M.; Ma, S. Influence of vertical result in web search examination. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval—SIGIR '15*; ACM: New York, NY, USA, 2015; pp. 193–202.

48. Mao, J.; Liu, Y.; Zhang, M.; Ma, S. Estimating credibility of user clicks with mouse movement and eye-tracking information. In *Natural Language Processing and Chinese Computing*; Zong, C., Nie, J.-Y., Zhao, D., Feng, Y., Eds.; Springer: Berlin, Germany, 2014; pp. 263–274.

49. Kim, J.; Thomas, P.; Sankaranarayana, R.; Gedeon, T. Comparing scanning behaviour in web search on small and large screens. In *Proceedings of the Seventeenth Australasian Document Computing Symposium on—ADCS '12*; ACM: New York, NY, USA, 2012; pp. 25–30.

50. Kim, J.; Thomas, P.; Sankaranarayana, R.; Gedeon, T.; Yoon, H.-J. Eye-tracking analysis of user behavior and performance in web search on large and small screens. *J. Assoc. Inf. Sci. Technol.* **2015**, *66*, 526–544. [CrossRef]

51. Kim, J.; Thomas, P.; Sankaranarayana, R.; Gedeon, T.; Yoon, H.-J. Pagination versus scrolling in mobile web search. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management—CIKM '16*; ACM: New York, NY, USA, 2016; pp. 751–760.

52. Kim, J.; Thomas, P.; Sankaranarayana, R.; Gedeon, T.; Yoon, H.-J. What snippet size is needed in mobile web search? In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval—CHIIR '17*; ACM: New York, NY, USA, 2017; pp. 97–106.

53. Gossen, T.; Höbel, J.; Nürnberger, A. Usability and perception of young users and adults on targeted web search engines. In *Proceedings of the 5th Information Interaction in Context Symposium on—IIiX '14*; ACM: New York, NY, USA, 2014; pp. 18–27.

54. Bilal, D.; Gwizdka, J. Children's eye-fixations on google search results. *Assoc. Inf. Sci. Technol.* **2016**, *53*, 1–6. [CrossRef]

55. Bataineh, E.; Al-Bataineh, B. An analysis study on how female college students view the web search results using eye tracking methodology. In Proceedings of the International Conference on Human-Computer Interaction, Crete, Greece, 22–27 June 2014; p. 93.

56. Kim, J.; Thomas, P.; Sankaranarayana, R.; Gedeon, T.; Yoon, H.-J. Understanding eye movements on mobile devices for better presentation of search results. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 2607–2619. [CrossRef]

57. Lagun, D.; Hsieh, C.-H.; Webster, D.; Navalpakkam, V. 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—SIGIR '14*; ACM: New York, NY, USA, 2014; pp. 113–122.

58. Lagun, D.; McMahon, D.; Navalpakkam, V. Understanding mobile searcher attention with rich ad formats. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management—CIKM '16*; ACM: New York, NY, USA, 2016; pp. 599–608.

59. Domachowski, A.; Griesbaum, J.; Heuwing, B. Perception and effectiveness of search advertising on smartphones. *Assoc. Inf. Sci. Technol.* **2016**, *53*, 1–10. [CrossRef]

60. Marcos, M.-C.; González-Caro, C. Comportamiento de los usuarios en la página de resultados de los buscadores. Un estudio basado en eye tracking. *El Prof. Inf.* **2010**, *19*, 348–358. [CrossRef]

61. González-Caro, C.; Marcos, M.-C. Different users and intents: An Eye-tracking analysis of web search. In Proceedings of the 4th ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011.

62. Rele, R.S.; Duchowski, A.T. Using eyetracking to evaluate alternative search results interfaces. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2005**, *49*, 1459–1463. [CrossRef]

63. Lu, W.; Jia, Y. An eye-tracking study of user behavior in web image search. In *PRICAI 2014: Trends in Artificial Intelligence*; Pham, D.-N., Park, S.-B., Eds.; Springer: Cham, Germany, 2014; pp. 170–182.

64. Kammerer, Y.; Gerjets, P. Effects of search interface and Internet-specific epistemic beliefs on source evaluations during Web search for medical information: An eye-tracking study. *Behav. Inf. Technol.* **2012**, *31*, 83–97. [CrossRef]

65. Schultheiss, S.; Sünkler, S.; Lewandowski, D. We still trust in Google, but less than 10 years ago: An eye-tracking study. *Inf. Res. An Int. Electron. J.* **2018**, *23*, 799.

66. Kammerer, Y.; Wollny, E.; Gerjets, P.; Scheiter, K. How authority related epistemological beliefs and salience of source information influence the evaluation of web search results—An eye tracking study. In Proceedings of the 31st Annual Conference of the Cognitive Science Society (CogSci), Amsterdam, The Netherlands, 29 July–1 August 2009; pp. 2158–2163.

67. Hofmann, K.; Mitra, B.; Radlinski, F.; Shokouhi, M. An eye-tracking study of user interactions with query auto completion. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management—CIKM '14*; ACM: New York, NY, USA, 2014; pp. 549–558.

68. Broder, A. A taxonomy of web search. *ACM SIGIR Forum.* **2002**, *36*, 3–10. [CrossRef]

69. Jiang, J.; He, D.; Allan, J. Searching, browsing, and clicking in a search session. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval—SIGIR '14*; ACM: New York, NY, USA, 2014; pp. 607–616.

70. Egusa, Y.; Takaku, M.; Terai, H.; Saito, H.; Kando, N.; Miwa, M. Visualization of user eye movements for search result pages. In Proceedings of the EVIA. 2008 (NTCIR-7 Pre-Meeting Workshop), Tokyo, Japan, 16–19 December 2008; pp. 42–46.

71. Guy, I. Searching by talking: Analysis of voice queries on mobile web search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval—SIGIR '16*; ACM: New York, NY, USA, 2016; pp. 35–44.

72. Klöckner, K.; Wirschum, N.; Jameson, A. Depth- and breadth-first processing of search result lists. In *Proceedings of the Extended Abstracts of the 2004 Conference on Human Factors and Computing Systems—CHI '04*; ACM: New York, NY, USA, 2004; p. 1539.

73. Aula, A.; Majaranta, P.; Räihä, K.-J. Eye-tracking reveals the personal styles for search result evaluation. *Lect. Notes Comput. Sci.* **2005**, *3585*, 1058–1061.

74. Gerjets, P.; Kammerer, Y.; Werner, B. Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learn. Instr.* **2011**, *21*, 220–231. [CrossRef]

75. Balatsoukas, P.; Ruthven, I. An eye-tracking approach to the analysis of relevance judgments on the Web: The case of Google search engine. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 1728–1746. [CrossRef]

76. Marcos, M.-C.; Nettleton, D.; Sáez-Trumper, D. A user study of web search session behaviour using eye tracking data. In *Proceedings of the BCS HCI 2012- People and Computers XXVI*; ACM: New York, NY, USA, 2012; Volume XXVI, pp. 262–267.

77. Nettleton, D.; González-Caro, C. Analysis of user behavior for web search success using eye tracker data. In *Proceedings of the 2012 Eighth Latin American Web Congress*; IEEE: Piscatway, NJ, USA, 2012; pp. 57–63.

78. Dickerhoof, A.; Smith, C.L. Looking for query terms on search engine results pages. In Proceedings of the ASIST Annual Meeting, Seattle, WA, USA, 31 October–4 November 2014; Volume 51.

79. Bhattacharya, N.; Gwizdka, J. Relating eye-tracking measures with changes in knowledge on search tasks. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications—ETRA '18*; ACM: New York, NY, USA, 2018; pp. 1–5.

80. Sachse, J. The influence of snippet length on user behavior in mobile web search. *Aslib J. Inf. Manag.* **2019**, *71*, 325–343. [CrossRef]

81. Cetindil, I.; Esmaelnezhad, J.; Li, C.; Newman, D. Analysis of instant search query logs. In Proceedings of the Fifteenth International Workshop on the Web and Databases (WebDB 2012), Scottsdale, AZ, USA, 20 May 2012; pp. 1–20.

82. Mustafaraj, E.; Metaxas, P.T. From obscurity to prominence in minutes: Political speech and real-time search. In Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, Raleigh, NC, USA, 26–27 April 2020.

83. Bilal, D.; Huang, L.-M. Readability and word complexity of SERPs snippets and web pages on children's search queries. *Aslib J. Inf. Manag.* **2019**, *71*, 241–259. [CrossRef]

84. Lewandowski, D.; Kammerer, Y. Factors influencing viewing behaviour on search engine results pages: A review of eye-tracking research. *Behav. Inf. Technol.* **2020**, 1–31. [CrossRef]