How We Perceive Search Engines

Leonardo Penna[(✉)] and Manuela Quaresma

LEUI - Laboratory of Ergodesign and Usability of Interfaces,
PUC-Rio University, Rio de Janeiro, Brazil
`leo.mep@gmail.com`, `mquaresma@puc-rio.br`

Abstract. This article presents a literature review related to users' perceptions about search engines. Its motivation was establish an information source upon a topic that directly affects people's interactions with these tools and currently is scattered in the literature. It was discussed impact generated in users' behavior by the confidence degree in the companies producing search engines and by credit given to algorithms responsible for selection and ordering of results. It was also analyzed the public view about impartiality, accuracy and reliability of these tools.

Keywords: Search engines · Search · Perception · Users · Results · Ordering · Ranking

1 Introduction

The rapid growth of Internet in early years of the 90s served as a catalyst for the development of tools that aggregate content, allowing people to move in a more orderly way in the virtual space. The increase from 130 to 23.500 websites between 1993 and 1995 [1] attested to the Web expansion potential and showed that the manual recovery of information would become unviable in a short time. The search engines have become the most successful response to this demand for support systems, by allowing users to find documents related to an interest group of keywords. Currently, they are the basis for experimentation and transit at virtual spaces, playing an equivalent role of an expert [2], which makes content indications that are relevant to subjects' questions. Google, for example, became the main starting point for students, both for searches of everyday life, as for academic researches [3].

Inserted in this context, this article presents a literature review related to users' perceptions about search engines. Its motivation was establish an information source upon a topic that directly affects people's interactions with these tools and currently is scattered in the literature. Using Google Scholar as search system of academic articles, it was sought in November 2014 publications with the term "search engines" or "google", associated with one or more of the following keywords: "users", "evaluation", "assess", "perceive", "perception" and "interpret". After review of the abstracts of the documents found, it was defined a first set of items relevant to the study. The second stage of the literature review was to search for all articles quoting one of the selected publications. This process of selecting documents and subsequent evaluation of quotes was performed iteratively until there were no new items related to the subject

researched. Finally, it was performed a content analysis of the main issues addressed in the informational mass found.

2 Confidence at Systemic Ordering

The essence behind how the search engines work lies within the selection algorithm and the results ordering related to the searches done by the user. Through this operation set, invisible to those who use the tool, the relevance concept adopted by these systems come to life. "As there is no independent metric for what actually are the most relevant search results for any given query, engineers must decide what results look 'right' and tweak their algorithm to attain that result (…)" [4].

An aspect related to how we perceive the search engines is the belief we have in their proper functioning. Through a survey conducted by Purcell et al. [5], it was observed that most users understand these tools as fair and unbiased information sources, especially among young people. The prevailing view is that the informational clipping performed by search engines is reliable and accurate. Added to this, the vast people majority are satisfied with services offered and more than half of respondents believe that these systems are improving over time.

This positive view associated to the search tools becomes clear when we analyze the people behavior facing the computational processes responsible for results. Several studies show us that users trust the sorting done by the tools and are willing to click at the first results [6–12], in addition to possessing a distinct evaluative look for items in listing's lower positions [9, 10].

In an experiment conducted by Pan et al. [6], 16 individuals — graduating students which highly trusted Google (7.9/10) and had vast experience using the tool (10 out of 10) — completed ten tasks comprising navigational and informational searches, while their eyes were recorded by an eye tracking device. In order to understand the ordering influence at the decision-making process, the results were manipulated. Each individual was given one of the three possible scenarios: normal, where the results page was the same as Google's; swapped, where the first item changed places with the second one; reversed, where the results were changed to be displayed from last to first. Aiming to contribute with the interpretation of the output data during the experiment, the researchers asked five people to decide the relevance of the search results. All the items found in two of the tasks were randomly evaluated.

The data gathered indicates that the ordering has a strong influence over how the user interacts with the results page. In general, the individuals viewed the top position more frequently and clicked at the first result most of the times. When exposed to the scenario where the first two results were swapped, the click count for the first item was, nonetheless, three times higher than the second one. Besides that, the top item from Google's ranking — shown as first place for the normal scenario, as second for the swapped scenario and as tenth for the reverse scenario — had a decreasing eye fixations number the lower it was displayed.

Despite the overwhelming trust in the results sorting, according to the authors, the consumption of the content presented by the search tool is not passive. When exposed to the reversed scenario, the users took longer to interact with the pages (10.9 s for

reversed, 6.5 s for normal, 5.8 s for swapped, $p < .05$), and the number of clicks was significantly lower compared to the other scenarios. It was also possible to notice an increased eye fixations number (30.0 reversed, 18.3 normal and 17.9 swapped, $p < .05$), a higher count of checked results (3.8 descriptions for reversed, 2.5 for normal and 2.7 for swapped, $p < .05$), and a larger quantity of description reevaluation (3.4 reversed, 2.2 normal and 1.9 swapped, $p < .05$). The interaction with results pages sorted inconsistently with the items' inherent relevance led to a higher awareness and caution by the individuals.

Statically foreseeing the different variables influence over the number of clicks among the viewed results, the list position had a slightly higher performance than the item's intrinsic relevance.

To assess whether data obtained in study performed by Pan et al. [6] would extend to other search engines or were exclusively related to Google, Lorigo et al. [7] conducted an experiment with very similar format to that used previously, replacing only the analysis target by Yahoo's tool. The reported results showed that both tools' behaviors is very similar, occurring a persistence of dependency on the result's ordering.

This relationship of trust between the users and the search engines can also be observed in a study by Balatsoukas and Ruthven [10], in which 24 college students (17–36 years of age, frequent or very frequent search engines' users) sought to satisfy a real informational need. Free to use any resource available on Internet, students had as restriction only a maximum time of 25 min to complete task. Their statements were analyzed, eye movements data were recorded (number of fixations and types of visualized components - titles, abstracts, url, etc.), and the displayed results' evaluations are stored (relevant, partially relevant and irrelevant). To complement information obtained during tasks execution, it was also conducted an interview at activity end. The researchers found that all participants used Google and during their interaction with the tool, tended to fix eyes longer on items at top of first results page, especially in the two highest positions. The influence of search engine's ordering was evidenced by variation in the relevance criteria used for evaluating results. On the first two items 11 criteria were used (e.g., topicality, quality and recency), while for the result in the page's last position (tenth) only four parameters were considered. In some intermediate positions, result's relation with the topic ceased to be the most important element in assessment, giving way to criteria such as the information and source quality (see Fig. 1).

One hypothesis to this behavior is the need for an extra incentive which compensates the inferior value attributed to lower listing's positions. In other words, to click on results that are not at top, individuals needs to know the source of information, consider it a good reference or have other evidences about displayed item quality.

Importantly, despite the evident favoring of results that appear in search engines' top positions, there isn't necessarily an extension of this behavior to selection process of reliable and quality contents. Salmerón et al. [12] conducted an experiment with 67 college students (average age of 22.27 years and with extensive experience in internet usage), in which they needed to research the topic "Reduction of greenhouse gas emissions" and point out the two most relevant pages in terms of information. All
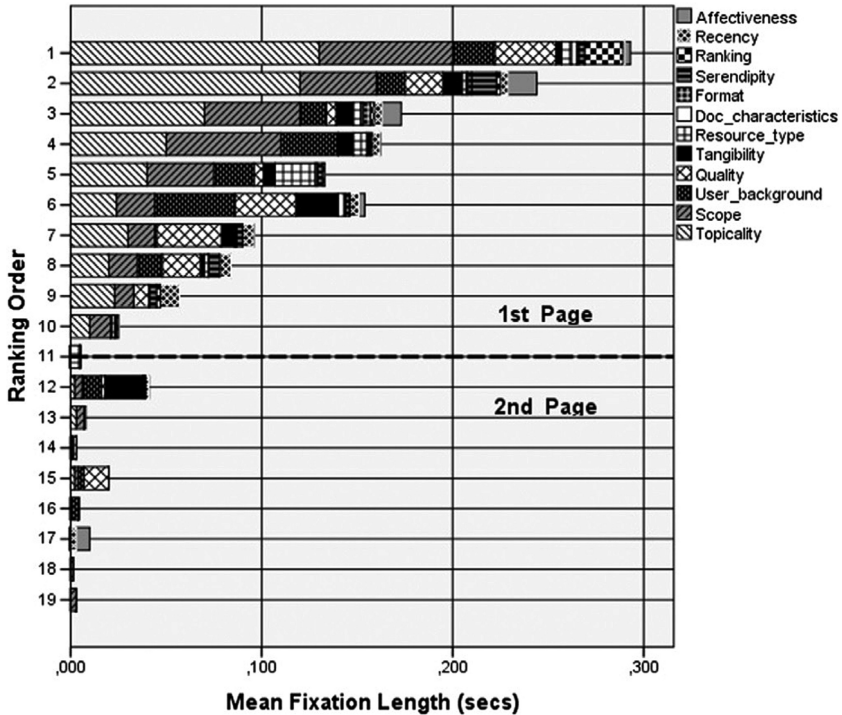
Fig. 1. Fixation in the selection criteria, according to the position (Source: [10], p. 1739)

participants received a list of 10 results that should be used in task. For half of them, items were showed in decreasing relevance order (default behavior at major search engines). The other half received an inverted list, that is, in ascending order. Although the subjects have clicked more over results at the top positions, the relevance and completeness assessment of content was similar in the two groups. Both chose an equivalent ratio of pages relevant/complete, relevant/incomplete and irrelevant.

3 Brands' Influence

As a comprehensive overview, we can classify brand as being the group of tangible representative elements — name, visual design, symbol — and the intangible components — values, concepts and personality — associated to an organization [13]. Its depiction is capable of inducting strong behavioral changes over consumers. Park, Harada, and Igarashi [14], for example, state that the brand perception of a product affected the user evaluation of the mental demand related to its use.

By directing attention to the search segment in internet environment, we can find data at literature that point out a strong relationship between brands and the information acquisition process. In an experiment conducted by Jansen, Zhang and Schultz [13]

with 32 individuals (ages 18–25), it was analyzed how brands affected the way users evaluate the results displayed by search engines. For such, the researchers simulated the search engines that were to be evaluated – Google, Yahoo! MSN Live Search, and AI^{2}RS (unknown to the public) –, deleting the results and leaving only the header and the footnote, areas where the organization's brand and visual identity are present. Subsequently, they added identical results to the modified versions for each company (see Fig. 2).

Each individual was exposed to four search scenarios, which had one of the altered results page as a starting point. The study was created assuring the participants would come into contact only once with each scenario and evaluated brand. During the activity, they were asked to communicate their actions and to evaluate the displayed links (3-point scale: irrelevant, somewhat relevant, relevant). If a link was not noticed, it was requested in a second step to the participant reopen the results page and evaluate the items which were not assessed during the task. Over the experiment, they also recorded the amount of viewed links, the number of clicked links, and the evaluation of the visited pages.

After analyzing the gathered data, the researchers identified two phenomena: (1) the brand perception affects the search engine's performance evaluation, even when the displayed results are not different; (2) the brand and its associated qualities change how the users analyze the results, affecting the number of viewed links, as well as the number of clicked items.
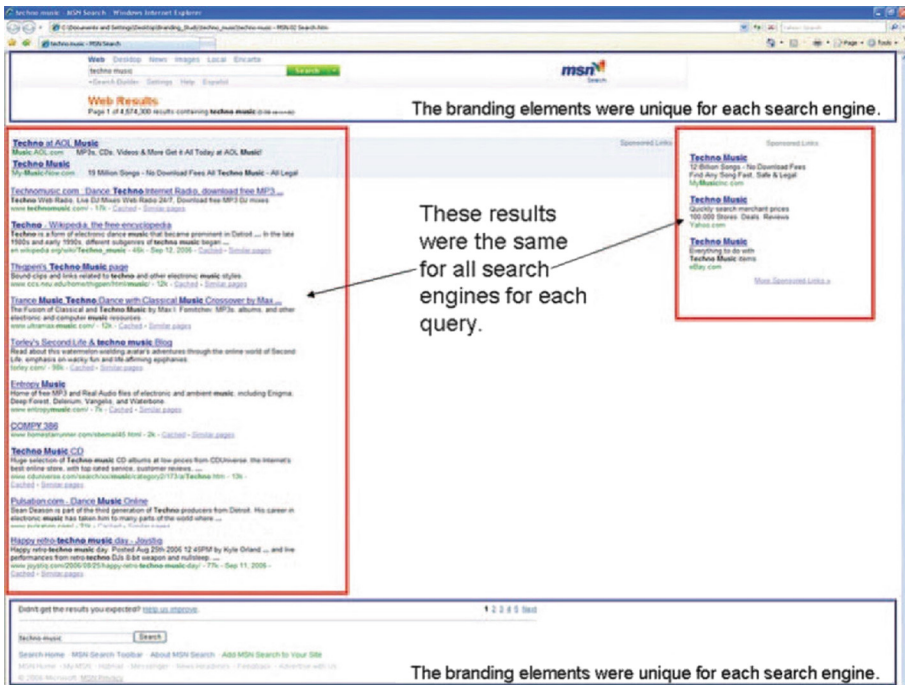

Fig. 2. Example of an experiment results page. (Source: [13], p. 1577)

3.1 Influence Over Performance Evaluation

Jansen et al. [13] gathered data showing that, although the results viewed by the individuals were constants, the link evaluation in each of the search engines presented significant differences. (Factorial ANOVA, p = 0.071). Judging the set of viewed links (organic and sponsored), Google had a higher rating than MSN Live Search, being considered 17 % more relevant than it competitor. Yahoo! and AI^2RS had 12 % higher evaluation than MSN Live Search. The fact that a tool without any public projection was superior than other belonging to a large company like Microsoft may indicate a negative brand view by consumers.

Although the study did not established which factors associated with brand are responsible for results evaluation differences, the relationship establishment between their findings and a later work by Jansen et al. [15] suggests that, since the tools have positive images, their performance will be better evaluated. This time, the researchers prepared a survey to understand the thoughts related to the brands of different search engines. With 207 respondents, the research asked participants to view 10 search engine logos – A9, AI^2RS (fake), Alltheweb, AOL Search, Ask.com, Dogpile, Google, MSN Live Search, Mahalo, and Yahoo! – and inform if they have ever used them, if they still used them, and what were their thoughts about them.

They verified that "Google has far and away the highest positive brand perception (…). Additionally, the depth of the positive sentiment is amazing (…). The term love was used by several participants to describe [it]" [15]. From all the participants, 87 % had a positive feedback, 12.6 % did not express any opinion, and only 0.5 % had a negative comments. There were no mixed or neutral perceptions about the tool. Only Yahoo! and Ask.com were also highly well rated, the former with 59.4 % of positive feedback and the latter with 52.7 %. MSN Live Search, despite being previously used by 54 % of the participants, got no good results. Microsoft's old engine received a number of bad reviews (27.1 %) higher than positive ones (22.2 %).

In the first experiment [13], evaluations of identical results by users was much more positive for Google and Yahoo!, especially when compared to MSN Live Search, which got a rating worse than the unknown search tool. In the 2012 research [15], similar results were found: Google and Yahoo! had a highly positive perception, whereas MSN Live Search got a very negative one. While it cannot establish a causal link between those two variables (i.e., a good perception implies a better feedback for the search engine), the data gathered in both studies seem compatible and encourage further investigation.

3.2 Influence Over How the Results are Analyzed

By statistically analyzing raw data about the links that were viewed and those clicked, Jansen et al. [13] were able to determine the interference exerted by changing brands of search engines. A significant difference was found regarding the number of viewed links (Factorial ANOVA, p = 0.022), with a prevalence of Yahoo! (40 % over AI^2RS) and Google (20 % over AI^2RS). "(…) When participants were viewing links, they favored the mainstream search engines (…) relative to the non-mainstream search engines. This may be because the participants were more trusting [in them]" [13]. In addition, was identified

a possible value difference related to each engine's results ordering. MSN Live Search and AI^{2}RS got a higher number of viewed links at the top positions, meaning that the users were cautious when they faced the displayed rankings from these tools.

As well as the link visualization, the search engine brands affected the number of clicked links (Factorial ANOVA, p = 0.045). The total from Yahoo! was 27 % higher than Google, 46 % superior than MSN, and 11 % larger than AI^{2}RS. When using the dominant tools, users seemed to delegate the task of finding relevant sites to system, relying little in the results evaluation and assigning high value to its positioning. This behavior leads to a higher number of less quality clicks. In less known search engines, the lack of trust makes the users more discerning about the results: they click less, choosing more relevant links.

4   Conclusions

The literature analysis related to uniqueness of our search engines perception showed that people have a wide confidence in systemic ordering. It was found that they consider these tools accurate and unbiased information sources [5] and they are predisposed to click in the first items of pages shown [6–12]. The ranking proposed by search engines exerts an influence on the links clicked slightly higher than result's relevance [6]. Furthermore, the number of criteria used in results evaluation at the top of page is greater than those considered for items in lower positions [9, 10].

It was also realized that individuals change their analysis and evaluation of results according to search tool, even if results are identical [13]. The judgment that they make about search engines performance, the number of links examined and the amount of clicks in results are dependent on their perception of search engines' brand [13]. In general, the dominants tools have higher rates for all mentioned variables. Apparently, in unfamiliar tools, the lack of trust decreases users' general interest in the results, but encourages a closer evaluative behavior in which users click less often and choose more relevant links [13].

It is necessary to point out that few academic publications mainly focuses on the topic covered by the article. Data relating to users' perception about search engines is often located in articles dealing with distinct topics, that make brief observations on the subject. Thus, the extensive sample space of available documents makes impossible a complete literature review and creates need for data selection to be explored. As a result, it was given an important step towards concatenation of scattered data in literature and it was obtained a research dealing with major issues related to subject, although it isn't a complete reflection of the existing production.

References

1. Gray, M.: Measuring the Growth of the Web (1995). Accessed from http://www.mit.edu/people/mkgray/growth/
2. Simpson, T.W.: Evaluating Google as an epistemic tool. Metaphilosophy **43**(4), 426–445 (2012). Accessed from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9973.2012.01759.x/full

3. Asher, A.D.: Search magic. discovering how undergraduates find information. In: American Anthropological Association, Annual Meeting (2011). Accessed from http://www.erialproject.org/wp-content/uploads/2011/11/Asher_AAA2011_Search-magic.pdf

4. Gillespie, T.: The relevance of algorithms. In: Gillespie, T., Boczkowski, P., Foot, K. (eds.) Media Technologies: Essays on Communication, Materiality, and Society, pp. 167–194. MIT Press, Cambridge (2014). Accessed from http://www.tarletongillespie.org/essays/Gillespie%20-%20The%20Relevance%20of%20Algorithms.pdf

5. Purcell, K., Brenner, J., Rainie, L.: Search engine use 2012 (2012). Accessed from http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf

6. 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. Med. Commun. 12, 801–823 (2007). doi:10.1111/j.1083-6101.2007.00351.x. Accessed from http://onlinelibrary.wiley.com/enhanced/doi/10.1111/j.1083-6101.2007.00351.x/

7. 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. 59, 1041–1052 (2008). doi:10.1002/asi.20794. Accessed from http://onlinelibrary.wiley.com/enhanced/doi/10.1002/asi.20794/

8. Bar-Ilan, J., Keenoy, K., Levene, M., Yaari, E.: Presentation bias is significant in determining user preference for search results—a user study. J. Am. Soc. Inf. Sci. 60, 135–149 (2009). doi:10.1002/asi.20941. Accessed from http://onlinelibrary.wiley.com/enhanced/doi/10.1002/asi.20941/

9. Balatsoukas, P. and Ruthven, I. (2010), The use of relevance criteria during predictive judgment: An eye tracking approach. Proc. Am. Soc. Info. Sci. Tech., 47: 1–10. doi:10.1002/meet.14504701145. Retrieved from http://onlinelibrary.wiley.com/enhanced/doi/10.1002/meet.14504701145/

10. 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. Tech. 63, 1728–1746 (2012). doi:10.1002/asi.22707. Accessed from http://onlinelibrary.wiley.com/enhanced/doi/10.1002/asi.22707/

11. Kammerer, Y., Gerjets, P.: How search engine users evaluate and select Web search results: the impact of the search engine interface on credibility assessments. Libr. Inf. Sci. 4, 251–279 (2012). Accessed from http://www.emeraldinsight.com/doi/abs/10.1108/S1876-0562%282012%29002012a012

12. Salmerón, L., Kammerer, Y., García-Carrión, P.: Searching the Web for conflicting topics: page and user factors. Comput. Hum. Behav. 29(6), 2161–2171 (2013). Accessed from http://www.sciencedirect.com/science/article/pii/S0747563213001465

13. Jansen, B.J., Zhang, M., Schultz, C.D.: Brand and its effect on user perception of search engine performance. J. Am. Soc. Inf. Sci. 60, 1572–1595 (2009). doi:10.1002/asi.21081. Accessed from http://onlinelibrary.wiley.com/enhanced/doi/10.1002/asi.21081/

14. Park, S., Harada, A., Igarashi, H.: Influences of personal preference on product usability. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 87–92. ACM, April 2006. Accessed from http://dl.acm.org/citation.cfm?id=1125475

15. Jansen, B.J., Zhang, L., Mattila, A.S.: User reactions to search engines logos: investigating brand knowledge of web search engines. Electron. Commer. Res. 12(4), 429–454 (2012). Accessed from http://link.springer.com/article/10.1007%2Fs10660-012-9101-0#