Are Ads on Google Search Engine Results Pages Labeled Clearly Enough?

The Influence of Knowledge on Search Ads on Users' Selection Behaviour

Dirk Lewandowski, Sebastian Sünkler, Friederike Kerkmann

Hamburg University of Applied Sciences, Germany
{dirk.lewandowki, sebastian.suenkler, friederike.kerkmann}@haw-hamburg.de

Abstract

In an online experiment using a representative sample of the German online population (n = 1,000), we compare users' selection behaviour on two versions of the same Google search engine results page (SERP), one showing advertisements and organic results, the other showing organic results only. Selection behaviour is analyzed in relation to users' knowledge on Google's business model, on SERP design, and on these users' actual performance in marking advertisements on SERPs correctly. We find that users who were not able to mark ads correctly selected ads significantly more often. This leads to the conclusion that ads need to be labeled more clearly, and that there is a need for more information literacy in search engine users.

Keywords: search engines; search engine results pages (SERPs); selection behaviour; representative user study; search engine advertising (SEA)

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th—15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 62—74.

1 Introduction

In the past years, we have been facing a discussion on the role search engines play for knowledge acquisition in society. The focus usually lies on Google, as it is by far the most dominant search engine in all countries except a few. Questions raised in this regard are, among others, what (types of) results users actually get to see when using Google (e.g., Noble, 2013; White & Horvitz, 2009), whether search engines are biased in general (Grimmelmann, 2010), and whether search engines in general and Google in particular should be made responsible for providing fair and unbiased results (Lewandowski, 2017).

Related to these questions, it is important to know how users actually interact with results presented by the search engines, first and foremost which results they select. The objective of the research presented in this paper is to find out whether users' knowledge on Google's business affects their clicking behaviour on the search engine results pages (SERPs). If so, i.e., if users who do not know that Google makes its revenue through advertisements on its results pages do select advertisements more frequently, then this would call for action either related to increasing users' information literacy and/or related to regulation.

To get reliable data and results regarding this, we conducted a large-scale online experiment, using a representative sample of the German online population.

The rest of this paper is structured as follows: In the next section, we give a brief overview on the elements presented on search engine results pages (SERPs), on selection behaviour on the SERPs, and on search engines' ad labeling practices. Then, we present our research questions and methods. After that, results are presented and discussed in a separate section each. We conclude with implications of our study and some suggestions for further research.

2 Literature review

In the following short literature review, we focus on the composition of search engine results pages, users' selection behaviour on these pages, and on ads labeling.

2.1 Search engine results pages

Results presentation on search engine results pages has changed in recent years. The simplest model of a search engine results page is a ranked list of document representations ("snippets") provided in response to a query. However, once search engines started displaying ads on SERPs, there were actually two ranked lists: the list of "organic results" and the list of advertisements. By adding results from vertical search engines such as news or video items and integrating them into the SERPs (known as Universal Search, cf. Taylor, Mayer & Buyukkokten, 2008), search engines moved away from plain ranked results lists and on to a richer presentation both of individual results as well as certain results types. Current search engine results pages go even further to additionally display factual information in addition to snippets. Factual information is shown in what are known as Knowledge Graph results (Drumond Monteiro & Aparecida Moura, 2014), satisfying at least some information needs directly on the results pages and representing a departure from the concept of a search engine being a tool for sending traffic (i.e., users) to external web pages.

For the purposes of the current study, we distinguish between the two results types *organic results* and *advertisements*. Other results types are omitted from the stimulus material.

2.2 Results selection

When looking at how users select results from a SERP, they are influenced mainly by the following factors:

1. Results position and reading behaviour: Users tend to click on results at or near the top of a results list (Joachims et al., 2005). For instance, a 2014 study based on 465,000 queries found that more than two thirds of all clicks go to the first five positions, and the result ranked first alone accounts for 31% of all clicks (Petrescu, 2014). Goel, Broder, Gabri-

lovich and Pang (2010) found that within Yahoo search, only 10,000 websites account for approximately 80% of results clicks. This clicking behaviour is due to the fact that, usually, lists of results are read from top to bottom.

2. Search engine relevance ranking algorithms are precision-based, i.e., they focus on presenting a few relevant results in the first several positions. Users have adapted to this kind of results ranking and therefore in most cases only consider the first few results (ibid.).
3. Due to screen resolutions and browser window sizes, SERPs can be divided into two areas, which are often referred to as "above the fold" vs. "below the fold" (Jansen & Spink, 2007), where the former refer to the results can be seen immediately without scrolling down. Users predominantly click on results shown above the fold. Therefore, it is important for content providers to take measures to make sure their content is listed in that area of the SERP.

2.3 Ad labeling

It is important to note that contextual, text-based ads ("sponsored links", "paid results") can be seen as one type of search result. They are similar to organic results in that they consist of a title, a short description and a URL, and they are also displayed on search engine results pages. Their design is also similar to organic results – the same colors or ones very similar to those used in organic results are used for headings, descriptions and URLs. Therefore, it seems reasonable to suppose that users may find it difficult to distinguish between the two results types.

Ads may be relevant to a query. What's more, their uniqueness compared to other forms of advertisements lies precisely in the fact that a user has already entered a query and thereby expressed his or her intent (cf. Battelle, 2005).

An early study (Fallows, 2005) found that only 38% of U.S. searchers were aware of the distinction between organic and paid results. The situation has surely changed since then, partly due to the U.S. Federal Trade Commission's guidelines on search engine ad disclosure (Sullivan, 2013a). However, the distinction between the two results types is still an issue, not only in general-purpose Web search engines, but also in many specialized vertical search engines (Sullivan, 2013b). Some industry studies strongly suggest that

the labeling of ads is not be clear enough (Bundesverband Digitale Wirtschaft, 2009; Charlton, 2013; Wall, 2012).

3 Research questions

The overall question guiding our research is whether users' knowledge on Google's business affects their clicking behaviour on the search engine results pages.

- *RQ1:* Does knowledge on the possibility of buying screen real estate on the SERPs (i.e., buying advertisements shown above the organic results) influence users' selection behaviour on the SERPs?
- *RQ2:* Does knowledge on the distinction between advertisements and organic results influence users' selection behaviour on the SERPs?
- *RQ3:* Does knowledge on Google's business model (i.e., selling ads) influence users' selection behaviour on the SERPs?
- *RQ4:* Does users' actual performance on distinguishing between organic results and paid advertisements on the SERPs influence their selection behaviour?

4 Methods

We conducted an online experiment using a representative sample of the German online population. The sample was built according to AGOF criteria ("Method – AGOF coverage currency", 2015) and consisted of 1,000 users. AGOF provides a standardized online coverage currency to measure the success of marketing tools. The online coverage currency is based on a Three-Pillar Model for Data mining and profiling by electronic measurement of page visits and page impressions, by on-site surveys on descriptive socio-demographic values and representative telephone surveys. The population includes Internet Users from the age of 10 years. In our study, each user was randomly assigned to either the experimental (ads) or control (no ads) condition.

The stimulus material was two screenshots of a Google search engine results page (fig. 1). Both had the exact same results and the same layout, the only difference being one having the first two results labeled as ads (yellow shading, info button), while the other version only had a list of organic results. Note that we only used organic results, i.e., the ads shown in the experimental condition are actually organic results, only with an ads labeling.

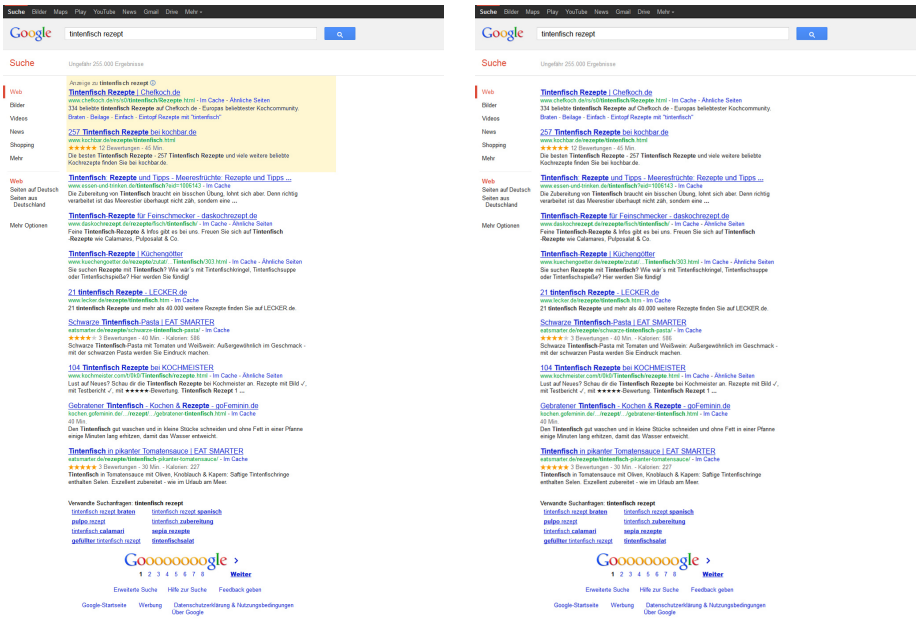

Fig. 1 Screenshots used as stimulus material

The participants were given the following search task: “Imagine you participate in a cooking competition where you should prepare fresh calamari. Your search query is ‘calamari recipe’. Which result(s) would you click on spontaneously?” [translated from German].

In our online experimental setting, participants were asked to mark the results they would select, i.e., we did not measure actual clicks but users labeled results as relevant. This allowed users to mark more than one result.

Each user was randomly assigned either to the experimental condition (ads) or the control condition (no ads). Additional data for the groupings used in the analysis below come from questions used in the data collection (detailed results on these can be found in Lewandowski et al., 2017). The questions relevant to the present study are the following:

- "Search engines are commercial Internet services, and therefore need to make money. Please describe in your own words how the search engine Google generates its revenues." (This question operationalizes RQ3.)
- "Is it possible to pay Google for preferably listing one's company on the search results pages, as an answer to a search query?" (This question operationalizes RQ1)
- "Is it possible to distinguish between paid advertisements and unpaid results on Google's search engine results pages?" This question operationalized RQ2.

The fourth and last grouping (operationalizing RQ4) is built on four tasks where users had to mark all advertisements or all organic results on screenshots of SERPs. For detailed results on this task, see Lewandowski, Kerkmann, Rümmele and Sünkler (2017). For the present study, we only distinguish between users who proved in these tasks that they were able to identify ads consistently and those who were not able to do so.

5 Results

In the following, we analyze the clicks on the first two results (the ads in the experimental condition and the first two organic results in the control condition, respectively).

5.1 Knowledge-based questions

First, we looked at relationships between users' knowledge on whether it is possible to buy screen real estate on the search engine results pages. There are three groups in this case: Those who say that it is possible (correct answer), those who say it is not possible (incorrect) and those how say they do not know. Looking at the selection behaviour in the two conditions, we do not find significant differences related to users' knowledge (cf. table 1). While users in the control condition select the first two results more frequently than users in the experimental condition, there are no significant differences between groups within the conditions.

Table 1: Results to the question "Is it possible to pay Google for preferably listing one's company on the search results pages, as an answer to a search query?"

Position	Experimental condition (ads)			Control condition (no ads)		
	Yes (n = 369)	No (n = 28)	Don't know (n = 103)	Yes (n = 364)	No (n = 36)	Don't know (n = 100)
1	38.5	42.9	36.9	59.1	63.9	55.0
2	25.7	17.9	33.0	38.7	41.7	44.0

* Differences between groups significant at $p \leq 0.05$

We found, however, significant results when it comes to users' knowledge on whether it is possible to distinguish between paid advertisements and organic results on the search engine results pages (table 2). In the experimental group, users who say that this is possible select the first result significantly more often. However, the group saying they do not know select the top results even more often. As expected, there are no significant differences in the control condition.

Especially the result that users who say that it is possible to distinguish between the results types select the top result (in this case, an ad) more often seems counterintuitive. However, this may be explained by users who say this is not possible selecting the results they think are no advertisements.

Table 2: Results to the question "Is it possible to distinguish between paid advertisements and unpaid results on Google's search engine results pages?"[+]

Position	Experimental condition (ads)			Control condition (no ads)		
	Yes (n = 217)	No (n = 90)	Don't know (n = 62)	Yes (n = 208)	No (n = 105)	Don't know (n = 51)
1	36.9*	32.2*	53.2*	59.6	54.3	66.7
2	23.5	23.3	37.1	37.0	38.1	47.1

* Differences between groups significant at $p \leq 0.05$
\+ Only participant who answered "yes" in the preceding question (cf. table 1)

The third knowledge-based question related to users' knowledge on how Google makes money from its search engine. We classified the answers into four groups: Correct answer (advertising), incorrect answer, partly correct answer (where advertising was mentioned, but other incorrect sources of revenue, as well), and "don't know" (where users admitted they did not

know). We found significant differences in the selection behaviour of the different groups only in the control conditions (table 3). Users who know how Google makes money choose the first position significantly more often than users without that knowledge. A likely explanation is that these users notice that there are no ads on the page and therefore regard the first result as trustworthy.

Table 3: Results to the question "How does Google generate its revenue?"

Position	Experimental condition (ads)				Control condition (no ads)			
	Correct (*n*=300)	Incorrect (*n*=51)	Partly correct (*n*=100)	Don't know (*n*=49)	Correct (*n*=306)	Incorrect (*n*=42)	Partly correct (*n*=106)	Don't know (*n*=46)
1	40.0	35.3	38.0	32.7	61.1*	52.4*	39.1*	62.3*
2	27.3	21.6	25.0	32.7	43.5	42.9	40.4	33.0

* Differences between groups significant at $p \leq 0.05$

5.2 Performance measures

We found significant differences between selection behaviour on the first two results between users who proved to be able to distinguish between organic results and ads and those who were not able to do so (table 4). The latter chose advertisements significantly more often (40.3% vs. 21.6% for the first results, and 28.3% vs. 13.7%, respectively). The numbers are quite impressive: Users who are not able to distinguish between the two results types choose ads around twice as often as users who are able to recognize the ads. As expected, there are no significant differences in the selection behaviour in the control condition.

Table 4: Actual performance (marking ads)

Position	Experimental condition (ads)		Control condition (no ads)	
	All areas labeled correctly (*n* = 51)	Not all areas labeled correctly (*n* = 449)	All areas labeled correctly (*n* = 45)	Not all areas labeled correctly (*n* = 455)
1	21.6*	40.3*	64.4	58.0
2	13.7*	28.3*	55.6	38.5

* Differences between groups significant at $p \leq 0.05$

6 Discussion

Our study revealed some surprising results, namely that users' self-reported knowledge on whether it is possible to pay for being shown on Google's SERPs did not affect their selection behaviour in the experimental condition (RQ1). We have to note, however, that we used *self-reported* measures, i.e., we do not know whether users saying that they know that it is possible to pay Google for being shown on the SERPs actually know how to spot paid results. Those who say it is possible to distinguish between ads and organic results select ads more often (RQ2). The likely reason is that they do so on purpose, as ads can be relevant to a search query (Jansen, 2007; Lewandowski et al., 2017). Users who know how Google generates its revenues selected the first results more frequently in the control condition (RQ3), which could be related to them trusting the search engine's ranking (Pan et al., 2007).

Maybe even more interesting, the performance measures based on distinguishing users into a group that actually *proved* to be able to distinguish between ads and organic results in a variety of tasks, and a group that was not able to do so, show a clear (and in some ways contradictory) result from the results discussed so far. They clicked on the ads approximately twice as often as the knowledgeable group. Through the experimental design of our study, we can rule out different relevance judgments as a reason for this, as users in both conditions saw exactly the same results.

Regarding methodology, the results of our study raise questions on the reliability of self-reported measures. Therefore, we argue for better using task-based user groupings than groupings based on questionnaires.

7 Conclusion

Our results have implications for search engine design and regulation, and for information literacy regarding search engines, as well. The results call for a clear labeling of advertisements on the search engine results pages. There has been a discussion on how ads should be labeled for at least 15 years now (Sullivan, 2013a), which has, however, not been based on a proper empirical basis. Our study contributes to filling this gap. However, it should be re-

garded as only a first step, as a limitation lies in that we used only one task, which was chosen because it is true-to-life and we assume all search engine users are able to choose relevant results to this task. However, further studies using tasks where the aim is to find high-quality information on professional tasks should be conducted.

Our results also call for more effort on helping users to become information literate when it comes to search engines. Efforts on increasing users' information literacy often focus on specialized information sources, and less on the tools that users are using on a daily basis. It is time for information literacy researchers and practitioners to focus on that blind spot.

Our study can only be a starting point on investigating users' selection behaviour when it comes to ads. Further research is needed, especially related to further results types, such as Universal Search results, which can be either ads or organic results.

References

Battelle, J. (2005): *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. London: Portfolio.

Bundesverband Digitale Wirtschaft (2009): Nutzerverhalten auf Google-Suchergebnisseiten: Eine Eyetracking-Studie im Auftrag des Arbeitskreises Suchmaschinen-Marketing des Bundesverbandes Digitale Wirtschaft (BVDW). http://www.hottracker.de/whitepaper/Eyetrackingstudie_2008_Ergebnisbericht.pdf

Charlton, G. (2013): 40% of consumers are unaware that Google Adwords are adverts. https://econsultancy.com/blog/62249-40-of-consumers-are-unaware-that-google-adwords-are-adverts

Drumond Monteiro, S., and M. Aparecida Moura (2014): Knowledge Graph and "Semantization" in Cyberspace: A Study of Contemporary Indexes. In: *Knowledge Organization*, 41 (6), 429–439.

Fallows, D. (2005): Search engine users: Internet searchers are confident, satisfied and trusting – but they are also unaware and naive. Pew Internet & American Life Project. Washington, DC: Pew Internet & American Life Project.

Goel, S., A. Broder, E. Gabrilovich, and B. Pang (2010): Anatomy of the long tail: Ordinary people with extraordinary tastes. In: *Proceedings of the third ACM international conference on Web search and data mining* (pp. 201–210). ACM.

Grimmelmann, J. (2010): Some Skepticism About Search Neutrality. In: *The Next Digital Decade: Essays on the Future of the Internet*, 435–459.

Jansen, B. J. (2007): The comparative effectiveness of sponsored and nonsponsored links for Web e-commerce queries. In: *ACM Transactions on the Web*, 1 (1), article 3.

Jansen, B. J., and A. Spink (2007): Sponsored search: is money a motivator for providing relevant results? In: *Computer*, 40 (8), 52–57.

Joachims, T., L. Granka, B. Pan, H. Hembrooke, and G. Gay (2005): Accurately interpreting clickthrough data as implicit feedback. In: *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 154–161). Salvador, Brazil: ACM.

Lewandowski, D. (2017): Is Google responsible for providing fair and unbiased results? In: L. Floridi and M. Taddeo (Eds.): *The Responsibilities of Online Service Providers*. Berlin, Heidelberg: Springer.

Lewandowski, D., F. Kerkmann, S. Rümmele, and S. Sünkler (2017): Search engine ad disclosures. [forthcoming]

Method – Method – AGOF coverage currency (2015). https://www.agof.de/studien/internet-facts/methode-internet/?lang=en

Noble, S. U. (2013): Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible. http://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/

Pan, B., H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka (2007): In Google we trust: users' decisions on rank, position, and relevance. In: *Journal of Computer-Mediated Communication*, 12 (3), 801–823.

Petrescu, P. (2014): Google Organic Click-Through Rates in 2014. https://moz.com/blog/google-organic-click-through-rates-in-2014

Sullivan, D. (2013a): FTC Updates Search Engine Ad Disclosure Guidelines After "Decline in Compliance." http://searchengineland.com/ftc-search-engine-disclosure-164722

Sullivan, D. (2013b): Revealed: The 17 Other Search Engines The FTC Warned Over Paid Ad Disclosures. http://searchengineland.com/revealed-17-other-search-engines-ftc-warned-168603

Taylor, B., M. Mayer, and O. Buyukkokten (2008): Interface for a universal search engine. United States Patent and Trademark Office.

Wall, A. (2012): Consumer Ad Awareness in Search Results. http://www.seobook.com/consumer-ad-awareness-search-results

White, R. W., and E. Horvitz (2009): Cyberchondria. In: *ACM Transactions on Information Systems*, 27 (4): Article No. 23. http://doi.org/10.1145/1629096.1629101