

In Google we trust? Domain knowledge and students' Web page selections when searching for a controversial topic

Ladislao Salmerón¹, Yvonne Kammerer², Ana Llorens¹, Pilar García-Carrión¹

¹ University of Valencia - Avd. Blasco Ibáñez, 21, 46010, Spain

² Knowledge Media Research Center - Konrad-Adenauer-Str. 40, 72072 Tübingen, Germany
lasalgon@uv.es, y.kammerer@iwm-knrc.de, a.cristina.llorens@uv.es, magarca3@alumni.uv.es

Abstract. This study investigated how prior domain knowledge and rank position in a search engine result page influence students' selections of Web pages of various trustworthiness levels. Results revealed that while learning about a controversial topic low prior knowledge students ($n = 9$) based their selection decisions of Web pages mostly on the respective page rank. In contrast, high knowledge students ($n = 11$) based their decisions more on Web page trustworthiness than on page rank. This supports the assumptions of a "superficial model" which proposes that low knowledge students' Web selections are mostly guided by superficial cues.

Keywords: Web page selection; domain knowledge; rank position; trustworthiness evaluations.

Introduction

Students in higher education often use the World Wide Web (WWW) as a source to learn about controversial topics, such as climate change. For example, they may be required to find information on the topic "Reducing greenhouse gas emissions" and to write a report on the issue. One of the main steps of this task involves the selection of relevant Web pages for further study. Due to the relatively low filters to publishing content on the WWW, however, Web pages for controversial topics may not only vary with regard to their relevance for a particular topic, but also with regard to their trustworthiness. Therefore, an important question is whether undergraduate students consider the relevance and trustworthiness of Web pages in their selection decisions while learning about a controversial topic.

Previous research has identified two main factors that might influence students' Web page selections: prior domain knowledge and page rank in a search engine results page (SERP). Text comprehension research suggests that the skills required to critically evaluate multiple sources develop with domain expertise (e.g., Bråten, Strømsø & Salmerón, in press; Rouet et al., 1997). In line with that, case studies of high school students' Web search behavior indicate that prior domain knowledge also facilitates relevance and trustworthiness evaluations of Web pages (e.g., MaKinster, Beghetto, & Plucker, 2002). Kammerer & Gerjets (2010a) who investigated evaluation processes during Web search for a controversial topic by means of eye-tracking methodology, found a positive relationship between university students' prior domain knowledge and inspection time of search results on SERPs and reference information on Web pages. Other recent studies which investigated trustworthiness evaluations during Web search for controversial topics through verbal reports, however, did not find any relationship between prior knowledge and trustworthiness evaluations (e.g., Mason, Boldrin & Ariasi, in press). A different line of research has explored the influence of the page rank in a SERP on students' Web page selections (e.g., Kammerer & Gerjets, 2010b; Pan et al., 2007). An eye-tracking study by Kammerer and Gerjets (2010b) showed that when top search results in a SERP list were of least trustworthiness and bottom search results of highest trustworthiness, most focus was still given to the top results and the most trustworthy pages were selected less often, as compared to a list where top search results were of highest trustworthiness.

An open question is how domain knowledge may interact with rank position to shape students' selections of Web pages of various trustworthiness levels. We anticipate two possible models to explain the relationship between these two factors. A "compensatory model" may propose that low

knowledge students will base their selections more on the trustworthiness of the Web page than on the list rank, because these students might strategically look for a reliable source to start with their study session. In contrast, a “superficial model” may consider that low knowledge students base their selection decisions regarding a Web page on superficial cues (cf. Rouet et al., in press). In case that rank position and trustworthiness level differ, thus they may be more inclined to select a Web page that occupies an upper position in the list. The assumptions of these models guided the study reported in the remainder of this paper.

Method

Thirty undergraduate psychology students from a large Spanish University participated in the experiment (mean age 22.2 years, 83% female). Participants’ task was to search the Web for pages on the topic of ‘Reduction of greenhouse gas emissions’ and to finally select the two pages that covered best the topic and thus could serve as basis for a report on this issue.

Participants were provided with 10 preselected Web pages retrieved from Google® with the query “Reduction of greenhouse gas emissions”. All pages included the terms in their titles, but not all were relevant for the task. For example, an irrelevant page was an online magazine informing about an agency that provided Spanish companies with certificates about their reduction of greenhouse gas emissions. We edited the Web pages to disable the hyperlinks, and to make them similar in length (607-725 words). In addition, eight graduate students with experience in text analysis rated the pages based on 1) the relevance to the topic “Reduction of greenhouse gas emissions” (relevance), 2) the degree to which the page covered the topic (completeness) and 3) the extent to which the page was trustworthy (trustworthiness), on scales from 1 (low) to 5 (high). Seven pages were considered rather relevant to the topic ($M = 3.8$), and three rather irrelevant ($M = 2.7$). From the relevant ones, three pages were considered to be more complete ($M = 4.1$), and four of them less complete ($M = 2.7$). Finally, from the three relevant and complete pages, one page was rated as more trustworthy ($M = 4.5$, page of the United Nations system’s work on climate change) than the other two ($M = 3.6$, personal blog and a page from an NGO).

On the SERP the Web pages were organized according to their topic relevance and completeness, as rated by the graduate students. Accordingly, the more relevant and more complete pages were presented in the top of the list. In addition, to test students’ trustworthiness evaluations the first two pages in the list were the less trustworthy ones (personal blog and the NGO page), and the third page the more trustworthy one (United Nations page). Students who only focus on rank position and not on page trustworthiness should more often select the first two pages. In contrast, students who base their decisions on Web page trustworthiness instead of page rank should select the third United Nations page more often. Less relevant web pages included sources of diverse trustworthiness levels.

Students proceeded as follows: First, data on demographics and prior domain knowledge were collected. Domain knowledge was assessed by means of a reliable multiple-choice test consisting of 17 items. Then, students could explore the 10 Web pages without time constraints by accessing them from the SERP. Afterwards they were required to select the two Web pages which they would like to use for further study.

Results

High knowledge students and low knowledge students were determined by a median split. In order to increase the difference between the two groups 10 participants which equaled the median were

excluded from further analyses, resulting in 9 low prior knowledge students and 11 high prior knowledge students ($t(18) = -8.66, p < .001$). With regard to students' selections, prior knowledge (PK) did not influence students' selections of the first and second Web page of the list (relevant, complete, but less trustworthy): $\chi^2(1, N = 20) = 0.02, ns$ (high PK: 36.36% selections; low PK: 33.33% selections), and $\chi^2(1, N = 20) = 0.14, ns$ (high PK: 36.36% selections; low PK: 44.44% selections), respectively. However, prior knowledge had an effect on the selection of the third relevant, complete and more trustworthy page of the list: $\chi^2(1, N = 22) = 4.79, p = .03$. High knowledge students selected this United Nations page significantly more often (81.88% selections) than low knowledge students (33.33% selections). Besides, there were no significant effects of prior knowledge on selection frequency of the lower ranked Web pages.

Conclusions

Results from Web page selections in our study support the superficial model, which considers that while learning about a controversial topic on the Web low prior knowledge students focus mostly on superficial cues, such as rank position on a SERP to select Web pages for further study (Rouet et al., in press). In contrast, students with high prior knowledge are more likely to base their selection on Web page trustworthiness instead of page rank. Thus, the results of this study concur with prior studies on text comprehension, which consider prior knowledge an important prerequisite to critically evaluate source trustworthiness. A potential reason for the absent relationship between prior knowledge and students' self-reports about Web trustworthiness in the Mason et al. study (2009) might be that trustworthiness evaluations could not be consciously verbalized by the students. Thus, to gain deeper insights into evaluation processes during Web search and into the interplay between prior domain knowledge and trustworthiness evaluations, future research might combine selection data, eye tracking methodology, and verbal protocols.

References

- Kammerer, Y., & Gerjets, P. (2010a, March). Searching the Web for science-related information: The impact of evaluation instructions, user characteristics, and search interface design on evaluations of information quality. Seminar on Computational Cognitive Modeling of Web-Navigation. Utrecht, The Netherlands.
- Kammerer, Y., & Gerjets, P. (2010b). How the interface design influences users' spontaneous trustworthiness evaluations of Web search results: Comparing a list and a grid interface. In C. Morimoto, & H. Instance (Eds.), *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications ETRA '10* (pp. 299-306). New York, NY: ACM.
- MaKinster, J.G., Beghetto, R.A., & Plucker, J.A. (2002). Why Cant I Find Newtons Third Law? Case Studies of Students Use of the Web as a Science Resource, *Journal of Science Education and Technology*, 11, 155-172.
- Mason, L., Boldrin, A., & Ariasi, N. (in press). Searching the web to learn about a controversial topic: Are students epistemically active? *Instructional Science*.
- Rouet, J.-F., Favart, M., Britt, M. A., & Perfetti, C. A. (1997). Studying and using multiple documents in history: Effects of discipline expertise. *Cognition and Instruction*, 15, 85-106.
- Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M. & Dinet, J. (in press). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learning and Instruction*.