How Authority-Related Epistemological Beliefs and Salience of Source Information Influence the Evaluation of Web Search Results – An Eye Tracking Study

Yvonne Kammerer (y.kammerer@iwm-kmrc.de)
Knowledge Media Research Center, Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Evelin Wollny (e.wollny@iwm-kmrc.de)
Knowledge Media Research Center, Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Peter Gerjets (p.gerjets@iwm-kmrc.de)
Knowledge Media Research Center, Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Katharina Scheiter (k.scheiter@iwm-kmrc.de)
Knowledge Media Research Center, Konrad-Adenauer-Strasse 40, 72072 Tuebingen, Germany

Abstract

In this study we examined how authority-related epistemological beliefs (AREBs) affect the evaluation of web search results presented by a search engine. Moreover, we investigated whether increasing the salience of source information on a search engine result page fosters students' evaluation processes, and whether the salience of source information moderates the effects of AREBs. Thirty university students participated in a web search experiment addressing a controversial search topic. Participants either used a standard Google search result list or an augmented search result list additionally containing source categories for each search result. Eye movements and mouse clicks were captured during information search. The results showed that the effects of participants' AREBs on their gaze behavior were moderated by the salience of source information. Furthermore, salience of source information affected participants' selection behavior, such that with additionally presented source categories they were more likely to select search results linked to web portals.

Keywords: web search; epistemological beliefs; search engines; evaluation processes; interface design; eye tracking; HCI

Introduction

In recent years, the World Wide Web (WWW) has evolved into one of the most important public sources for science-related contents offering convenient and rapid access to vast amounts of information. Particularly for domains of personal concern such as medicine and healthcare using the web has achieved great popularity. At the same time, however, the web imposes new challenges onto searchers due to its high complexity and heterogeneity of information. Search engines help to find information on the web, but searchers are still required to evaluate and select search results for further inspection of corresponding websites. Thus, search result descriptions displayed on search engine result pages (SERPs) play a critical role in guiding search processes as searchers rely on their evaluations of these descriptions to decide whether to visit the web pages they represent or not.

Evaluating topicality and information quality during web search

In cognitive science, a central approach used to explain and predict search result selection as well as other types of web search behavior is the Information Foraging Theory (Pirolli & Card, 1999). The theory assumes that the selection of hyperlinks (e.g., from a SERP), is determined by their so called "information scent". Information scent is defined as the semantic similarity between a cognitive representation of the current information needs and an external representation of cues on a display (e.g., keywords in the title, abstract, or URL of a search result). The stronger the semantic overlap between the cues contained in a particular search result and the user's information needs (i.e., the stronger the topical fit), the stronger is the information scent of the search result. The information scent of a particular search result determines its likelihood for being selected. Thus, according to Information Foraging Theory, search result selection is mostly guided by evaluating the topicality of web information.

In line with this theory, thinking-aloud data from web search experiments on science-related topics revealed that university students strongly focus on the topical fit (i.e., connection to the task) when evaluating search results returned by a search engine (e.g., Google). Evaluation criteria beyond topicality with regard to the quality of information (e.g, credibility, trustworthiness, reliability, or accuracy) were uttered rather seldom (Brand-Gruwel, van Meeuven, & van Gog, 2008; Gerjets, Kammerer, & Werner, subm.).

Although these findings are in line with the Information Foraging Theory, they are nevertheless rather astonishing if one takes into account that anyone can publish any information on the web. As a result, the web is characterized by a high heterogeneity of information sources, varying not only with regard to topical fit, but also with regard to authority and credibility. For instance, when it comes to information on complex and contradictory scientific topics, not only scientific and other institutions, but also journalists, companies, and laypeople can act as information providers.

Contrary to traditional information sources like printed publications, there is hardly any quality assurance on the web. As a consequence quality of web information can vary enormously depending on the authority of the information source, with scientific websites usually providing the most credible and commercial websites the least credible information.

In this paper, we address two potential reasons for the dominance of topicality-related evaluation criteria and the neglect of information quality: First, the lack of quality-related evaluation processes during web search might go back to deficits on the side of the searchers with regard to their epistemological beliefs. Second, the interface design of SERPs might lead to a neglect of information quality because quality-related source information is usually not displayed in a salient way. Both explanations will be further elaborated in the subsequent sections.

Authority-related epistemological beliefs

In order to evaluate the quality of web information searchers need to consider how credible a source of information is, how certain and consistent with other sources the information itself is, and how strongly the information might be influenced by the interests of the information provider. According to Hofer (2004) this kind of reasoning is closely connected to a person's epistemological beliefs (EBs), that is, to one's personal beliefs about the nature of knowledge and knowing. More precisely, with regard to evaluating a source of information the sophistication of one's beliefs in epistemic authorities (cf. Schommer, 1990) seems to be important. This dimension of EBs, in the remainder of this paper referred to as AREBs (authority-related epistemological beliefs), focuses on beliefs about the validity of the source of knowledge. It ranges from naïve beliefs that all knowledge is provided by an omniscient authority and should not be questioned (authority believers) to sophisticated beliefs that knowledge can be derived through one's own reasoning processes (authority disbelievers). Accordingly, we expect that searchers' AREBs influence their information evaluation during web search. For example, searchers who consider the publishers of web information to be epistemic authorities might see no need to critically evaluate the source of information.

In line with this reasoning, not only AREBs, but EBs in general have been shown to affect information seeking behavior during web search. In a thinking-aloud study with students searching the web for information on a scientific topic, Hofer (2004) found that students expressing naïve EBs (knowledge is absolute and unchanging and provided by authorities) conducted the search task in a brief and perfunctory way, without pursuing additional sources or reflecting on the credibility and accuracy of the sources they inspected. On the other hand, students with more sophisticated EBs (knowledge is relative, contextual and changing, and derived by reason) were more likely to actively seek for more recent sources and to pursue informed strategies for searching.

Similarly, the results of an interview study by Whitmire (2003) on undergraduate students' web information seeking behavior in a class project showed that students' EBs affected their search techniques, their ability to recognize authority, and their evaluation of the information itself. Whereas students with more sophisticated EBs used a variety of search techniques, students with more naïve EBs did not. Moreover, students with more sophisticated EBs considered themselves more capable of critiquing and evaluating the information that they encountered during the search process. Furthermore, they did not reject conflicting information during document collection and they were more capable to differentiate between highly authoritative sources and rather obscure ones.

Recently, Mason and Ariasi (2008) showed by using eye movements as indicators for visual attention and cognitive processing, that depending on their EBs university students visually attended to different parts of a web page and visually inspected different kind of web pages with varying intensity. However, this type of more fine grained studies focusing on the impact of EBs on specific cognitive processes during the evaluation of search results or the information extraction from web pages have been conducted only very rarely up to now.

Based on studies reported in this section we assume that in particular searchers' AREBs will influence their cognitive processes and their respective gaze behavior during the inspection and evaluation of SERPs. As there is no research on the impact of EBs on the evaluation of search results displayed on SERPs yet, this issue will be addressed in the present study.

Quality-related source information

Search result descriptions used in standard search environments like Google are mostly confined to topical information, whereas quality-related source information is sparse and non-salient. Thus, the interface design of standard search engines might not provide sufficient affordances for users to spontaneously engage in quality-oriented evaluation processes when selecting search results during web search for further inspection. It can be assumed that providing additional quality-related source information on SERPs might enable searchers to evaluate the quality and trustworthiness of the corresponding web pages more accurately. However, so far there are only a few empirical studies that have investigated how quality-related information affects the evaluation of search results.

With regard to the online news service Google News, Sundar, Knobloch-Westerwick, and Hastall (2007) could show that the availability of quality-related cues for news item (i.e., information on the source and the recency of a story, and on the number of related articles) affects the subjective evaluation of news leads. These results are related to our research questions because there are several similarities between news reading and searching for complex science-related information

Ivory, Yu, and Gronemyer (2004) enhanced the Google search engine interface with additional quality-related cues for each search result indicating the number of graphical ads, the number of words, and the estimated quality of the corresponding web page. They used simple fact-finding tasks to demonstrate that adding this type of quality-related cues improved participants' ability to select appropriate search results. From their pattern of results it can be assumed that for more complex information search tasks, quality-related cues might even have a greater impact.

This assumption is supported by findings from Rose, Orr, and Kantamneni (2007) who conducted a series of experiments with rather complex search scenarios in which searchers had to answer questions about the quality of search results on the basis of search result descriptions. Search result descriptions were experimentally manipulated to make sure that they differed with regard to specific attributes (e.g., length, text choppiness, and cues indicating a source category). The most interesting finding with regard to the present paper was that providing cues about the source category of a web page (e.g., a corporate homepage or a blog), influenced whether searchers trusted the information on the web pages.

In sum, the results of these studies indicate that augmenting SERPs in a way that not only topical information but also cues about the quality of information is included might provide substantial affordances for searchers to engage in quality-related evaluation processes. In the present study we will investigate how additional cues about the source of web information will influence the evaluation of web search results and how the availability of these cues might moderate the effects of searchers' AREBs on evaluation processes.

Research Questions

In this study we assume that two factors influence whether users engage in quality-related evaluations of search results presented by a search engine. First, searchers' epistemological prerequisites with regard to their view about the nature of knowledge and knowing, and second, the provision of quality-related source information by means of cues integrated in the interface design of a SERP.

Hence, we examined three research questions: (1) Do searchers' AREBs affect their evaluation of search results on SERPs? (2) Does the salience of quality-related source information affect searchers' evaluation processes on SERPs (3) Does the salience of quality-related source information moderate the relationship between searchers' AREBs and their evaluation of search results on SERPs?

Experiment

In order to investigate participants' evaluation processes on SERPS, we tracked participants' eye movements to gain detailed insights into their cognitive processing (cf. Rayner, 1998) and we used log files to record their consequent search result selections. Beyond logging overt interactions with the search interface, eye tracking methodology seems to be very promising because it allows reconstructing every search result that was looked at and that has been evaluated – independent from its selection or rejection (cf. Brumby & Howes, 2008). In contrast to previous studies, we avoided explicit questions about the perceived quality and trustworthiness by means of questionnaires or interviews because these methods are prone to create artifacts (cf. Gerjets et al., subm.).

Method

Participants and Design. Thirty participants (12 male, 18 female; mean age 25.90 years, *SD* = 3.58) from different majors at the University of Tuebingen, Germany, participated in this experiment for either course credit or payment. Participants had normal or corrected to normal vision.

As a first independent variable participants' *AREBs* were assessed and used as a continuous factor (see "Measures" for details), with the continuum ranging from authority disbelievers (participants strongly believing that knowledge can be derived through one's own reasoning processes; sophisticated EBs) to authority believers (participants strongly believing that knowledge it provided by authority and should not be questioned; naïve EBs).

As a second independent variable the *Source Salience* on the SERPs was experimentally manipulated. In the Low Source Salience condition, standard Google search result descriptions were used displaying only a title, an abstract, and a URL for each search result with search terms printed in bold as topical cues (see Figure 1a). In the High Source Salience condition, the salience of source information was increased, such that in addition to the URL of each search result, source category cues were presented in bold green that indicated to which of five different source categories a search result belongs ("Science/Institutions", "Portals/Advisors", "Journalism/TV", "Readers'Comments", and "Shops/Companies"). Additionally, search terms were not printed in bold to decrease topical salience (see Figure 1b).

Participants were randomly assigned to one of the two Source Salience conditions.

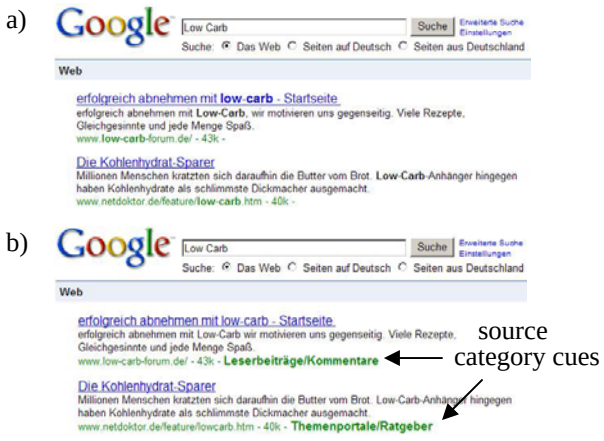

Figure 1: a) standard SERP with Low Source Salience and b) augmented SERP with High Source Salience.

Task and Materials Participants were presented with a fictitious request from an overweight friend, who asks for advice because she wants to loose weight by changing her diet. Participants were asked to conduct a 20 minutes web research to make an informed decision between low fat and low carb diets in order to recommend one of the two diets.

For their web research about this controversially discussed topic, participants were provided with three prearranged Google-like SERPs with ten search results each retrieved for the search terms "low fat", "low carb", and "low carb + low fat". Participants could access 30 web pages corresponding to the list of search results. The collection of search results and web pages for each of the three SERPs reflected the given heterogeneity of information sources available online. All three SERPs included web sites provided by scientific and other institutions (e.g., universities), web advisors and portals (e.g., medical or health portals), journalists (e.g., online magazines), laypeople (e.g., forums or blogs), and industry and companies (e.g., online shops for nutrition or pharmaceutics). Note that the different source types were balanced across all positions of the SERPs.

For the High Source Salience condition each of the search results was assigned to one of the five source categories.

Measures. To assess the independent factor AREBs a translated and revised version of the Omniscient Autority scale of the Epistemic Beliefs Inventory (EBI; Schraw, Dunkle & Bendixen, 1995) was used, comprising 4 items which had to be rated on 5-point scales (5 = highly agree). A sample item is "People who question authority are trouble makers."

As dependent variables we assessed participants' gaze behavior and selection behavior.

For the analysis of participants' gaze behavior (recorded with a Tobii 1750 remote eye tracker) so-called areas of interest (AOIs) were defined manually on the 30 search results. AOIs are precisely specified areas of an object, in our case a search result on a SERP, for which eye tracking parameters are aggregated. Each of the search result descriptions (including title, abstract, and URL) was defined as a single polygonal AOI (Search Result-AOI). Identical Search Result-AOIs were used for the two Source Salience conditions. For the High Source Salience condition, additional rectangular AOIs were defined on each of the 30 source category cues (Category-AOIs).

As a first dependent variable the mean total dwell time (in milliseconds) on Search Result-AOIs belonging to one of the five categories was measured by dividing the total time for which participants inspected the search result descriptions of a source category by the number of search results available in this category. In addition, for the High Source Salience condition the mean total dwell time on the Category-AOIs was assessed.

For the analysis of participants' selection behavior, i.e. participants' mouse clicks on search result links in order to access a webpage, we analyzed the selected search results with regard to what source category they belonged to by dividing the sum of selected search results of a source category by the total number of selected search results.

Procedure Participants were tested in individual sessions of approximately 1 hour. Before starting with the web search experiment participants were asked to fill in a computer-based questionnaire to assess control variables and AREBs. Subsequently, participants received instructions about the web search experiment and were calibrated on the eye tracking system. Then, they underwent a training task (equivalently constructed as the subsequent main task) for approximately three minutes to get acquainted with the web search environment. After the training task, participants were given the instruction for the main task including the fictitious request of their friend. Eye movements, screen recordings, and mouse clicks were captured during the entire 20 minutes task performance. Subsequent to the search task participants were required to decide which of the two diet methods they would recommend to their friend. However, only process measures during web search were analyzed to address our research questions.

Results

Gaze behavior measures with regard to Search Result-AOIs and selection behavior measures were analyzed by means of MANCOVAs with an interaction term with *Source Salience* as a fixed factor and *AREBs* as a continuous factor. Dependent variables were grouped according to the five source categories of the search results. To determine the direction of interaction effects correlations were computed.

With regard to participants' *gaze behavior* the overall MANCOVA showed a significant main effect for Source Salience (*Pillai's Trace* = .42, $F(5, 22) = 3.15$, $p = .03$) and a significant interaction between Source Salience and AREBs (*Pillai's Trace* = .40, $F(5, 22) = 2.88$, $p = .04$) on mean total dwell time on Search Result-AOIs. There was no main effect for AREBs ($F(5, 22) = 1.67$, *ns*).

In univariate analyses, the multivariate main effect of Source Salience could be traced back to search results belonging to the source category "Shops/Companies" (*Pillai's Trace* = .14, $F(1, 26) = 4.10$, $p = .05$). Participants in the High Source Salience condition with a mean total dwell time of $M = 1.46$ seconds ($SE = 0.23$) per search result paid less attention to search result descriptions labeled with "Shops/Companies" than participants in the Low Source Salience condition ($M = 2.12$, $SE = 0.23$). The multivariate interaction effect was corroborated univariately for search results of all but one source category ("Portals/Advisors": $F(1, 26) = 4.97$, $p = .04$; "Journalism/TV": $F(1, 26) = 4.27$, $p = .05$; "Readers' Comments": $F(1, 26) = 12.37$, $p < .01$; "Shops/Companies": $F(1, 26) = 10.64$, $p < .01$; not for "Science/Institutions").

For High Source Salience correlation results revealed a significant *negative* association between AREBs and mean total dwell time for the three source categories "Portals/Advisors" ($r = -.61$, $p = .02$), "Journalism/TV" ($r =$

-.61, p = .02), and "Readers' Comments" (r = -.58, p = .02). Thus, when additional source categories were presented, the more sophisticated participants' AREBs were, the longer were their mean total dwell times on the search result descriptions belonging to these categories. No correlation was revealed for "Shops/Companies" search results. In contrast, for Low Source Salience correlation results showed significant *positive* correlations between AREBs and mean total dwell time for search results of the categories "Readers' Comments" (r = .56, p = .03) and "Shops/Companies" (r = .57, p = .03). No correlations were revealed for search results of the other three categories. Figure 2 shows these interaction effects between Source Salience and AREBs for the different source categories.

Additionally, we analyzed the correlation between participants' mean total dwell time on Category-AOIs and their AREBs. Correlation results revealed a significant *negative* correlation for three of the five source categories, namely for "Portals/Advisors" (r = -.62, p = .01), "Journalism/TV" (r = -.53, p = .04), and "Readers' Comments" (r = -.54, p = .04). Thus, the more sophisticated participants' AREBs were, the longer were their mean total dwell times on the source category cues presented.

With regard to participants' *selection behavior* an overall MANCOVA revealed no significant main effect of Source Salience on the category distribution of selected search results (F(5, 22) = 1.27, *ns*), nor did it show a significant main effect for AREBs or a significant interaction between Source Salience and AREBs (both Fs < 1).

Nonetheless, a univariate main effect of Source Salience was obtained for the portion of selected search results belonging to the category "Portals/Advisors" (F(1, 26) = 4.20, p = .05). In the High Source Salience condition significantly more of the search results participants selected belonged to this source category (High Source Salience: M = 32.91%, SE = 2.63; Low Source Salience: M = 25.22%, SE = 2.68). Overall, during the 20 minutes web search participants clicked on M = 13.67 search results in the High Source Salience condition and on M = 13.60 in the Low Source Salience condition, whereas they visually inspected nearly all search results (M = 27.73 and M = 28.20).

Discussion

The purpose of the reported study was to investigate the effects of searchers' AREBs and of the salience of quality-related source information on SERPs on searchers' evaluation of search results during web research on a controversial topic. Furthermore, we examined, whether the salience of source information moderates the relationship between searchers' AREBs and their evaluation processes with regard to search results on SERPs. Evaluation processes on SERPs were analyzed in terms of gaze and selection behavior.

With respect to AREBs no main effects on searchers' gaze and selection behavior were obtained.

With respect to the salience of source information the gaze data revealed that searchers with additionally presented source category cues paid less attention to search result descriptions linked to commercial web pages. It seems that there is no need for searchers to further inspect these search result descriptions as the source category "Shops/Companies" immediately indicates that the corresponding web pages might not be suitable for a science-related information search. Furthermore, concerning participants' selection behavior the findings show that searchers with additionally presented source cues were more likely to select search results belonging to the category "Portals/Advisors", that is, search results linked to medical and health portals. Web pages from this type of source might have been expected to contain important information for participants' web research. Additionally, the term "advisors" itself might have prompted participants to select these search results with the expectation to receive an advice concerning their search task on the corresponding web page.

Moreover, complex interaction effects were found between AREBs and salience of source information concerning gaze behavior, with differences in the mean total dwell times for authority believers and disbelievers depending on source salience and the type of search result.

Gaze behavior on SERPs with high source salience. On SERPS with additionally presented cues that indicate the source category, the sophistication of participants' AREBs

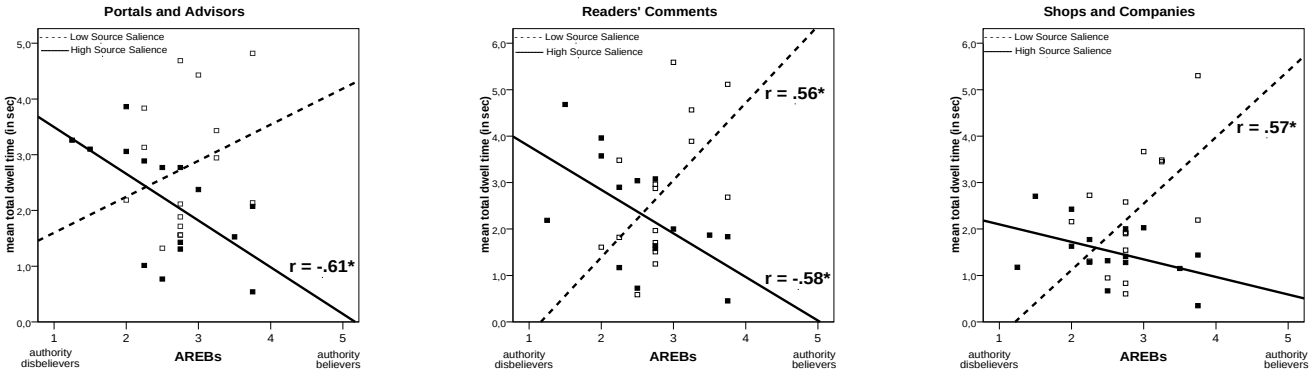

Figure 2: Significant interactions between AREBs and Source Salience on mean total dwell time on search result descriptions belonging to the categories "Portals/Advisors" (left), "Readers' Comments" (middle), and "Shops/Companies" (right).
Note: Interaction pattern for "Journalism/TV" is identical to the visualization of "Portals/Advisors".

was associated with a more thorough inspection of those source category cues and search result descriptions that belonged to the categories "Portals/Advisors", "Journalism/TV", and "Readers' Comments". Interestingly, particularly these three category labels do not provide clear indications about the quality of corresponding web pages. Thus, authority disbelievers seem to focus their attention on source categories where they might be inclined to analyze the trustworthiness of the source by themselves. Authority believers, on the other hand, do not focus their attention on these ambiguous categories.

Gaze behavior on SERPs with low source salience. On standard SERPs, for search results belonging to the three source categories "Science/Institutions", "Portals/Advisors", and "Journalism/TV" no relation between AREBs and gaze behavior was found. However, for the other two categories "Readers' Comments" and "Shops/Companies", that is, search results linked to forums or blogs and commercial websites, participants inspected the search result descriptions the shorter, the more sophisticated their AREBs were. A possible explanation is that authority disbelievers might be able to identify such search results as being of rather low quality by having only a quick look on the search result descriptions (e.g., the URLs).

Conclusion

Summarizing the results of this study, searchers visually inspected different kind of search results with varying intensity depending both on their AREBs and on the salience of source information. An increased salience of source information on SERPs stimulated authority disbelievers to engage in evaluation processes with regard to more ambiguous search results in terms of the trustworthiness of the source. This supports the assumption that both sophisticated AREBs and SERPs containing quality-related source information are important factors for adequate quality-related evaluation processes on SERPs.

Log file analysis revealed a rather simple relationship between the salience of source information and the selection of specific search results, namely that portals were selected more often on SERPs with source category cues. Portals seem to be a good choice given their good trade off between reliable information on the one hand and comprehensible information on the other hand. Beyond that, eye tracking analyses revealed complex interaction effects between AREBs and the interface design of SERPs on searchers' visual inspections of search results. Hence, in order to gain more detailed insights into searchers evaluation processes, eye tracking proved to be the method of choice.

In summary, AREBs seems to play an important role for evaluation processes on SERPs depending on the source information included in the search results. Furthermore, redesigning the interface of SERPs by additionally providing quality-related source information seems to have an impact on searchers' selection behavior. However, further research is needed in order to shed light on the complex interplay between AREBs and salience of source information.

References

Brand-Gruwel, S., Van Meeuwen, L., & Van Gog, T. (2008). The use of evaluation criteria when searching the WWW: An eye-tracking study. *Proceedings EARLI Special Interest Group Text and Graphics: Exploiting the opportunities - Learning with textual, graphical, and multimodal representations,* 34-37. Tilburg, NL.

Brumby, D. P., & Howes, A. (2008). Strategies for guiding interactive search: An empirical investigation into the consequences of label relevance for assessment and selection. *Human-Computer Interaction, 23*, 1-46.

Gerjets, P., Kammerer, Y. & Werner, B. (subm.). *Methodological issues in investigating metacognitive evaluation processes during web search.* Manuscript submitted for publication.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*, 43-55.

Ivory, M. Y., Yu, S., & Gronemyer, K. (2004). Search result exploration: A preliminary study of blind and sighted users' decision making and performance. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems.* ACM, New York, NY, 1453-1456.

Mason, L., & Ariasi, N. (2008, September). Critical thinking during web searching: Tracking students' evaluation of sources and information through eye movements and allocation of visual attention. Paper presented at the first meeting of the Research Network "Use of (external) representations in mathematical and scientific reasoning and problem solving: Analysis and improvement". Leuven, Belgium.

Pirolli, P., & Card, S. K. (1999). Information Foraging. *Psychological Review, 106*, 643-675.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rose, D. E., Orr, D., & Kantamneni, R. G. P. (2007). Summary attributes and perceived search quality. *In Proceedings of the 2007 International Conference on the World Wide Web 2007*, 1201-1202.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498-504.

Schraw, G., Dunkle, M. E., & Bendixen, L. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*, 523-538.

Sundar, S. S., Knobloch-Westerwick, S., & Hastall, M. R. (2007). News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology, 58*, 366-378.

Whitmire, E. (2003). Epistemological beliefs and the information-seeking behavior of undergraduates. *Library & Information Science Research, 25*, 127-142.