Search and Justification Behavior During Multimedia Web Search for Procedural Knowledge

Georg Pardi
Leibniz-Institut für Wissensmedien
Tübingen, Germany
g.pardi@iwm-tuebingen.de

Yvonne Kammerer
Leibniz-Institut für Wissensmedien
Tübingen, Germany
y.kammerer@iwm-tuebingen.de

Peter Gerjets
Leibniz-Institut für Wissensmedien
Tübingen, Germany
p.gerjets@iwm-tuebingen.de

ABSTRACT

In an eye-tracking study, N = 38 participants performed two procedural-knowledge search tasks by using a mockup multimedia search engine results page (SERP). By presenting both conventional websites and videos as results on the SERP, we aimed at examining the role of the modality of information resources in individuals' retrieval behavior as well in their final recommendation of one most suitable information resource. Across both tasks, the results of this study indicate that participants who finally recommended a video resource spent a greater proportion of time inspecting video results on the SERP as well as on the video resources themselves. Furthermore, participants' written justifications for the recommended information resource revealed that in both tasks about one third of the participants mentioned the modality of the information resource to justify their recommendation decision. Our findings indicate that the modality of information resources at least to some extent plays a role during web search for procedural learning resources.

CCS CONCEPTS

Information systems → Information retrieval
Information systems → World Wide Web
Human-centered computing → Human computer interaction

KEYWORDS

Search Engine, eye-tracking, search as learning, modality, procedural knowledge, videos

ACM Reference format:

Georg Pardi, Yvonne Kammerer, and Peter Gerjets. 2019. Search and Justification Behavior During Multimedia Web Search for Procedural Knowledge. In *WebSci'19 Companion*, June 30–July 3, 2019, Boston, MA, USA, 4 pages. ACM, New York, NY, USA. https://doi.org/10.1145/3328413.3329405.

WebSci '19 Companion, June 30-July 3, 2019, Boston, MA, USA

ACM ISBN 978-1-4503-6174-3/19/06…$15.00
https://doi.org/10.1145/3328413.3329405

1 INTRODUCTION AND BACKGROUND

The world wide web offers a huge amount of information presented in various modalities, such as texts, images, and videos, made accessible via search engine result pages (SERP). Thus, unsurprisingly, the web has become a major medium for learning, i.e., to acquire knowledge and skills. One of the most popular online information resources currently are videos, which are easy and convenient to consume. Hence, Google and most other search engines have adapted their SERPs to a format where the user has a broad choice between information resources of different modalities [2]. Several information seeking models [e.g., 4, 8] describe the processes of information seeking on the web (i.e., defining the information need, selecting, processing, and integrating information resources) in order to solve an information problem at hand. However, to the best of our knowledge none of these models addresses the fact that web searchers can choose between different types of modalities (such as text, images, video, or audio). In the present paper, we argue that web searchers do not only have to define their information need, but also to decide what information resources of which kind of modality they want to retrieve.

On SERPs, the presentation of information resources, according to Lewandowski [6], can be classified in four major search results types: organic results, advertisements, vertical (or specialized) search results, and knowledge-graph results. In this paper, the focus lies on organic and vertical search results. Organic results are webpages that are indexed by the search engine's general web index. In contrast, vertical search results come from specialized databases, such as the search engine's video index, news index, or image index. Such search results can be retrieved by choosing the specialized search tabs, but vertical results are also integrated into the ranked list of organic results (known as Universal Search)[1]. [6]. While vertical search results, particularly in the case of video results, directly indicate the type of modality, for organic search results linking to conventional websites it is usually unclear whether they are purely text-based or also contain images or even video-elements. A few studies recently examined effects of these

[1] Since June 2018, video, news, and image search results are displayed in carousel-formatted boxes in the Google SERP. However, when the study was conducted in April 2018 they were typically presented in an interleaved way in the SERP.

different results types being presented in a SERP in an interleaved way on users' viewing and selection behavior. For example, [7] found that vertical search results in general attract the gaze behavior of participants. However, even in a SERP comprising vertical search results the search result position still seems to have an enormous impact on the viewing and selection process [9].

In the present paper, we aimed at investigating to what extent information seekers consider whether a search result links to a video or to a conventional website both during the retrieval process when selecting search results from the SERP and when inspecting information sources, and in their final selection of one most suitable information resource. Moreover, as some individuals might prefer videos, whereas others might rather prefer conventional websites, we also wanted to examine, whether such preferences are already observable during the retrieval process.

Furthermore, we chose to focus on web searches for procedural knowledge, which have not been given much attention to in prior research (for an exception see [5]). Anderson and Krathwohl [1] define procedural knowledge as knowledge concerning "how to do something, methods of inquiry, and criteria for using skills, algorithms, techniques and methods". In addition, for our work we further distinguish between sensori-motor procedural knowledge (i.e., performing observable movements, such as tying a knot) and cognitive procedural knowledge (i.e., solving purely cognitive procedures, such as mathematical problems) (cf. [3]).

In sum, the goal of the present study was to answer the following three research questions: First, how much time do individuals spend on information resources of different modalities when searching for procedural-knowledge *(RQ1)*? Second, which kind of information resource (video vs. website) do they preferably choose, when being asked to recommend one most suitable information resource for learning purposes, and how do they justify this recommendation *(RQ2)*? And third, (how) do video recommenders differ from website recommenders in their retrieval behavior *(RQ3)*?

2 METHOD

2.1 Participants

Participants were 38 undergraduate students (32 female, M = 23.76 years, SD = 3.10) from different majors at a German university. For their participation, students were rewarded with 8€. Participants reported on a 5-point scale from 1= "very rarely" to 5= "very often" to frequently use the internet (M = 4.76, SD = 0.59) and search engines (M = 4.37, SD = 0.75).

2.2 Tasks

In the experiment, each participant had to perform the following two procedural knowledge web search tasks: one about how to tie a Figure-eight knot (sensori-motor procedural knowledge), and one about the function of the while loop in Python (cognitive procedural knowledge). The final purpose of each task was to choose one most suitable information resource that participants would recommend to a friend who wants to learn about these issues. This fictitious friend was said to have no prior knowledge about the respective topics. For each search task, participants were provided with a mockup Google SERP with eight search results linking to four videos and four websites. Participants could access all eight information resources via the Firefox browser within an overall time limit of 8 min. per task. After their recommendation, participants were asked to provide a written justification for it. Once having completed the first topic, they continued with the second topic in the same procedure. The order of the two topics was counterbalanced between participants.

2.3 Materials

The starting point for each task was a mockup Google SERP. The constructed SERP consisted of a header with the query box and predefined search terms (which could not be altered), the core of the SERP with eight search results, and the footer without any navigation options. The used search terms in the query box were for the while-loop task "while Python" and for the knot task the German word for Figure-eight knot ("Achterknoten"). The eight search results consisted of four YouTube videos and four website results, which were presented in an alternating order (for an example screenshot, see Figure 1). For half of the participants, a video result was presented first (for both tasks), for the other half a website result. So, in total we constructed two SERPs for each task, which differed only in the sequence of video and website results.

The average length of the videos was M = 4.80 min (SD = 1.66) for the while-loop task, and M = 2.59 min (SD = 1.03) for the knot task. The websites for the while-loop task, on average, contained M = 696.75 words (SD = 792.16) and M = 10.75 depictions of code segments (SD = 9.78). The websites for the knot task comprised M = 527.75 words (SD = 402.75) and M = 13.25 (SD = 2.99) photo images. The provided videos and websites were locally saved HTML copies of real search results. However, we disabled all further links on the websites and videos and removed advertisements. All interactive features of videos such as start, stop, and the timeline were still usable. All of the eight information resources contained relevant information about the procedure, but differed in their presentation quality.

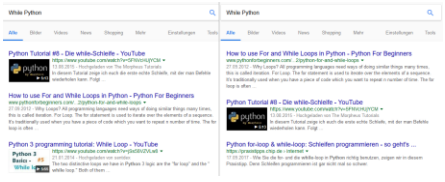

Figure 1: Excerpt of the SERP for the while-loop task (video first, left; website first, right)

To achieve this variety, in preparing the experimental materials, a real Google search was performed (website results were searched within the organic Google results; for video results the Google video index was used). From the obtained results, four

different videos and four different websites were selected for each topic (3 domain experts for each topic were consulted who assured the correctness of the explained procedures for all selected information resources). The order of the results was defined based on their original ranking position in the real searches.

2.4 Procedure

Participants were tested in individual sessions in a lab setting. First, they filled out a questionnaire about demographics and their internet and search engine usage. Participants were then positioned in front of an SMI remote eye tracking system attached to a 15.5-inch laptop (1920x1080 px). Each task started with a 9-point calibration followed by a written instruction of the task displayed in the browser. Participants were then forwarded to the SERP to perform the task. They were free to navigate between the SERP and the eight information resources linked to the search results. Time was limited to 8 min per task, to ensure that participants had to apply a selection strategy instead of just accessing all results. Participants could also end the task earlier. Subsequently, they were asked to select the most suitable information resource for their friend. After having made the selection, they were redirected to an HTML site showing their selected information resource and an open text box with the instruction to give a detailed justification for their selection. There was no time or character limit for the text box.

2.5 Dependent Measures

Navigation behavior. Based on logfile data, for each task we analyzed the number of visited video and website resources, the total time spent viewing these resources, and the total time spent on the SERP.

Viewing behavior on SERP. Areas of interests (AOIs) were defined around each of the eight search results to determine for how long (in ms) a participant was looking at a search result. For each task, the total fixation time on video search results and on website search results was determined.

Recommended information resource. For each task, we determined the number of participants who recommended a video resource or a website resource, respectively.

Written justification of recommended information resource. To investigate the reasons for recommending an information resource, participants' written justifications were analyzed for each task. Our developed coding scheme included the following justification criteria: (a) usefulness, (b) modality, (c) hedonic aspects, (d) credibility, and (e) scope. *Usefulness* was coded for statements that the presented information was useful for solving the respective task or presented it in a helpful and comprehensive way. Notions regarding the advantage of either videos or static text and images were coded as *modality*. Statements about emotional or entertaining aspects were coded as *hedonic aspects*. *Credibility* was coded for statements dealing with the assessment of the credibility of the information source, such as the perceived expertise or trustworthiness of the presenter in a video or the website. *Scope* was coded for statements that the information resource did not contain too much or unnecessary information. Two independent raters familiar with the search tasks and materials as well as with the coding scheme coded all 76 (38 per task) written justifications. For each written justification they coded whether or not a participant had addressed a criterion at least once. Divergent codings were resolved through discussion.

3 RESULTS

3.1 Navigation and Viewing Behavior

Analyses of the logfile data revealed that participants, on average, spent M = 5.88 min (SD = 2.20) on the while-loop task and M = 5.51 min (SD = 2.19) on the knot task, of which 54.77 sec (SD = 26.87) for the while-loop task and 45.70 sec (SD = 23.17) for the knot task, respectively, were spent on the SERP. With regard to the modality of information resources, during the while-loop task, participants spent M = 2.27 min (SD = 1.30) on website resources and M = 2.69 min (SD = 1.73) on video resources. Likewise, during the knot task, participants spent M = 2.16 min (SD = 1.54) on website resources and M = 2.60 min (SD = 1.74) on video resources. In the while-loop task, participants, on average, visited M = 3.32 (SD = 0.96) website and M = 2.63 (SD = 1.34) video resources, and in the knot task M = 3.16 (SD = 1.00) website and M = 2.79 (SD = 1.38) video resources. Moreover, eye-tracking analyses revealed that during the while-loop task, participants fixated (i.e., visually inspected) video search results for M = 19.54 sec (SD = 11.00) and website results for M = 21.12 sec (SD = 10.38). For the knot task, video search results were fixated for M = 17.78 sec (SD = 9.74) and website search results for M = 17.07 sec (SD = 8.78).

3.2 Recommended Information Resource

In the knot task, 24 participants (i.e., 63.16%) recommended a video resource. In the while-loop task, 16 participants (i.e., 42.11%) recommended a video resource. Whether the first search result on a SERP was a website or a video did not affect these recommendations (knot: $\chi^2(1) = 0.45$, $p = .501$; while-loop: $\chi^2(1) = 0.43$, $p = .511$).

3.3 Written Justification for Recommendation

The most frequently mentioned reason for recommending a resource was with 100% for the while-loop task and with 92.1% for the knot task the criterion *usefulness*. An exemplary statement of this criterion is: *"[…] the explanation was clear and easy"*. The second most frequently mentioned criterion was *modality*, which was mentioned in 34.2% of the justifications for the while-loop task (6 video recommenders and 7 website recommenders) and 39.5% of the justifications for the knot task (11 video recommenders and 4 website recommenders). Most of the *modality* statements addressed general advantages or disadvantages of static or dynamic representations. An exemplary statement addressing perceived advantages of videos is: *"Within a video all processes are shown. In addition, you only have to concentrate on watching, because the person from the video explains the steps simultaneously. Besides, you can stop the video at any time*

if it's going too fast." In contrast, an exemplary statement addressing advantages of static images in websites is: "*The advantage of images is that I can look at them as long as I want to, in order to knot the rope, and I don't have to press pause or go back as in videos.*" An exemplary statement addressing perceived advantages of textual information in websites is: "*Text in this case is better than video, because it is easier to read the text several times to understand what you have read, with video it is tedious to always start from the beginning.*" Regarding the criterion *hedonic aspects*, no respective statements were found for the while-loop task. For the knot task, 18.4% of the participants (only video recommenders) mentioned hedonic aspects such as: "*It's fun to learn the Figure-eight knot when watching the video*". The criterion *credibility* was mentioned by 13.16% of the participants for the while-loop task and by 15.79% of the participants for the knot task. Typical examples were statements about the source, e.g.: "*This is a video of the fire department. They are real professionals […].*" The criterion *scope* was mentioned for the while-loop task by 18.4 % of the participants and for the knot task by 10.5% of the participants (e.g., "*The explanation does not contain any unnecessary information.*").

3.4 Differences Between Video and Website Recommenders

Finally, we analyzed whether participants who recommended a video versus a website resource differed in their viewing and navigation behavior during the search tasks. Specifically, we analyzed (a) the proportion of time participants spent fixating video (vs. website) search results and (b) the proportion of time they were staying on video (vs. website) resources, as a function of whether they had recommended a video or a website resource. For the while-loop task, video recommenders spent proportionally more time fixating video search results (M = 51.8% (SD = 7.8) than website recommenders (M = 43.9%, SD = 7.0), $t(36) = 3.28$, $p = .002$, and also proportionally more time viewing video resources (M = 64.1%, SD = 15.4) than website recommenders (M = 40.9%, SD = 23.8), $t(35.60) = 3.65$, $p = .001$. For the knot task the same pattern occurred, with video recommenders having spent proportionally more time fixating video search results (M = 55.4%, SD = 8.4) than website recommenders (M = 41.6%, SD = 5.3), $t(36) = 5.53$, $p < .001$, and also proportionally more time viewing video resources (M = 65.4%, SD = 18.4) than website recommenders (M = 25.4%, SD = 21.3), $t(24.26) = 5.88$, $p < .001$.

4 DISCUSSION

The results of the present paper indicate that individuals who are asked to search for procedural knowledge on the web and who are presented with video and website results in an alternating order on SERPs, to a considerable extent access video resources. Moreover, particularly for the sensori-motor procedural task, about two thirds of the participants recommended a video resource for learning the respective procedure (while for the cognitive procedural task it were only 42%). This might be explained by the fact that for sensori-motor procedures it is of particular importance to see how the procedure is being performed. Furthermore, participants' written justifications revealed that for both tasks about one third of the participants mentioned the modality of the information resource to justify their recommendation decision. To conclude, the modality of search results and information resources, respectively, at least to some extent seems to play a role in individuals' selection decisions when searching for procedural knowledge on the web. Moreover, consistently for both tasks, the results indicate that participants who recommended a video resource already inspected video search results on the SERP and the respective video resources for a greater proportion of time than website recommenders. Thus, it seems that individuals might have a preference for one or the other type of information resource (i.e., videos vs. conventional websites) already during the search process.

5 CONCLUSION AND FUTURE WORK

This short paper provides first evidence that the modality of search results plays a certain role during web search for procedural learning resources, as indicated by both individuals' retrieval behavior and their written justifications. Thus, such observed preferences for particular kinds of information resources during web search might be used as predictors to provide individuals a personalized composition and structure of SERPs. Yet, further research is needed that aims at replicating our results with larger samples, a larger number and greater variety of tasks, and when searching on the open web instead of being provided with a predefined environment. In addition, in this context, the effects of carousel-formatted video or image boxes should also be examined. Moreover, an important next step will also be to investigate, how individuals' search behavior relates to their learning outcome (e.g., their own performance of the procedures).

REFERENCES

[1] Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman.*

[2] Azzopardi, L., Thomas, P., & Craswell, N. (2018). Measuring the utility of search engine result pages: an information foraging based measure. In *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 605-614). ACM, New York, NY, USA.

[3] Bétrancourt, M., & Benetos, K. (2018). Why and when does instructional video facilitate learning? A commentary to the special issue "developments and trends in learning with instructional video". *Computers in Human Behavior*, *89*, 471-475.

[4] Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*, 1207-1217.

[5] Eickhoff, C., Teevan, J., White, R., & Dumais, S. (2014). Lessons from the journey: a query log analysis of within-session learning. In *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, New York, NY, USA.

[6] Lewandowski, D. (2017). Users' understanding of search engine advertisements. *Journal of Information Science Theory and Practice*, *5*, 6-25.

[7] Liu, Z., Liu, Y., Zhou, K., Zhang, M., & Ma, S. (2015, August). Influence of vertical result in web search examination. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 193-202). ACM, New York, NY, USA.

[8] Marchionini, G. (2019). Search, sense making and learning: closing gaps. *Information and Learning Sciences*, *120*(1/2), 74-86.

[9] Marcos, M. C., Gavin, F., & Arapakis, I. (2015). Effect of snippets on user experience in web search. In *Proceedings of the XVI International Conference on Human Computer Interaction* (pp. 471-478). ACM, New York, NY, USA.