Supporting Information Search by Older Adults

Herre van Oostendorp
Utrecht University
Utrecht, The Netherlands
h.vanoostendorp@uu.nl

Saraschandra Karanam
Utrecht University
Utrecht, The Netherlands
s.karanam@uu.nl

ABSTRACT

Using cognitive models of web-navigation to generate support has long been a topic of research. In this paper, we address two limitations in this area. First, these models have so far been used to generate support for navigation within a website and not for interaction with a search engine. Second, very few studies have looked at the usefulness of such model-generated support for older adults who are known to be less efficient than younger adults. An experiment with 24 younger and 24 older adults on six simple and six difficult information search tasks was conducted. Results showed that the semantic relevance of queries showed a decreasing trend across reformulations for older adults and remained constant for younger adults, indicating that as older adults reformulated, they produced queries that were further away from the target information, which could be the reason for their lower efficiency. Based on these outcomes, two types of model-generated support mechanisms for interaction with a search engine are proposed, one which visually highlights the most relevant search result given a query and the other which monitors the average semantic relevance of search results for a given query and warns the user if it falls below a threshold.

CCS Concepts

•**Information systems** → **Personalization;** •**Human-centered computing** → **HCI theory, concepts and models;**

Keywords

Aging; Cognitive Model; Support; Information Search

1. INTRODUCTION

A typical information search task is usually carried out in two phases: first, using a search engine and formulating queries relevant to one's goal (search by query) and second, navigating on websites opened by clicking on one or more of the search results (search by navigation) [18]. Both phases involve several cognitive processes such as memory, attention, problem solving, comprehension and decision making. These cognitive processes are known to be affected by one or more cognitive factors such as aging. But the focus

of information retrieval systems so far has been only on retrieving relevant results and largely follow a one-size-fits-all model without taking into account the variations caused by the cognitive factors. Unlike traditional information retrieval systems, cognitive models of web-navigation take into account all the cognitive processes and the cognitive factors that influence those cognitive processes involved in an information search task to predict information search performance in a precise and automated way. The focus of these cognitive models is on the process that leads to target information and not only on retrieving the correct answer. Because of the elaborate description of the processes underlying information search, these cognitive models have been used effectively to generate automated support for improving navigation performance of users. However, they have never been used to support interaction with a search engine. Also, they have not been tested on older adults who are known to be less efficient than younger adults due to the natural decline in their cognitive capacities. These two aspects form the main focus of this paper. We give a brief description of the two cognitive models of web-navigation relevant to our work, summarize the outcomes of recent experiments that used navigation support based on these models, report related work on aging and search performance, conduct our own experiment investigating information search performance of older adults and propose two support tools motivated by the outcomes of the experiment.

1.1 Cognitive Models of Web-Navigation

CoLiDeS or Comprehension-based Linked Model of Deliberate Search developed by Kitajima et al. [15] is based on Information Foraging Theory [20] and connects to the Construction-Integration Reading Model of Kintsch [14]. In CoLiDeS it is assumed that the evaluation and selection of hyperlinks is based on information scent. Information scent is operationalized as the semantic similarity between the user goal and each of the hyperlinks. Based on the semantic similarity values, the model predicts that the user is most likely to click on that hyperlink which has the highest semantic similarity value. This process is repeated for every new page until the user reaches the target page. CoLiDeS uses Latent Semantic Analysis (LSA) introduced by Landauer et al. [16, 17], to compute the semantic similarities. LSA is a machine-learning technique that builds a semantic space representing a given user population's understanding of words, short or whole texts by applying statistical computations, singular value decomposition and represents them as a vector in a multidimensional space of about 300 dimensions. The cosine value (+1 if identical and 0 if unrelated) between two vectors in this representation gives the measure of the semantic relatedness. It has been shown that higher semantic relevance between two texts relates to higher overlap in the meanings associated with those two texts. The higher the LSA value of a hyperlink is, the higher the probability of a user clicking on that hyperlink is. These similarity

values are based on many contexts in which given words do and do not appear. The LSA extraction method of meaning appears to be a powerful predictor of human meaning-based judgements [16, 17], although, of course, individual deviations will occur. A further important consideration in taking into account the individualistic aspect is that we start with the queries that are formulated by the participants. These form individual-based conceptions of the task.

CoLiDeS+ [9] further augments CoLiDeS by incorporating the concept of path adequacy and navigation strategies as a complement to the concept of information scent. It defines path adequacy as the semantic similarity between the user goal and the navigation path. If information scent and path adequacy do not increase during navigation, a latent impasse is said to have occurred and CoLiDeS+ invokes backtracking strategies: i.e., backtracking to other regions within the same page and eventually to the previously visited pages. When both CoLiDeS and CoLiDeS+ were tested on real websites, CoLiDeS+ was found to not only locate the target page more often than CoLiDeS but also reach closer to the target page than CoLiDeS, whenever the target page was not located [10].

1.2 Support based on Cognitive Models of Web-Navigation

The CoLiDeS model has been used to build a tool for designers called Auto-CWW that has been successfully applied in finding usability problems, by predicting hyperlinks that would be unclear or confusing to users [1]. The designers could then make use of this to make corrections in the website's hyperlinks. Juvina & Van Oostendorp [24] used predictions made by CoLiDeS+ to provide suggestions in real time during a navigation session in two modalities: auditory and visual. Participants who received support were significantly faster and more accurate in answering questions and were less disoriented compared to participants who did not receive any support. Similar outcomes were found by later studies with extended versions of CoLiDeS [11, 23]. The ScenTrails system developed by [18] integrates searching and browsing into one interface by utilizing the query terms generated by users and the website architecture to highlight relevant hyperlinks both on the search results page and the intermediate pages leading to the target page. ScenTrails enabled users to find information significantly more quickly than by either searching alone or browsing alone. This system, however, assumes knowledge of the entire collection of websites to generate support.

Please note that all the above studies were focused on support for navigation within websites and not interaction with a search engine and that is what we will examine in this paper. Also, they did not focus on the usefulness of model-generated support for older adults. A recent study by Chin & Fu [5] on age differences in navigation performance within a website, showed that providing recommended links based on browsing histories helped older adults in improving their performance. This study neither involved interaction with a search engine nor the support generated was from a cognitive model. We will finally look at the impact of aging on information search performance.

1.3 Aging and Information Search Performance

Older adults are known to be less effective than younger adults in solving information search tasks on the web due to the natural decline in their motor skills and fluid intelligence involving processing speed, cognitive flexibility or ability to switch processing strategies, attentional control and visuospatial span [7, 8, 25]. Many studies have shown that older adults generate less queries, use less keywords per query, reformulate less, spend longer time evaluating the search results, spend more time evaluating the content of

websites opened from SERPs, switch less number of times between SERPs and websites and find it difficult to reformulate unsuccessful queries [6, 19, 21]. However, older adults are known to adapt their search and navigation strategies exploiting their higher crystallized knowledge to compensate for their decreased fluid capabilities when performing information search tasks using the Internet.

Youmans et al. [26] conducted a study in which older adults were found to follow a more structured and methodical approach to information search such as careful selection of query terms and spending longer time evaluating the search results whereas younger adults were found to follow a more impulsive strategy that involved more switches between search results and websites and more clicks on irrelevant search results. A more recent study by [3] found significant differences in the amount of time allocated to the three phases of information search: planning, controlling and evaluating in relation to age and task difficulty. Younger adults were found to control their strategy more than older adults enabling them to perform better especially at difficult and impossible tasks. In another study [4], it was found that older adults were found to do less exploration (number of search results opened for any given query) and more exploitation (number of hyperlinks within those websites visited) in terms of spending longer time and viewing more information compared to younger adults. They even found that the older adults were adaptive in adjusting the two processes depending on the difficulty of the task.

2. RESEARCH QUESTIONS

What was overlooked in the aging research in relation to information search so far is the semantic aspect of the queries themselves. In this paper, we examine the content of the queries during reformulations in relation to age and difficulty of the task. In summary, the main research questions of the study were:

(a) How does the semantic relevance of a query with the target information sought vary in relation to (i) age and (ii) difficulty of a task? (i) Based on the outcomes of the study by [3], in which younger adults were found to be better at the controlling phase of information search (modifying the query and strategy if required, which resulted in their higher performance) compared to older adults, we hypothesize that the semantic relevance of queries generated by younger adults would be higher than that of older adults. (ii) The above mentioned Construction-Integration Model of text comprehension [14] states that comprehension of text happens cyclically in two phases: *construction phase* during which a representation of a new piece of text and all possible meanings of the text are generated and *integration phase* during which using reader's prior knowledge and context, a single coherent meaning is selected by filtering out all those representations that do not fit the context. We assume that this model also applies to information search and we hypothesize that solving difficult search tasks successfully urges users to activate a greater number of related concepts from Long Term Memory into Working Memory [14]. Assuming that these activated concepts are included as terms in the search queries formulated by users, we hypothesize that the semantic relevance of a query with the target information would be higher for difficult tasks compared to simple tasks.

(b) Can we propose support tools for interacting with a search engine based on the outcomes of the semantic relevance of queries?

3. EXPERIMENT

We briefly report here the results of an experiment conducted to collect behavioral data from participants and check if the results are in line with prior outcomes from aging-related literature. Next to that we report the main issue of this study: the content of the queries during reformulations in relation to age and task difficulty and the support tools based on the outcomes of the analysis.

3.1 Method

3.1.1 Participants

24 young participants (16 males and 8 females) ranging from 18 to 31 years ($M = 22.7$, $SD = 3.31$), and 24 old participants (14 males and 10 females) ranging from 65 to 88 years ($M = 73.58$, $SD = 6.74$) participated in the study.

3.1.2 Design

We followed a 2 (Age: Young vs. Old) X 2 (Task Difficulty: Simple vs. Difficult) mixed design with age as between-subjects variable and task difficulty as within-subjects variable.

3.1.3 Material

The experiment was conducted with twelve simulated information search tasks [2]: six simple and six difficult, all from the domain of health. For simple tasks, participants in most cases could find the answer easily either in the snippets of the search engine results or in one of the websites referred to by the search engine results. For difficult tasks, users had to frame queries using their knowledge and understanding of the task, the answer was not easily found in the snippets of search engine results and often they had to evaluate information from multiple websites. The tasks were all presented in Dutch.

3.1.4 Procedure

Participants first completed a demographic questionnaire in which they were asked details about their age, gender, familiarity with search engines (on a Likert scale of 1(A bit) to 4 (Very Much)) and computer experience (number of years). Based on these self-reported ratings, old participants (M=2.63, SD=1.01) were found to be significantly less experienced with search engines than young participants (M=3.75, SD=0.68) $t(46) = 4.52$, p<.001. Old participants were significantly longer experienced with computers (M=22.7 years, SD=9.43) than young participants (M=15.33 years, SD=4.09) $t(46) = 3.1$, p<.005. They were next presented with a fluid intelligence test: a computerized version of the Trail Making Test [22]. Older participants took significantly longer to finish this test ($t(46)$=5.13, p<.001) than younger participants indicating that they had significantly lower fluid abilities compared to young participants.

After a break of five minutes, participants were presented with twelve information search tasks (six simple and six difficult) in a counter balanced order. They were first shown the task and then directed to the home page of Google's search engine. They were not allowed to use any other search engine. Participants could enter queries as they normally would on any browser and the corresponding search results appeared. The task description was made available to the participant at all times in the top left corner. An empty text box was provided in the top right corner for the participant to enter his/her answer. All the queries generated by the users, the corresponding search engine result pages and the URLs opened by them were logged in the backend using Visual Basic.

3.2 Analysis of Search Outcomes

3.2.1 Measures

Task-completion time is computed from the moment of opening a browser and typing in the first query to the moment of answering the question. This includes the time it takes to type in queries, evaluating search results, clicking on one of the search results, evaluating the content of websites opened from the search results and finally typing the answers.

Accuracy is measured as 1, 0.5 or 0 depending on whether the participant's answer was correct (in which case the score is 1) or partially correct (in which case the score is 0.5) or wrong (in which case the score is 0).

Data for only those tasks for which the participants correctly completed the tasks was included in the analyses. 37 data points out of a total number of (12 tasks X 48 participants) = 576 data points (6.4%) were therefore dropped. For the two dependent variables, a 2 (Age: Young vs. Old) X 2 (Task Difficulty: Simple vs. Difficult) mixed ANOVA was conducted with age as between-subjects variable and task difficulty as within-subjects variable.

3.2.2 Results

Task-completion time Figure 1a shows the mean task-completion time in relation to age and task difficulty. The main effect of age was not significant (p>.05). The main effect of task difficulty was significant $F(1,46) = 90.07$, p<.001. The interaction of age and task difficulty was also significant $F(1,46) = 11.63$, p<.001. We repeated the above analysis separately for the time spent on SERPs as well as the time spent on websites. For the time spent on SERPs, results showed that there was a significant main effect of age $F(1,46) = 18.13$, p<.001 and a significant main effect of task difficulty $F(1,46) = 74.31$, p<.001. Only the interaction of age and task difficulty was significant ($F(1,46) = 6.24$, p<.05) for the time spent on websites. All the other effects were not significant (p>.05).

These outcomes indicate that difficult tasks demanded significantly more time (total time and time spent on SERPs). There was no significant difference in the time spent on websites between simple and difficult tasks. In terms of total task-completion time, old participants were found to take significantly more time than young participants when performing simple tasks. Also, when the time spent on SERPs and the time spent on websites are analyzed separately, old participants were found to spend significantly more time than young participants on SERPs (for both simple and difficult tasks) and significantly less time than young participants on websites when performing difficult tasks.

Accuracy As can be seen from Figure 1b, the main effect of age was significant $F(1,46) = 17.18$, p<.001. The main effect of task difficulty was significant $F(1,46) = 196.4$, p<.001. The interaction of age and task difficulty was not significant (p>.05). Accuracy of difficult tasks was significantly lower than that of simple tasks and the accuracy of old participants was significantly lower than that of young participants.

These results concerning task-completion time and task accuracy are in-line with the prior findings from aging-related literature and provide more evidence to the fact that old participants are less efficient than young participants.

3.3 Analysis of Content of Search Queries

We will use the notion of semantic relevance to analyze the content of search queries. Semantic relevance was used in the past as a metric to evaluate the content of hyperlink texts (to predict navigation behavior on websites) [1, 9] or the content of the search results on the SERPs (to predict interaction behavior with search
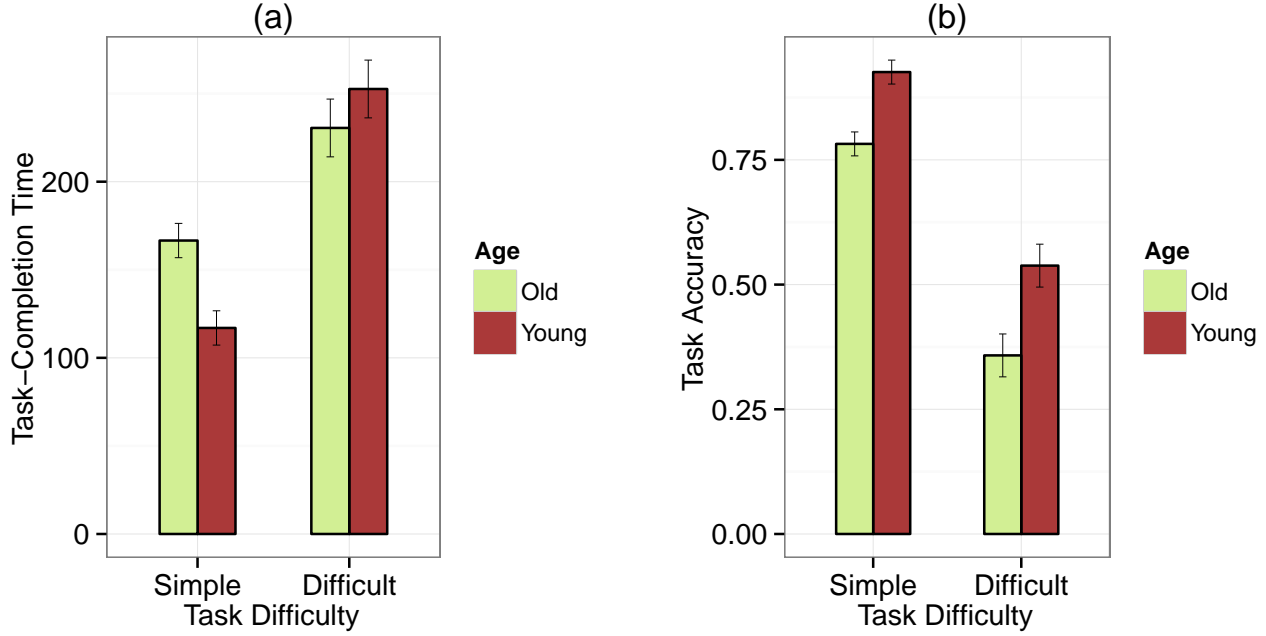
Figure 1: Analysis of search performance in terms of (a) task-completion time, (b) accuracy in relation to age and task difficulty.

engines) [13]. In both cases, semantic relevance was computed in relation to the user goal and not with the target information. In this study, we compute the semantic relevance of the queries with the *target* information sought. The main reason is that we know in current context rather precisely what the correct answer (i.e. the target information) is to a given task which is normally not the case, and we want to know how closely related this target information is to the query that is formulated by the participant.

3.3.1 Measures

Semantic Relevance of Query (SRQ): For each task and each query corresponding to that task, semantic relevance was computed between the query and the target information sought using LSA. We used 70,000 Dutch documents (consisting of 60% newspaper articles and 40% medical and health related articles) as a corpus to create a semantic space in Dutch. The LSA values were then computed between a query and each of the headings and search snippets of search results from the corresponding SERP using this space. For a detailed procedure to compute semantic similarities, refer [12]. This is repeated for all queries of the task and a mean LSA value is computed. This is repeated again for all the tasks of a participant and finally for all the participants. This metric gives us an estimation of how close the queries participants generated in terms of semantic similarity, were to the target information.

3.3.2 Results

Semantic Relevance of Query: A 2 (Age: Young vs. Old) X 2 (Task Difficulty: Simple vs. Difficult) mixed ANOVA was conducted with age as between-subjects variable and task difficulty as within-subjects variable and mean semantic relevance of query as dependent variable. It is evident from Figure 2 that the main effect of task difficulty was significant $F(1,46) = 9.47$, $p<.005$. The mean semantic relevance of queries for difficult tasks was significantly higher than that for simple tasks. The main effect of age and the interaction of age and task difficulty were not significant ($p>.05$).
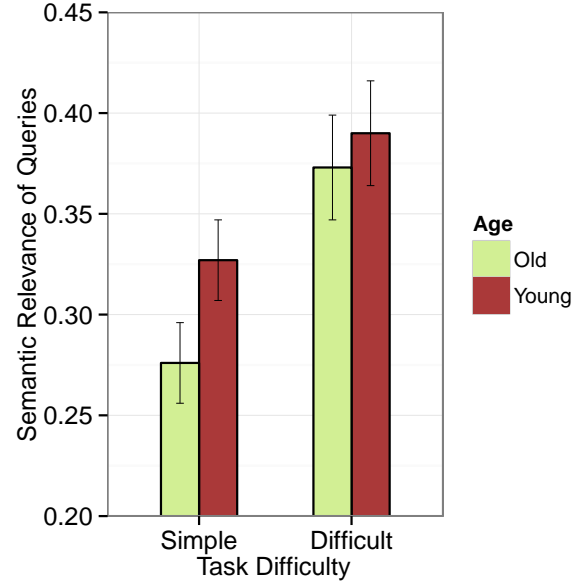

Figure 2: Mean semantic relevance of queries with target in relation to age and task difficulty.

We next analyzed the mean semantic relevance of the queries with the target information at a more granular level by looking at each reformulation cycle separately. First cycle corresponds to the first query, second cycle corresponds to the second query, third cycle corresponds to the third query and so on. Mean semantic relevance (SRQ) was computed for all the queries of all the tasks of a particular type (simple and difficult separately), generated by young and old participants in each reformulation cycle. To achieve
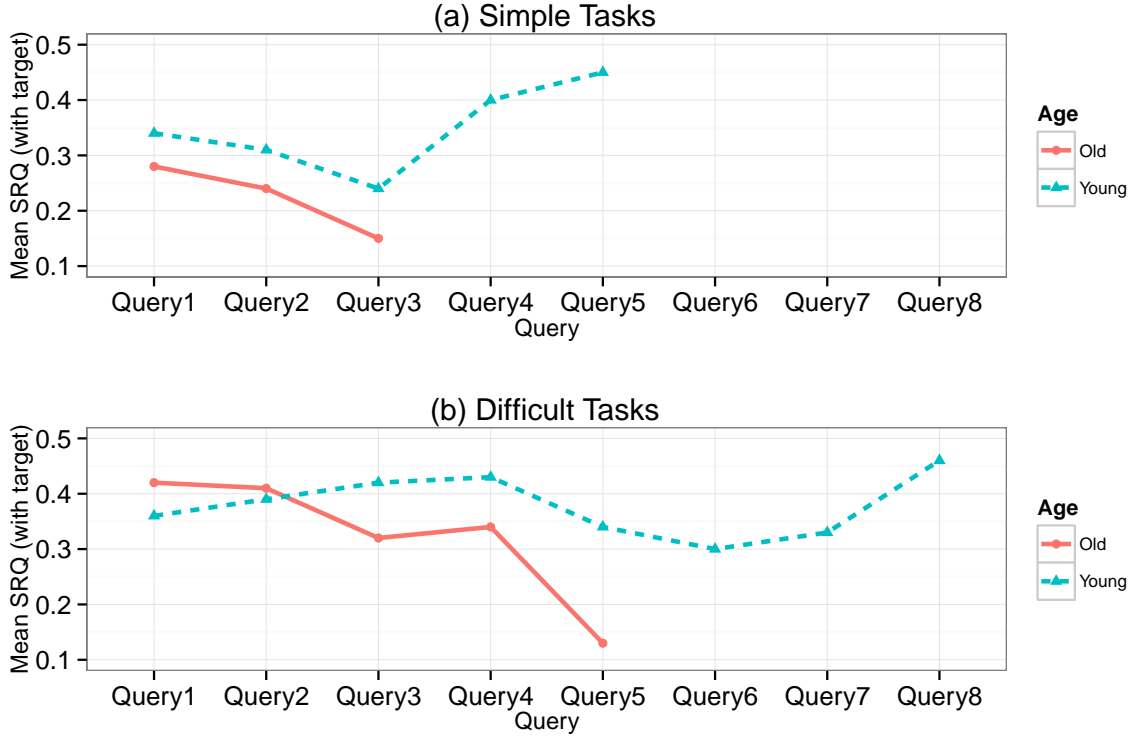
Figure 3: Mean semantic relevance of queries with target at each reformulation cycle for (a) simple and (b) difficult tasks.

higher reliability, only those cycles for which there were at least 5 queries (per reformulation cycle), were considered. By doing so, only 2.8% of data was excluded from the analysis. The resulting graphs are shown in Figure 3. We can observe from Figure 3 that young participants reformulated much longer than old participants.

We conducted 4 different one-way ANOVAs (for simple old, difficult old, simple young and difficult young) with cycle (reformulation cycle 1 vs 2 vs 3 and so on as repeated measures) as independent variable and mean LSA value of the SERP generated by queries in a reformulation cycle as dependent variable. The main effect of cycle on mean LSA value was weakly significant $F_{(2,56)} = 2.55$, $p<.09$, for simple old and not ($p>.05$) for others. For simple old, post-hoc tests showed that the mean LSA value of the SERPs in reformulation cycle 3 was significantly less ($p<.05$) than that in the first cycle. There was also a significant linear trend $F_{(2,56)} = 5.03$, $p<.05$, for simple old and a weaker trend for difficult old $F_{(1, 66)} = 3.52$, $p<.07$, indicating that as old participants reformulated, they produced queries which were further away from the target information.

We next checked the difference in mean LSA value of the SERPs between the starting (first query) and ending points (last query) of a search session. We found no significant difference in the starting points of young and old participants, for both simple and difficult tasks. However, at the ending points, young participants ($M=.45$, $SD=.45$, for simple tasks and $M=.48$, $SD=.40$ for difficult tasks) end up significantly higher than old participants ($M=.14$, $SD=.17$ for simple tasks and $M=.12$, $SD=.14$ for difficult tasks), $t(20) = 2.3$, $p<.05$ (for simple tasks), $t(8) = 2.28$, $p=.06$ (for difficult tasks), indicating that the semantic relevance of queries with target information decreases with further reformulations for old participants whereas it remains constant for young participants.

4. SUPPORT TOOLS

Based on the behavioral outcomes and the analysis of the content of search queries during reformulations, we come up with the design and methodology of constructing two types of automatic tools that can support interaction with a search engine.

4.1 Support Tool 1

This tool is intended to reduce the amount of effort involved in the process of selecting a search result from the SERP. We found that though older adults spend much longer time than younger adults in evaluating the SERPs, their accuracy of finding the target information is still less than that of younger adults (See Figure 1). One of the possible reasons could be that they are unable to differentiate between a relevant and a non-relevant search result as efficiently as the younger adults do. We propose a support tool that visually highlights the most relevant search results for a given query, as shown in Figure 4. This methodology was successfully used in the past to provide navigation support within websites [9, 11, 23]. We propose to extend the same methodology to generate support for interaction with a search engine.

Given a query, semantic relevance is computed between the query and each of the search results on the basis of LSA. The search result with the maximum semantic relevance is highlighted with a green arrow as shown in Figure 4. This form of support would enable older adults to spend less mental resources in differentiating between a relevant and a non-relevant search result which in turn would lead to better accuracy. One can even think of providing the next-best search result (in terms of semantic relevance to the query) only on demand, i.e., if the user is not satisfied with the most relevant search result, he/she will have the possibility of requesting the support tool to reveal the next-best search result.
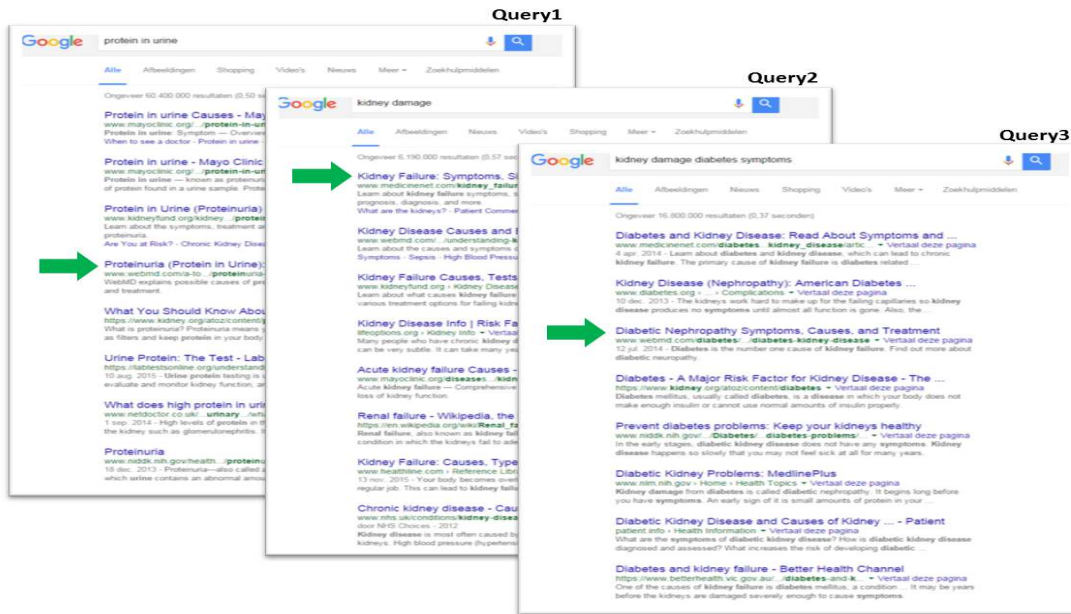
Figure 4: Design of a support tool for interacting with a search engine that highlights the most relevant search result for a given query.

Figure 5: Design of a support tool for interacting with a search engine that monitors the average semantic relevance of the user's queries to the goal information

4.2 Support Tool 2

Above tool is focused on supporting to select the best search result, next tool focused on improving the queries. This support tool is intended to ensure that a user does not digress too far away from the goal information in the form of irrelevant queries. Analysis of the content of search queries showed that the semantic relevance of the queries generated by younger adults remained constant and that of the queries generated by older adults decreased as they reformulated further. In other words, older adults were going further away in semantic distance from the target information as they reformulate, leading probably to lower accuracy (See Figure 3). To address this problem, we propose a second support tool that monitors - based

on the LSA value - the average semantic relevance of the SERPs with the goal information derived from the query and warns the user when it falls below a threshold as shown in Figure 5. This form of support would indicate to the users that their search results on a page are not relevant enough and they can use this information to take corrective actions such as generating a more relevant query.

5. CONCLUSIONS AND DISCUSSION

In this study, we conducted an experiment with 24 younger and 24 older adults on 12 information search tasks (six simple and six difficult) to analyze the content of the search queries during reformulations. The behavioral outcomes were in-line with prior literature on aging: difficult tasks demanded significantly more time (total time and time spent on SERPs). There was no significant difference in the time spent on websites between simple and difficult tasks. In terms of total task-completion time, older adults were found to take significantly more time than younger adults when performing simple tasks. When the time spent on SERPs and the time spent on websites are analyzed separately, older adults were found to spend significantly more time than younger adults on SERPs (for both simple and difficult tasks) and significantly less time than younger adults on websites when performing difficult tasks. Accuracy of difficult tasks was significantly lower than that of simple tasks and the accuracy of older adults was significantly lower than that of younger adults. Several cognitive factors could be responsible for the age effects we found, such as reduced fluid ability or less familiarity with search engines. On both we found in this study older adults to perform less well than younger participants. These patterns of results provide a useful reference point for our main research questions: how does the semantic relevance of a query vary in relation to age and task difficulty?

We did not find evidence to support our first hypothesis: there was no significant difference between the semantic relevance of queries generated by younger and older adults. However, a more granular analysis showed that the semantic relevance of the queries generated by younger adults remained constant and that of the queries generated by older adults decreased as they reformulated further. Analysis of the content of search queries during reformulations confirmed our second hypothesis: the semantic relevance of queries with the target information was found to be significantly higher for difficult tasks compared to simple tasks.

Based on the above outcomes we proposed two support tools for older adults which could potentially enhance their performance. First, we found that though older adults spend much longer time evaluating the SERPs, they are still less accurate than the younger adults. One of the possible reasons for this could be that they are unable to differentiate between a relevant and a non-relevant search result as efficiently as younger adults do. Therefore, a support tool that visually highlights the most relevant search result given a query was proposed.

Second, we found that the semantic relevance of queries generated by older adults decreased as they reformulate. This led us to propose our second tool that monitors the average semantic relevance of the search results on a page in relation with the goal information (the query) and warns the user when it falls below a pre-defined threshold.

It is important to stress that automatic, online and dynamic support tools for information search and navigation can be constructed based on the computations of LSA values or semantic relevance of queries. It is not our aim to build a robust, fully functional and scalable support tool. However, we are currently working on implementing a demonstrable proof of concept to test the usefulness of the tool and the principles behind construction of such tools.

7. REFERENCES

[1] M. H. Blackmon, D. R. Mandalia, P. G. Polson, and M. Kitajima. Automating usability evaluation: Cognitive walkthrough for the web puts lsa to work on real-world hci design problems. In T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, editors, *Handbook of Latent Semantic Analysis*, pages 345–375. Lawrence Erlbaum Associates Mahwah, NJ, 2007.

[2] P. Borlund and P. Ingwersen. The development of a method for the evaluation of interactive information retrieval systems. *Journal of documentation*, 53(3):225–250, 1997.

[3] A. Chevalier, A. Dommes, and J.-C. Marquié. Strategy and accuracy during information search on the web: Effects of age and complexity of the search questions. *Computers in Human Behavior*, 53:305–315, 2015.

[4] J. Chin, E. Anderson, C.-L. Chin, and W.-T. Fu. Age differences in information search: An exploration-exploitation tradeoff model. In *Proceedings of the Human Factors and Ergonomic Society (HFES 2015)*, pages 85–89, 2015.

[5] J. Chin and W.-T. Fu. Age differences in exploratory learning from a health information website. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3031–3040. ACM, 2012.

[6] A. Dommes, A. Chevalier, and S. Lia. The role of cognitive flexibility and vocabulary abilities of younger and older users in searching for information on the web. *Applied Cognitive Psychology*, 25(5):717–726, 2011.

[7] J. L. Horn. The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in. In *Aging and cognitive processes*, volume 8, pages 237–278. Springer Science & Business Media, 2012.

[8] J. L. Horn and R. B. Cattell. Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26:107–129, 1967.

[9] I. Juvina and H. van Oostendorp. Modeling semantic and structural knowledge in web navigation. *Discourse Processes*, 45(4-5):346–364, 2008.

[10] S. Karanam, H. van Oostendorp, and W. T. Fu. Performance of computational cognitive models of web-navigation on real websites. *Journal of Information Science*, 42(1):94–113, 2016.

[11] S. Karanam, H. van Oostendorp, and B. Indurkhya. Towards a fully computational model of web-navigation. In *Modern Approaches in Applied Intelligence*, pages 327–337. Springer, 2011.

[12] S. Karanam, H. van Oostendorp, and B. Indurkhya. Evaluating colides+ pic: the role of relevance of pictures in user navigation behaviour. *Behaviour & Information Technology*, 31(1):31–40, 2012.

[13] S. Karanam, H. van Oostendorp, M. Sanchiz, A. Chevalier, J. Chin, and W. T. Fu. Modeling and predicting information search behavior. In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics*, page 7. ACM, 2015.

[14] W. Kintsch. *Comprehension: A paradigm for cognition*. Cambridge University Press, 1998.

[15] M. Kitajima, M. H. Blackmon, and P. G. Polson. A comprehension-based model of web navigation and its

application to web usability analysis. *People and Computers*, pages 357–374, 2000.

[16] T. K. Landauer, P. W. Foltz, and D. Laham. An introduction to latent semantic analysis. *Discourse processes*, 25(2-3):259–284, 1998.

[17] T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch. *Handbook of Latent Semantic Analysis*. Mahwah,NJ: Erlbaum, 2007.

[18] C. Olston and E. H. Chi. Scenttrails: Integrating browsing and searching on the web. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(3):177–197, 2003.

[19] R. Pak and M. M. Price. Designing an information search interface for younger and older adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(4):614–628, 2008.

[20] P. Pirolli and S. Card. Information foraging. *Psychological Review*, 106(4):643–675, 1999.

[21] T. L. Queen, T. M. Hess, G. E. Ennis, K. Dowd, and D. Grühn. Information search and decision making: Effects of age and complexity on strategy use. *Psychology and Aging*, 27(4):817–824, 2012.

[22] E. Strauss, E. M. Sherman, and O. Spreen. *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press, USA, 2006.

[23] H. van Oostendorp and S. Aggarwal. Modeling and supporting web-navigation. *Journal of Interaction Science*, 3(1):1–12, 2015.

[24] H. van Oostendorp and I. Juvina. Using a cognitive model to generate web navigation support. *International Journal of Human-Computer Studies*, 65(10):887–897, 2007.

[25] J.-J. Wang and A. S. Kaufman. Changes in fluid and crystallized intelligence across the 20-to 90-year age range on the k-bit. *Journal of Psychoeducational Assessment*, 11(1):29–37, 1993.

[26] R. J. Youmans, B. Bellows, C. A. Gonzalez, B. Sarbone, and I. J. Figueroa. Designing for the wisdom of elders: age related differences in online search strategies. In *Universal Access in Human-Computer Interaction. User and Context Diversity*, pages 240–249. Springer, 2013.