

Modeling individual differences in information search

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

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

ABSTRACT

A number of cognitive processes are involved in the process of information search on the Internet: memory, attention, comprehension, problem solving, executive control and decision making. Several cognitive factors such as aging-related cognitive abilities, domain knowledge, spatial ability and need for cognition, etc. in turn influence either positively or negatively these cognitive processes. Traditional click models from information retrieval community that predict user clicks do not fully take into account the effect of the above cognitive factors. We propose to exploit the capabilities of computational cognitive models to simulate the effects of cognitive factors on information search behavior. In this direction, we present some ideas how to incorporate these factors into a computational cognitive model called CoLiDeS+. Preliminary analysis of our ideas on modeling and predicting individual differences in information search due to age and domain knowledge show promising outcomes.

Author Keywords

Computational Cognitive Models; Information Search;
Cognitive Factors; Individual Differences

ACM Classification Keywords

H.1.2. Information Systems: Human factors; H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Internet is transforming how we communicate, shop, get health-related advice, plan a vacation, entertain ourselves and all in all how we do just about anything. It is pervasive, ubiquitous and becoming an indispensable part of human life today. However, a number of barriers still exist preventing large scale adoption and optimal usage of the Internet. One of the reasons for the slow adoption is that a number of cognitive processes such as memory, attention, problem solving, comprehension, decision making and executive control are involved when searching for information on the Internet. These cognitive processes, in turn, are known to be affected by one or more cognitive factors such as age [5], domain knowledge [23], spatial ability [15], need for cognition [22] etc.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IHCI '16, December 07-09, 2016, Mumbai, India

© 2016 ACM. ISBN 978-1-4503-4863-8/16/12...\$15.00

DOI: <http://dx.doi.org/10.1145/3014362.3014363>

Technologies that enable us to find information from the Internet, such as search engines, however, assume a homogeneous user and follow a one-size-fits-all model. Many experiments examining the relationships between the cognitive factors that influence the cognitive processes underlying information search and the performance of users have been conducted by researchers from the domain of cognitive psychology. However, using laboratory studies to understand these relationships requires participation from real users, who are not always available. Also, as the number of factors in the experiment increase, the complexity of the experiment increases and the number of experiments required to investigate all possible relationships also increase. This is not only expensive but also time consuming and difficult to scale up. Models, on the other hand, allow us to simulate user behavior without performing the experiment, which then only serves as a verification of the model. Modeling and simulation of user behavior during information search has therefore been an active area of research in the information retrieval community. Many click models to simulate and predict user behavior during information search have been proposed [7]. Except for a few models like [30, 27, 14], most of them do not consider variations caused by cognitive factors. Moreover, they provide only limited *process* description. Our focus therefore in this paper, would be on computational cognitive models. The focus of computational cognitive models is on the process that leads to the target information and they are therefore more capable of providing opportunities to incorporate behavioral differences due to variations in cognitive factors. In this paper, we propose some preliminary ideas that can be incorporated into computational cognitive models to simulate the behavioral differences in information search performance due to the variations in cognitive factors such as age, domain knowledge, spatial ability and need for cognition. The main research question of this study was: would a cognitive model that adapts itself to the variations in a cognitive factor predict individual differences in information search behavior (caused by that cognitive factor) more accurately than a cognitive model that does not?

The remaining part of the paper is organized as follows: in the first half, we describe the process of information search and the cognitive processes underlying it. We introduce two computational cognitive models CoLiDeS [18] and CoLiDeS+ [15], which are based on well tested theories of cognitive psychology and cognitive science on information search and give a process view of information search. We list some of the cognitive factors and their impact on information search performance. For each cognitive factor, we discuss how its ef-

fect on information search performance could be incorporated into CoLiDeS+. In the second half of the paper, we evaluate our ideas to incorporate individual differences in information caused by two cognitive factors age and prior domain knowledge. Evaluation is based on the number of matches between model predictions and the actual user behavior. Details of an experiment conducted to collect actual behavioral data and the outcomes of modeling are described. The paper concludes by summarizing the main outcomes of this study and providing potential directions for future work.

INFORMATION SEARCH

A typical process to search information on the Internet proceeds as follows: the user opens a search engine URL such as <http://google.com> with a question or a goal in mind. Based on the understanding of the goal and his/her prior knowledge, the user formulates a search query. The search engine retrieves the top-10 (and more) most relevant documents and presents them to the user (commonly called search engine result pages or SERPs, henceforth). The user evaluates one or more of these results shown by the search engine in order to find the search result that will lead to the target information. At this stage, based on the evaluation of SERPs, the user can take two actions: either click on one of the search results or reformulate the search query to get a fresh set of SERPs. If the user does the latter, the process repeats. But if s/he does the former, i.e., if the user clicks on any one of the search results, then a website corresponding to that search result is opened. At this stage again, user has two choices: either come back to the SERPs or click on more hyperlinks within the opened website and evaluate its content more deeply. This process goes on until either the user finds the information s/he was looking for or is frustrated and leaves (refer also the model by Sharit et al. [26]).

Several cognitive processes such as memory (activation of background knowledge or re-activation of information seen before), attention (understanding the visual layout of the website or search engine result pages), comprehension (evaluating the relevance of search results, understanding the content of websites), problem solving (information can be found in multiple locations and there could be multiple paths leading to them), decision making (choosing a relevant search result) and executive control (reformulating unfruitful queries, backtracking to earlier pages, comparing new information with what was found earlier) are involved during an information search process. We shall see in the later sections how variations in cognitive factors lead to variations in one or more of these cognitive processes. In the next section, we briefly describe two computational cognitive models CoLiDeS and CoLiDeS+.

COGNITIVE MODELS

Based on theories of cognitive psychology and cognitive science, the main goal of cognitive models of information search is to use well-tested cognitive mechanisms to characterize information search behaviour.

CoLiDeS

CoLiDeS, or Comprehension-based Linked Model of Deliberate Search, developed by Kitajima et al. [18] explains user

navigation behaviour on websites. It divides user navigation behavior into four stages of cognitive processing: parsing the webpage into high-level schematic regions, focusing on one of those schematic regions, elaboration / comprehension of the screen objects (e.g. hypertext links) within that region, and evaluating and selecting the most appropriate screen object (e.g. hypertext link) in that region. The focus of the modeling is on the navigation process based on the evaluation and selection of links (the fourth phase in the CoLiDeS model). CoLiDeS is based on Information Foraging Theory [24] and connects to the Construction-Integration theory of text-comprehension [17]. The notion of *information scent*, defined as the estimate of the value or cost of information sources represented by proximal cues (such as hyperlinks), is central to CoLiDeS. It is operationalized as the semantic similarity between the user goal and each of the hyperlinks. The model predicts that the user is most likely to click on that hyperlink which has the highest semantic similarity value with the user goal, i.e., the highest information scent. This process is repeated for every new page until the user reaches the target page. CoLiDeS uses Latent Semantic Analysis (LSA, henceforth) introduced by [20] to compute the semantic similarities. LSA is an unsupervised machine learning technique that employs singular value decomposition to build a high dimensional semantic space using a large corpus of documents that is representative of the knowledge levels of the target user group. The cosine value (+1 if identical and 0 if unrelated) between two vectors in this high dimensional space 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. The CoLiDeS model described above has been successful in simulating and predicting hyperlink selections made by users during navigation [18]. The model has also been successfully applied within the domain of Human-Computer Interaction for finding usability problems, by predicting links that will be unclear to users [1, 2, 19]. However, please note that CoLiDeS does not incorporate information from hyperlinks already clicked by a user in a session to predict which hyperlink he/she would click next in the same session. Also, backtracking behavior, which is known to be quite common among users on the Internet [8], is not implemented.

CoLiDeS+

CoLiDeS+ [15] shares the main theoretical foundations (Construction-Integration theory of text-comprehension [17] and Information Foraging Theory [24]) on which it is based with its predecessor CoLiDeS [18]. CoLiDeS+ further augments CoLiDeS and makes it more consistent with its theoretical assumptions by incorporating context and backtracking strategies. For example, when reading a text, it has been shown that contextual information helps users in comprehending new incoming sentences better, especially those with potentially multiple interpretations [4]. Analogously, when interacting with a search engine or navigating on a website, users often encounter information that is varying in its degree of ambiguity.

CoLiDeS+ takes a task description as input and assumes it to be a representation of the user goal. The model parses a web-page into several regions and a particular region is focused on (e.g., a set of hyperlinks). The set of hyperlinks are comprehended (based on how semantically similar to the user's goal they are) and one hyperlink (the one that is most similar to the user's goal) is selected. This opens a new web-page and if it is not the page with the target information, the cycle is repeated. Until this step, CoLiDeS+ runs exactly in the same fashion as CoLiDeS. However, CoLiDeS+ retains in memory the selected links which are used, starting from the second cycle, to compute the *navigation path*. Thus, the navigation path is defined as the sequence of hyperlinks clicked by a user at any given moment. CoLiDeS+ also computes a new metric called *path adequacy* (PA). Path adequacy is defined as the semantic similarity between the user goal and the navigation path. CoLiDeS+ incorporates contextual information using path adequacy as follows: only if the information from an incoming hyperlink increases in information scent (i.e., the semantic similarity with the user-goal), CoLiDeS+ considers it for selection. If it does not increase in information scent, path adequacy is checked. If path adequacy increases, then the incoming hyperlink is selected even when it does not increase in information scent. In other words, first semantic similarity is locally evaluated based on information scent, and only when it is not satisfying, a more effortful evaluation of the context is performed by checking the path adequacy. If path adequacy also does not increase, a latent impasse is said to have occurred and CoLiDeS+ invokes a backtracking strategy. It first considers hyperlinks on the same page with lower information scent value (also called the next-best strategy). It then considers backtracking to other regions within the same page and eventually to the previously visited pages. CoLiDeS+ stops when the user declares the current page is the page with the target information. Figure 1 shows a schematic diagram of the steps involved in CoLiDeS+.

Both CoLiDeS and CoLiDeS+ models have been developed to describe the navigation path of participants within websites but it has been shown that they can also be applied to the interaction with search engines [16]. In the context of interaction with a search engine, the models consider each SERP as a page of a website. And each of the search engine results as a hyperlink within a page of a website. The problem of predicting which search result to click on is now equivalent to the problem of predicting which hyperlink to click within a page of a website. Therefore, the process of computing information scent and path adequacy and predicting which search engine result to click remains the same. The user-generated query is used as a representation of local goal or the understanding of the user at any point of time and semantic similarity values between the query and search results are computed from it. Our focus in this paper is also on modeling individual differences in user interactions with a search engine due to variations in cognitive factors and we will use the CoLiDeS+ model. More specifically, this paper focuses on modeling individual differences in the search results clicked by users, for a given query and its corresponding SERP, due to variations in cognitive factors.

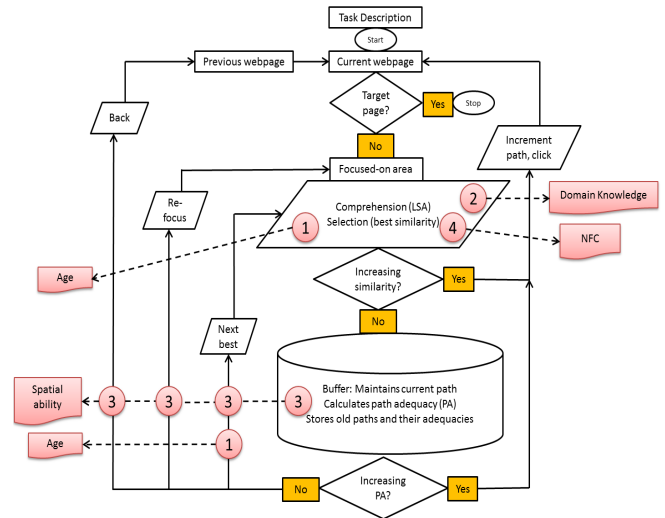


Figure 1: Schematic diagram of steps involved in CoLiDeS+ (reproduced from Van Oostendorp & Juvina [15]). Shaded circles indicate the locations where individual differences are involved (1: Age, 2: Domain Knowledge, 3: Spatial Ability and 4: Need for Cognition). See also later in text references to these locations.

INFLUENCE OF COGNITIVE FACTORS

In this section, we will describe the influence of the following cognitive factors: age, domain knowledge, spatial ability and need for cognition on the cognitive processes underlying information search. Next to the description of the influence of each cognitive factor, we present some preliminary thoughts on how the variations in information search behavior caused by the differences in cognitive factors can be incorporated into the computational cognitive model CoLiDeS+.

Age

Aging leads to a natural decline in motor skills and fluid intelligence involving processing speed, cognitive flexibility or ability to switch processing strategies, attentional control and visuospatial span [13, 29]. However, crystallized knowledge (vocabulary skills and knowledge in a specialized domain such as e.g. health) seems stable or even increases with age.

Some of these cognitive abilities directly influence the cognitive processes underlying information search resulting in lower efficiency of older adults on information search tasks. For example, lower processing speed could lead to longer time in evaluating search results or hyperlinks on a website, difficulty in switching strategies could lead to difficulty in reformulating unsuccessful queries or difficulty in getting out of an unsuccessful path, lack of attentional control could lead to inefficient handling of relevant and irrelevant search results or hyperlinks and finally lower visuospatial span could mean less efficient exploration of the search result page or a website.

Many empirical 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 [25, 11]. Older adults were found to allocate less resources to exploration (fewer keywords, fewer clicks on search results etc.) and more resources to exploitation (longer time on a search result page, deeper navigation into websites opened from the search results) compared to younger adults [6]. Based on these age-related effects, we think that the steps involving selection of a hyperlink or a search result and backtracking in CoLiDeS+ can get affected (marked with ① in Figure 1). In the context of interacting with a SERP, selection of a search result is primarily based on its information scent or the semantic similarity between the query and the search result. And backtracking is based on the number of times the next-best strategy is invoked when both the information scent and the path adequacy of a search result do not increase. By running the model at different values of the two parameters, it is possible to simulate differences in the number and the choice of search results clicked by a user.

Prior Domain Knowledge (PDK)

Users with high domain knowledge have more appropriate mental representations characterized by more relevant concepts, higher activation values and stronger connections between different concepts in the conceptual space compared to users with low domain knowledge [17]. Users with high domain knowledge, therefore can comprehend the search results and the content of the websites better and evaluate the relevancy more thoroughly than users with low domain knowledge [17].

Variations in information search behaviour due to the differences in domain knowledge of users have been well researched and documented in the cognitive psychology community [10, 23]. In the study by [23], domain experts were found to find more correct answers in shorter time and via a path closer to the optimum path than non-experts. This difference was stronger as the difficulty of the task increased. Higher domain knowledge enables a user to formulate more appropriate queries and comprehend the search results and the content in the websites better, which in turn, enables them to take informed decisions regarding which hyperlink or a search result to click next. Domain experts are also known to evaluate search results more thoroughly and click more often on relevant search results compared to non-experts. This is because their higher domain knowledge enables them to differentiate between a relevant and a non-relevant search result better [10]. Therefore, the step involving comprehension and evaluation of a hyperlink or a search result in CoLiDeS+ can get affected by variations in prior domain knowledge of users (marked with ② in Figure 1).

To simulate the differences in the comprehension and evaluation of relevancy of search results due to the differences in the domain knowledge of users in CoLiDeS+, we can use the semantic space in LSA. A semantic space in LSA is an approximate representation of a given user population's knowledge. It is possible to create two semantic spaces that reflect two different user population's knowledge levels. Based on this idea, it is possible to create two semantic spaces having low (general semantic space) and high (specialized semantic

space) amount of domain specific information (such as health) to represent respectively the low and high knowledge levels of users in that particular domain.

Spatial Ability

Spatial ability is the ability to form visual or mental representations, process spatial information and remember the spatial relationships among different objects. Visuo-spatial skills are of great importance for success in solving many tasks in everyday life: for instance, using a map to navigate through an unfamiliar city, understanding the layout of a new building, merging into high speed traffic etc. Since the structure or architecture of the information space of the Internet can be represented spatially as a graph, it is plausible that the spatial ability of a user to mentally form an equivalent representation of the information space has an influence on search performance.

Users with high spatial ability would form a more accurate mental representation of the hyperspace that they are navigating through, compared to users with low spatial ability. They would be more aware of their current location in the hyperspace and therefore would be able to navigate themselves forward and backward with little effort. Whereas, users with low spatial ability usually do not understand their current location in the hyperspace and most often get lost.

Spatial ability was found to correlate with the number of revisits made by users to already visited pages, number of times the back button is used etc [28]. Many studies have repeatedly shown that spatial ability correlates highly with information search and navigation performance [15]. We think that the steps involving backtracking to already visited hyperlinks, to other regions on the same page and to previous pages in CoLiDeS+ will get affected by spatial ability (these steps are marked with ③ in Figure 1).

To simulate differences in number of times users revisit already visited pages or the number of times they use the back button, due to differences in their spatial ability, we can vary the frequency and the depth of backtracking behavior in CoLiDeS+. Depth of backtracking measures the number of levels a user would backtrack. We assume that setting a high value for both frequency and depth of backtracking would model the information search behavior of users with high spatial ability and setting a low value would model the information search behavior of users with low spatial ability. We can also vary the amount of context that is taken into consideration when computing path adequacy in CoLiDeS+. If we assume that users with high spatial ability are able to utilize more contextual information than users with low spatial ability, we can vary the number of hyperlinks from the preceding session that are used to compute path adequacy. The higher the number of links is, the greater the amount of context is. Therefore, we propose that we should use a higher number of links (greater amount of context) for users with high spatial ability and a smaller number of links (smaller amount of context) for users with low spatial ability.

Need for Cognition

Cohen et al. [9] describes Need for Cognition (NFC, henceforth) as “a need to structure relevant situations in meaningful, integrated ways”. If this need is not satisfied, it can result in feelings of tension and deprivation that can lead to “active efforts to structure the situation and increase understanding” (p. 291). It is an indication of the urge to make sense of the world.

Users with high NFC are more likely to put efforts to clear all ambiguities and uncertainty in the presented information compared to users with low NFC. Therefore, we expect that NFC will have influence on the cognitive processes of comprehension, decision making and executive control during information search. Users with high NFC would evaluate information more thoroughly and take more informed decisions on clicking, reformulating and backtracking than users with low NFC.

This personality trait or cognitive style was found to have an influence on how users perform information search activities [12, 21, 22]. In a study by [22], users with high NFC were found to spend more time on accessing relevant information, prefer formal sources for reliable information to informal ones, select up-to-date resources, value high-quality information, use advanced search options for formulating queries, prefer to select information sources with complex and multidimensional contents rather than simple ones. In other studies, users with high NFC were found to explore more information from multiple sources [12] and evaluate available information more thoroughly than users with low NFC [21]. Therefore, we think that the most important step in CoLiDeS+ that can be affected by variations in NFC is the step involving comprehension of a hyperlink or a search result (marked with ④ in Figure 1).

The computational cognitive model as introduced draws analogies between the cognitive processes that govern text comprehension and the cognitive processes during information search. Therefore, to simulate what goes on inside human working memory when comprehending a hyperlink text or a search result, these models use the process of elaboration (or activation of associated prior knowledge). The process of elaboration simulates the cognitive processes of activation of semantically related terms to a piece of text that is present in our working memory through a spreading activation mechanism. These elaborations are known to assist in better comprehension of what is being read [17]. Using LSA, the models can extract words that are close in semantic similarity to the target text. By varying the threshold value of the semantic similarity, the number of words extracted and their degree of semantic similarity to the target text can be varied. We conjecture that this threshold value should be high for users with high NFC and low for users with low NFC. A higher threshold value would extract only those words which are highly similar to the target text and avoids words which are less similar, thereby, ensuring less ambiguity. A lower threshold value would extract words which are further away in semantic similarity to the target text and therefore leads to greater ambiguity.

EXPERIMENT

Above, we have explained how variations in age, domain knowledge, spatial ability and need for cognition could be incorporated into the cognitive modeling. In this section, we evaluate our ideas to model individual differences in information search behavior caused by variations in *two* cognitive factors: age and prior domain knowledge. More specifically, we evaluate our ideas to model individual differences in the search results clicked by users, for a given query and its corresponding SERP. The evaluation is conducted by comparing the model-predicted search results with the search results clicked by actual users. If a model-predicted search result is also clicked by an actual user, we consider it a match. The higher the number of such matches is, the higher the efficacy of the model is.

We first describe the details of an experiment conducted to collect actual behavioral data from participants. In the experiment, we will examine the impact of age (young vs. old), prior domain knowledge (low vs. high) and task difficulty (simple vs. difficult). We next describe the outcomes of simulations of CoLiDeS+ based on the ideas proposed for each cognitive factor.

Method

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.

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.

Material

The experiment was conducted with twelve simulated information search tasks [3]: 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 instance, for a task like “*What is the main function of sweat glands?*”, user could use the words like “*function of sweat glands*” from the task description itself as queries. One can easily find the answer “*regulation of body temperature*” within the snippets of the corresponding search results without having to click on any of them. 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. As an example, for the task “*Elbert, 76 years old has been suffering for few years from burning sensation while passing urine. He passes urine more often than normal at night and complains of a feeling that the bladder is not empty completely. Lately, he also developed acute pain in the hip, lower back and pelvis region. He also lost 12 kilos in the last 6 months. What problem could he be suffering from?*”, users had to formulate multiple queries such as “*kidney stones pain in the back*”, “*burning sensation*

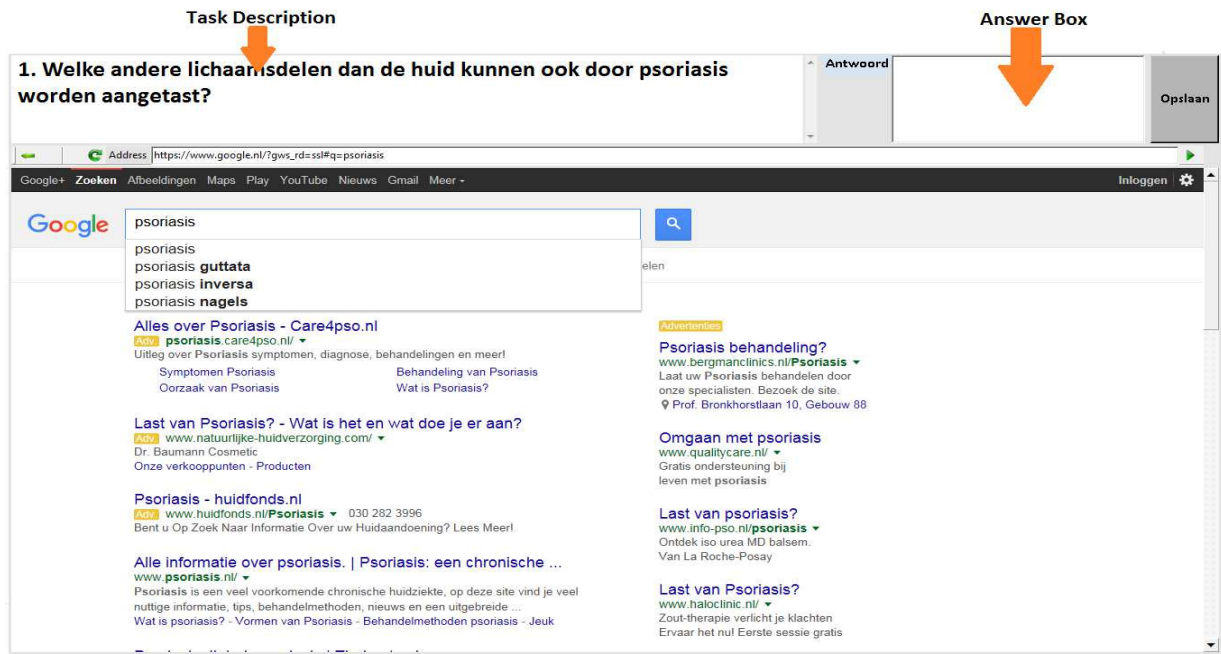


Figure 2: Interface showing the main screen in which the information search tasks are solved.

when urinating”, “urinary infection” to find the answer. The answer to this task “prostate cancer” is also not found easily in the snippets of the search results of the queries, unless the query is very specific. The tasks were all originally presented in Dutch and therefore participants used queries also in Dutch.

Procedure

Participants first did 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 prior domain knowledge test on the topic of health consisting of 12 multiple choice questions. The score on the prior domain knowledge test gives us an indication of the amount of prior knowledge on the topic of health. For the prior domain knowledge test, participants were presented with 12 questions (one followed by the other) along with four answers for each question presented as multiple choice options. For each question, the participants had to choose an option from the four alternatives. There was only one possible correct answer for each question. Correct choices were scored 1 and wrong choices were scored 0. Thus the maximum possible score on this test is 12 and the minimum possible score is 0.

After the prior domain knowledge test, participants were allowed a break of five minutes. They were then presented with

twelve information search tasks (six simple and six difficult) in a counter balanced order. Participants were first shown the task and then directed to the home page of Google’s search engine. Participants were not allowed to use any other search engine. We show in Figure 2 the main screen of our interface that participants used while solving the information search tasks. 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. We analyzed the performance in terms of time spent on SERPs, number of clicks on SERPs and accuracy.

Analysis of the cognitive factor: Age

We first analyze the behavioral differences between younger and older adults to check if they are in line with aging-related literature. Next to it, we run simulations of CoLiDeS+ and compute the match between model predictions and actual user behavior.

Behavioral Analysis

Task-completion time is computed as the total time spent on search engine result pages. It starts from the moment of opening a browser and includes the time taken to generate a query and the time spent on search engine result pages and excludes the time spent on websites.

Number of clicks is the total number of clicks made by a participant for each task on search engine result pages.

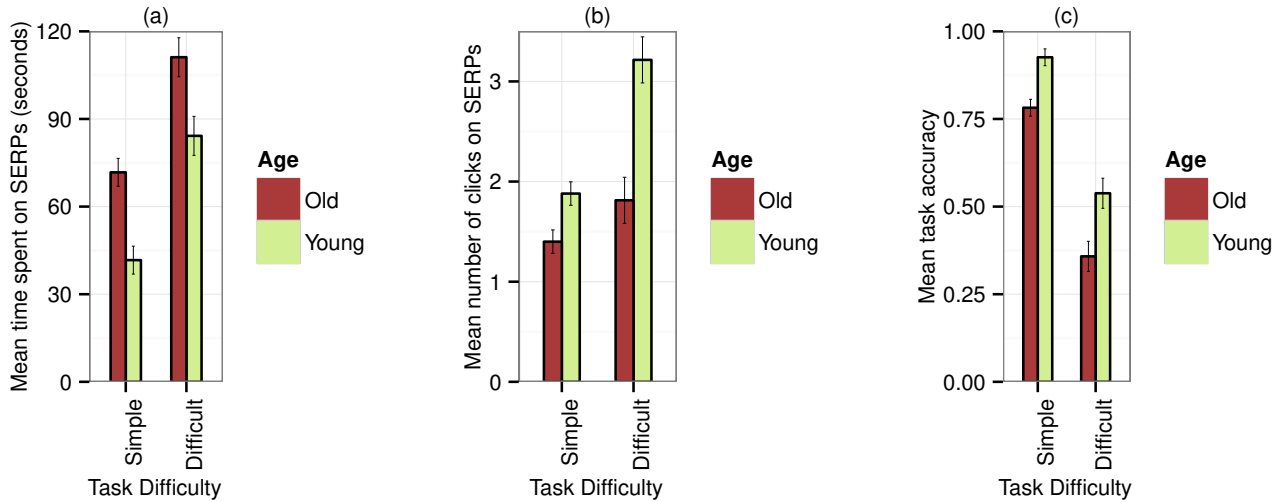


Figure 3: Analysis of search performance in terms of (a) time spent on SERPs, (b) clicks on SERPs and (c) accuracy in relation to age and task difficulty.

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 of only those tasks was included in the analysis for which the participants successfully completed the tasks. 37 data points out of a total number of (12 tasks X 48 participants) = 576 data points (6.4%) were therefore dropped. For all the three 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.

Results

Task-completion time Figure 3a shows the mean task-completion time in relation to age and task difficulty. The main effect of age was significant $F(1,46) = 18.13, p < .001$. The main effect of task difficulty was significant $F(1,46) = 74.31, p < .001$. The interaction of age and task difficulty was not significant ($p > .05$).

Number of clicks Figure 3b shows the relationship between age and task difficulty for mean number of clicks. The main effect of age was significant $F(1,46) = 18.0, p < .001$. The main effect of task difficulty was significant $F(1,46) = 42.94, p < .001$. The interaction of age and task difficulty was also significant $F(1,46) = 12.0, p < .001$.

Accuracy As can be seen from Figure 3c, the main effect of age was significant $F(1,46) = 17.18, p < .001$. The main effect of task difficulty was significant $F(1,46) = 196.4, p < .001$. Interaction of age and task difficulty was not significant ($p > .05$).

Summarizing the main behavioral outcomes, we found that for difficult tasks, significantly more time was spent on SERPs, significantly more clicks were made on SERPs and the accuracy was significantly lower compared to simple tasks.

Younger adults were found to spend significantly less time on SERPs and click significantly more often on SERPs, especially for difficult tasks. The accuracy of older adults was significantly much lower compared to the accuracy of younger adults. These results concerning time spent on SERPs, clicks on SERPs and task accuracy are in-line with the prior findings from aging-related literature and provide more evidence to the fact that older adults are less efficient than younger adults.

Modeling Analysis

For analyzing model behavior, we focus only on difficult tasks to simplify presentation of data. Also, the individual differences due to the cognitive factor age are generally known to be more prominent for difficult tasks. We ran simulations using CoLiDeS+ under ten possible combinations of each of the two parameters: the number of times the next-best strategy is used (0 to 9) and the minimum LSA value of a search result (0 to 0.9) on six difficult information search tasks. In order to evaluate the model simulations, we computed the number of matches between the model predictions and the actual user clicks from the behavioral experiment.

As younger adults are more impulsive, follow an explore-more and exploit-less strategy, we expect that they apply more often the next-best strategy and click more often on search results with lower LSA value compared to older adults. Therefore, we expect the difference in the mean number of matches between younger and older adults to increase as the number of times the next-best strategy is applied, is increased and the minimum LSA value of a search result is decreased.

Results

A repeated measures ANOVA with age as between-subjects variable and the number of times the next-best strategy is applied as within-subjects variable and mean number of matches as dependent variable was conducted. The main effect of age was significant $F(1,46) = 9.5, p < .005$, indicating that the

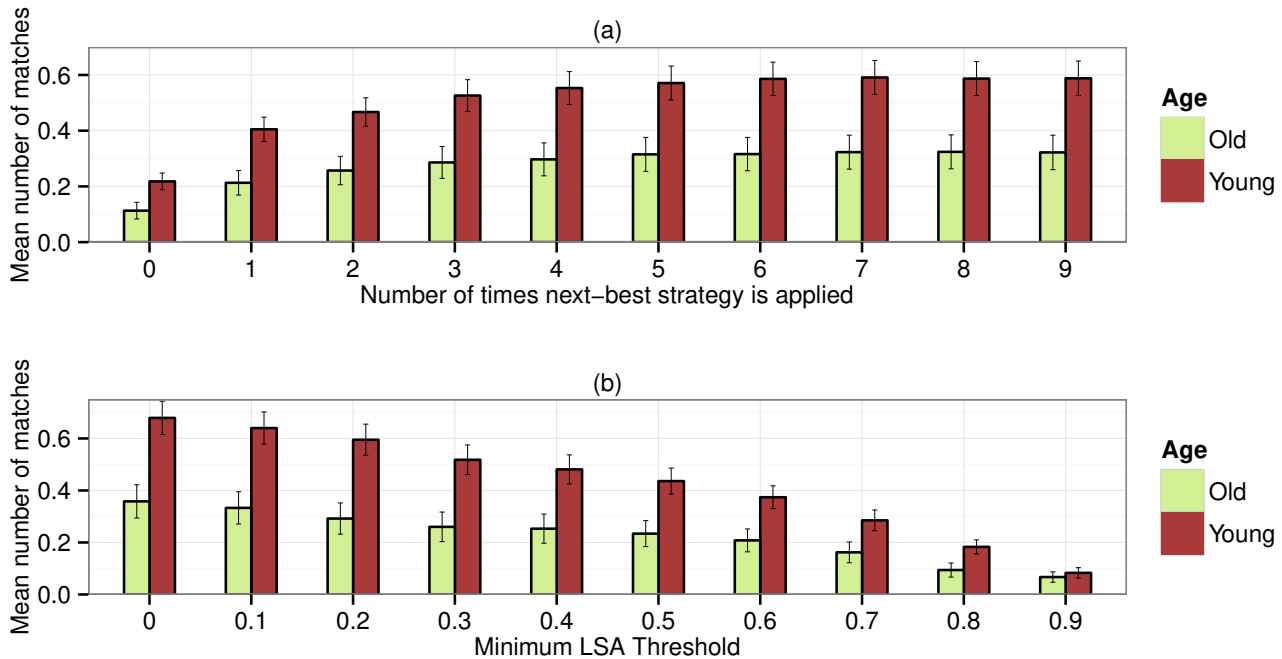


Figure 4: Mean number of matches in relation to variations in (a) number of times next-best strategy is applied and (b) minimum LSA value of a search result.

model matched behavior of younger participants significantly better than that of older participants (See Figure 4(a)). The main effect of number of times the next-best strategy is applied was significant $F(1,46) = 62.25, p < .001$. As the number of times the next-best strategy is applied increased, the match between the model and the actual user behavior also increased. Also, the interaction of the number of times the next-best strategy is applied and age was significant $F(1,46) = 4.76, p < .05$. Post-hoc tests show that the rate of increase in match was higher for young participants compared to old participants.

A similar ANOVA with the minimum LSA value of a search result as within-subjects variable was conducted. The main effect of age was significant $F(1,46) = 10.2, p < .005$ (See Figure 4(b)). The model matched behavior of younger participants significantly better than that of older participants. The main effect of the minimum LSA value of a search result was significant $F(1,46) = 74.39, p < .001$. As the minimum LSA value of a search result decreased, the match between the model and the actual user behavior increased. Also, the interaction of the minimum LSA value of a search result and age was significant $F(1,46) = 9.13, p < .001$. Post-hoc tests show that the rate of increase in match was higher for young participants compared to old participants when LSA decreased.

These outcomes are in-line with our expectations. Most importantly, these results imply that the match between the model and the actual behavior varies significantly in the expected direction with the number of times the next-best strategy is applied and the minimum LSA value of a search result that is used as a threshold in the model. We are currently performing

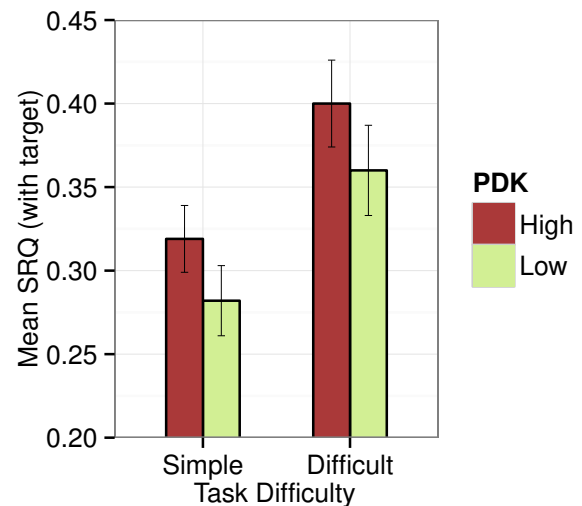


Figure 5: Mean semantic relevance of queries with target information in relation to PDK and task difficulty

more analyses to understand the optimum parameter values for younger and older adults.

Analysis of the cognitive factor: Prior Domain Knowledge

We first analyze the behavioral differences between low and high domain knowledge users to check if they are in line with literature. We divided the participants into two groups of high (25 participants) and low (23 participants) prior domain

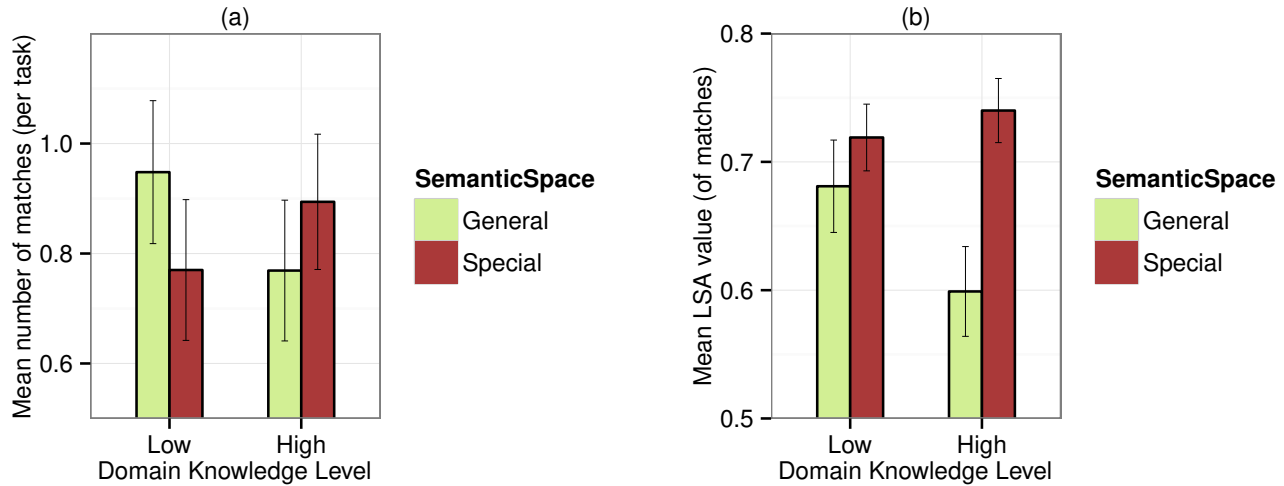


Figure 6: (a) Mean number of matches (per task) and (b) Mean LSA value (of matches) in relation to Semantic Space and Prior Domain Knowledge (PDK).

knowledge (PDK) by taking the median score on the prior domain knowledge test. In this analysis, we omit the variable age. Next to it, we run simulations of CoLiDeS+ and compute again the match between model predictions and actual user behavior.

We could not find significant differences in task-completion time, number of clicks or accuracy between the two groups. Therefore we examined the impact of prior domain knowledge on the *semantic relevance of queries* (SRQ) generated by the participants. As users with high domain knowledge have more appropriate mental representations characterized by more relevant concepts, higher activation values and stronger connections between different concepts in the conceptual space compared to users with low domain knowledge, we expect that the semantic relevance of queries generated by them with target information would be much higher than that of users with low prior domain knowledge.

Behavioral Analysis

Semantic Relevance of Query (SRQ) is computed as the semantic relevance between the query and the target information sought using LSA. This metric gives us an estimate of how close in semantic similarity the queries generated by the participants are to the target information. So in general, the higher the SRQ value is, the more relevant the query is.

Results

Semantic relevance of query: A 2 (PDK: Low vs. High) X 2 (Task Difficulty: Simple vs. Difficult) mixed ANOVA was conducted with PDK as between-subjects variable and task difficulty as within-subjects variable. The main effect of task difficulty was significant $F(1,46) = 9.35, p < .005$. See Figure 5. The semantic relevance of queries with target information was significantly higher for difficult tasks compared to simple tasks. The main effect of PDK was close to conventional significance $F(1,46) = 3.68, p = .06$, indicating that participants with high prior domain knowledge generated queries with significantly

higher semantic relevance to target information compared to participants with low prior domain knowledge. The interaction of PDK and task difficulty was not significant ($p > .05$). These results provide evidence to the outcomes in prior work [10].

Modeling Analysis

Just like we did with the cognitive factor age, for analyzing model behavior, we focus only on difficult tasks to simplify the presentation of data. Also, the individual differences due to the cognitive factor prior domain knowledge are generally known to be more prominent for difficult tasks.

We collated two different corpora (a general corpus and a special corpus, each consisting of 70,000 articles in Dutch) varying in the amount of medical and health related information (because our behavioral experiment used information search tasks from the domain of health). The *general corpus*, representing the knowledge of low domain knowledge users had 90% news articles and 10% medical and health related articles whereas the *special corpus*, representing the knowledge of high domain knowledge users had 60% news articles and 40% medical and health related articles. After removing all the stop words, these two corpora were used to create two semantic spaces: a *general* semantic space using the general corpus and a *special* semantic space using the special corpus. We ran CoLiDeS+ simulations on six difficult information search tasks using both general and special semantic spaces.

Results

A 2 (Semantic Space: General vs. Special) X 2 (Prior Domain Knowledge (PDK): Low vs. High) mixed model ANOVA was conducted with semantic space as within-subjects variable and prior domain knowledge as between-subjects variable and mean number of matches as dependent variable. The main effects of semantic space and prior domain knowledge were not significant ($p > .05$). However, the interaction of semantic space and prior domain knowledge was significant $F(1,46) = 7.5, p < .01$ (Figure 6a).

The number of matches between the model and behavior of participants is a relevant variable to examine the efficacy of modeling. Also, the mean LSA value of a match can be an useful indicator.

A similar 2 (Semantic Space: General vs. Special) X 2 (Prior Domain Knowledge (PDK): Low vs. High) mixed model ANOVA was conducted with semantic space as within-subjects variable and prior domain knowledge as between-subjects variable and mean LSA value of a matched search result as dependent variable. The main effect of semantic space was significant $F(1,41) = 8.88, p < .005$. The main effect of prior domain knowledge was not significant ($p > .05$). The interaction of semantic space and prior domain knowledge was tending towards significance $F(1,41) = 2.9, p < .09$ (Figure 6b).

These results showed that the efficacy of the modeling of user interaction in terms of the number of matches (Figure 6a) and the mean LSA values of the matched search results (Figure 6b) is significantly higher with the special semantic space compared to the general semantic space for *high* domain knowledge participants while for *low* domain knowledge participants it is the other way around. It is important to note that these interaction effects are lost when semantic space is not used as a factor in the analysis. That is, if we would not have used semantic space as a factor, we would have concluded that there is no difference in model performance between the participants with high and low domain knowledge levels. This would have been a hasty conclusion because we saw that when we included semantic space as a factor in the analysis, there was an effect of PDK, but it was dependent on the type of semantic space. Overall, our outcomes suggest that by incorporating differentiated domain knowledge levels into computational cognitive models, it is possible to simulate differences in the search results clicked by users due to differences in domain knowledge.

CONCLUSIONS AND DISCUSSION

In this paper, using a computational cognitive model called CoLiDeS+, we presented ideas to simulate variations in information search behavior (more specifically, variations in the search results clicked by a user for a given query and its corresponding SERP) due to variations in cognitive factors such as age, domain knowledge, spatial ability and need for cognition. The main research question of this study was: would a cognitive model that adapts itself to the variations in a cognitive factor predict individual differences in information search behavior (caused by that cognitive factor) more accurately than a cognitive model that does not? Preliminary analysis of our ideas on predicting individual differences due to age and domain knowledge effects showed promising outcomes. Concerning *age*, by varying the number of times the next-best strategy is applied by CoLiDeS+ and the minimum LSA value of a search result, we were able to simulate the variations in the number of search results clicked by younger and older adults on SERPs. It appeared that the difference in the mean number of matches between the model-predicted clicks and the actual user clicks between younger and older adults increased as expected as the number of times the next-best strategy is in-

creased (Figure 4a) and as the minimum LSA value of a search result decreased (Figure 4b). To simulate the differences in evaluating the relevancy of search results due to differences in *prior domain knowledge*, we used two different semantic spaces that varied in the amount of medical and health related information (the general semantic space had lower medical and health related knowledge than the special semantic space) to compute LSA values corresponding to information scent and path adequacy. The efficacy of the modeling in terms of the number of matches between the model-predicted clicks and the actual user clicks (Figure 6a) and the mean LSA values of the matched search results (Figure 6b) was found to be higher with the special semantic space compared to the general semantic space for high domain knowledge participants while for low domain knowledge participants it was the other way around. Consequently, it appeared important to adapt the semantic space to the prior domain knowledge of the participants. These modeling outcomes indicate that indeed a cognitive model that adapts itself to variations in cognitive factors would predict information search behavior more accurately than a cognitive model that does not.

More experiments are required to empirically verify the ideas presented in this paper corresponding to spatial ability and need for cognition. We do not claim that the ideas presented in this paper fully explain the variations in information search behavior due to the variations in cognitive factors, but we hope that they could be starting points for further research. The influence of each of the cognitive factors discussed in this paper on the psychological processes during information search can be simulated in more than one location in the model (Figure 1). Furthermore, it is not clear, at this moment, if the influence on one location in the model is more significant than the other. Also, there exist other cognitive factors (not discussed in this paper) such as internet experience, gender and interface characteristics which can also have an influence on information search behavior.

ACKNOWLEDGMENTS

This research was supported by Netherlands Organization for Scientific Research (NWO), project MISSION (464-13-043).

REFERENCES

1. Marilyn Hughes Blackmon, Muneo Kitajima, and Peter G Polson. 2005. Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 31–40.
2. Marilyn Hughes Blackmon, Dipti R Mandalia, Peter G Polson, and Muneo Kitajima. 2007. Automating usability evaluation: Cognitive walkthrough for the web puts LSA to work on real-world HCI design problems. *Handbook of latent semantic analysis* (2007), 345–375.
3. Pia Borlund and Peter Ingwersen. 1997. The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation* 53, 3 (1997), 225–250.

4. Raluca Budiu and John R Anderson. 2004. Interpretation-based processing: A unified theory of semantic sentence comprehension. *Cognitive Science* 28, 1 (2004), 1–44.
5. Aline Chevalier, Aurélie Dommes, and Jean-Claude Marquié. 2015. Strategy and accuracy during information search on the Web: Effects of age and complexity of the search questions. *Computers in Human Behavior* 53 (2015), 305–315.
6. Jessie Chin, Evan Anderson, Chieh-Li Chin, and Wai-Tat Fu. 2015. Age differences in information search: An exploration-exploitation tradeoff model. In *Proceedings of the Human Factors and Ergonomic Society (HFES 2015)*. 85–89.
7. Aleksandr Chuklin, Ilya Markov, and Maarten de Rijke. 2015. Click Models for Web Search. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 7, 3 (2015), 1–115.
8. Andy Cockburn and Bruce McKenzie. 2001. What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies* 54, 6 (2001), 903–922.
9. Arthur R Cohen, Ezra Stotland, and Donald M Wolfe. 1955. An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology* 51, 2 (1955), 291–294.
10. Michael J Cole, Xiangmin Zhang, Chang Liu, Nicholas J Belkin, and Jacek Gwizdka. 2011. Knowledge effects on document selection in search results pages. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 1219–1220.
11. Aurelie Dommes, Aline Chevalier, and Sarah Lia. 2011. 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 (2011), 717–726.
12. Shuk Ying Ho. 2005. An exploratory study of using a user remote tracker to examine web users’ personality traits. In *Proceedings of the 7th International Conference on Electronic Commerce*. ACM, 659–665.
13. John L Horn. 2012. The Theory of Fluid and Crystallized Intelligence in Relation to Concepts of Cognitive Psychology and Aging in. In *Aging and Cognitive Processes*. Vol. 8. Springer Science & Business Media, 237–278.
14. Botao Hu, Yuchen Zhang, Weizhu Chen, Gang Wang, and Qiang Yang. 2011. Characterizing search intent diversity into click models. In *Proceedings of the 20th International Conference on World Wide Web*. ACM, 17–26.
15. Ion Juvina and Herre van Oostendorp. 2008. Modeling semantic and structural knowledge in Web navigation. *Discourse Processes* 45, 4-5 (2008), 346–364.
16. Saraschandra Karanam, Herre van Oostendorp, Mylène Sanchiz, Aline Chevalier, Jessie Chin, and Wai Tat Fu. 2015. Modeling and predicting information search behavior. In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics*. ACM, 7.
17. Walter Kintsch. 1998. *Comprehension: A paradigm for cognition*. Cambridge University Press.
18. Muneo Kitajima, Marilyn H Blackmon, and Peter G Polson. 2000. A comprehension-based model of Web navigation and its application to Web usability analysis. *People and Computers* (2000), 357–374.
19. Muneo Kitajima, Marilyn Hughes Blackmon, and Peter G Polson. 2005. Cognitive architecture for website design and usability evaluation: Comprehension and information scent in performing by exploration. In *HCI International 2005*, Vol. 4. L. Erlbaum Associates Mahwah, NJ.
20. Thomas K Landauer, Danielle S McNamara, Simon Dennis, and Walter Kintsch. 2007. *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.
21. Jingjing Liu and Xiangmin Zhang. 2008. The effect of need for cognition on search performance. *Proceedings of the American Society for Information Science and Technology* 45, 1 (2008), 1–12.
22. Heidar Mokhtari, Mohammad-Reza Davarpanah, Mohammad-Hossein Dayyani, and Mohammad-Reza Ahanchian. 2013. Students’ need for cognition affects their information seeking behavior. *New Library World* 114, 11/12 (2013), 542–549.
23. Sophie Monchaux, Franck Amadieu, Aline Chevalier, and Claudette Mariné. 2015. Query strategies during information searching: Effects of prior domain knowledge and complexity of the information problems to be solved. *Information Processing & Management* 51, 5 (2015), 557–569.
24. Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological Review* 106, 4 (1999), 643.
25. Tara L Queen, Thomas M Hess, Gilda E Ennis, Keith Dowd, and Daniel Grühn. 2012. Information search and decision making: Effects of age and complexity on strategy use. *Psychology and aging* 27, 4 (2012), 817.
26. Joseph Sharit, Mario A Hernández, Sara J Czaja, and Peter Pirolli. 2008. Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. *ACM Transactions on Computer-Human Interaction (TOCHI)* 15, 1 (2008), 3.
27. Si Shen, Botao Hu, Weizhu Chen, and Qiang Yang. 2012. Personalized click model through collaborative filtering. In *Proceedings of the fifth ACM International Conference on Web Search and Data Mining*. ACM, 323–332.
28. Franklin P Tamborello II and Michael D Byrne. 2005. Information search: the intersection of visual and semantic space. In *CHI’05 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1821–1824.

29. Jing-Jen Wang and Alan S Kaufman. 1993. Changes in fluid and crystallized intelligence across the 20-to 90-year age range on the K-BIT. *Journal of Psychoeducational Assessment* 11, 1 (1993), 29–37.
30. Qianli Xing, Yiqun Liu, Jian-Yun Nie, Min Zhang, Shaoping Ma, and Kuo Zhang. 2013. Incorporating user preferences into click models. In *Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management*. ACM, 1301–1310.