The impact of expertise on query formulation strategies during complex learning task solving: A study with students in medicine and computer science

Cheyenne Dosso[1], Lynda Tamine[2], Pierre-Vincent Paubel[1], and Aline Chevalier[1]

[1] University of Toulouse, CLLE-CNRS, 05 Allée Antonio Machado 31058 Toulouse, France
cheyenne.dosso@univ-tlse2.fr
[2] University of Toulouse, IRIT-CNRS, 118 Route de Narbonne 31062 Toulouse, France
lynda.lechani@irit.fr

Abstract. This study focus on queries formulation strategies when expert users in a medical or computer science domain solved complex tasks. Ten medical students and ten computer science students had to perform four fact-finding search tasks (two simple tasks and two inferential tasks) and six learning tasks (two exploratory, two decision-making and two problem solving tasks) in these two domains. Results showed that non-experts used more terms from task statement to build their queries than experts did. Experts often produced new keywords than non-experts did. Specifically, computer science experts used more keywords not specific to the domain knowledge whereas medical experts used specific domain keywords to formulate queries. These results are a beginning to better understand how users are searching to learn when they are using Internet but further ergonomics studies have to more explore this subject to create search systems adapted to Search as Learning activity.

Keywords: Search as Learning, Task Complexity, Expertise Domain, Query Strategies.

1 Introduction

When users are engaging in information search (IS) activity with a search engine, they are often faced with a lot of information they have to with regard to their objectives, prior knowledge to achieve their search goals. (Sharit, Taha, Berkowsky, Profita, & Czaja 2015). Recently, many researchers have emphasized the importance of improving current search systems in a learning context because users often pursue an overall objective of acquiring new knowledge (Gwizdka, Hansen, Hauff, He, & Kando, 2016). To develop systems that fit users' learning objectives, it is important to consider user's individual characteristics involve in this activity. Among them, the level of prior domain knowledge play a central role. If many studies focused on the role of prior domain knowledge on IS (Monchaux, Amadieu, Chevalier, & Mariné, 2015; Sanchiz et al., 2017a), in the Searching as Learning approach the impact of this variable need a better understanding specifically on users' search behavior when they are solving complex learning tasks. To this end, in the present study, we focus on query formulation strategies with regard to the level of prior domain knowledge of users (i.e. experts vs non-experts) and the type of search tasks to be performed (fact-finding vs learning tasks). First, we present related work concerning the information search activity, particularly concerning query formulations and the expertise effects. In the second part, we present the method used to study formulation strategies of users. Then the results are presented and we finish on the discussion with limits and the perspectives of further researches.

2 Related work

The cognitive model of IS initially developed by Sharit et al. (2015) describe IS into three main stages: 1. *Planning:* users have to build a mental representation of the search goal from task statement and their prior domain knowledge. 2. *Evaluation of information*: users have to compare their search goal stored in working memory with information from Search Engine Results Pages (SERPs). 3. *Depth processing and navigation*: From links selected just before, users access to information content (i.e. web pages, PDF...). They may decide to process information more precisely by comparing their goal with content or to navigate within several web pages. To plan the activity, to evaluate and to process information content, prior domain knowledge allows building a more consistent mental representation of the search problem, to be more relevant in the select of links from SERPs and for analyzing the web pages content (Sanchiz, Amadieu, & Chevalier, 2020). At the level of query formulation strategies during IS, expert users are more efficient than non-expert ones. They formulate more queries (Monchaux et al., 2015; Sanchiz et al., 2017a) and longer ones than non-experts do (Hembrooke, Granka, Gay, & Liddy, 2005; Tamine & Chouquet, 2017). They also produce more new keywords (Monchaux et al., 2015 ; Sanchiz, Chevalier & Amadieu, 2017b) linked to domain vocabulary (Sanchiz et al., 2017a; Tamine & Chouquet, 2017; O'Brien, Kampen, Cole, & Brennan, 2020), than non-experts, who need to use the task statement to build their queries (Sanchiz et al., 2017a, 2017b). This is a major problem for non-expert users because they have to get to relevant content to be able to learn new knowledge, but search engine results pages depend on queries content (Vakkari, 2016). In this way, the activity of search as learning should be more critical for users without prior domain knowledge because they do not have specific vocabulary to formulate queries allowing to get new information.

In addition, the information search activity depends on characteristics of tasks to be performed. For instance, complexity of search task is often manipulated at the level of task goal (e.g., Monchaux et al., 2015; Sanchiz et al., 2017a, 2017b). For learning tasks, literature describes several complexity level. Firstly, the exploratory learning tasks, allowing users to gain knowledge about a topic (Marchionini, 2006). These tasks are open-ended because several sub-goals may be carried on by users. Then, there were decision-making tasks, in which the final goal is to select the better solution among several possibilities and make the better decision (Campbell, 1988). The sub goals consist to compare a set of information leading to decision-making. This task is also open-ended, because several answer are acceptable. To make choice, users have to define criteria related to the goal to be achieve with regard to their level of knowledge. Finally, there is problem-solving task, which needs the elaboration and the creation of a new set of information. Users have find the better path to achieve the goal, which is clearly specified (Campbell, 1988) and they have to re-use retrieved information.

The objective of this study is to better understanding how users formulate queries depending on their level of prior domain knowledge according to the search and learning task complexity.

2 Method

2.1 Variables

Independents variables. *IV1-Prior domain knowledge level* (experts vs non-experts) as between-subject factor. *IV2-Task type* (simple, exploratory learning, decision-making, problem solving, inferential) as within-subject factor. *IV3-Task domain* (medical vs computer science) as within-subject factor.

Dependents variables. All dependents variables were recorded per search session (i.e. to complete one task, user have to do one search session from the first query produced to the close of navigator).

DV1a-Total number of new queries produced and DV1b-Query length: DV1a corresponds to the total number of new queries submitted to the search box. If user submitted the same query during his/her search session, only first production was computed. For DV1b, query length was calculated as a mean of all queries produced divided by all keywords produced per search session.

DV2-Total number of keywords used from tasks corresponds to the total number of keywords used, which were terms contained in the task statement per search session.

DV3a-Total number of new keywords produced by users from not specific vocabulary and DV3b-Total number of new keywords produced by users from specific vocabulary. DV3a corresponds to the total new keywords not related to medical or computer science domain produced by users per search session. DV3b corresponds to the total new keywords related to medical or computer science produced by users per search session.

2.2 Participants

Twenty participants performed the experience: ten in computer science (6 males, 4 females) and ten in medical (5 males, 5 females). The age of participant was ranging from 20 to 32 years old (M= 24.6 *SD*= 3.3), for computer science (M= 23.8 *SD*= 3.2) and for medical (M= 25.4 *SD*= 3.4). All of them were students in master degree, five for computer science and two for medical, or were PhD students, five in computer science and 8 in medical. We selected students with similar level of information search to avoid its influences on information search activity (Sanchiz et al., 2020). We controlled this variable through pre-test online distributed from Qualtrics XM plateform, which contained a self-efficacy scale in information search (Rodon & Meyer, 2018). The total score was calculated from the ten items proposed with a 4-point likert scale. There were no significant differences between two groups ($t(18)= 1.24$, $p > .05$, computer science students (M= 33.5 *SD*= 5) and medical students (M= 30.8 *SD*= 4.83). We also controlled the level of prior domain knowledge, through a self-report 5-point Likert scale of prior domain knowledge. There was a significant difference for the computer science knowledge self-report ($t(18)= 5.5$, $p < .001$), where computer science students reported to have higher knowledge (M= 4.2 *SD*= 0.42) than medical students (M= 2 *SD*= 1.2). The reverse was obtained for medical knowledge ($t(18)= -11.5$, $p < .001$), for which medical students indicated to have higher knowledge (M= 4.2 *SD*= 0.6) than computer science students (M= 1.3 *SD*= 0.5). Participants also had to complete a knowledge questionnaire in the two domains, with ten questions for each domain and 5 possible answers per question (one right, three wrongs and one "I do not know"). Concerning the computer science knowledge test ($\alpha= .94$), computer science students obtained better score (M= 6.8 *SD*= 2.2) than medical students (M= 0.5 *SD*= 0.7) ($t(18)= 8.8$, $p < .001$). We also found a significant difference ($t(18)= -12.1$, $p < .001$) for the medical knowledge test ($\alpha= .84$): medical students (M= 5 *SD*= 0.7) had better scores than computer science students (M= 0.5 *SD*= 1).

2.3 Procedure

The study was in two stages. First, participants received a first mail containing a link to the pre-test online (i.e. demographic information, age, level and domain of studies, knowledge self-report scale, ten question tests, self-efficacy scale in information search). Participants had to sign a free and informed consent. Second, given COVID-19 crisis, we scheduled an appointment with each participant to provide him/her the experimental material (i.e. general instructions, USB key containing the software for experiment that allowed retrieving logs during search session and the instructions with task statements). At the end of the experiment, all participants received a gift-card of 15 euros. During search sessions, participants solved ten tasks: 4 fact-finding search tasks (i.e. simple task and inferential task in each domain, computer science and medical domain), and 6 learning search tasks (i.e. exploratory learning, decision-making and problem solving task in each domain too). Some examples of task statements per domain are introduced in Table 1.

Table 1. Examples of task statements per domain.

Task type	Examples of task statement
Simple	*Medical* - What is the value of severe hyponatremia?
Exploratory	*Computer science* - You want to learn more about "Big Data".
Decision-making	*Medical* - An 83-year-old woman had an unremitting stroke 5 months ago. On the stroke assessment, atrial fibrillation was discovered. She has fallen 3 times in the last two months. Should anticoagulant treatment be started? After evaluating the risk-benefit ratio of starting anticoagulant treatment or not, select the management that seems best for you and justify your choices.
Problem solving	*Computer science* - As part of your job interview, you will be asked to create a resource that allows you to transcribe a text written in textos language into a text written in a well-trained language. With information collected on the internet, propose a general but precise methodology that shows your assets and motivates the employer to hire you.
Inferential	*Medical* - A very young person comes in for consultation and presents a sudden and transient onset dermatitis. By observing the lesions, we note the presence of papules. At the rest of the clinical examination, adenopathy are found. In your opinion, what does this patient suffer from?

3 Results

For each dependents variables, we performed an ANOVA (repeated measures) on three independents variables: 1. Prior domain knowledge level (experts vs non-experts) as between-subject factor; 2. Task type (simple, exploratory learning, decision-making, problem solving, inferential) as within-subject factor; 3. Task domain (medical vs computer science) as within-subject factor. When ANOVA was significant, we performed Scheffe post-hoc. All results with means and standard deviations are presented below.

Concerning the total number of new queries produced and their length, none significant effect appeared (*ps* > .05).

For the number of keywords from task statement, the ANOVA was not significant for the expertise ($p > .05$) nor the interaction between expertise and the task type ($p >$ .05). But, the ANOVA was significant for the interaction between expertise and task domain ($F(1,18)= 37$, $p < .001$, $\eta_p^2= 0.70$). Computer science experts used more keywords from the statement for medicine tasks (M= 4.3 *SD*= 3.1) than computer science tasks (M= 3.04 *SD*= 3.1) with $p < .001$. On the contrary, medical experts used more keywords from the statement for computer science tasks (M= 4.22 *SD*= 2.73) than medical tasks (M= 2.8 *SD*= 2.8) with $p < .001$. The interaction between expertise and task type and domain was significant ($F(4,72)= 12.4$, $p < .001$, $\eta_p^2= 0.41$). For decision-making tasks, medical experts used more keywords from statement in computer science (M= 7.4 *SD*= 2.8) than for medicine decision-task (M= 1.9 *SD*= 1), with $p < .001$. In addition, medicine experts used more keywords from statement in computer science than computer science experts (M= 3.8 *SD*= 2.44; $p < .05$). For the decision-making task in medicine, computer science experts used more keywords from the statement (M= 5.7 *SD*= 3.3) than medical experts ($p < .001$). For problem solving task in medicine, experts in computer science (M= 8.3 *SD*= 3.43) and medicine (M= 6.1 *SD*= 2.3) use more keywords from the statement than when solving the computer science task, with experts (M= 2.8 *SD*= 1.6) and medical experts (M= 2.8 *SD*= 0.8) using $p <.001$ for both comparisons. For the inferential task in computer science, medical experts used more keywords from the statement (M= 7.1 *SD*= 1.9) than they did for the inferential task in medicine (M= 2.6 *SD*= 1.6), with $p < .001$.

For the number of not specific new keywords, the ANOVA did not reveal any significant effect of expertise, nor interaction between expertise and task type ($p < .05$). The ANOVA was significant for expertise × task domain interaction ($F(1,18)=6.43$, $p < .05$, $\eta_p^2= 0.3$): Computer science experts produced more not specific keywords ($p = .001$) in computer science tasks (M= 2.9 *SD*= 2.8) than in medical tasks (M= 1.12 *SD*= 2.8). They also produced more not specific keywords ($p < .05$) than medical experts did in computer science tasks (M= 1.5 *SD*= 1.8) and in medical tasks (M= 1.10 *SD*= 1.8) with, $p < .05$.

For the interaction between expertise, type and domain, ANOVA indicated a significant effect ($F(4,72)= 2.8$, $p < .05$, $\eta_p^2= 0.13$). Post-hoc analysis showed that for inferential tasks only, computer science experts produced more not specific keywords (M= 5.6 *SD*= 3.53) than they did in the medicine task (M= 1.10 *SD*= 1), with $p < .001$. Finally, computer science experts produced more not specific keywords in the inferential computer science task than did medical experts in the inferential medicine task (M= 1.2 *SD*=1.5), with $p < .001$.

Concerning the main effect of expertise on the number of specific keywords, the ANOVA was significant ($F(1,18)= 7.11$, $p <.05$, $\eta_p^2= 0.30$). Medical experts produced more specific keywords (M= 1.2 *SD*= 1.71) than computer science experts (M= 0.6 *SD*= 1.21). The ANOVA was not significant for expertise × task type interaction ($p > .05$). The interaction between expertise and task domain was significant ($F(1,18)= 23.8$, $p < .001$, $\eta_p^2= 0.57$): Medical experts formulated more specific domain keywords ($p < .05$) in medical tasks (M= 1.90 *SD*= 1.80) than computer science experts in their task domain (M= 0.80 *SD*= 1.20). They also formulated more ($p < .001$) than when solving computer science tasks (M= 0.54 *SD*= 1.71) and more ($p < .001$) than computer science experts completing medical tasks (M= 0.44 *SD*= 1.26).

4 Discussion and perspectives

The present experiment did not show any significant effect of expertise on number of queries and their length. In contrast, significant differences appeared concerning the keywords produced. Non-experts used more task statement words when these ones were not from their domain, whereas experts when solving tasks in their domain used fewer keywords from task statements. This effect was particularly true for decision-making and inferential tasks. In addition, computer science experts tended to produce more not specific new keywords (i.e. common language) in computer science than in medical tasks. In contrast, medical experts tended to formulate more queries with more domain specific words related to a high vocabulary in medicine when solving tasks in medicine than the non-experts. Results showed that experts users translated easier learning task goals to others terms, whereas non-experts needed to rely on statements. However, the generation of new keywords tended to be more not specific in computer science and more specific in medicine. This difference may be explained by the fact that computer science vocabulary (e.g. software, programming…) are words fallen in the everyday language whereas medicine words are more specific to this domain.

One main limit of this study is the sample size. Currently, we are retrieving more data from more participants in computer science and medicine. In addition, to study only query formulation strategies is not enough to understand relationships between search behavior and learning. To bring deepen result interpretations, we will analyze the relevance of the outcomes (answers) provided as well as variables from questionnaires completed before and after each tasks (e.g. expected and perceived difficulty, self-perception of answer quality). Finally, further studies should investigate the expertise domain on search abilities during search as learning. The aim is to link search and learning variables to determine difficulties experienced by no-experts users and how experts users do to perform better than no-experts users during complex learning tasks? These studies will allow proposing new web navigational supports for users who are searching complex information out of their domain of knowledge.

5 Acknowledgments

The French National Research Agency (ANR), CoST-Modelling Complex Search Tasks (ANR-18-CE23-0016), supported this research.

6 References

Campbell, D. J. (1988). Task Complexity: A Review and Analysis. *Academy of Management Review*, *13*(1), 40-52. https://doi.org/10.5465/amr.1988.4306775

Gwizdka, J., Hansen, P., Hauff, C., He, J., & Kando, N. (2016). Search as Learning (SAL) Workshop 2016. *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, 1249-1250. https://doi.org/10.1145/2911451.2917766

Hembrooke, H. A., Granka, L. A., Gay, G. K., & Liddy, E. D. (2005). The effects of expertise and feedback on search term selection and subsequent learning. *Journal of the American Society for Information Science and Technology*, *56*(8), 861-871. https://doi.org/10.1002/asi.20180

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, *49*(4), 41-46. https://doi.org/10.1145/1121949.1121979

Monchaux, S., Amadieu, F., Chevalier, A., & Mariné, C. (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), 557-569. https://doi.org/10.1016/j.ipm.2015.05.004

O'Brien, H. L., Kampen, A., Cole, A. W., & Brennan, K. (2020, mars). The Role of Domain Knowledge in Search as Learning. *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, 313-317. https://doi.org/10.1145/3343413.3377989

Rodon, C., & Meyer, T. (2018). Self-efficacy about information retrieval on the web across all domains : a short measure in French and English. *Behaviour & Information Technology*, *37*(5), 430-444. https://doi.org/10.1080/0144929x.2018.1449252

Sanchiz M., Amadieu F., & Chevalier A. (2020). An Evolving Perspective to Capture Individual Differences Related to Fluid and Crystallized Abilities in Information Searching with a Search Engine. In W.T. Fu, & H. van Oostendorp (Eds.). *Understanding and Improving Information Search*, Human–Computer Interaction Series. Springer, Cham, 71–96. https://doi.org/10.1007/978-3-030-38825-6_5

Sanchiz, M., Chin, J., Chevalier, A., Fu, W. T., Amadieu, F., & He, J. (2017a). Searching for information on the web: Impact of cognitive aging, prior domain knowledge and complexity of the search problems. *Information Processing & Management, 53*(1), 281-294. https://doi.org/10.1016/j.ipm.2016.09.003

Sanchiz, M., Chevalier, A., & Amadieu, F. (2017b). How do older and young adults start searching for information? Impact of age, domain knowledge and problem complexity on the different steps of information searching. *Computers in Human Behavior*, *72*, 67-78. https://doi.org/10.1016/j.chb.2017.02.038

Sharit, J., Taha, J., Berkowsky, R. W., Profita, H., & Czaja, S. J. (2015). Online Information Search Performance and Search Strategies in a Health Problem-Solving Scenario. *Journal of Cognitive Engineering and Decision Making*, *9*(3), 211-228. https://doi.org/10.1177/1555343415583747

Tamine, L., & Chouquet, C. (2017). On the impact of domain expertise on query formulation, relevance assessment and retrieval performance in clinical settings. *Information Processing & Management*, *53*(2), 332-350. https://doi.org/10.1016/j.ipm.2016.11.004

Vakkari, P. (2016). Searching as learning: A systematization based on literature. *Journal of Information Science, 42*(1), 7-18. https://doi.org/10.1177/0165551515615833