



Wen, L. and Ruthven, I. and Borlund, P. (2006) The effects on topic familiarity on online search behaviour and use of relevance criteria. In: Proceedings of the 28th European Conference in Information Retrieval (ECIR 2006). Springer.

<http://eprints.cdlr.strath.ac.uk/2750/>

This is an author-produced version of a paper published in Proceedings of the 28th European Conference in Information Retrieval (ECIR 2006). This version has been peer-reviewed, but does not include the final publisher proof corrections, published layout, or pagination.

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in Strathprints to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profitmaking activities or any commercial gain. You may freely distribute the url (<http://eprints.cdlr.strath.ac.uk>) of the Strathprints website.

Any correspondence concerning this service should be sent to The Strathprints Administrator: [eprints@cis.strath.ac.uk](mailto:eprints@cis.strath.ac.uk)

# The Effects on Topic Familiarity on Online Search Behaviour and Use of Relevance Criteria

Lei Wen<sup>1</sup>, Ian Ruthven<sup>1</sup> and Pia Borlund<sup>2</sup>

<sup>1</sup> Department of Computer and Information Sciences  
University of Strathclyde, Glasgow, G1 1XH, United Kingdom  
amanda\_wen37@hotmail.com, ir@cis.strath.ac.uk

<sup>2</sup> Department of Information Studies  
Royal School of Library and Information Science, DK-9000 Aalborg, Denmark  
pb@db.dk

**Abstract.** This paper presents an experimental study on the effect of topic familiarity on the assessment behaviour of online searchers. In particular we investigate the effect of topic familiarity on the resources and relevance criteria used by searchers. Our results indicate that searching on an unfamiliar topic leads to use of more generic and fewer specialised resources and that searchers employ different relevance criteria when searching on less familiar topics.

## 1 Introduction

Searchers of online resources make a variety of judgments on the material retrieved by Information Retrieval (IR) systems. In evaluations of IR systems these judgments are often simplified into decisions of relevance: are the documents relevant or not to the searcher? However, studies of how people search for information, and the assessments they make on retrieved material, show that this is overly simplistic and people instead make a range of decisions on material they encounter whilst they are searching. These decisions are generally known as *relevance criteria*, reasons people give for assessing material as relevant. There have been many studies of relevance criteria uncovering either which criteria are used in searching [1], how criteria change over time [7] and how these criteria relate to the search problem being undertaken [8].

Previous work [e.g. 4, 5, 6] also shows that topic familiarity can affect a searcher's search strategy and the type of information they believe they will require. In this paper we present the results of a pilot study to investigate the degree to which a searcher's familiarity with a topic affects their use of resources and their ability to predict which relevance criteria will be important to them in a search.

## 2 Methodology

18 participants took part in this study, all of whom were postgraduate students in Masters' level courses at the University of Strathclyde. Each participant was asked to

search on basis of two simulated work task situations [2] both of which had a common structure in which the participant was asked to find 10 or more good resources (books, web pages, or articles) that they would recommend to a colleague interested in that topic. The major difference between the two tasks was the task *topic*.

The first task, given to all participants, asked the participants to find information on the topic of 'multimedia information retrieval'. As all participants took part in a compulsory course in Information Retrieval as part of their degree course, this was a topic with which we expected them to be familiar. For the second task the topic varied between participants. Participants were asked at the start of the study which of three domains (psychology, history, or architecture) was the least familiar to them and they were then given a search task based on that topic. 11 participants chose architecture as the topic of which they had least knowledge, 2 chose history and 5 chose psychology.

Simply asking the participants which topic is least familiar does not guarantee that they are unfamiliar with the topic. However, as part of the study the participants were asked to assess the familiarity of each search topic on a 5-point scale (1 being 'very unfamiliar', 5 being 'very familiar'). The average rating given to the familiar task was 4.1, and 2.1 for the unfamiliar task. Using a Wilcoxon Test we found a statistically significant difference between the participants' responses showing that the participants felt more familiar with the topic 'multimedia information retrieval'.

For both search tasks the participants were asked to find a number of resources. As the study was aimed at understanding the effects of topic familiarity on the participants' existing search behaviour we placed no restrictions on the participants search strategies: the participants could search in any way they felt comfortable and could use any search engine, database or visit any website they wished. The only restriction we placed on their searching was that they were not allowed to ask for recommendations from the investigators or to ask for opinions on the resources they found. The participants were restricted to 15 minutes on each search task and the study took place in the departmental laboratory which was the main laboratory for the participants and hence is a familiar place to carry out work tasks.

As well as the effects of topic familiarity on search behaviour we were also interested in how topic familiarity would affect people's relevance criteria: would people with less familiarity use different criteria in assessing relevance? To investigate this we gave the participants a list of 12 relevance criteria drawn from the criteria described in [1]. We selected the relevance criteria that could reasonably be applied to any search task and which were mentioned as being the most frequent in [1]. The relevance criteria we chose were *depth/scope/specificity*, *accuracy/validity*, *clarity*, *currency* (recent information), *tangibility* (material contains specific facts), *quality of sources*, *accessibility* (information is available), *verification a.* (information is consistent with other information and *verification b.* the participant agrees with the information), *affectiveness a.* (the participant enjoys reading the material and *affectiveness b.* the document is interesting) and *background experience* (participant is familiar with the topic or source of the document).

The criteria were presented in a list with a simple English description of each criterion as part of the initial study interview rather than before each search task. Each participant was asked to note which criteria they felt would be important to them in assessing material for different types of search task including familiar and unfamiliar tasks. There was a slight, although not significant, difference in the number of criteria

chosen per task (average of 6 criteria for familiar task vs. 5.6 criteria for unfamiliar task) and a similar distribution of criteria to both tasks. The main difference was that all participants rated the criterion *background experience* as being important for the familiar task whereas only 4 participants rated this as important for the unfamiliar task. *Currency* and *verification b.* were also more important for the familiar task.

### 3 Findings

The effects of topic familiarity in this study can be seen in two areas of searching: the use and selection of resources and how searchers employed relevance criteria. The participants used more formal resources and search engines when searching on the familiar task than the unfamiliar task (average 2.6 resources in the familiar task vs. average of 3 for the unfamiliar task) with a higher use of domain specific resources on the familiar task, e.g. Google Scholar, or ACM Digital Library, than on the unfamiliar task where more generic resources were used such as Yahoo, or Wikipedia. The participants' reasons for using these generic resources were that they helped either structure their searching, e.g. Yahoo hierarchies, or provided more information on the topic which helped the participant assess the quality of retrieved material, e.g. Wikipedia. This is in line with work by Michel [5] who found that topically experienced searchers are better at assessing the relevancy of retrieved objects.

There was a strong relationship between familiarity and the number of sources recommended: participants recommended significantly more documents on the unfamiliar task than on the familiar task (average 9.9 familiar task vs. 12.6 unfamiliar task). This relates to the work of Byström and Järvelin [3] who found that low task familiarity increases the complexity of a task and that, as task complexity increases, the need for more sources of information increases. The increase in number of sources found could also be due to an increased number of sources available. It is difficult to assess the number of sources available for a given topic on the Internet but we note that 15 of the 18 participants recommended more sources for the unfamiliar task.

In section 2 we explained that the participants were asked to predict which relevance criteria they would use in assessing material. After searching we asked the participants to view the recommended resources and discuss *why* they would recommend them to a colleague. This was to elicit whether the criteria they did use in assessing relevance were the ones that they predicted would be important. For familiar tasks the predicted relevance criteria were generally similar to the ones used in assessing the retrieved material. For unfamiliar tasks, however, they were often not the same and criteria such as depth/scope, or accuracy could not be easily employed. Although the participants were confident before searching that these criteria were important features of the information they sought, these were criteria that could not easily be judged on an unfamiliar topic. Instead, participants compensated for their lack of topic familiarity by using other criteria. For example, *tangibility* was employed as a substitute measure for *accuracy* because the extent to which specific information was provided was seen as a measure of the reliability of information.

The use of a criterion such as *accessibility* was also not straightforward. On unfamiliar tasks participants would recommend resources that were less accessible,

e.g. a book from Amazon rather than a freely available book from the University Library, because generic sites such as Amazon gave more information on the resources. As one participant noted “Amazon represented more book details...even a few paragraphs from each chapters, this information is very helpful to evaluate a book which I have never seen before”. The University library, on the other hand, only supplies limited bibliographic information. The degree to which information is consistent with other information (relevance criterion *verification a*) was more important for unfamiliar tasks because participants reported that in early search stages they were forced to learn about a topic whilst searching and consistent information helped the process of learning about a topic.

#### **4. Conclusions**

This paper reports on a pilot test to investigate the effects of topic familiarity on search behaviour. We found that topic familiarity can affect the number and type of resources selected by searchers and the ability of a searcher to use relevance criteria.

#### **Acknowledgements**

We gratefully acknowledge the help of our participants and the comments made by the referees. The research is supported by The Royal Society Research Project *Task-centred evaluation methodologies for interactive information retrieval*.

#### **References**

1. Barry, C. and Schamber, L.: Users' Criteria for Relevance Evaluation: a Cross-situational Comparison. *Inf. Proc. & Manag.* 34. 2/3. (1998) 219–236.
2. Borlund, P.: The IIR Evaluation Model: a Framework for Evaluation of Interactive Information Retrieval Systems. *Inf. Res.* 8.3. (2003) <http://informationr.net/ir/8-3/paper152.html>
3. Byström, K. and Järvelin, K.: Task Complexity Affects Information Seeking and Use. *Inf. Proc. & Manag.* 31. 2. (1985) 191–213.
4. Hsieh-Yee, I.: Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experiences Searchers. *J. Am. Soc. Inf.* 44. 3. (1993) 161–174.
5. Michel, D.: What is Used During Cognitive Processing in Information Retrieval and Library Searching? Eleven Sources of Search Information. *J. Am. Soc. Inf.* 45. 7. (1994). 498–514.
6. Serola, S. and Vakkari, P.: The Anticipated and Assessed Contribution of Information Types in References Retrieved for Preparing a Research Proposal. *J. Am. Soc. Inf.* 56. 4. (2005) 373-381.
7. Vakkari, P. and Hakala, N.: Changes in Relevance Criteria and Problem Stages in Task Performance. *J. Doc.* 56 .5. (2000) 540–562.
8. Yuan, X-J, Belkin, N. J. and Kim, J-Y.: The Relationship Between ASK and Relevance Criteria. In: Järvelin, K., Beaulieu, M., Baeza-Yates, R., and Myaeng, S. H. (eds): *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. (2002) 359–360