

Creating Exploratory Tasks for a Faceted Search Interface

Bill Kules

The Catholic University of America
School of Library and Information Science
Washington, DC
kules@cua.edu

Robert Capra

University of North Carolina at Chapel Hill
School of Information and Library Science
Chapel Hill, NC
rcapra3@unc.edu

ABSTRACT

In this paper we describe a process for creating and evaluating exploratory tasks for a faceted search interface. We used the tasks in an eye tracking study of a faceted library catalog search interface. We report on user perceptions of the tasks. The method is intended to be extensible to generate exploratory tasks for other types of interfaces and domains.

INTRODUCTION

Designing exploratory search tasks is an important, but challenging, requirement for successfully evaluating exploratory search interfaces. When creating any type of search task, there is a challenge of creating a realistic, representative task. When creating exploratory search tasks there is an additional burden of actually inducing an **exploratory** search. This high level goal of doing an exploratory search drives how users interpret the tasks, their relevance, and the results (Kules & Shneiderman 2008).

We set out to explore interfaces for exploratory search in a library Online Public Access Catalog (OPAC). Specifically, we were interested in studying facet use in exploratory search in a faceted-OPAC system such as the one currently in use at North Carolina State University (<http://www.lib.ncsu.edu/catalog/>).

Creating well-grounded, realistic exploratory search tasks was one of the primary challenges of the study design. Exploratory search tasks in a library catalog are a form of what librarians call “subject searches”. We differentiate exploratory tasks because a subject search in a catalog can take place at any stage of the search process, whereas exploratory search describes the high level goal of the task. In this work, we explicitly situate the subject search at the early stage of the overall search and design tasks that induce subject search driven by a high level scenario. To create such tasks, we first needed to operationalize exploratory search for this study. Second, we needed to construct a concrete set of tasks that were appropriate for the system being used.

Operationalizing exploratory search

Exploratory tasks inherently have uncertainty, ambiguity and discovery as common aspects (White, Kules, et al.; Marchionini 2006). The searcher may not know the domain well and the information need may be ambiguous or imprecise. In addition, exploratory search typically requires

retrieving multiple results to achieve the objective. This suggests several operational characteristics for exploratory search tasks:

- Answers are not found on the first interaction
- Searchers interact with the results and/or reformulate their queries
- Searchers search for multiple items

We used these characteristics to drive the development of our search tasks based on topics mined from actual usage logs of the North Carolina State University (NCSSU) OPAC.

Desirable characteristics of exploratory tasks

The literature suggests a number of desirable characteristics for exploratory search. Marchionini (2006) lists exploratory tasks, characterizing them as either learning-oriented or investigative. This suggests that the high-level scenario should be described so that it involves learning or investigation. Kuhlthau (1991) describes six stages of search and predicts various types of searcher interaction. Early stages are characterized by uncertainty.

Task complexity refers to the degree of predeterminability of task performance (Byström and Järvelin, 1995). Some tasks are well established and understood (known), while others are more unique and less understood (genuine decision tasks). Problem structure, task complexity and prior knowledge have an interconnecting impact when searching. “The more complex the task, the more ill-structured it is, and the less prior knowledge the actor has.” (Vakkari 1999).

Borlund (2000) advises that simulated situations include: “i. A situation which the test persons can relate to and in which they can identify themselves; ii. A situation that the test persons find topically interesting, and; iii. A situation that provides enough imaginative context in order for the test persons to be able to relate and apply the situation.”

Kules & Shneiderman (2003) used four simulated work tasks for journalists constructed around an exploratory search task to evaluate a faceted web search interface, drawing on Yee et al. (2003), which included open-ended tasks that were constructed with similar objectives.

This brief review suggests that exploratory search tasks should:

- Indicate uncertainty, ambiguity in information need and/or need for discovery.

- Suggest a knowledge acquisition, comparison, or discovery task
- Be an unfamiliar domain for the searcher
- Provide a low level of specificity about:
 - The information necessary for their search
 - How to find the required information
 - How to recognize the required information
- Be a situation which the test persons can relate to and in which they can identify themselves
- Be a situation that the test persons find topically interesting
- Be a situation that provides enough imaginative context in order for the test persons to be able to relate and apply the situation.

Not all of these are practical or feasible, however. For example, in our study, we constrained the searchers to use the faceted OPAC, which indicates a very specific direction for “how to find the required information.” Also, we may not be able to control for prior knowledge. Instead it may be more practical to measure it and analyze that factor.

TASK CONSTRUCTION

We followed a two-step approach to create the exploratory search tasks used in our study. First, we mined log files from the NCSU OPAC for topics that met a series of criterion. Second, we plugged the topics extracted from the log analysis into “task templates” that we designed to motivate an exploratory search. Each of these steps is described in more detail in the sections below.

Topic extraction from log data

We had the benefit of partnering with NCSU on this study and thus had access to several days of anonymized log data from their OPAC. The log files provided a list of queries issued to the OPAC. For each query, the keyword string and list of facets selected was recorded in the log. While we could not determine a searcher’s exact intent from the data, we looked for searches that had characteristics of exploratory searches based on the operationalized characteristics outlined in the first section of this paper. Additionally, we were interested in facet usage, so we included it as a criterion when examining the log data. We scanned the log files looking for searches in which:

- *Answers were not found on the first interaction* – we looked for searches with multiple page views.
- *Searchers used facets* – we looked specifically for the use of one or more of three facets: Subject, Region, and Time Period. We selected these three because they were commonly used and they were not specific to NCSU’s library system.
- *Searchers interacted with the results and/or reformulate their queries* – we looked for searches in which facets were added to an original query.
- *Searchers searched for multiple items* – again we looked for searches with multiple page views.

These criteria indicate that the user did not find the results on their first interaction and either reformulated the search or interacted with the results. We disregarded instances where a facet chosen was either identical or similar to the search terms (for example: a search for “cotton management” modified by selecting the Subject facet “cotton”). We also disregarded instances where the user needed to use a “show more” option to see additional facet values because we wanted to focus on facets visible from the initial results page. For example, from one log file, we observed queries for the search term “British History” with the facets “History” (subject) and “Twentieth Century” (time period). From these log entries, we developed a candidate topic “British History”. We intentionally included facets in the task creation process, because our goals were to study searcher behavior in this context.

Mining the log data for searches that involved multiple interactions could lead to searches that were problematic rather than exploratory. For example, a bad interface, or poor match between facets and the task could lead to multiple interactions. The refinement step described below should help address such tasks.

Plugging the topic into a task template

To help achieve the goal that the exploratory search tasks motivate consideration of multiple items, we developed a task template that involved finding multiple items – which the specific candidate topics could be plugged into. The objective of the template was to situate the participant in a familiar situation in which multiple items would need to be found. Since we recruited participants from a university population, we used a task that involved writing a paper for a class..The basic form of the template is shown below:

*Imagine that you are taking a class called _____.
For this class, you need to write a paper on the topic _____.
Use the catalog to find two possible topics for your paper. Find three books for each topic.*

Based on prior experience creating exploratory search tasks, we asked participants to find specific target numbers of topics and books.

Task refinement

Once candidate tasks were created, we refined them by conducting a set of searches related to the topics on the NCSU OPAC. The purposes of this step were to: 1) clarify the wording of the task, 2) insure that the task was not too easy to qualify for use in an exploratory search, and 3) make sure that the task benefited from using facets (since facet use was a focus of our study). To do this, refined the tasks such that:

- Facet values matched one or more terms in the task; either exactly or a semantically close term
- The first 10 results did not answer the task. If the task was too easy, it would not require exploratory search.

- The facets were useful without having to click the “show more” link for the facet.

Using the example started in the previous section, we found that the query “British History” resulted in many relevant results in the top ten results returned. We then explored other topics that could be added to make the topic more challenging. By looking at the facets presented in the OPAC, we found that by adding the topic of “Colonies”, the task met our criterion. Thus, the final topic was “the relationship between Great Britain and its Colonies in the Twentieth Century”. The tasks generated by the process were reviewed by library science experts, and then pilot tested and further refined with three participants.

Resulting exploratory search tasks

Using the process described in the preceding sections, we developed four exploratory tasks (see A–D below). We also used two known-item tasks (E and F) based on a previous NCSU study. This was to permit comparisons with that study. The final tasks used in the study are given below.

A. Imagine you are taking a class called “Feminism in the United States”. For this class you need to write a research paper on some aspect of the U.S. feminist movement, but have yet to decide on a topic. Use the catalog to find two possible topics for your paper. Then use the catalog to find three books for each topic so that you might make a decision as to which topic to write about.

B. Your professor wants you to write a paper comparing the textile industry in three countries in three different continents. Use the catalog to find three countries which have a textile industry about which books have been written. Find three books for each country.

C. Imagine you are taking a class titled “Great Britain and its Colonies in the Twentieth Century”. For this class you need to write a research paper on some aspect of the relationship between Great Britain and its Colonies in the Twentieth Century but you have yet to decide on one. Use the catalog to find two possible topics for your paper. Then use the catalog to find three books for each topic so that you might make a decision as to which topic to write about.

D. You are taking a class called “History of the Olympic Games” for which you need to write a research paper. You have yet to decide on a specific topic for this paper. Use the library catalog to explore possible topics and find two. Then find at least three books for each so that you might make a decision as to which topic to write about.

E. Your professor has suggested that your group begin your project on Conservation and Biological Diversity by looking up background information in a book titled Firefly encyclopedia of trees.

F. You are working your way through the Harry Potter books and are ready to read the next one on your list, titled “Harry Potter and the Goblet of Fire”.

METHODS

Our broader goals in this research were to investigate facet use in exploratory search when using a library OPAC. Generating a set of well-grounded, representative tasks that

would induce exploratory search was a significant challenge in the study design. As part of the study, we included metrics and measures to give us feedback on the tasks to see if we had achieved our goals for task creation. In this section, we present details of the study as they relate to evaluating the tasks.

Twenty-one participants were recruited from the University of Maryland at College Park (UMD) to participate in this study. Of these, data was successfully collected from 18 (two sessions were unsuccessful due to system problems and we were unable to calibrate the eye tracker for one participant). The testing system was a web-based, faceted OPAC interface based on a modified version of the North Carolina State University library catalog of over 1.8 million titles. The study was conducted in the Human-Computer Interaction Lab at UMD using a computer equipped with an eye-tracker. Results related to the eye-tracker are outside the scope of this paper and will be reported elsewhere. Data was collected about the searches issued, the results selected, and the facets used for each task.

The participants were shown a 90 second video demonstration of the interface. They then conducted six short searches motivated by the tasks, completed a questionnaire and provided a retrospective verbal report while viewing screen video of their searches with their gaze pattern overlaid. The exploratory tasks were presented first, followed by the known item tasks. Within each task type, presentation order was counterbalanced to minimize order and learning effects. In between each task, participants completed a questionnaire with five questions about their experience. All responses were given as ratings on 5-point Likert-type scales (anchors shown in parenthesis):

1. How familiar were you with this subject when you began this task? (1 = not familiar at all, 5 = very familiar)
2. How difficult was it to accomplish this task? (1 = very difficult, 5 = very easy)
3. I am confident that I fulfilled the task asked of me. (1 = strongly disagree, 5 = strongly agree)
4. To what extent did completing this task involve finding a single item versus finding multiple items? (1 = single item, 5 = multiple items)
5. To what extent did you change what you were looking for based on the results you found? (1=not at all, 5=a lot)

Additionally, at the end of the session, we asked users to perform a card sort to group the six tasks according to what tasks they thought were most similar.

RESULTS

For the exploratory searches, none of the participants found their answer(s) on their first interaction – they all interacted with multiple pages.

Perceptions of tasks

Table 1 shows the averages and standard deviations (in parenthesis) of the participants' perceptions of the

exploratory and known item tasks based on the five questions asked after each task. Participants were slightly more familiar with the known item tasks and found them somewhat easier. They were also slightly more confident that they had accomplished the indicated task. Participants clearly differentiated between the number of items that each task required (single vs. multiple). They also changed what they were looking for more for the exploratory tasks.

	Exploratory n=72 avg (stdev)	Known-item n=36 avg (stdev)
1. Familiarity	2.6 (1.39)	3.0 (1.80)
2. Difficulty **	4.0 (0.91)	4.9 (0.23)
3. Confidence **	4.2 (0.94)	4.8 (0.80)
4. Single/Multiple **	4.2 (0.92)	1.4 (1.15)
5. Changed goal **	3.3 (1.33)	1.1 (0.40)

** significant difference found between exploratory and known-item at $p < 0.001$ using two-tailed T-test with $\alpha = 0.05$

Table 1. Overall Perception Ratings

Card sorting the tasks

We wished to learn whether participants perceived the exploratory tasks as similar to each other and different from the known item tasks, so we asked them to group the tasks “and put the ones that are the most alike together into groups.” Of the 17 participants who completed this step, all 17 put the two known item tasks (E & F) in their own group. Nine of the participants grouped tasks A, C, and D together, placing B separately. Three put A, B, C, D all together. The remainder had various grouping of A, B, C, D. When asked about task B, the explanations focused on the geographic nature of the task and the fact that it asks for books instead of topics, as the other three do. We anticipated the strong distinction between exploratory and known-item, but the sub-distinction of tasks within the exploratory set was unexpected and suggests that participants considered the geographic/topical and books/ideas differences to be important aspects of the nature of the tasks.

Limitations

Our operational definition of exploratory search was fairly narrowly tailored to the goals of this study. Future work should incorporate additional dimensions. Task complexity, in particular, is an important dimension – multiple levels of complexity in the task descriptions could be evaluated to determine what levels of complexity induce exploratory search behavior. Only one high level scenario was used for the task template. A broader range of scenarios should be explored and tailored to more directly fit test participants, consistent with Borlund’s (2000) recommendations for simulated work tasks.

DISCUSSION AND CONCLUSION

Overall, the tasks achieved our objectives. Based on the participants’ perceptions of the tasks, we believe that our procedure for task generation led to well-grounded, realistic

tasks that did elicit exploratory search behavior for the exploratory tasks. The exploratory tasks met the desired characteristics we outlined as goals: relatively low initial topic familiarity, require multiple items to be considered, and some ambiguity as to the final answers (as indicated by the confidence and changed goal measures). The difference in task B suggests that searchers differentiate between the indicated object (books vs paper topics) and by the nature of the facets (topical vs geographic).

This paper suggests a principled way of task building that incorporates consideration of the dimensions of the task, then building and refining the task description while taking into account both the broader dimensions of exploratory search and the pragmatics of the particular search system and collection technique. We hope that this task development strategy is a first step toward making tasks more comparable across studies.

ACKNOWLEDGEMENTS

This research was supported by a grant from Catholic University and in part by a grant from the NSF/Library of Congress (IIS 0455970). Thanks to Matt Banta for extensive assistance in implementing the study. We thank Tito Sierra, Jason Casden, and Joe Ryan at NCSU for development of the faceted OPAC interface, log analysis and contributing to the task construction. Thanks also to Doug Oard and members of the UMD HCIL for the use of their facilities and eye-tracker.

REFERENCES

- Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, **56**(1), 71-90.
- Byström, K. & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, **31**(2), 191 - 213.
- Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*, **42**(5): 361-371.
- Kules, B. and Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing & Management*. **44**(2): 463-484.
- White, W. Kules, B., Drucker, S., schraefel, m.c. (2006). Introduction. *Commun. ACM* **49**(4): 36-39.
- Vakkari, P. (1999). Task complexity, problem structure and information actions. Integrating studies on information seeking and retrieval. *Information Processing & Management*, **35**(6), 819-8
- Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proc. SIGCHI*, Ft. Lauderdale, FL (pp. 401-408). New York: ACM Press.