

BEYOND TEXT QUERIES AND RANKED LISTS: FACETED SEARCH IN LIBRARY
CATALOGS

Xi Niu

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Doctor of Philosophy degree in the School of Information and Library Science.

Chapel Hill
2012

Approved by:

Bradley M. Hemminger

Gary Marchionini

Barbara M. Wildemuth

Jane Greenberg

Bill Kules

©2012
Xi Niu
ALL RIGHTS RESERVED

ABSTRACT

XI NIU: Beyond Text Queries and Ranked Lists: Faceted Search in Library Catalogs
(Under the direction of Dr. Bradley Hemminger)

Since the adoption of faceted search in a small number of academic libraries in 2006, faceted library catalogs have gained popularity in many academic and public libraries. This dissertation seeks to understand whether faceted search improves the interactions between searchers and library catalogs and to understand ways that facets are used in different library environments. Interactions under investigation include possible search actions, search performance, and user satisfaction. Faceted catalogs from two libraries, the University of North Carolina at Chapel Hill (UNC-CH) Library and the Phoenix Public Library, are chosen as examples of two different facet implementations.

To observe searchers in natural situations, two log data sets with over 3 million useful records were collected from the two libraries' servers. Logs were parsed, statistically analyzed, and visualized to gain a general understanding of the usage of these faceted catalogs. Two user experiments were conducted to further understand contextual information, such as the searchers' underlying motivations and their perceptions. Forty subjects were recruited to search different tasks using the two different catalogs.

The results indicate that most searchers were able to understand the concept of facets naturally and easily. Compared to text searches, however, faceted searches were complementary and supplemental, and used only by a small group of searchers. When

browsing facets were incorporated into the search, facet uptake greatly increased. The faceted catalog was not able to shorten the search time but was able to improve the search accuracy. Facets were used more for open-ended tasks and difficult tasks that require more effort to learn, investigate, and explore. Based on observation, facets support searches primarily in five ways. Compared to the UNC-CH Library facets, the Phoenix Library facets are not as helpful for narrowing the search due to both its essential and lightweight facet design. Searchers preferred the Book Industry Standards and Communications (BISAC) subject headings for browsing the collection and specifying genre, and the Library of Congress subject headings (LCSH) for narrowing topics. Overall, the results weave a detailed ‘story’ about the ways people use facets and ways that facets help people employ library catalogs.

The results of this research can be used to propose or refine a set of practical design guidelines for designing faceted library catalogs. The guidelines are intended to inform librarians and library information technology (IT) staff to improve the effectiveness of the catalogs to help people find information they need more efficiently.

DEDICATION

This dissertation is dedicated to my parents, Chunlin Niu and Xiaodong Yang, and to my husband, Bing Xie.

ACKNOWLEDGEMENTS

Pursuing a Ph.D. degree in a second language ranks among the proudest activities I have accomplished. I would not have finished without the support of family, mentors, and friends.

When I knew little about research, Brad Hemminger introduced me to the field of information seeking behavior. He showed me a path that allowed me to apply my skills in data analysis and interest in technology into the practical research of understanding how people seek information for their own reference. He is an excellent hands-on advisor, offering me suggestions, directions, and financial support as well.

My committee have helped and challenged me along the way. Barbara Wildemuth has constantly provided me valuable advice and useful resources. I feel she is just an email away when I need her. Bill Kules has consistently challenged me to sharpen my thinking and clarify my writing. He flew from Washington D.C. to Chapel Hill every time of my important moments along the program. Jane Greenberg has been a mentor since my involvement in one of her projects. She has recommended me numerous papers and books, and also provided me detailed advice on the dissertation. Gary Marchionini has encouraged me to deeply explore fundamental human-computer interaction issues. As the dean with a busy schedule, he was able to find time advising me and serving on my committee.

My appreciation also goes to my family. As the only child of my aging parents, who live in China, I have been studying and working seven thousand miles away from them. During the six years, they offered me endless emotional, personal, and financial support. They always encouraged me when I questioned my ability. It is their unconditioned love that keep me move forward. I would like to thank my beloved husband, Bing Xie, as well. I know it is tough to be a husband of a Ph.D. student. Getting the degree means little housework, little entertainment, and always quiet time after dinner. Without his continued support and love, I could not have accomplished this degree.

I also want to thank some of my colleagues, such as Ben Pennell and Jill Sexton, the UNC library staff, who shared with me the transaction logs from their servers and helped me set up the experimental interfaces; Jesse Haro, the web service manager of the Phoenix Public Library, who volunteered to send me the data and the specification of their library catalog; Cory Lown, whose master's paper benefited me a lot; and Felix Portnoy, Sarah Ramdeen, and Wan-Ching Wu, my PhD fellows, who helped me grading the search performances, recruiting subjects, and transcribing the interviews. Without their help, my dissertation research would be impossible.

TABLE OF CONTENTS

1. Introduction.....	1
1.1 Problem Statement.....	2
1.2 Research Questions.....	6
1.3 Significance of the Study.....	10
1.4 Introduction to the University of North Carolina-Chapel Hill Library Catalog and the Phoenix Public Library Catalog.....	13
2. Literature Review.....	18
2.1 Context: Information Seeking Behavior in Online Library Catalog Environments.....	18
2.1.1 Brief History of Online Public Access Catalogs (OPACs).....	18
2.1.2 Search Behavior.....	21
2.1.3 Ways People Search Using OPACs.....	26
2.2 Facet Theory and Faceted Search.....	30
2.2.1 Facet Theory and Faceted Classification.....	30
2.2.2 Faceted Search.....	31
2.3 Academic Research on Faceted Search.....	33
2.3.1 Well-known Faceted Search Projects.....	33
2.3.2 Faceted Search Used in Library Catalogs.....	37
2.3.3 Empirical Studies on Faceted Search Systems.....	42

3. Research Methods	46
3.1 Transaction Log Analysis	46
3.1.1 Data Collection and Description	47
3.1.2 Data Processing.....	47
3.1.3 Data Analyzing	50
3.2 User Experiments.....	55
3.2.1 Subject.....	55
3.2.2 Within-subject Design	56
3.2.3 Procedure	58
3.2.4 Training.....	60
3.2.5 Tasks	61
3.2.6 Instruments.....	63
3.2.7 Variables and Data Collecting	64
3.2.8 Data Analysis.....	66
3.3 Summary	67
4. Results.....	68
4.1 Log Analysis Results	68
4.1.1 General Search Statistics.....	68
4.1.2 Facet Usage.....	79
4.1.3 Facets to Support Formulating or Reformulating a Search	83
4.1.4 Summary	86
4.1.5 Discussion	86
4.2 User Experiment 1: Faceted Search vs. Non-faceted Search.....	91

4.2.1 Characteristics of the Subjects	91
4.2.2 Search Time	92
4.2.3 Accuracy	97
4.2.4 Users' Perceptions	100
4.2.5 The Role Facets Play during the Search	106
4.2.6 Summary	115
4.2.7 Discussion	115
4.3 User Experiment 2: UNC catalog facets vs. Phoenix catalog facets	123
4.3.1 Characteristics of the Subjects	124
4.3.2 Search Time	125
4.3.3 Accuracy	129
4.3.4 Users' Perceptions	133
4.3.5 The Role Facets Play during the Search	137
4.3.6 Summary	144
4.3.7 Discussion	145
5. Conclusions and Discussion	152
5.1 Conclusions	152
5.2 Benefits of Facets	153
5.3 Limitations of Facets	155
5.4 Search Context: Laboratory vs. Naturalistic Setting	158
5.5 Individual Differences	159
5.6 Information Barriers in Library Catalogs	159
5.7 Design Principles for Faceted Library Interfaces	161

5.8 Limitations	164
5.9 Future Work	166
Appendix A. Consent Form	168
Appendix B. Questionnaires	171
Appendix C. Interview Questions.....	173
Appendix D. Sources of Questionnaire and Interview Questions	174
Appendix E. Tasks	176
Appendix F. Subject Training Scripts and Tasks.....	180
Appendix G. Email Exchange with Subjects.....	183
Appendix H. Receipts	185
References.....	186

LIST OF TABLES

Table 3.1 Log dataset	48
Table 3.2 Summary of action codes (both fine-grained and coarse-grained codes)	51
Table 3.3 Basic Experiment Flow	59
Table 3.4 Task types in library and information science literature	62
Table 3.5 Task types and examples	63
Table 3.6 Variables and data collecting from the experiments	65
Table 4.1 Coarse code of actions	76
Table 4.2a Popular facet list by frequency for the UNC catalog	82
Table 4.2b Popular facet list by frequency for the Phoenix catalog	82
Table 4.3 Quantitative measures for formulating/reformulating a search	83
Table 4.4 GLM regression results for search time	95
Table 4.5 Factorial result for search time.....	95
Table 4.6 GLM regression results for search accuracy	98
Table 4.7 Factorial results for search accuracy	99
Table 4.8 Codes and the number of times each was mentioned	104
Table 4.9 GLM regression results for search time	127
Table 4.10 Factorial results for search time	127
Table 4.11 GLM regression results for search accuracy	131
Table 4.12 Factorial results for search accuracy	132

LIST OF FIGURES

Figure 1.1 UNC Library Catalog (http://search.lib.unc.edu)	13
Figure 1.2 Phoenix Public Library Catalog (http://www.phoenixpubliclibrary.org)	15
Figure 2.1 Model of search in context (Järvelin & Ingwersen, 2004)	22
Figure 2.2 Exploratory search components (Marchionini, 2006)	25
Figure 2.3 Facets for a laptop search	32
Figure 2.4 Collection of environmental data from the National Aeronautics and Space Administration (NASA).....	34
Figure 2.5 Hierarchical facet navigation in Flamenco.....	35
Figure 2.6 Relation Browser	36
Figure 2.7 Interface of North Carolina State University’s faceted library catalog	39
Figure 2.8 Interface of Queens Library faceted library catalog.....	39
Figure 2.9 Interface of Georgia Public Libraries faceted library catalog	41
Figure 3.1 Process for visualizing transaction logs	53
Figure 3.2 Example of a visualized search session.....	54
Figure 3.3 Interfaces used in Experiment 1	57
Figure 3.4 Interfaces used in Experiment 2	58
Figure 3.5 Online answer collector interface screenshots	60
Figure 4.1 Distribution of session lengths based on number of actions	69
Figure 4.2 Distribution of action codes (a) in an original scale, and (b) in an adjusted scale	71
Figure 4.3 Beginning of a search: Text or facet.....	75
Figure 4.4 Clusters of (a) the UNC catalog searchers, and (b) the Phoenix catalog searchers.....	77
Figure 4.5 Action distribution for each cluster (a) UNC, and (b) Phoenix.....	78
Figure 4.6 Breakdown of <i>FacetOperation</i> for (a) UNC, and (b) Phoenix.....	80

Figure 4.7 Subjects' experience with (a) online library catalogs, (b) the UNC library catalog, and (c) searching with categories	92
Figure 4.8 Comparison of search time results for (a) interface, (b) task type, and (c) level of task complexity	93
Figure 4.9 Search time comparison between the two interfaces.....	96
Figure 4.10 Comparison of search accuracy results for (a) interface, (b) task type, and (c) level of task complexity	97
Figure 4.11 Search accuracy comparison between the two interfaces.....	100
Figure 4.12 Users' ratings for the two interfaces	101
Figure 4.13 Users' ratings of task complexity (Question 1 in the post-search questionnaire) ...	102
Figure 4.14 Facet usage for task type and level of task complexity	107
Figure 4.15 Examples: Facet for appending specific information.....	109
Figure 4.16 Examples: Facets for eliminating information	110
Figure 4.17 Examples: Facets for exploring	111
Figure 4.18 Example: Facets for reformulating text queries	112
Figure 4.19 Examples: Facets for browsing	114
Figure 4.20 Subjects' experience with (a) online library catalogs, (b) the UNC library catalog, and (c) categories	124
Figure 4.21 Comparison of search time results in terms of (a) interface, (b) task type, and (c) level of task complexity	126
Figure 4.22 Search time comparison between the two interfaces.....	128
Figure 4.23 Comparison of search accuracy in terms of percentage between (a) interface, (b) task type, and (c) level of task complexity	130
Figure 4.24 Interaction effect of interface and task complexity on search accuracy.....	131
Figure 4.25 Search accuracy comparison between the two interfaces.....	133
Figure 4.26 Users' ratings for the two interfaces	133
Figure 4.27 Users' ratings of task complexity (Question 1 in the post-search questionnaire) ...	134

Figure 4.28 Facet usage for task type and level of task complexity (Phoenix only)	138
Figure 4.29 Examples: Subjects tended to browse often	139
Figure 4.30 Example: Browse might increase search time.....	141
Figure 4.31 Example: Subjects tended to try one language at a time	141
Figure 4.32 Example: Subjects tended to use BISAC subject headings for genre information .	142
Figure 4.33 Example: Subjects tended to put the year information into the search box	143
Figure 5.1 Before (a) and after (b) adding facets to library catalogs.....	160

1. Introduction

Mankind by nature is an information consumer. As information becomes more and more ubiquitously available, various search technologies are in demand to facilitate the access to information and to learn about the world. A current search system must go beyond the traditional query-response and ranked list paradigm to incorporate the increase in human searching behavior, such as filtering, browsing, and exploring, in addition to simple look-up. Modern search engine technology already does a reasonable job of tackling the problem of what library scientists call *known-item search*, in which the user knows which documents to search for, or at least knows about certain aspects of the documents. In contrast, comparably mature tools for exploratory search, where the information needs and target documents may not even be well-established, are not well developed (Tunkelang, 2009). In addition, in order to organize search results, traditional search systems usually display results by in a single list ranked by relevance. Information seekers, however, often require a user interface that organizes search results into meaningful groups in order to better understand and utilize the results (Hearst, 2006).

Faceted search, which categorizes and summarizes search results, is a way to extend ranked lists. It also helps mitigate difficulties in query formulation and incorporates browsing into the search process. Faceted search is widely used in both commercial web search engines and library catalogs. Faceted classification, a classic theory in library science of knowledge representation developed in the 1930s by Ranganathan, overcomes the rigidity of traditional bibliographic classifications by offering a flexible, multidimensional view of

knowledge. Since 2006, facet theory has been actively used in information retrieval (IR) and employed to create numerous faceted search systems. Faceted search systems map the multidimensional classification of knowledge presentation level into multiple access points of knowledge access level. The central concept derived from early facet theory is that the facets are “clearly defined, mutually exclusive, and collectively exhaustive aspects” of knowledge (Taylor, 1992). In many current faceted search systems, however, the overlap of facets may occur, and the facets may not be exhaustive.

1.1 Problem Statement

In the field of interactive information retrieval (IR), it is well known that typical users are not adept at expressing their information needs through queries. They tend to input very few words when using web-based search engines (Jansen, Spink & Saracevic, 2000) and rarely take advantage of advanced search features when searching online public access catalogs (OPACs) (Lau & Goh, 2006). Poorly constructed queries provide inadequate evidence for the IR system to infer the user’s information needs, thereby leading to unsatisfying results. Typically, users must go through several iterations of formulating and reformulating queries before they reach the documents they need. Belkin’s (1980) concept of *anomalous state of knowledge* (ASK) describes the problem of “non-specifiability of information need”, which suggests the dilemma faced by users who are often ill-equipped to articulate their information needs.

Understanding ways to bridge users’ ASK(s) and the information source(s) becomes an essential problem of interactive IR. Currently, users are provided direct access to an unprecedented number of electronic information sources, yet most users have difficulty utilizing the full capacity of the substantial amount of information that IR systems offer.

Concepts such as relevance feedback (Rocchio, 1971), term suggestion (Belkin, 2000), and query expansion techniques (Magennis & Rijsbergen, 1997; Harman, 1988) are potentially effective ways to enrich the query and lead to improved performance in the IR task. Research into query assistance also suggests that longer queries lead to higher satisfaction and less iteration in an interactive web searching environment (Belkin et al., 2003).

In addition to the query elicitation techniques mentioned above, Borgman, Hirsh, Walter and Gallagher (1995) found that people find browsing easier than producing search queries. *Browsing* is defined as movement in a connected space (Kwasnik, 1992). Users who have only vague information or little knowledge of the document collections face considerable difficulty in specifying effective queries. Browsing provides an alternative to querying, allowing the user to navigate through the documents, using one document to find another. In browsing, people are scanning information items, omitting irrelevant ones and occasionally picking up new relevant data. By browsing within document collections, searchers can better define their information needs, become familiar with the information space and finally find ways to locate relevant information. Browsing has become an increasingly subtle searching activity in information-seeking research (Ingwersen & Wormell, 1989; P. Noerr & K. Noerr, 1985). According to Ellis (1989), browsing is “semi-directed or semi-structured searching” and an important part of standard information searching. He further suggests that browsing features should be made available in automated search systems to accommodate searchers’ browsing behavior that traditionally would have taken place physically in a library.

Switching from query formulation to post-query searching, conventional search systems usually display results by a ranked list generated by some type of algorithm. The

goal of the algorithm is to return to the user the most relevant documents first. Displaying results in an ordered list scheme is the most common way to present the search results, but is not without shortcomings. Long web search results lists can be difficult to browse and navigate, and people typically view only the first results page (Jansen & Pooch, 2001). Additionally, the problems of heterogeneous data, scale, and non-traditional information formats reflected in the documents, as well as the fact that search engines increasingly are integrated as a component of complex information management processes, not just stand-alone systems, demand new modes of system response to a query. Returning search results in terms of categories could be more helpful than providing just a ranked list in order to find the result (Hearst, 2006).

One approach to mitigate such difficulties in query formulation and to incorporate browsing into searching and organize search results into groups is the faceted approach, a form of information organization and access that was developed in the 1930s by Indian mathematician and librarian, Ranganathan. Facets are “clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject” (Taylor, 1992). The basic idea of faceted search is to display the indexing vocabularies (facets) of different dimensions as links to the users. The indexing vocabularies also are called *faceted metadata*, generated by either humans (English, Hearst, Sinha, Swearingen, & Lee, 2002; Frei and Jauslin, 1984) or algorithms (Chen, Yim, Fye, & Schatz, 1995). Metadata is defined as structured data that record contextual and behavioral information of an object (Greenberg, 2003). Faceted metadata offers users browsable structures in addition to the current search results and incorporates navigation into the search process.

In recent years, faceted search has grown to be a well-accepted approach for commercial websites (Breeding, 2007). Since the adoption of faceted search by the North Carolina State University (NC State) D. H. Hill Library in early 2006, faceted library catalogs have gained attention and popularity in many academic and public libraries. Faceted navigation has become one of the features used to describe next-generation catalogs.

Faceted search can be used in various online search systems, such as general web search engines, online databases and OPACs. OPACs are the special search systems that serve as catalogs for libraries and have a separate history from that of web search engines. The early OPACs were simply electronic replicates of card catalogs, and offered only basic search capabilities. The rapid development of web technology and the fast growth of bibliographic utilities have led to the development of OPACs as web-based rather than isolated entities. In addition to their web technology, OPACs also incorporate features such as web interfaces and search capabilities derived from web search engines and e-commerce websites.

This study investigates ways that people use facets in library catalogs and ways that facets help people find information they need. The advantages of faceted search over traditional search have become increasingly evident in commercial settings. Yet little is known about the effectiveness of faceted search as it pertains to library catalogs. Even less is known about the effect of different facet implementations on people's ways of interacting with the catalogs. This dissertation research is designed to investigate the effectiveness of faceted search in faceted library catalogs. From a system developer's perspective, faceted search may be helpful to users in at least three ways: to refine the initial search by drilling down to a specific facet, to gain a better understanding of the result set by browsing different

facets, and to discover hidden items by flattening out the depth of the original result set. However, it remains an empirical research question as to whether this facet feature really helps users to locate information they need and whether users are able to take advantage of this feature in the library environment.

1.2 Research Questions

Although the general application of facets to the information search process is of interest, this study's scope is much narrower, and examines library catalogs as the search environment. Thus, all the research questions are addressed within this scope.

This research has three purposes: (1) to attain a general understanding of the ways people search and employ faceted library catalogs, (2) to investigate whether faceted search interfaces improve the search using library catalogs, and (3) to explore the ways facets are used in different OPAC environments.

Based on these three research goals, three groups of specific research questions are posited. The first group of research questions (Questions 1, 2, and 3) primarily concerns the general search statistics and overall search patterns found when using faceted library catalogs. Depending on the information collections, local requirements, audience, and IT developments, faceted navigation is implemented differently for different libraries. Dozens of excellent faceted library catalogs can be found in both academic and public libraries. Two representative library catalogs were selected for investigation in this study: the University of North Carolina at Chapel Hill (UNC-CH) Library catalog and the Phoenix Public Library catalog (referred to in this paper as the UNC catalog and Phoenix catalog, respectively). Both

of these systems offer high quality facet implementation and use the same faceted search application, Endeca¹.

The first three specific research questions are:

1. What are the possible interactions people employ with faceted library catalogs, and how many of these interactions are faceted interactions?
2. Do search sessions segregate naturally into clusters according to their search behavior?
3. Do searchers change the way they formulate or reformulate their search with facets compared to without facets?

Question 1 aims to understand the general search process and overall search statistics in terms of using faceted catalogs. Question 1 addresses factors such as the length (duration) of a typical search session, the types of actions people might conduct, the proportion of facet-related operations, ways people initiate a search, and identification of frequently used facets or facet combinations. These factors are examined quantitatively by descriptive statistics. The results are compared and contrasted across the two library catalogs.

According to previous studies, searchers tend to divide into different groups according to their search behavior (for example, Chen & Cooper, 2002; Wolfram, Wang, & Zhang, 2009). Question 2 investigates the search session level and examines whether search sessions naturally aggregate into certain types of search groups. Question 2 also investigates the distribution of different search groups and makes comparisons across different groups.

Question 3 examines people's interactions beyond the single action or search session level. It focuses on the search tactical level to investigate the ways people formulate or reformulate their searches. The query formulation/reformulation tactic generally refers to the

¹ Endeca is a software company that provides a faceted search platform.

methods and strategies adopted by searchers for issuing a query for a search system. These tactics include how many keywords searchers enter into the search box (query length), how many search iterations they go through (number of query submissions), how many records they click through on a retrieved result list (number of items viewed), and how deeply they explore the retrieved result list (record depth). It is believed that people adjust and alter their search tactics based on new feature(s) added to the system (Capra, Marchionini, Oh, Stutzman, & Zhang, 2007; Kules, 2006). This study aims to understand whether and to what extent users might change the ways they formulate or reformulate their searches based on the availability of facets. Question 3 also addresses the search tactics beyond just query formulation tactics and includes more general search tactics that are enabled by facets. The literature (Kules, 2006) and preliminary transaction log analysis suggest that searchers might adapt their search tactics when using facets. They might use facets as a way of organizing their search order, overview, and backup plan, or they might completely ignore them. They might also explore more deeply or issue shorter queries for later refining. This study identifies the general tactics that have been made possible by adding the facet feature to library catalogs and also compares the different search tactics that are enabled by these different facet implementations.

The second group of research questions (Questions 4 and 5) concerns the search comparisons between faceted and non-faceted interfaces to determine if the faceted interfaces improve the search. Different versions of the UNC faceted catalog, each employing different manipulations, serve as both the faceted and the non-faceted interfaces. Two specific research questions are:

4. Does a faceted search improve the search performance and user satisfaction?

5. What is the role that facets play during a search?

Question 4 explores the premise that categorizing search results into facets could improve search performance and search satisfaction. *Search performance* is defined as the search time and search accuracy of the search. Less time and higher accuracy indicates better performance. User satisfaction is an affective measure that reflects users' perceptions about the faceted search. Based on the literature, user satisfaction is broken down into search process satisfaction and search result satisfaction. The former can be measured by how enjoyable and how intuitive the search process is while the latter can be measured by how relevant the result is and how easy it is to find results. This study examines whether, to what extent, and under what circumstances, faceted search would improve or hinder the library search.

Question 5 aims to extend beyond statistics and numbers and go inside the search process itself. It aims to reveal, both quantitatively and qualitatively, the search 'story' with an emphasis on the reasons people use facets, the ways they use facets, and the ways facets help searchers find the information they need.

The third purpose of the study involves the ways that facets are used in different faceted catalogs. In this research, two library catalogs are chosen for the investigation, the UNC catalog and the Phoenix catalog. This group of research questions (Questions 6 and 7) explores the ways people use facets differently and the ways that this different facet implementation affects people's interaction style. The specific research questions are:

6. Is there any difference in terms of search performance and user satisfaction between the two library catalogs (UNC and Phoenix)?

7. Do people use facets differently in the two faceted catalogs? If so, in what ways?

Question 6 aims to compare the search performance and user satisfaction between this study's two library catalogs. *Search performance* and *user satisfaction* are defined the same as in Question 4. However, different from Question 4, the contrast here is not a quantitative comparison, because many confounding factors are found in the two different library catalogs, which may make the quantitative comparison less powerful.

Question 7 addresses the reasons and the ways that facets are used in different library settings, both quantitatively and qualitatively. Facet usage is broken down by task type and complexity level to investigate the circumstances in which facets are most needed. Qualitatively, this question explores the reasons people would use facets, narrow a search or browse a collection, and explores also the ways that facets are incorporated in their search activities.

In summary, this study sets out to investigate ways people use facets, particularly in library catalogs, and investigates a variety of factors. The independent variables are interfaces, task types, complexity levels, and external settings. The dependent variables are the users' performance, satisfaction, facet usage, search characteristics, and query formulation and reformulation tactics.

1.3 Significance of the Study

This research provides a basic understanding of the ways that people use facets in library catalogs. The searchers' behaviors are investigated in terms of different but comparable factors: non-faceted and faceted interfaces, academic and public library settings, different types of search tasks (close-ended and open-ended, simple and complex), and natural and controlled laboratory settings. Through these different situations and contexts, the

understanding of facet use is more thorough and comprehensive than has been possible heretofore.

Historically, online catalogs have been criticized as being hard to use. According to some researchers, such as Borgman, Warren, and Hildreth (Borgman, 1996; Hildreth 1991; Hildreth, 2001; Warren, 2000), OPAC designs do not incorporate sufficient understanding of search behaviors. The ability of OPAC systems to analyze query terms and correctly interpret users' information needs is still far from perfect. This study explores people's search behaviors using OPACs. The findings have implications for OPAC developers to design more responsive systems for different patrons. Particularly, insights into ways people use facets can guide library technical staff in making facets more effective in helping people find the information they need.

Subject headings and classification numbers have been notoriously difficult to understand for average users. Knutson (1991) suggests that inadequate subject access is one of the reasons that many items in large academic libraries are rarely, if ever, checked out. Libraries need to modify current subject cataloging practices to make more items accessible to users. One of the benefits of the faceted library catalogs is that they decouple and repackage these subject headings and expose them to users. This study pays special attention to the use of the subject facet to investigate if users are able to take advantage of this benefit. In addition, because the Phoenix catalog incorporates two types of subject headings, LCSH² and BISAC³, a comparison between these two subject heading systems would have implications for catalogers to better tailor the subject headings for users.

² LCSH is maintained by the United States Library of Congress and comprises a controlled vocabulary of subject headings in bibliographic records.

³ The BISAC subject headings list is a standard used to categorize books based on content.

This research suggests that an OPAC might take advantage of features found in web search engines or commercial websites in order to serve people better. Many OPAC users in the library setting, especially in academic libraries, are also likely to be web search engine users and would bring their mental models and web search engine experiences to OPACs (Young & Yu, 2004). The results of this study may provide suggestions to libraries that need to benefit from web search engines instead of being overwhelmed by them. The findings also would have implications for developing collaborative relationships between library services and commercial web search engines.

The faceted search concept is a case whereby an old theory is used in a new context and with a new twist. Faceted search is an application of classic facet theory to an online, digital environment, although the specific definition of *facets* has been tailored to better fit the new context. The faceted search concept demonstrates the importance of theories and concepts in the field of information science (IS). Many technological innovations and revolutions are, in fact, the application of fundamental concepts and theories in the field.

This study also has implications for researchers who need to think about search systems beyond just the conventional query-response and ranked list paradigm. An effective search system must go beyond traditional implementations in order to incorporate human search behaviors, such as filtering, browsing, and exploring, in addition to simple look-up.

Another contribution of this work is its innovative methodology that analyzes data derived from transaction logs of modern, faceted search-and-browse OPACs. This innovative method includes user identification rules, differently grained coding schemes, and automated data processing scripts. In addition, the VUTL (visualization for understanding transaction logs) supports the exploration and analysis of large amounts of log data. The framework for

automatically converting transaction log data into visualizations, and the open software tool, could be used for other libraries.

1.4 Introduction to the University of North Carolina-Chapel Hill Library Catalog and the Phoenix Public Library Catalog

In this research, the University of North Carolina at Chapel Hill (UNC-CH) Library catalog and the Phoenix Public Library catalog are chosen as two cases of the two primary streams of library catalogs and faceted implementation.

The UNC Library catalog (<http://search.lib.unc.edu>) is representative of OPACs in leading research universities with large collections and comprehensive resources. It uses Endeca faceted search (as of 2008) on top of its traditional catalog system to enhance search performance.

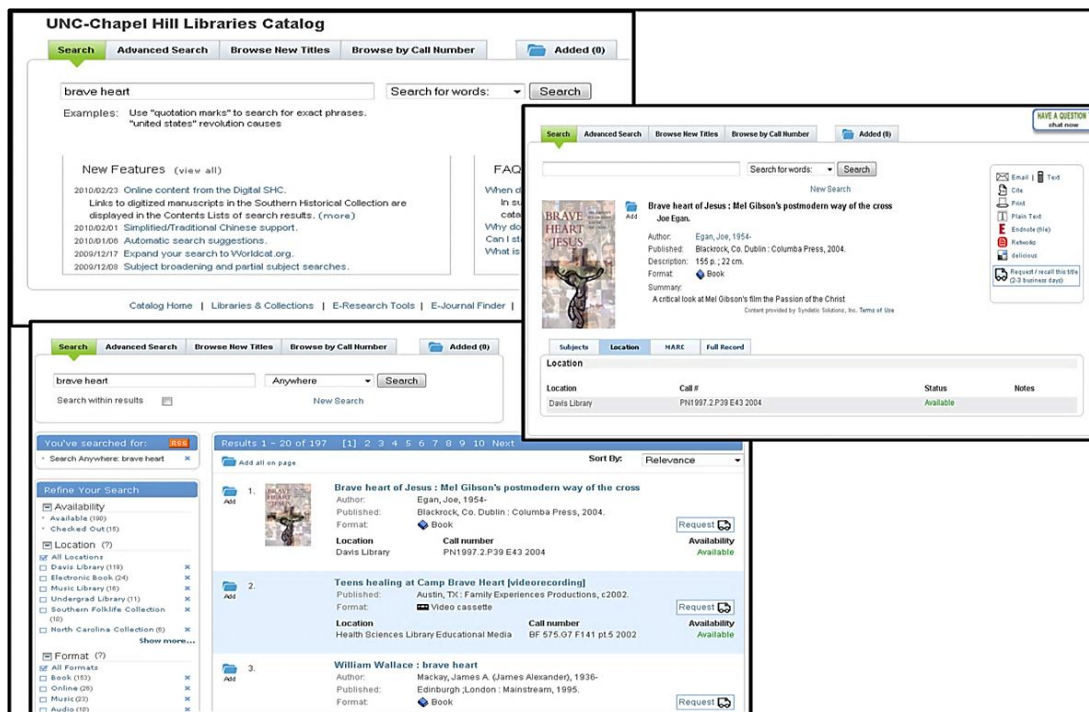


Figure 1.1 UNC Library Catalog (<http://search.lib.unc.edu>)

The top left is the opening search page. The bottom center image is the results page with facets along the left. The top right is a sample record page.

On the UNC-Chapel Hill Libraries Catalog homepage, a *Quick Search* box is available for searching against the catalog. Only a text search for five fields (keyword, title, journal title, author, and subject) is allowed at this point. The real catalog entry point is through the link 'catalog' below the *Quick Search* box. On the catalog page, four tabs (*Search*, *Advanced Search*, *Browse New Titles*, and *Browse by Call Number*) provide four options to start a search. The first one, labeled *Search*, has a similar function as *Quick Search*. *Advanced Search* looks much like many advanced search pages offered by other OPAC systems. *Browse New Titles* lets users choose a particular period of time for new titles. *Browse by Call Number* displays the Library of Congress and National Library of Medicine call number ranges to give users options to browse by subjects.

Whichever option the user chooses to begin searching, the next page the user sees is similar. The result list page is accompanied by options for refining the current search. These options are displayed on the left-hand side as a series of different facet categories for users to refine their search. The commonly used facets are expanded by default to display the preset values. The breadcrumb feature appears as a rectangle above all the facet categories, listing the current search parameters. Clicking the *x* icon next to a parameter removes that search parameter. It is important to note that if the user executes a text search from the result page, all of the previous text search parameters and any facets are removed, effectively restarting a new search. Numerous features can be found on the result list, including book cover images and brief descriptions of items. Clicking through a particular item brings up the item display page. Many new features have been added to this page, such as the table of contents, full machine-readable cataloging (MARC) records, and other features as offered by separate tabs on the page.

The Phoenix Public Library is a county-level comprehensive public library with vast holdings. The Phoenix catalog (<http://www.phoenixpubliclibrary.org/default.jsp>) is adopting Polaris ILS with an Endeca faceted search platform on top of it. In addition to offering faceted search capabilities, Phoenix has licensed the BISAC subject headings for browsing its holdings. Thus, the Phoenix catalog uses two systems of subject headings, LCSH, which is common for libraries in North America, and the BISAC subject headings, which has been widely adopted by the book industry and commercial bookstores. Phoenix's rationale for licensing the BISAC subject headings is the assumption that the BISAC subject headings are more user-friendly and intuitive for average searchers in a public library setting.

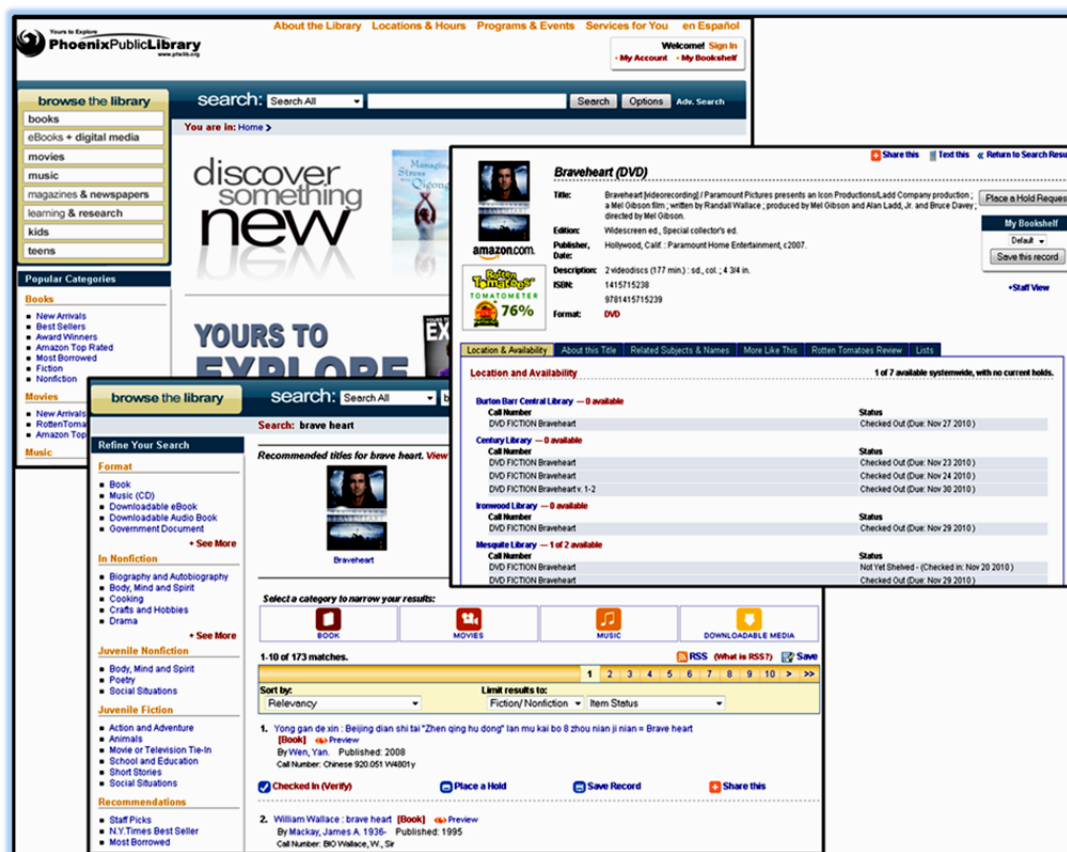


Figure 1.2 Phoenix Public Library Catalog (<http://www.phoenixpubliclibrary.org>)

The top left is the opening search page. The bottom center image is the results page with facets along the left. The top right is a sample record page.

The Phoenix Public Library homepage itself is the catalog page. The search box is prominently displayed at the top-center of the page. Along the left side, the top rectangle with an off-white background contains the general browsing categories, such as books, e-books, movies, music, etc. The searcher can drill down several levels along the browsing structure until reaching a result set. Below the off-white rectangle are more browsing options based on the library circulation data, such as new arrivals, most borrowed, etc. These options also include circulation data from commercial bookstores, such as Amazon top-rated and Rotten Tomatoes top picks. The searcher usually has two choices when starting a search: either enter a query in the search box or browse the categories. Whichever the user chooses, he/she will eventually arrive at a result list page on which the results are accompanied by facets for refining the current search. Thus, the Phoenix catalog has two sets of facets, one for browsing and the other for refining searches. The results page offers similar functions as offered by the UNC results page, such as the breadcrumb feature, sorting functions, and item cover images. Other interface elements also make the Phoenix homepage appear different from most library homepages, such as the slide show of popular book covers, the link to Twitter, and so-called *really simple syndication* (RSS) feeds.

The significant difference between the Phoenix and UNC faceted interfaces is the support for faceted browsing. Phoenix has the facet hierarchies available from the starting point of a search. Users are able to reach the result set by browsing the hierarchies without typing any text query. The UNC interface, however, does not have hierarchies for browsing. Another difference is their respective metadata sources. Generally speaking, the UNC Library is highly dependent on MARC records and the LCSHs to generate metadata. These metadata have better authority control than those in a public library. In contrast, Phoenix has

more varied sources of metadata and subject headings. The commonality is that both catalogs are leading high quality facet implementations, and both are implemented with the Endeca platform.

Despite the differences, the main search page and results page are similar for both catalogs, whereby the facet categories for narrowing a search are displayed on the left side beside the result list.

2. Literature Review

Given that the research problem of this study is to examine the ways people use facets in conjunction with text searches in a faceted OPAC system, this literature review aims to survey the existing research on information seeking behavior in an OPAC environment, facet theory and faceted search, and previous academic research into the topic of faceted search.

Section 2.1 starts with a review of information seeking behavior in the setting of OPACs. Section 2.2 moves to the foundation of faceted search, i.e., facet theory and faceted classification. Then, Section 2.3 surveys some well-known research projects on faceted search systems, which includes faceted library catalogs, and also reviews the empirical research into ways that people search through a faceted system.

2.1 Context: Information Seeking Behavior in Online Library Catalog Environments

The body of literature that concerns information seeking behavior is quite large, and some of it focuses on a particular kind of information system. The focus of this study is OPACs because this research focuses on ways that people search through faceted library catalogs.

2.1.1 Brief History of Online Public Access Catalogs (OPACs)

A library catalog is an organized set of bibliographic records that represents the holdings of a particular collection and/or resources accessible in a particular location (Taylor,

2006). The two major reasons to use catalogs are for retrieval and inventory purposes. Library catalogs can assume different forms: book catalogs, card catalogs, microform catalogs, CD-ROM catalogs, and online catalogs (OPACs). The latter form is currently prevalent in libraries in the United States, and is the focus of this review.

Early online catalog systems appeared in the late 1970s and early 1980s and are considered to be the first generation of OPACs. These early systems tended to replicate card catalogs but in a digital environment, and contained the same bibliographic information as library cards and provided some access points. Using a dedicated terminal or telnet client, users could search a handful of pre-coordinate indices and browse the resulting display in much the same way they had previously navigated the card catalog. Most of these early catalogs required an exact match between the user's input and the bibliographic record, thereby reducing the recall rate. Users seemed inclined to conduct known-item searches on an OPAC.

The second-generation OPACs are catalogs with more user-friendly systems than the first-generation ones and are still found in many libraries. Such OPACs include more sophisticated features, such as keyword searching on titles and other fields within the bibliographic record, Boolean matching, browsing functions, and ancillary functions. About the same time that these second-generation catalogs began to emerge, libraries began to develop applications to automate purchasing, cataloging, and circulation of books and other library materials. These applications, known as an integrated library system (ILS) or library management system, treated the OPAC as one module of the whole system.

Since the 1990s, rapid advances of computer and communication technologies and the fast growth of bibliographic utilities and networks have led to the development of

OPACs. The internet and, more specifically, the web undoubtedly have made OPACs remotely accessible and widely available, and web-based OPACs began to emerge in the late 1990s. In addition to web technology, these OPACs incorporated other new features, such as online resources, book covers, hyperlinks, and other features aimed at improving the interface. Despite the migration from catalogs to web interfaces, the underlying indices and exact-match Boolean search found in most library catalog systems, however, did not advance much beyond the second-generation catalogs. Web OPACs are considered to be advanced second-generation OPACs that serve as a gateway to resources held not only by a particular library but also other linked libraries, and further to regional, national and international resources (Babu and O'Brien, 2000).

Since the emergence of web OPACs, the major developments in OPAC technology have stabilized. Meanwhile, the industry outside of libraries has developed different types of web-based IR systems. Web search engines, such as Google, and popular e-commerce websites, such as Amazon.com, provide simple yet powerful search systems. As the internet has become more and more accessible to people, OPAC users have grown more and more accustomed to these websites and search engines. As such, they began to express increasing dissatisfaction with library catalog systems. This dissatisfaction has led in recent years to the development of newer, often termed *next-generation*, catalogs that have brought back wide attention to OPAC research.

These next-generation catalogs use more advanced search technologies than their previous counterparts, including in particular, faceted search, and features aimed at greater user interaction and participation with the system, including some web 2.0 technology, such as tagging, reviewing, and RSS feeds. The collaboration of TLC, a library automation

vendor, and Endeca, a software company that provides search applications, has served as a catalyst for the emergence of faceted library catalogs. One example is the NC State University library, which acquired Endeca's Information Access Platform (IAP) software in 2005 and started implementation of the new catalogs in early 2006.

2.1.2 Search Behavior

In order to investigate information seeking behaviors in an OPAC environment, the situational nature of information behaviors and search activities needs to be understood. Järvelin and Ingwersen (2004) produced a model for searching context (Figure 1), which suggests that searching behavior is composed of multiple layered contexts wherein *Information Retrieval* is the most narrowly focused, *Information Seeking* is a larger context, and both are set within an even larger purview of *Work Task*. Information Retrieval, as the smallest context in the model, represents the actions, usually keyword searches, by which users find relevant documents to match their query. Searchers may perform a series of Information Retrieval actions as part of broader Information Seeking tasks. One or more Information Seeking tasks are situated within the Work Task (or personally motivated goal), and are associated with the *Socio-Organizational and Cultural* context, as described by the model.

This study situates searching activities in the context of Järvelin and Ingwersen's *Information Seeking* (Figure 2.1) because this focus is the primary lens for faceted search systems.

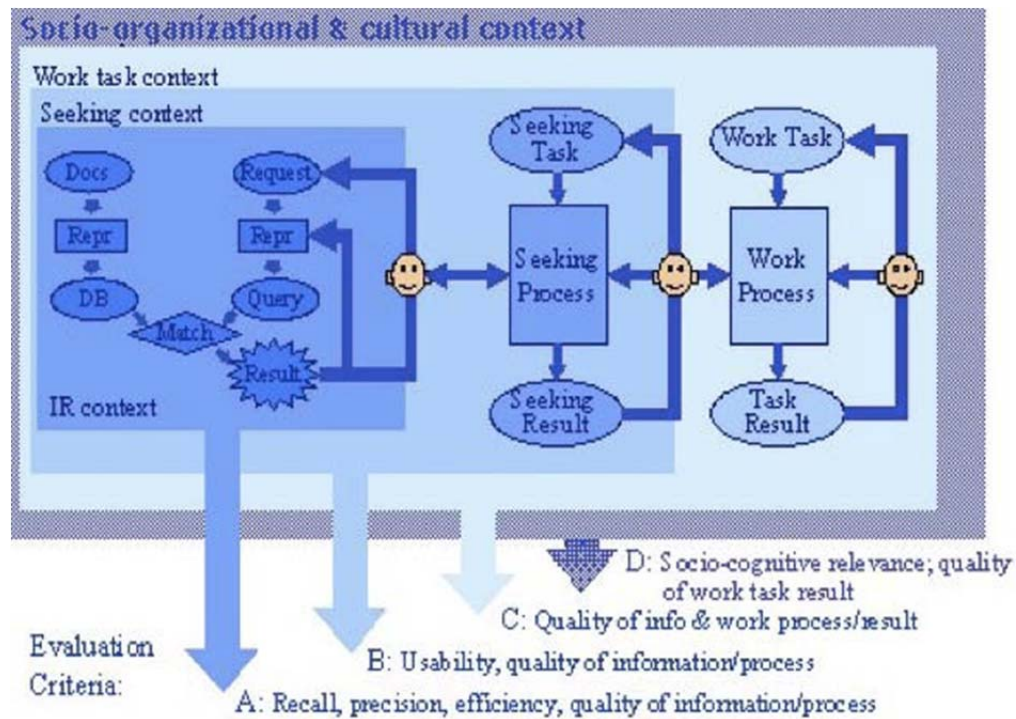


Figure 2.1 Model of search in context (Järvelin & Ingwersen, 2004)

At the information-seeking (IS) level, search systems usually function beyond the query-result-evaluation cycle typically seen in IR systems. The IS search systems have more features that support IS tasks, such as search history mechanisms for multiple-session searches, tagging mechanisms for grouping a set of documents to address a larger information need, overviews of collections, and browsing structures. Evaluations of systems that support IS tasks typically focus on assessing the quality of information acquired by users relative to the information need, rather than some system-oriented metrics, such as precision and recall, in the context of IR.

The following information provides types of information activities within the context of IS.

Searching and browsing

Searching and browsing represent two basic activities in IS. Searching is the most common and the most identified information activity of users. In searching, users express their information need in query terms that are understandable by the system, and then the users examine the results returned by the system until the target is found. In browsing, people are scanning information items, omitting irrelevant ones and occasionally picking up relevant ones. When browsing, each new information scent that is gathered can provide new ideas, suggest new directions, and change the nature of the information need (Bates, 1989). Browsing is an increasingly subtle searching activity in IS research (e.g., Ingwersen and Wormell, 1989; Noerr & Noerr, 1985). Ellis (1989) suggests that browsing features, e.g., contents pages, lists of cited works, and subject terms, should be made available in automated catalog systems to accommodate searchers' browsing behaviors that usually occur physically in the library.

Focused searching

It is usually the case that people need to do some post-query searching after viewing the result set returned by an initial query. These post-query searches require system support for query specification and refinement, selection of search results, and post-query navigation paths. Thus, people may get a clear sense of their information targets and the trails to follow. Faceted navigation is one way to support post-query refinement in that it offers users the ability to extend the query by slicing a large result set down to a smaller size through controlled vocabularies, or even expanding the result set in a structured way.

The motivation behind the need for post-query interaction is the inability of systems to fully understand the information needs of their users (White & Roth, 2009). However,

even if the search engine is able to understand a user's query well and return exactly the information that is sought, given a well-specified query, situations may still arise where users are unable to express their information need. In reality, people are observed to have a style of interaction referred to as *orienteering* (O'Day & Jeffries, 1993). The initial query and initial result set might be only partially relevant to the searcher. Through post-query interaction, people are transported to multiple result sets where they may be able to attain the complete set of information they need. Post-query navigation trails extracted from search logs exhibit traits of orienteering behavior (White and Drucker, 2007).

Another need for supporting post-query interaction lies in the inversely proportional relationship between precision and recall. An over-specified query may gain a high precision rate for the result set, but may hurt the recall, and many related but non-core documents might be excluded. On the other hand, an under-specified query may have good recall, but at the price of precision. To strike a balance between precision and recall, it is likely that users will find information from multiple result sets rather than from a single one, necessitating post-query interaction as a way of navigating the result sets.

Exploratory search

With more and more online information accessible to searchers, they are no longer satisfied with simply conducting a quick, look-up search. In addition to known-item, fact-finding searches, exploratory searching is another common type of search conducted by current library users. In addition, exploratory searching is an important use case for faceted search.

Exploratory searchers utilize a combination of searching and browsing behaviors to navigate through and to information that helps them to develop powerful cognitive

capabilities and leverage their newly acquired skills to address open-ended, persistent, and multifaceted problems (White & Roth, 2009). According to White and Roth, exploratory searches comprise broader searching activities than traditional look-up searches, and include exploratory browsing, berry-picking, information foraging, results comparing, etc.

People who conduct exploratory searches generally 1) have vague information needs, 2) are unsure about the ways to satisfy their information needs, and 3) are unfamiliar with the information space. Exploratory searching usually involves complex situations. The problem context and the definition of the search task often are ill-structured, which requires searchers to clarify their search during the search process. Multiple information resources, including some partially relevant and irrelevant ones, are needed to satisfy the search task. In addition, information needs are always fluid and developing. Marchionini (2006) identifies two key components of the exploratory search: learning and investigation. In his proposed model (Figure 2.2), he depicts three search activities – look-up, learn, and investigate – and highlights exploratory search as related especially to the learning and investigating activities. The overlapping ‘clouds’ of the three search activities suggest that some activities may be embedded in others, and that no clear boundary exists between them.

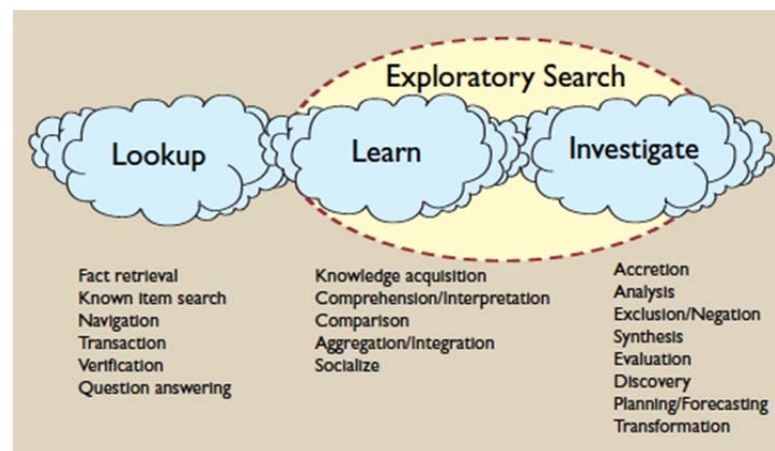


Figure 2.2 Exploratory search components (Marchionini, 2006)

2.1.3 Ways People Search Using OPACs

Basically, people conduct two types of searches when they use OPACs. One is the known-item search where the user wants to locate information about a specific item (e.g., author, title, and publication year). The other type of search is a subject search for a topic under a LCSH or other subject headings. Many researchers have examined the distribution of OPAC searches between the two types, and the results vary considerably. Sometimes, no clear boundary is found between the two search types.

Researchers are in general agreement that the known-item search type is less problematic than a subject search (Large and Beheshti, 1997). Research has shown that author and title searches are the most common search fields for known-item searches (Cochrane & Markey, 1983; Lewis, 1987). Compared to a known-item search, a subject search is much more open-ended, which may be popular, but is also problematic. Tolle and Hah (1985) found that subject searching is the most frequently used and the least successful of the search types. Hunter (1991) reports that 52% of all searches were subject searches, and 63% of those had zero hits. For a subject search, users need to know how to express their information need as subject “aboutness”, how to map the subject “aboutness” to the controlled vocabulary of a LCSH, and how to re-conduct a search if no records, too many records, or irrelevant records are retrieved after the first attempt. These requirements may account for the fact that subject searching is being replaced by keyword searching. Knutson (1991) suggests that inadequate subject access is one of the reasons that many items in large academic libraries are rarely, if ever, checked out, and that libraries need to modify current subject cataloguing practices to make more items accessible to users.

Online catalogs have been criticized as being hard to use because their designs do not incorporate sufficient understanding of searching behaviors (Borgman, 1996). The ability of OPAC systems to analyze query terms and correctly interpret a user's information needs is still far from being perfect. For example, Large and Beheshti (1997) report that users encounter many problems in choosing suitable search terms to represent their subject interests. Some people enter very broad terms and then feel overwhelmed by the amount of results returned (Hunter, 1991). Some subjects enter very specified queries by pasting long phrases or sentences directly into the search box. Sit (1998) states that users' difficulties include finding subject terms to enter, using non-distinctive words, over-specification (e.g., a query that is too long), reducing results, and increasing results. Additional user difficulties include complex command syntax (e.g., Janosky, Smith, & Hildreth, 1986), scrolling through large retrieval sets and selecting appropriate database fields and keywords (e.g., Ensor, 1992; Yee, 1991), predicting the results of various search algorithms (e.g., Chen & Dhar, 1990), using multiple databases (e.g., Yee, 1991), error-recovery processes (Peters, 1989; Yee, 1991), and information comprehension and location in displays (Janosky, Smith, & Hildreth, 1986; Yee, 1991). Therefore, a serious need exists to establish a closer working relationship between systems designers and users to develop useful IR systems. According to Warren (2000), the general design of the Urica OPAC system, for example, actually hindered rather than helped users in their search process. From the library organization perspective, difficulties might come from the restriction of the bibliographic records that are the basis for the catalog. O'Brien (1990) states that users do not necessarily understand the subject headings and classification numbers due to their artificial nature.

Borgman (1996) developed a three-layer framework of knowledge needed for successful OPAC searching: (1) conceptual knowledge for translating an information need into a searchable query, (2) semantic knowledge for how and when to use system features to implement a query, and (3) technical and basic computing skills. Borgman (1986b) concludes that people might have problems with each of the three layers. However, conceptual problems are more similar across types of systems than semantic and technical problems. Conceptual problems are essential because “only when the conceptual aspects of searching were understood could the user exploit the system fully and effectively.” On the other hand, technical problems seem to be more common among novice catalog users.

People tend to use short queries when they search through OPACs. The most common length is one or two terms (Jones, Cunningham, & McNab, 2000; Lau & Goh, 2006; Mahoui & Cunningham, 2001; Wallace, 1993). People rarely use operators such as *AND*, *OR*, or *NOT*, and tend to use simple queries, although it is assumed by the system designer that the correct use of search operators would increase the effectiveness of the searches (Eastman & Jansen, 2003; Jansen & Pooch, 2001; Lau & Goh, 2006). The overall field of information-searching through OPACs has grown large enough to support investigations into demographic-based groups, for example, children (Borgman et al., 1995; Hutchinson, Druin, & Bederson, 2007; Solomon, 1993), older adults (Sit, 1998), and university staff and students (Conaway, Budd, & Kochtanek, 1995).

Many research studies on OPACs include failure analysis in which a *failed search* is typically defined as a search that matches no documents in the collection (Jones et al., 2000). Generalizing from several studies, approximately 30% of all searches result in zero results. The failure rate is even higher, at 40%, for subject searches, as reported by Peters (1993).

However, there is disagreement on the definition of *failed search* among researchers. Large and Beheshti (1997) state that not all zero hits represent failures, and not all hits represent successes. Some researchers also define an upper number of results for a successful search (e.g., Cochrane & Markey, 1983). Like the definition of *search failure*, the reasons for search failures also vary considerably in the literature. Large and Beheshti (1997) suggests that some of the failed searches are in fact helpful ones that could lead users to relevant information if users had more perseverance to look beyond the first results page rather than terminating the search.

Another stream of research reports feelings and reactions to OPAC searches through questionnaires and/or interviews. Satisfaction with search results often serves as a metric of utility (Hildreth, 2001). Measures, such as the wording ‘easy to use’ and ‘confusing to use’ (Dalrymple & Zweizig, 1992), or a high to low scale have been employed (Nahl, 1997) to assess user satisfaction. Many researchers have challenged the validity of using satisfaction and perception as evaluation measures for search systems. For example, Hildreth (2001) found no association between users’ satisfaction and their search performance. He found that users often express satisfaction with poor search results and further investigated the phenomenon of false positives, which inflated assessments of the systems.

The availability of web technology and the appearance of web search engines in the 1990s have had a significant effect on OPACs. Jansen and Pooch (2001) report that 71% of web users use search engines. Many OPAC users in the library, especially in academic libraries, are also likely to be web search engine users, and bring their mental models and web search engine experience to OPACs (Young & Yu, 2004). Luther (2003) states in her study, “Google has radically changed users’ expectations and redefined that experience of

those seeking information.” Furthermore, users tend to prefer a single search box type interface that conceptually allows them to perform a metasearch over all the library resources rather than performing separate searches (Hemminger, Lu, Vaughan, & Adams, 2007). “Users appear to be using the catalog as a single hammer rather than taking advantage of the array of tools a library presents to the user” (Young and Yu, 2004). Despite the popularity of web search engines, Muramasts and Pratt (2001) report that users commonly do not understand the ways search engines process their queries, which leads to poor decisions and dissatisfaction with some search engines. Young and Yu (2004) believe that the same lack of understanding applies to OPACs. Features of web search engines and/or some online commercial websites could raise the bar for library catalogs; however, OPACs typically do not offer some of the features of web search engines and online commercial book stores (e.g., Amazon, Barnes & Noble). Such features include: free-text (natural language) entry, automated mapping to controlled vocabulary, spell checking, relevance feedback, relevance-ranked output, popularity tracking and browsing functions (Young & Yu, 2004). “Search inside the book”, i.e. full text searching, as implemented by Amazon, Google Books, and some web search engines, is another feature that OPACs have not incorporated.

2.2 Facet Theory and Faceted Search

In order to understand the details of faceted search, the foundations of facet theory and faceted classification must be discussed. Then, the application of facet theory in the online digital environment, i.e., faceted search, is examined.

2.2.1 Facet Theory and Faceted Classification

The notion of a *facet* is the central concept to the facet theory that was initiated by Ranganathan, an Indian mathematician and librarian. In facet theory, each characteristic

(parameter) represents a *facet*. After Ranganathan, other researchers have contributed their summaries and understanding of *facets*. According to Taylor (1992), *facets* are “clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject.” Hearst (2006) defines *facets* as categories that are a set of meaningful labels organized in such a way as to reflect the concepts relevant to a domain. In many current online faceted search systems, overlap of *facets* may occur, and the *facets* may not be exhaustive.

2.2.2 Faceted Search

Faceted search is the application of classic facet theory in the online digital environment. It is the combination of free, unstructured text search, with faceted navigation. White and Roth (2009) describe faceted search interfaces as interfaces that seamlessly combine keyword searches and browsing, allowing people to find information quickly and flexibly based on what they remember about the information they seek. Faceted interfaces can help people avoid feelings of ‘being lost’ in the collection and make it easier for users to explore the system. According to Ben-Yitzhak et al. (2008), a typical user’s interaction with a faceted search interface involves multiple steps in which the user may 1) type or refine a search query, or 2) navigate through multiple, independent facet hierarchies that describe the data by drill-down (refinement) or roll-up (generalization) operations. Bast and Weber (2006) loosely define a faceted search interface as one that, in addition to showing ranked results for keyword queries as usual, organizes query results by categories. Figure 2.3 illustrates a website with a dynamic presentation of facets when searching for a laptop. The facets for a laptop are price range, manufacturers, screen size, memory size, and so on.

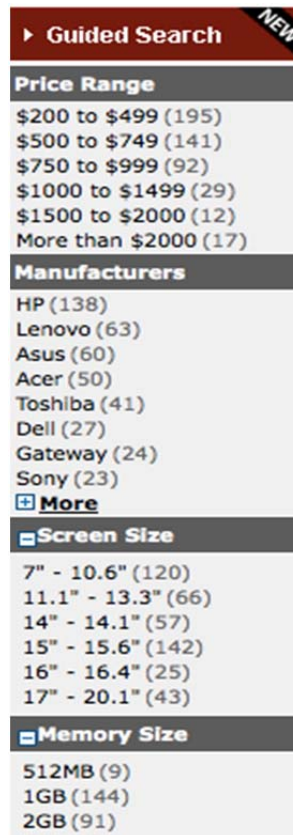


Figure 2.3 Facets for a laptop search

Faceted search enables users to explore a subject in terms of its different dimensions. Although keyword searches usually bring about a ranked result list, in faceted searches, users may filter the result set by specifying one or more desired attributes of the dimensions. The faceted interface gives users the opportunity to evaluate and manipulate the result set, typically to narrow its scope (White & Roth, 2009). It is important to recognize that primary attributes of “faceted search” as referred to in this work, are the interactive filtering along these multiple dimensions of information. And these dimensions do not formally adhere to facet theory definitions (for instance facets like date and time period are overlapping and not mutually exclusive). Yet, in the mainstream literature, and in this work, these interfaces will be referred to as “faceted interfaces” supporting “faceted search”. Faceted search also gives

users flexible ways to access the contents. Navigating within the hierarchy builds up a complex query over sub-hierarchies. As White and Roth (2009) describe, the approach reduces mental work by promoting recognition over recall and suggesting logical but perhaps unexpected alternatives, while avoiding empty result sets. Meaningful categories support learning, reflection, discovery, and information finding (Kwasnik, 1999; Soergel, 1999). The counts next to facet labels give users a quantitative overview of the variety of data available, thereby hinting at the specific refinement operations that seem most promising for targeting the information need(s) (Ben-Yitzhak et al., 2008).

2.3 Academic Research on Faceted Search

This section introduces some important academic projects on faceted search and faceted library catalogs, and then enumerates some empirical studies on this subject.

2.3.1 Well-known Faceted Search Projects

The query previews developed by Shneiderman and his colleagues (Doan, Plaisant, Shneiderman, & Bruns, 1997) probably serve as the catalyst for the current interest in faceted search. According to Shneiderman, query previews allow users to specify the parameters that generate visually displayed results. Figure 2.4 shows the changes before and after selection of a geographic attribute, in this case, North America. The preview bar at the bottom of the map as well as the attributes above it update responsively. Users are able to obtain a sense of the overall collection and alleviate zero-hit queries. The left side of Figure 2.4 displays summary data on preview bars. Users learn about the holdings of the collection and can make selections over a few parameters (in this case geographic locations, environmental parameters, and the year). The right side of Figure 2.4 displays the updated bars (in less than

100 msec) when users select an attribute value (in this case, North America). The results bar at the bottom shows the total number of selected datasets.

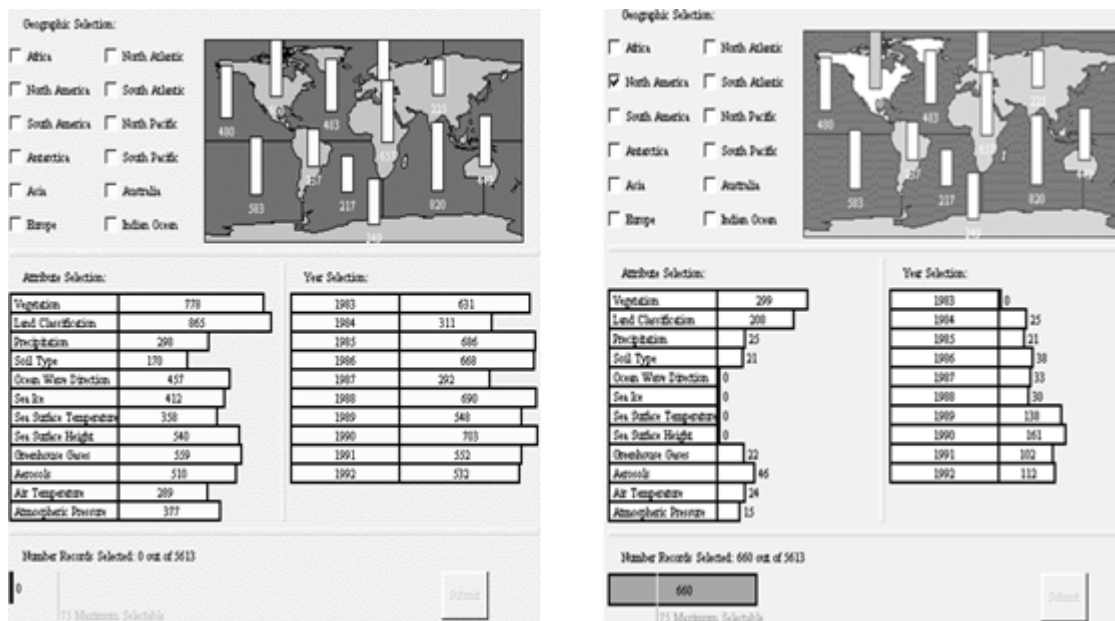


Figure 2.4 Collection of environmental data from the National Aeronautics and Space Administration (NASA)

The Flamenco Project led by Hearst at the University of California, Berkeley, represents almost a decade of work on developing faceted search tools and performing usability studies. (*Flamenco* is derived from *flexible information access using metadata in novel combinations*.) The lead researcher of Flamenco, Marti Hearst, explicitly credits the query previews by Shneiderman in the work of the Flamenco Project and situates Flamenco's interface as a form of a query preview (Hearst et al., 2002). Flamenco allows users to navigate by selecting facet values. In the example shown in Figure 2.5, the retrieved images are the results of specifying a value from *Locations*. The matching images are displayed and grouped by the facet values from *People*.

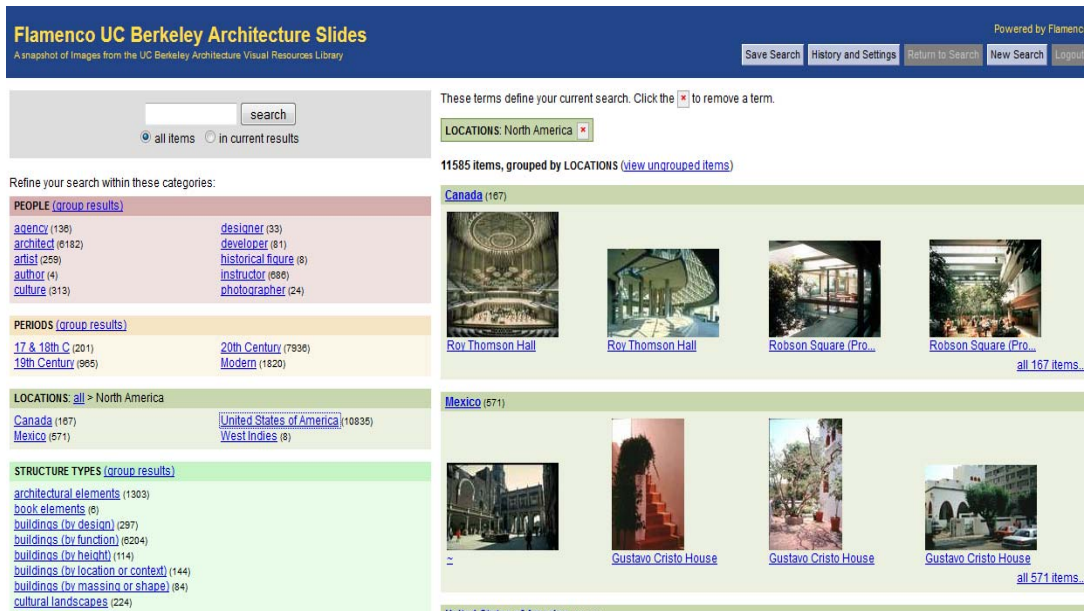


Figure 2.5 Hierarchical facet navigation in Flamenco

As described by Hearst (2006), the interface aims to support flexible navigation, seamless integration with directed (keyword) searches, fluid alternation between refining and expanding, avoidance of empty results sets, and at all times retaining a feeling of control and understanding. A usability study by Yee, Swearingen, Li, and Hearst (2003) indicates that users are more successful at finding relevant images and report higher subjective measures than the traditional search interface.

The so-called *relation browser* (RB) is a generic search interface that can be applied to a variety of data. The RB is a tool developed by the Interaction Design Lab at the University of North Carolina at Chapel Hill for understanding relationships between items in a collection and for exploring an information space (Capra & Marchionini, 2008; Marchionini & Brunk, 2003; Zhang & Marchionini, 2005). The project, originally developed for the United States Bureau of Labor Statistics, has been through a number of major design revisions. The most recent version is displayed as Figure 2.6. In Figure 2.6, 1 and 2 support

multiple facet views; 3 supports multiple result views; 4 indicates the current query display and control; and 5 and 6 show the full-text search and search within results.

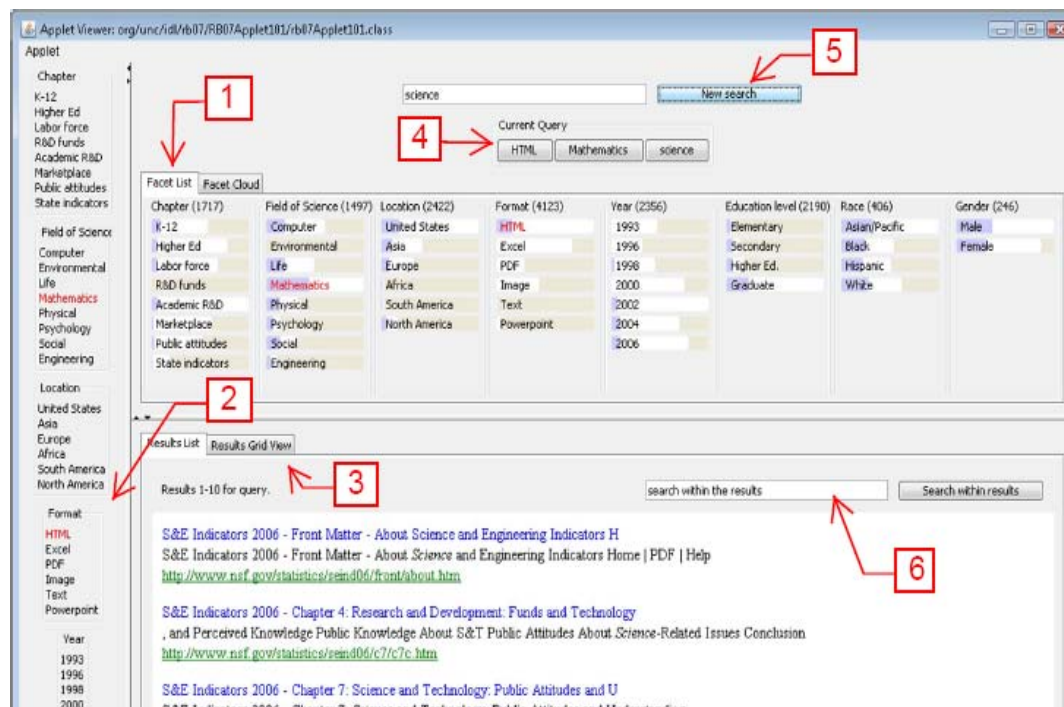


Figure 2.6 Relation Browser

The RB combines simple text search and facet navigation as a way to refine the search. It provides searchers with a small number of facets (topic, time, data format) with a manageable size of values in each facet. Users can easily move between searching and browsing strategies. The current text query is displayed at the top of interface, and the current incorporated facet values are highlighted in red and shown below the current text query. Mouse-over capabilities allow users to explore relationships among the facets and attributes, and dynamically generate results as the mouse slides over them. One of the issues of RB lies in its dependence on dynamic client-side graphics to update the interface in real time. Scalability would be a problem for client applications if billions of records must be processed instantly.

Faceted search concepts can also be applied to the field of personal information management, where people acquire, organize, maintain, retrieve, and use information items (Jones, 2007). Information overload makes re-finding and re-using personal ‘stuff’ similar to information discovery. Using facets in generic IR systems allows for pre-filtering personal information. A series of research studies has been conducted by Microsoft Research on applying facets to personal information management. *Phlat* (Cutrell, Robbins, Dumais, & Sarin, 2006) and *Stuff I’ve Seen* (Dumais, & Horvitz, 2003) are two examples found in this series.

2.3.2 Faceted Search Used in Library Catalogs

Since 2006, a small number of academic libraries have implemented faceted navigation on their online catalogs. Among them are McMaster University Library (Hamilton, Ontario, Canada), State University Libraries of Florida, NC State University Library (Raleigh, North Carolina), and WorldCat. In recent years, faceted navigation has grown to be a well-accepted approach and has been applied as a standard technique on commercial websites for many years (Breeding, 2007). Since the adoption of faceted search by the NC State University Library in early 2006, faceted library catalogs have gained popularity in many academic and public libraries. Many library automation vendors and software companies have produced applications for facets (e.g., Endeca, AquaBrowser, Encore, Primo, Smart Library System, OPAC GiB, etc.), and some programmers and librarians have worked together to develop open source faceted integrated library systems (ILS) (Evergreen, Koha, Solr, VuFind, etc.).

Endeca, a well-known company for providing facet search applications to e-commerce sites, started the implementation of facet browsing in their catalog. Figure 2.7

presents the interface of NC State's library catalog, which acquired the Endeca applications in 2005. This new generation of library catalog gives its users both relevance-ranked keyword search results and rich facet metadata previously trapped in MARC records to enhance collection browsing and search refinement. The faceted metadata are grouped into subject, genre, format, location, author, etc. A user may enter the text query in the query box as a starting point and then click one attribute of facets from the left-hand box to filter the result set. An empty query in the query box will generate the results for the whole collection held by the library, organized by a set of facets. In addition to simple text search mode combined with facet browsing, users also can select other search modes, for example to browse through new titles that have been recently cataloged by the system, and to scan through the LCSH.

AquaBrowser is another world leader application in visual faceted search that connects to heterogeneous data sources. AquaBrowser can be found in public, academic and special libraries around the United States and the world. AquaBrowser motivates users to explore the library's content by incorporating various common search behaviors. Its unique 'search, discover, refine' methodology provides features that help users quickly and easily uncover relevant results. Figure 2.8 captures a screenshot from Queens Library, which implements AquaBrowser as its search solution. This OPAC's facet implementation is similar to that of the NC State University catalog, except that the facet panel is placed on the right side. Another major difference is the word cloud on the left side that explores associations between the current query and other vocabularies as a query recommendation tool. Another development is the separation of collections according to item type, i.e., books, music, movies, etc.

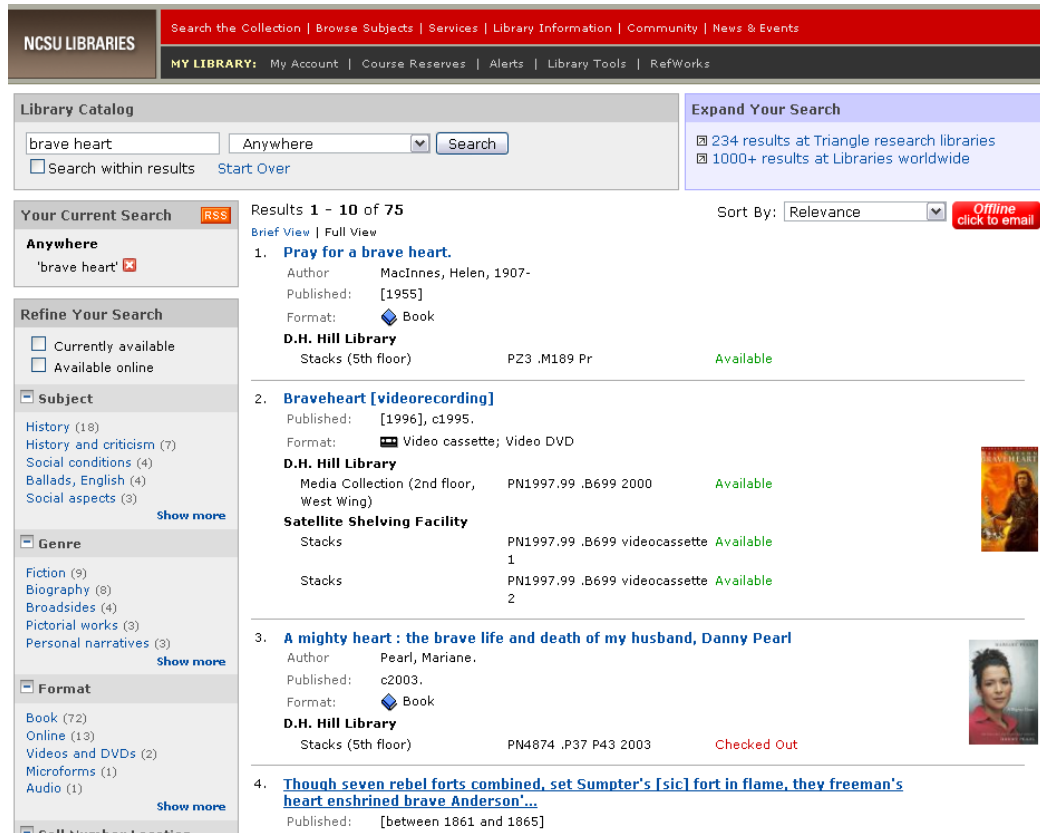


Figure 2.7 Interface of North Carolina State University's faceted library catalog

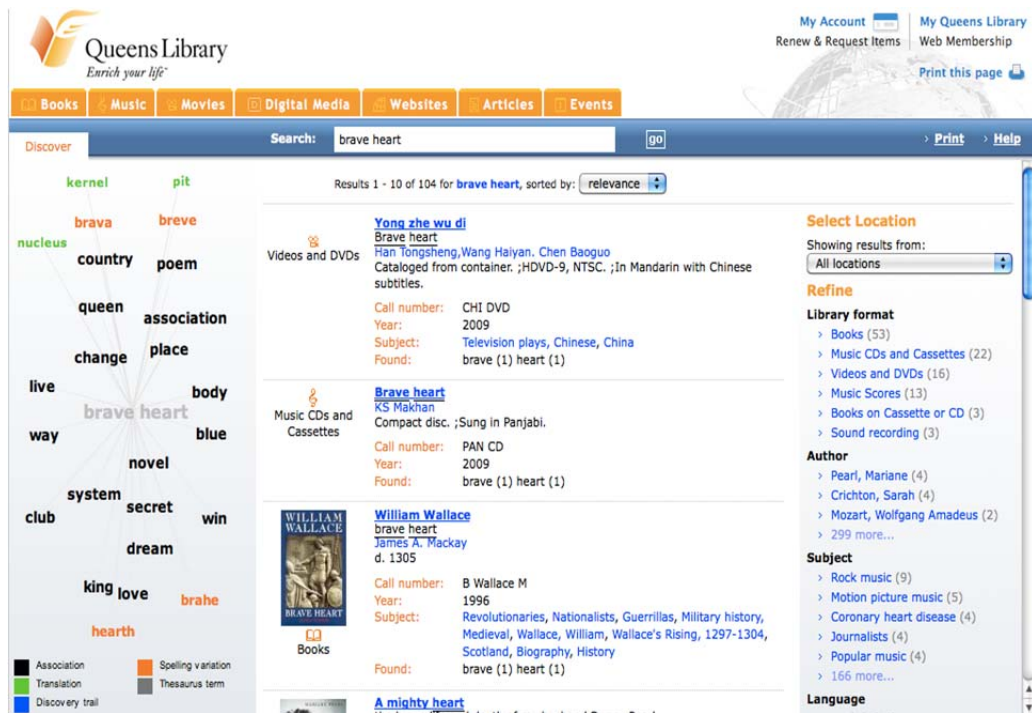


Figure 2.8 Interface of Queens Library faceted library catalog

Encore is another popular commercial application for faceted library catalogs. In addition to faceted navigation and relevance ranking, it also presents tag clouds, popular choices, and recently added suggestions. Encore even makes use of user contributions as a tool for discovery by incorporating community participation features, such as tagging.

Primo is an Ex Libris offering that aims to revitalize the library environment by creating next-generation interfaces. According to Ex Libris, Primo provides services for searching as well as delivering access to all of the library's resources, whether those resources are maintained and hosted locally or need to be accessed remotely. In addition to relevance ranking and faceted browsing, Primo indexes data from sources such as Syndetic Solutions, Blackwell, Amazon and others to provide additional access points when searching. It also includes features that are popular in e-commerce websites, such as user-supplied reviews, recommendations based on what others who viewed the same item selected, and grouping similar results. Primo also includes dictionaries and thesauri to provide search suggestions and structured lists as part of the search process.

In addition to commercial search solutions for faceted OPACs, some open source catalogs have been developed by programmers and librarians. These catalogs aim to be next-generation catalogs and regard facet searching as one of their major features. Also, open source OPACs are more cost-effective than proprietary ones, so many libraries choose to use open source solutions mainly for their affordability. Although users of open source OPACs may experience difficulties with installation and incomplete documentation, they are modestly more satisfied than users of proprietary OPACs (Riewe, 2008). Some common open source OPACs are Evergreen, Koha, Solr, VuFind, etc.

Evergreen is an open source ILS developed in 2004 by a consortium of public libraries in Georgia. Currently, as the largest open source ILS, scalability has not been an issue for Evergreen, which accommodates 1.8 million patrons and a collection of 9 million items (Riewe, 2008). Since 2007, the Michigan Libraries Consortium has installed Evergreen. British Columbia, which installed Evergreen in three libraries, estimated a huge cost reduction compared to proprietary ILS. Figure 2.9 presents an Evergreen interface from the Georgia Public Libraries. The most important feature of this faceted navigation is that it uses facets created from results to submit new searches instead of refining existing searches. Thus, each facet serves as a query recommendation. In addition, the facet hierarchy allows for third-level facet values, which is different from most faceted navigations that allow only a two-level hierarchy.

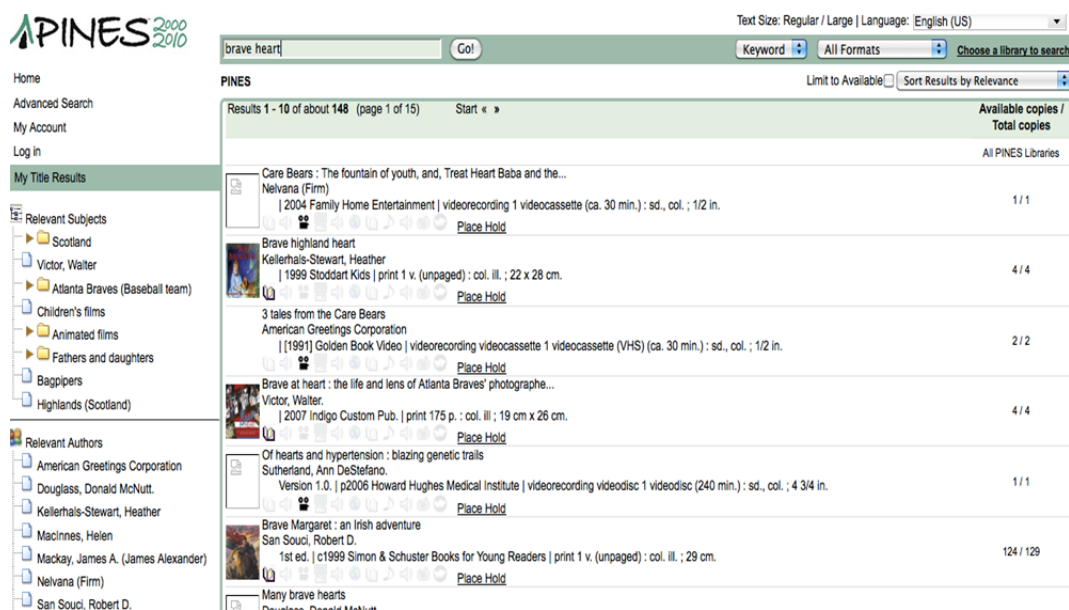


Figure 2.9 Interface of Georgia Public Libraries faceted library catalog

This section provides a comprehensive, but not necessarily exhaustive, overview of some well-known faceted search projects, either for general purposes, personal information

management, or library catalogs. Despite the differences among the implementations, most faceted search systems offer users two-level faceted metadata for refining the text search or browsing the whole collection. Most systems allow a single choice of facet value under the same facet and multiple choices of facets. Overall, the facet feature has provided more powerful search assistance for users than was available prior to the introduction of facet searches.

2.3.3 Empirical Studies on Faceted Search Systems

Few empirical studies have been published on faceted navigation in OPACs (Olson, 2007). Especially in North America, most research into faceted systems has been commercial, and proprietary reports generally are not published (La Barre, 2007). However, a small stream of research is available that has been conducted by either system implementers or interactive IR researchers and examines the effectiveness of various faceted interfaces.

Studies suggest that users take advantage of facets or categories if these options are presented during the search process (Antelman, Lynema, & Pace 2006; Lown, 2008). Antleman et al.'s log analysis (2006) of the NC State University faceted library catalog suggests that approximately 30% of searches involve post-search refinements from the facets on the results page. Lown's follow-up analysis (2008) indicates that faceted searches account for 15% to 18% of all requests. Users employ facets to help refine the search (Hearst, 2000), sharpen a vague query or formulate a new query (White & Roth, 2009) and browse the whole information collection (Shneiderman, 1994). For the dimension (facet) usage, according to Antelman et al. (2006), dimension use does not exactly parallel dimension placement in the interface. *LC Classification* is the most heavily used facet, followed closely

by *Subject: Topic*, and then *Library, Format, Author*, and *Subject: Genre*. Query test results indicate that 68% of the top results in Endeca were judged to be relevant, whereas 40% of the top results in traditional catalogs were judged to be relevant. This finding suggests a 70% better performance for the Endeca catalog than the traditional catalogs.

Empirical research into faceted search interfaces often use two common methods to study the effectiveness of faceted search interfaces: large-scale log analysis and comparative user studies (Kules, Capra, Banta, & Sierra, 2009). Some studies use a combination of the two methods (e.g., Antelman et al., 2006; Tang, 2005). Log analysis employs server logs to examine users' interaction with the system and constitutes the most common research method in this field. Comparative user studies complement transaction log analysis in that they capture the context information for users' interaction with the system by directly observing the users' behaviors and actions. Most empirical research into faceted search systems incorporates user studies as one of the data collecting methods (e.g. English et al., 2002; Olson, 2007). Käki (2005) provides a good example of a user study that employs a within-subject design, balanced task sets, time limitations, pre-formulated queries, cached results pages, and limited access to results documents. Beyond the two common research methods mentioned, Kules et al. (2009) adopt eye tracking, stimulated recall, and interviews to investigate important aspects of gaze behavior in a faceted search interface. The top ten gaze transitions derive from the eye-tracking data that indicate what the searchers look at in the interface and suggest the specific part or component of the interface that plays an important role. Olson (2007) conducted qualitative research on 12 humanities Ph.D. students at the dissertation level. He found that nine of the participants reported finding materials that they had not found in their previous use of the traditional catalog interface. Tang (2005)

employed a naturalistic, longitudinal research method to overcome the artificial nature of the laboratory environment and to investigate ways the user's mental model of the tool evolved over time. The most striking part of her study is the regression models that were applied to users' satisfaction with the results to understand the effect of interactions between two independent variables: users' 'problematic situations' and query submission methods.

User studies, also called *usability testing*, generally involve measuring how well test subjects respond in four areas: performance, accuracy, recall, and emotional response. Performance and emotional response are the two frequently examined measures for testing a faceted search system. Performance is often operationalized as the amount of time required for people to complete basic tasks. For example, Shneiderman (1994) compared the response time of the dynamic query approach over a natural language query facility and a '10-page paper listing'. The counterbalanced within-subjects design found a statistically significant speed advantage for the dynamic queries over the other two methods. English et al. (2002) used a histogram to show the median task completion times for each task on two faceted interfaces, Matrix View and SingleTree View. Matrix View allows users to select multiple terms from facets in any order and can have the items grouped under any facet, whereas SingleTree View allows users only to drill down to subcategories of the current category and does not allow them to select terms from more than one facet. English et al. found that times were longer for Matrix than for SingleTree. Emotional response is usually collected through post-search questionnaires to measure the participants' perception of the system. For example, Tang (2005) found that user satisfaction was highest with MeSH browsing when searching for unfamiliar topics, but was poor for familiar topics. Most users gave positive comments for the classification display. Kules et al. (2009) confirm the users' perception that

they are slightly more familiar with and more confident about the known-item tasks. English et al. (2002) asked the participants to rate the preference of the two interfaces. As a result, Matrix was generally rated higher over SingleTree interface, both for specific tasks and overall.

Time as a measurement is a point of discussion, as initiated by Capra et al. They suggest that time might not be a suitable measure for exploratory tasks. Completing an exploratory task quickly may suggest that a search system does not provide support for investigating and exploring. This finding is backed up by the Kammerer, Narin, Pirolli, and Chi' study (2009) results that suggest that the participants who used the MrTaggy interface spent more time and produced better reports than participants who used other interfaces. Time, in this case, is a positive measure for the system.

3. Research Methods

This research seeks to understand searching behavior with regard to faceted OPACs and is an empirical study in the field of information seeking behavior. In light of the research questions posed in Chapter 1, the combined use of transaction log analysis (TLA) and user experiments best suits the purposes of the research methods. Although TLA is a non-intrusive, inexpensive way of collecting large amounts of data from a great number of users, it fails to capture any information about the context in which the event occurs (Sheble & Wildemuth, 2009). Contextual information includes user demographics, motivations, satisfaction levels, etc. The user experiments complement the limitations inherent of TLA by providing such missing contextual information. In addition, user experiments allow researchers more control over the task complexity, search process, and definition of search success/failure. However, the limitations of user experiments, such as the limited amount of data collected and artificial problematic scenarios, may be complemented by TLA.

3.1 Transaction Log Analysis

The first proposed method is to analyze logs, i.e., the recorded interactions of the general public, which are collected from library servers. Transaction logs are electronic records of human interactions with the system that are recorded by machine. TLA (transaction log analysis) is the process of analyzing these log data in order to obtain useful information. TLA leads to improved understanding about the interactions that have occurred during a search episode between the faceted OPACs and the searchers.

In this proposed study, transaction logs were collected from the UNC and Phoenix OPAC servers. The logs are information-rich, not only because of the large amounts of data they contain, but also in that the logs can capture the users' behaviors both in a naturalistic setting and in a laboratory environment. The following sections introduce TLA as it is employed in this research, and include log collecting, processing, and analyzing.

3.1.1 Data Collection and Description

The UNC library transaction logs were collected from the library's Apache Server by the library IT staff. The logs were generated by Apache itself. A log file contains all the requests for a single day. It is customary for the library IT staff to store approximately three months of log files. Therefore, the author for this study requested log data for every available three-month period between October 1, 2010 to March 31, 2011; the log data were sent to the author in the form of CDs.

The Phoenix library transaction logs were sent to the author by the Phoenix library webmaster via email attachments. Each data package contained one month's worth of logs. Along with the data, the webmaster also sent site specifications that enabled the author to understand the log parameters. The Phoenix logs in this study match the time window of the UNC logs, i.e., October 1, 2010 to March 31, 2011. The two available datasets are described in Table 3.1.

3.1.2 Data Processing

After collecting the data, the author continued to process the data and found scalability to be a challenge. Processing two datasets with more than 3 million records required significant computing and storage resources. However, even more challenging was

analyzing the data, making sense of the data, and understanding the ways people interacted with the library catalogs from the data.

Table 3.1 Log dataset

Log Dataset	Time Frame	Size	Available Fields
UNC library Apache server logs	10/01/2010– 3/31/2011 180 days	1,556,707 useful records 406,794 sessions	IP address /date,time/URL /referrer URL /user agent
Phoenix Apache server logs	10/01/2010-03/31/2011 180 days	1,585,788 useful records 239,523 sessions	IP address /date,time/URL /referrer URL /user agent

Data cleansing

Because transaction logs are used to study human interactions, the logs must reflect the actions of real people rather than web crawlers or other automated processes that act as a person making requests of servers. In addition to crawlers, other unwanted data include duplicate transaction logs. Consequently, methods were needed to eliminate as much automated crawler activity as possible from the logs. Due to technical problems with the Apache server, some transaction logs had duplicates. A script was written to eliminate all the duplicate records.

Data parsing

Server logs are suited for automated analysis because they contain structured data wherein each log record includes the same set of information. For parsing plain text logs into meaningful pieces of data, Perl was chosen for its strong support of regular expressions. A series of Perl scripts was written to accomplish the entire processing. Rather than combining several scripts into a single script, maintaining several scripts facilitated error checking and

later modification. Detailed information regarding these scripts is described in Niu, Lown, and Hemminger (2009).

One challenge in processing the logs in this study was the large size of the two datasets, each of which held millions of records. Much computing and storage resource was needed. As a result, all the data were processed on one of the servers which belonged to the School of Information and Library Science at UNC-CH and then stored in a MySQL database on another School server.

Because the transaction logs are derived from the library server, which is open to the public, the recorded interactions reflect the behaviors of a large number of individuals mixed together and are ordered according to time stamp. Identifying individual users from logs is of great importance to the study and, therefore, the concept of ‘session’ is introduced. One session is assumed to be conducted by one person for one search task, and sessions are independent of each other. In the literature, the definition of session never achieves agreement and suffers problems accordingly. Generally, researchers use times that range from 5 to 60 minutes as a ‘session’. (For more information about this issue, see Chau, Fang, & Sheng, 2005; Göker & He, 2000; Hert & Marchionini, 1997; Marchionini, 2002; Mat-Hassan & Levene, 2005; Silverstein, Henzinger, Marais, & Morics, 1999). Most of the methods are merely a rough and imprecise way to separate user sessions. This dissertation employs Lown’s definition (2008) of session as a series of consecutive requests from the same IP address with no periods of inaction greater than 30 minutes.

Data coding

An ‘action’ refers to a user’s interaction with the system. In most cases, a *transaction* represents a single action. One of the essential tasks in this study is to code log entries as

particular individual actions that were part of the search process. Due to the large amount of data, the coding work had to be done automatically rather than manually. A Perl script was used to code each log record by comparing its URL request and the referrer's URL request in order to understand the user's action. In some cases, logs might be coded incorrectly because some actions were cached on the local machine or proxy servers, and thus were not captured by the logs. However, given the enormous number of records, the percentage of incorrectly coded records was rather low, with an estimate of less than 1%⁴, which is acceptable for the analysis.

Granularity is a major concern in adopting coding schema. According to Wildemuth and Moore (1995), detailed coding schemes are too fine-grained to make statistical analysis effective, although a coarser scheme would not provide enough detail. For this study, both fine-grained coding and coarse-grained coding schema are used to complement each other. Table 3.2 summarizes the available action codes and their descriptions.

3.1.3 Data Analyzing

After the log data were processed using Perl scripts, they were ready for generating some basic and descriptive results to address Research Question 1. In order to answer Questions 2 and 3, some in-depth analyzes were needed to identify the search clusters and to understand the role that facets play in helping to formulate or reformulate the search. An established clustering technique and an innovative visualizing method are proposed.

Clustering

In order to investigate whether the search sessions derived from the transaction logs naturally segregate into groups, a clustering method is proposed to detect groups of

⁴ This is a rough estimate from a sample of 100 search sessions.

Table 3.2 Summary of action codes (both fine-grained and coarse-grained codes)

UNC grained codes	coarse	UNC grained codes	fine	Description	Phoenix grained codes	fine	Phoenix grained codes	coarse
TextSearch		SingleTermText		Submit a single word query in the text search box	SingleTermText		TextSearch	
		MultipleTermText		Submit a multiple-word query in the text search box	MultipleTermText			
		SwitchTextField		Switch the search field (e.g. author, title) using the same text query as before	SwitchTextField			
		MultipleFieldText		Submit a query with multiple search boxes in different fields, typically on advanced search page	MultipleFieldText			
		BlankText		Submit an empty query in the search box	BlankText			
		BooleanSearch		Submit a query in the Boolean search box	N/A			
FacetOperation		AddFacet		Click on a facet value to incorporate it to the current search. For UNC library, <i>AddFacet</i> is for refining the search only. For Phoenix, could be either search refining or collection browsing	Refine		FacetOperation	
					Browse			
		RemoveFacet		Click <i>x</i> next to the chosen facet value to remove it	RemoveFacet			
		ShowMoreFacet		Click <i>more</i> under a facet group to show more values	ShowMoreFacet			
		RefineYears		Under <i>Publication Year</i> facet group, manually type the starting and ending years and submit	N/A			
		OpenFacet		Click the + next to a facet group to show the values of the facet	N/A			
OptionSelection		CloseFacet		Click the - next to a facet group to hide the values of the facet	N/A			
		BeginSimpleTextSearch		Open the <i>Search</i> tab to begin a simple text search	N/A			
		BeginNewTitlesSearch		Open the <i>Browse New Titles</i> tab to begin a new title search	N/A			
		BeginCallNumberSearch		Open the <i>Browse by Call Number</i> tab to begin a call number search	N/A			
		BeginAdvancedSearch		Open the <i>Advanced Search</i> tab to begin a multiple fields text search	BeginAdvancedSearch		OptionSelection	
		N/A		Open the <i>books</i> tab to search in book collection	BookSearch			
		N/A		Open the <i>movies</i> tab to search in movie collection	MoviesSearch			
		N/A		Open the <i>music</i> tab to search in music collection	MusicSearch			
RecordOperation		N/A		Open the <i>downloadables</i> tab to search in media collection	DownloadablesSearch			
		N/A		Open the <i>magazines & newspapers</i> tab to search in book collection	MagNewsSearch			
		ViewRecord		Click on a record link to view details about the record	ViewRecord		RecordOperation	
		NextPage		Click on a page number or the next button on the result page to view the next page	NextPage			
		SortResult		Choose the options (relevance, publication year ...) from the <i>Sort by</i> drop down menus to sort the result list	SortResult			
FollowupAction		N/A			SaveItem			
		FollowupAction		Click on a link provided within a record to find related records	FollowupAction		FollowupAction	
Refresh		Refresh		Click refresh button of the browser	Refresh		Refresh	

that is based on which clusters were identified represents the distribution of actions within the session.

In statistics, cluster analysis is a technique of creating groups that fit observations so that within-group observations are alike and between-group observations are different. Two general types of clustering methods are available: hierarchical and non-hierarchical. A limitation of the hierarchical method is its large computational costs for a large data set. A drawback of the non-hierarchical method is that the number of clusters is predetermined information. Without knowledge of the underlying structure, it is difficult to determine the number of clusters *a priori*. To reduce the computational costs and avoid subjectively predefining the number of clusters, a hybrid approach that combines the hierarchical and non-hierarchical methods was employed in this study. This hybrid method is derived from Chen and Cooper's study (2001) and can be described briefly as the following steps:

1. Identify and calculate the characteristics for each search session. (In this study, the characteristics of search sessions are the proportions of each action code within a search session.)
2. Classify search sessions into groups using a non-hierarchical clustering method (SAS procedure FASTCLUS).
3. Cluster the resulting groups further using the hierarchical clustering method (SAS procedure CLUSTER with Ward's algorithm).
4. Graph and visualize the clustering results.
5. Interpret the nature of the session characteristics associated with each cluster and label these clusters.

Visualizing

This research proposes a visualization technique, referred to as VUTL (visualization for understanding transaction logs). The technique, based on the flowchart model, utilizes standard web programming tools (PHP, HTML, SQL) to convert standard library transaction logs automatically into graphical representations that may be visualized in web browsers. The resulting visualizations support the exploration and analysis of large amounts of log data produced by library catalog interfaces, so that individual and group behaviors may be studied. The visualization takes advantage of the human visual system to understand complicated patterns quickly and to facilitate high-quality human judgment. This technique is applied to logs from both the UNC Library and the Phoenix Public Library in this study.

The visualization was implemented by the following two steps: 1) Store the transaction logs into the MySQL database, and 2) retrieve the records from the database using PHP scripts and convert the records into a predefined HTML graphical representation. The whole process is illustrated in Figure 3.1.

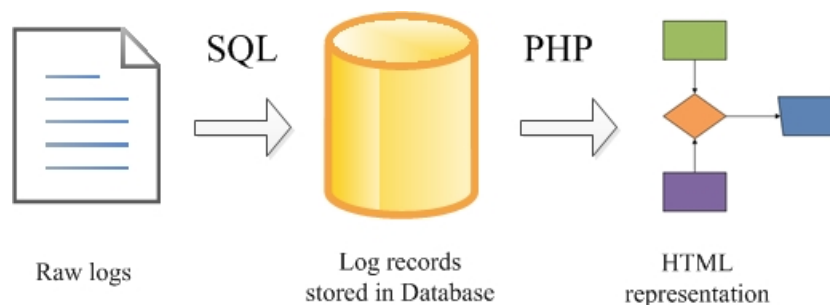


Figure 3.1 Process for visualizing transaction logs

Figure 3.2 shows a randomly chosen search session that is visualized using this technique. The action sequence, types of actions, and their durations are all displayed visually. Figure 3.2 shows that each box represents a single action, and that each row of

boxes denotes one search iteration under the same keywords. The number in each box is the time (in seconds) that the person spent on the particular action. Therefore, the visualized representation allows an understanding of the searcher's actions in a time order.

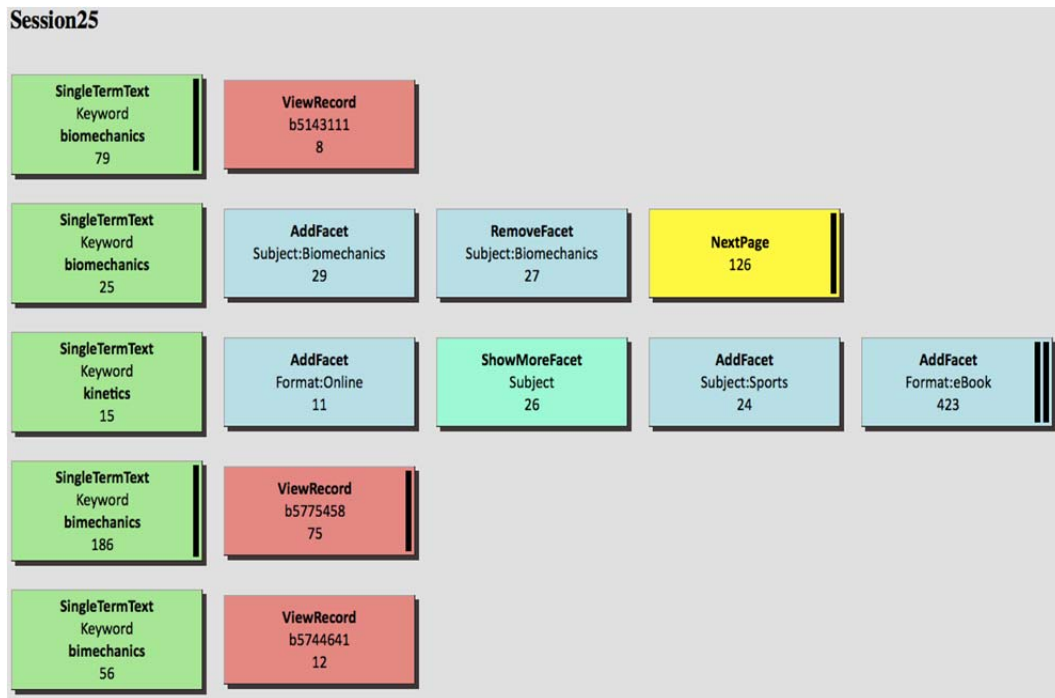


Figure 3.2 Example of a visualized search session

The primary benefit of this VUTL method is its ability to provide a visual and intuitive presentation for a large amount of machine-generated server log data. Furthermore, the automation of the visualization process allows the input of a large amount of data. This work is unique in that it provides open source software for generating a graphical representation of users' interactions customized to a specific search system, in this case a faceted library catalog. Although VUTL is derived from the log data of a library catalog, it is flexible enough to be adapted to many other information systems, such as web search engines, online databases, and other OPACs (Niu & Hemminger, 2010).

3.2 User Experiments

As mentioned before, from log data only, the author was unable to understand the searchers' underlying motivations and feelings. Information collected from user experiments complements TLA because such information captures the search context that includes users' demographics, motivations, and satisfaction levels. User experiments also provide control over the investigation in terms of controlled task complexity, defined search success/failure, etc.

In order to address the second and the third group of research questions, as well as to match the library catalogs from which the logs were collected, two user experiments were conducted. The first experiment compares a non-faceted and a faceted library interface to understand whether facets improve the search, and the second experiment compares two different faceted library interfaces to reveal the effects of different facet implementation on ways that people search. The experimental designs for the two experiments are similar and include the recruitment of study participants, the procedure, balanced task sets, and various instruments used to collect data. The designs are described as follows.

3.2.1 Subject

In this study, the subjects' prior searching experiences need to be controlled. The subjects were expected to have similar experiences with the UNC and the Phoenix libraries. Freshmen were assumed to have limited experience with both libraries and, therefore, were chosen to be the experiment subjects. In addition to representing novice users, freshmen also had little possible bias with either version of the library interfaces. The study recruited forty freshmen across the UNC campus. Subjects were recruited via campus-wide email solicitation. Twenty-four of the subjects participated in the first experiment and the other 16

subjects participated in the second experiment. Each subject received \$30 as compensation for the time devoted to the experiment.

3.2.2 Within-subject Design

Within-subject measurements are more efficient than between-subject measurements in terms of the use of subjects (Wildemuth & Cao, 2009). Within-subject measurements allow the direct comparison of the same individual between two study interventions; therefore, they provide powerful statistical results and minimize the variation in individual characteristics. A common disadvantage, however, is that if the same cases are used, unwanted learning effects may exist between the two interactions. Latin square experimental design is a way to arrange the order of tasks in order to counterbalance the order effect on the study. Therefore, a counterbalanced within-subject design was chosen for this research as an optimal way to design the two experiments.

The purpose of the first experiment is to examine whether facets improve search performance and user satisfaction. Different manipulations were conducted on the UNC faceted catalog to prepare the non-faceted and faceted interfaces. The facet feature was removed from the standard catalog interface, and the resulting interface became the non-faceted interface. To produce the faceted interface, the advanced search option was removed to focus the interaction on facet usage and more tightly control other interactions to allow for better evaluation of facet usage (as shown in Figure 3.3).

Each of the twenty-four subject completed 4 close-ended and 4 open-ended tasks for each interface (i.e., $(4 \text{ tasks} + 4 \text{ tasks}) * 2 \text{ interfaces} = 16 \text{ tasks in total}$). Within-subject measurements were used, and each subject experienced both interfaces. The orders of the

interfaces and the tasks were controlled by Latin square design to reduce the order bias. The experiment time for each subject was roughly two hours.

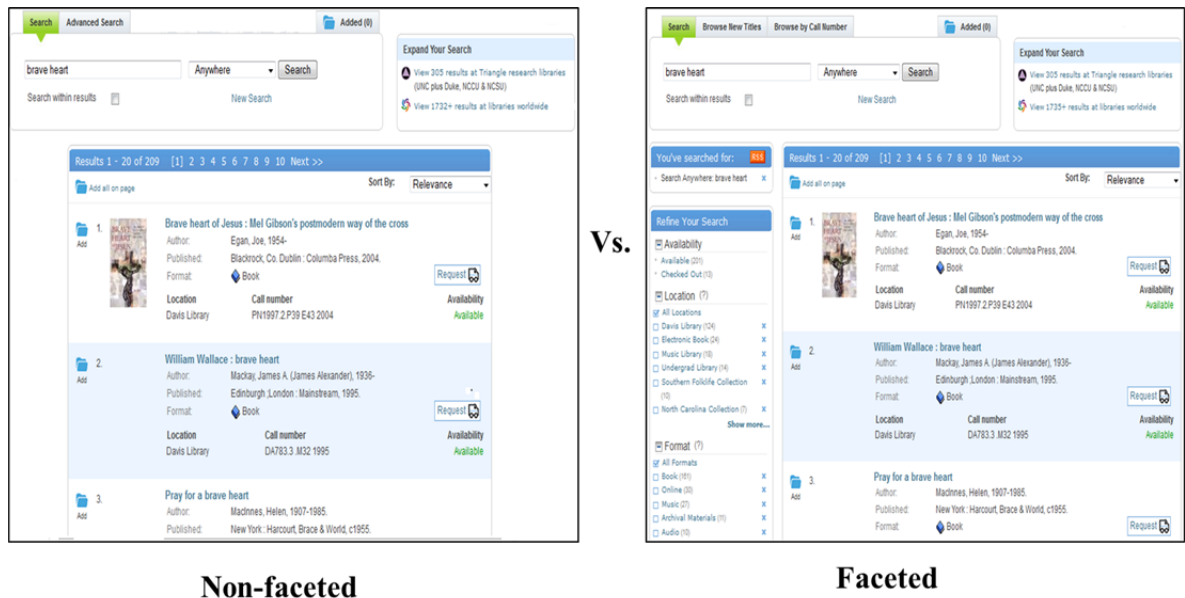


Figure 3.3 Interfaces used in Experiment 1

The second experiment aimed at comparing across the two faceted interfaces (the UNC and the Phoenix catalogs) to understand searchers' performance, satisfaction, and ways they used facets in the different library settings (Figure 3.4). Sixteen subjects conducted 4 close-ended tasks and 4 open-ended tasks for both catalogs (i.e., (4 tasks + 4 tasks) * 2 interfaces = 16 tasks in total). The experiment design is similar to that of the first experiment, and the experiment time was also approximately two hours.

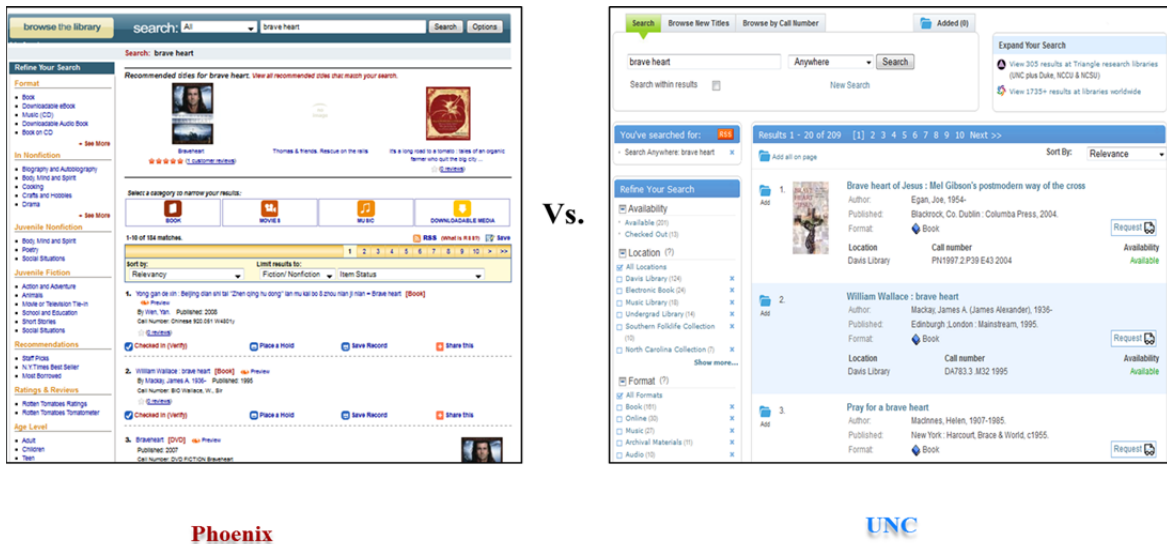
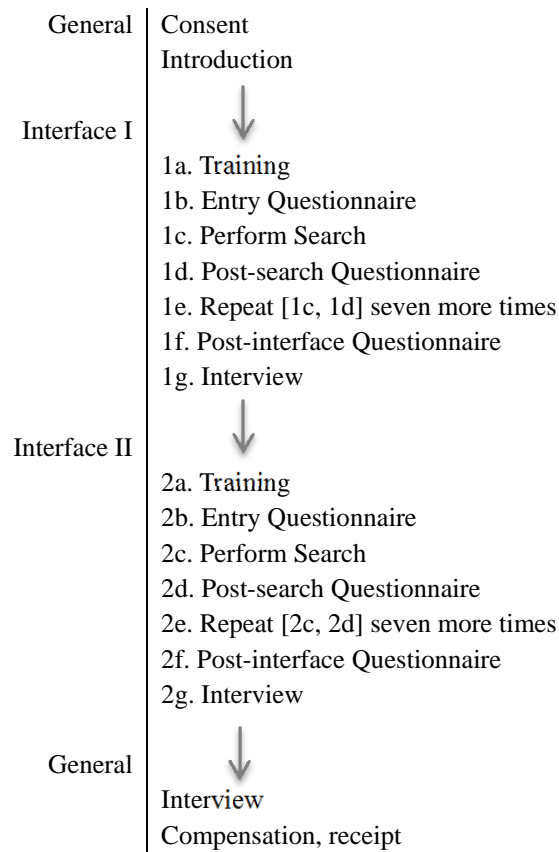


Figure 3.4 Interfaces used in Experiment 2

3.2.3 Procedure

After arriving at the laboratory, the subjects were introduced to the study and completed a consent form. Then, the subjects were presented with the first search interface and trained with the basics of that interface. Next, they performed eight tasks on this first interface. After the subjects finished searching for each task, they completed a post-search questionnaire. After the subjects finished searching for each interface, a post-interface questionnaire and a semi-structured interview were presented to the searcher. After a brief break, the subject was introduced to the second interface, whereby the process was exactly the same as for the first interface. After the second interface tasks were finished, the subjects were debriefed about the research purpose and given a USD \$30 honorarium. The individual sessions lasted about two hours. The procedure flow for the experiments is described in Table 3.3.

Table 3.3 Basic Experiment Flow



The consent form, entry questionnaire, post-search questionnaire, post-interface questionnaire, and the interview scripts were all adapted from previous established user experiments (Kelly, Cushing, Dostert, Niu, & Gyllstrom, 2010; Kules, 2006; Ramdeen & Hemminger, 2012) and modified to fit the current study needs. They are found in Appendix A - E and are also available at: <http://www.unc.edu/~xiniu/>.

In order to follow the experiment flow, the user must log in to an answer collecting system, which was developed by the author using PHP script as part of this dissertation project. Several screenshots of the system are displayed in Figure 3.5.

Introduction

- Welcome to our experiment!
- We are inviting you to be involved in the evaluation of library interfaces. The purpose of this study is to evaluate the usability of the library catalog interfaces. Your evaluation will help us improve the usability for others.
- Your participation in the study is voluntary.
- Detailed information including the informed consent form is available in the informed consent form.

Login

User ID

Task 5

You want to watch some recently released movies; you are to find three movies that have been received in last month by the UNC "Media Resources Center".

Note: please place your answers to the questions in the table below, taken from the library catalog webpage. Do not put "N/A" if you can't find an item, please put "N/A".

Answers:

Post-search Questionnaire

1. How easy was it to find the relevant item(s) for the topic?

very difficult					very easy
1	2	3	4	5	

2. How satisfied are you with the search results?

very unsatisfied					
1	2	3	4	5	

3. How enjoyable was the search process?

very frustrating					
1	2	3	4	5	

Post-Interface Questionnaire

	Strongly Disagree				Strongly Agree
	1	2	3	4	5
The interface was intuitive.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interface makes the searching process enjoyable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interface makes it easy to find relevant items.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interface makes it quick to find relevant items.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 3.5 Online answer collector interface screenshots

3.2.4 Training

There has been an ongoing debate on whether and how to train subjects before the ‘real’ experiment. Previous studies indicate that different amount and different method of training can affect how people search (for example, Grossman & Fitzmaurice, 2010; Kules & Capra, 2011). To observe subjects’ natural behavior and not bias their performance too much, it is important to not overtrain subjects on the variables being tested, in this case facet use. At the same time, it is important to do sufficient training to reduce performance inconsistencies due to learning effects, and to be sure subjects are aware of parts of the interface (facets). In this research, we balance these two objectives with an ‘appropriate’ amount of training. Each subject received two stages of training: the first stage is an interactive stage with two practice tasks. The searcher may ask questions and the author may make suggestions as to how he/she

might proceed; the second stage is an independent stage with three practice tasks for which the searcher needs to use his/her own judgments. The training was result-oriented, which means the training was finished once the target items of the tasks were found. No effort was made to train the searchers in the basic principles and terminologies of library catalogs. The training tasks are attached in Appendix E.

3.2.5 Tasks

The literature on OPAC studies suggests that people primarily conduct two types of searches using OPACs (Hancock-Beaulieu, 1990). One is the known-item search where the user wants to find a specific item using information such as author, title, and publication year. In contrast, another type of search frequently conducted by users is the subject search, which is conducted on a topic using either a keyword or a subject heading. Known-item searches and subject searches can also be called close-ended and open-ended searches, respectively, because the former has a definite target document and the latter has more open-ended target documents. As far as complexity is concerned, people conduct both simple and complex searches using library catalogs. Table 3.4 presents some definitions of these search types obtained from relevant literature.

In this research, two types of tasks are proposed: close-ended (known-item) and open-ended (subject search). A description of each type is given in Table 3.5. While in the literature many terms are used to describe the complexity of the tasks, for consistency in this study we use “simple” and “complex” for the complexity of the tasks. We operationalize this complexity through two levels of ‘fuzziness’ of the task description. The fuzzier the description, the more complex the task. To reduce the subjectivity of predefined task

complexity, the searchers' perceived complexity was collected via the post-search questionnaire. A task example is provided for each type in Table 3.5.

Table 3.4 Task types in library and information science literature

Task Type	Definitions	Attribution
Simple	"common conditions, required few answers, and were relatively well-defined"	Sharit, Hernandez, Czaja & Pirolli, 2008, p. 13
	"required only the use of one or two commands and followed explicitly from a single-step example of system operation"	Borgman, 1986a, p. 53
Complex	"generally dealt with uncommon conditions, required many answers, and were more ill-defined"	Sharit, Hernandez, Czaja & Pirolli, 2008, p. 13
	"one that required the user to extrapolate from the examples given and to apply some problem-solving skills"	Borgman, 1986a, p. 53
Known-item	"find a piece of information known to exist. The search scope is so narrowly focused and specific that every searcher should have the same criteria in evaluating the relevance of the retrieved information"	Kim, 2001, p. 238
	"the user wants to find out from the catalog whether a specific item about which some bibliographic information is known (e.g., author, editor, or title) is listed"	Large & Beheshiti, 1997, p121
Subject	narrowly defined as a subject heading search	
	"searches conducted on a topic where the user wants to locate in the catalog all items that deal with the topic"	Large & Beheshiti, 1997, p121
Open	"There is no one exact answer and searchers must develop acceptable responses. There may be many relevant sources, and searchers may have to study them and perhaps combine information"	White & Iivone, 2001, p. 723
	"ill-structured problems, where the information required for accomplishment cannot be determined in advance"	Bilal, 2002, p. 1171
Closed	"exact answers are wanted", "Searchers have little discretion in judging correct answers or choosing alternatives"	White & Iivone, 2001, p. 723
	"simple, well defined, and have structured problems. They can be routine information processing tasks with elements that are predetermined (the user knows them)"	Bilal, 2002, p. 1171

For the first experiment, there are two sets of tasks, A and B, with 8 tasks each. Each subject needs to experience A with an interface and B with the other interface. The order of the task sets and the combination with the interfaces were counterbalanced to reduce the order bias. The detail about the experiment randomization is available at <http://www.unc.edu/~xiniu/randomization.html>.

For the second experiment, a set of the same or similar tasks, C and D (similar to A and B) were used. Set C is almost the same as Set A except a few tasks and Set D is nearly the same with Set B with a few exceptions. The study randomization is similar to the first experiment. All the four task sets are displayed in Appendix D.

Table 3.5 Task types and examples

Task Type	Close-ended	Open-ended
	Example	Example
Complexity		
Simple	This task type asks users to find one specific item based on some specific information, such as author, title, publisher, etc. There is only one absolutely correct answer.	This task type asks users to find three items on a particular topic. There might be more than three correct items.
The task description has three exact pieces of information	You want to find the classic piece of “To kill a mockingbird”, published by “HarperCollinsPublishers” in 1995.	You want to watch some recently released movies; you are to find three movies that have been received in last month by the UNC "Media Resources Center".
complex		
The task description has one exact piece and two fuzzy pieces of information	You are developing an online questionnaire and need some entry-level guidance on PHP programming. A professor recommended a book to you, but all you can remember is the author's last name started with "Vas", like Vassemeier, Vastuck. And that the book was published in 2009.	You want to find three science fiction books that have stories involving robots published after 2000.

3.2.6 Instruments

The two experiments were conducted on a Dell OptiPlex 980 desktop equipped with Windows 7 Enterprise, 4GB RAM and a 3.33 GHz Intel processor. The desktop also was equipped with a Morae Recorder.

Morae (<http://www.techsmith.com/morae.html>) is commercial software developed by TechSmith. Morae is used for recording desktop activity on the user’s computer and employs

camera video of the user. It also creates a synchronized index of events that occur behind the scenes in applications and in the operating system. It has three major components: 1) Morae Recorder, for recording onscreen and keyboard events, 2) Morae Observer, installed on another computer for researchers to observe the screen video and audio of the computer being researched, and 3) Morae Manager, for researchers to analyze the recordings after the experiment.

Morae Recorder was installed on a Dell desktop computer and used to capture the contents of the computer screen, including the screen text, keystrokes, and mouse clicks, as well as the audio data via a web camera connected to it. The camera is a Logitech QuickCam® Orbit AF.

Morae Observer was installed on a Lenovo ThinkPad X201 laptop for the researcher to observe events on the Dell computer directly. The Observer and Recorder were connected through the University LAN.

The web browser used for the experiment answer collecting system is Internet Explorer 9, which is compatible with the Morae Recorder.

3.2.7 Variables and Data Collecting

The variables of interest for both experiments can be summarized into three groups: user demographics, user performance, and user satisfaction. Table 3.7 summarizes the groupings of the variables.

The entry questionnaire is the main source that was used to gather user demographics. Participants were asked about their age, gender, occupation, past search experience with OPACs and the faceted display.

The answer collecting system is where users entered their finding results. Therefore, it is the primary source for collecting the performance data. The search time was derived by calculating the time differences for a particular task from the transaction logs. As for search accuracy, a grading rubric was developed for each of the task types, which was reviewed by two independent judges. The interrater agreement between the two judges was high (kappa of 0.89⁵). The differences were reviewed together by the two judges until an agreement was reached.

User perception is concerned primarily about search process perception, as captured by several 5-point Likert variables in the post-search and post-interface questionnaires. User perception data also were gathered in the concluding interview by asking users their feelings and perceptions after they experienced a particular interface.

Table 3.6 Variables and data collecting from the experiments

Variable group	Variable	Collecting methods
User demographics	<ul style="list-style-type: none"> • Age • Gender • Past searching experience with OPACs • Past searching experience with facets 	<ul style="list-style-type: none"> • Entry Questionnaire
User performance	<ul style="list-style-type: none"> • Accuracy • Time 	<ul style="list-style-type: none"> • Answer collecting system • Transaction logs
User perception	<ul style="list-style-type: none"> • Level of complexity • Satisfaction • Enjoyment • Other feelings 	<ul style="list-style-type: none"> • Post-search questionnaire, • Post-interface questionnaire, • Concluding semi-structured interview
Facet usage	<ul style="list-style-type: none"> • Facet operation percentage • Why use facets • How use facets 	<ul style="list-style-type: none"> • Transaction logs • Concluding semi-structured interview

⁵ Calculated using an online Kappa calculator, <http://justusrandolph.net/kappa>.

Facet usage is addressed quantitatively in terms of facet operation percentage broken down by task type and task complexity. Facet usage also is addressed qualitatively by investigating the reasons and the ways people use facets. Transaction logs and interviews were analyzed to answer this group of questions.

The transaction logs collected for both experiments could be used to triangulate with the transaction logs harvested from the 'wild', i.e., the generic users of the libraries. Thus, users' search behaviors in a controlled laboratory environment were compared to those in a naturalistic setting.

3.2.8 Data Analysis

The independent variables for the experiments are interface, task type, and level of task complexity while the dependent variables are search time, search accuracy, and users' ratings on the interface intuitiveness and enjoyment.

T-test, also referred to as Student's t-test, was conducted to assess whether the means of two groups were statistically different from each other. General linear model (GLM) regression analysis was applied to investigate the effect and the interaction effect of the independent variables on the dependent variables. The statistical significance level is .05 for all the statistical tests. All the data were imported to SAS 9.2 package for statistical analyzing.

In order to analyze the qualitative data collected during the interview at the end of the experiment, subject responses were manually coded and organized by similarity and then by theme.

3.3 Summary

This study proposes a hybrid of user experiments and transaction log analysis to investigate people's information search behavior with faceted library catalogs. The two research methods complement each other in that one is designed for a controlled laboratory environment with a limited amount of data, and the other is designed for a naturalistic setting with a large amount of data. Detailed descriptions of the two proposed user experiments are reported in this section, and a summary of the proposed analytical methods for transaction logs also is presented.

4. Results

This chapter presents the results gleaned from the different research methods used in this study. First, the overall search statistics and patterns gained from the transaction logs are described. Then, search performance and users' ratings for the first experiment are presented. The role that facets play during the search is discussed as well. Finally, the results from the second experiment that compare and contrast the two different library catalogs are presented.

4.1 Log Analysis Results

The transaction logs provide a quantitative overview of the search patterns, user groups, facet usage, and behavioral changes employed as factors in this study. Log analysis is adopted to obtain a general understanding of the ways that people interact with faceted library catalogs.

4.1.1 General Search Statistics

The two datasets were collected from the UNC and Phoenix library catalog servers. In total, 3,142,495 records representing 646,317 search sessions were analyzed.

Session length

Session length in this study is defined as the number of actions within the particular session. Most search sessions were brief, with a single-action session being the most frequent type of session. Using a consistent session boundary identification method, the author was able to contrast the session lengths of the two catalogs. Figure 4.1 presents a graph of the

distribution of the session lengths. (Due to the low percentage ($<0.1\%$), sessions longer than 30 actions are not displayed to make the visualization clearer).

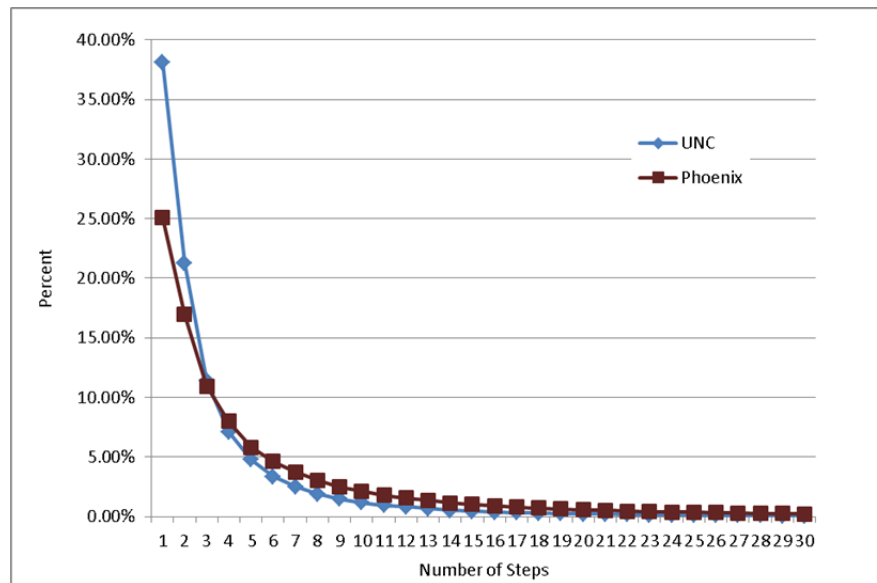


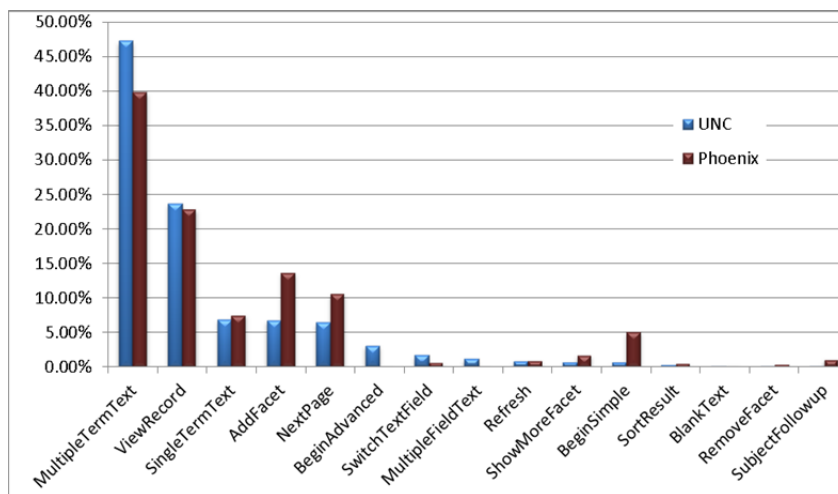
Figure 4.1 Distribution of session lengths based on number of actions

Figure 4.1 shows that both curves appear to follow the power law distribution with a ‘long tail’ at the end of each curve. This graph suggests that, for both the UNC catalog and the Phoenix catalog, the vast majority of search sessions are short, with only a few actions. Sessions with 15 actions or more are very rare (less than 1%). One-action sessions are the most frequent type of search session, accounting for 38% of the UNC data and 25% of the Phoenix data. That is, more than one-third of the UNC catalog searchers and approximately one-fourth of the Phoenix patrons finished their searches with only one interaction with the catalog. These searches are so-called ‘hit and run’ sessions where the users quickly finish the search with or without the information they wanted to find. This result confirms that of previous research that indicates that users’ interactions with search tools are brief, and users tend not to put much effort into their search behavior (Spink et al., 2001). The slight

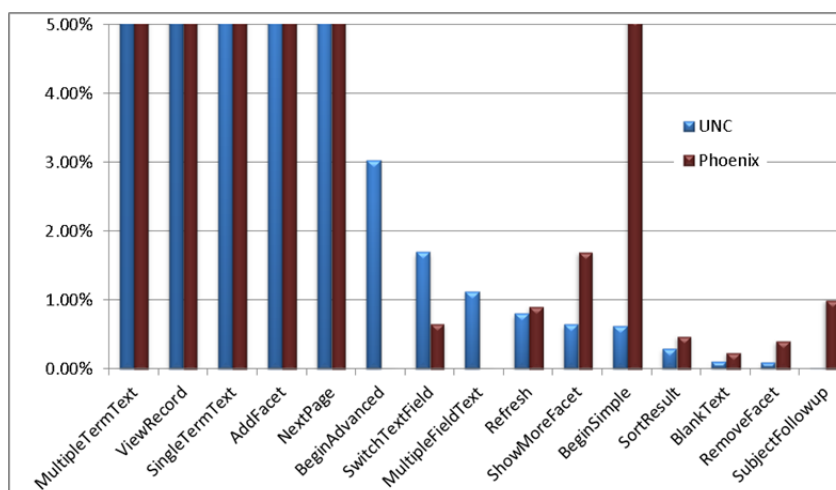
difference between the two curves, demonstrates the more frequent use of 1-2 action sessions in the UNC catalog, while the Phoenix curve shows slightly more 3-15 action sessions. These slight differences result in the lower average number of search actions for the UNC catalog, between these otherwise very similar curves. While this suggests that people tend to conduct slightly more actions using the Phoenix catalog than with the UNC catalog, this might be due to the catalog differences or due to patron differences.

Action distribution

An *action* is one step of user interaction with the catalog. In most cases, one transaction record represents a single action. Despite different implementations of the UNC and Phoenix catalogs, interactions are for the most part common across the two catalogs, which allows for similar coding of actions across the two catalogs. Table 3.2 shows that 15 actions are common to both catalogs. Seven actions are unique to the UNC catalog and six are unique to the Phoenix catalog. Common actions account for approximately two-thirds of the overall possible actions. Figure 4.2 shows the contrast of the common action distributions.



(a)



(b)

Figure 4.2 Distribution of action codes (a) in an original scale, and (b) in an adjusted scale

By and large, the distributions are fairly consistent across the two libraries with a few exceptions. For both catalogs, *MultipleTermText* and *ViewRecord* are the most frequent actions, accounting for 47.3%/39.9% for *MultipleTermText* and 23.7%/22.9% for *ViewRecord*. This finding suggests that the traditional query-response search is the primary mode of interaction. This finding is consistent with Lown's (2008) finding that most requests are text searches and page views. Some actions, such as *SortResult*, *BlankText*, *RemoveFacet*, and *SubjectFollowup*, were rarely conducted in this study. These infrequent actions typically were associated with the novel features that were assumed to be helpful to users. For instance, *SubjectFollowup* is an action where users can use one result item to find other related items by clicking the current item's subject headings or authors. However, most users did not take advantage of this new and sophisticated feature.

The most striking difference is for adding a facet. For the UNC catalog, *AddFacet* accounts for 6.9% of the actions, whereas for the Phoenix catalog, *AddFacet* accounts for

13.7%, which is nearly double that of the UNC catalog. Data analysis reveals that many facet actions that were being incorporated were browsing actions without any text query. The main difference between browsing and refining facets lies in the purpose of the facets: one is for browsing the collection and the other is for narrowing an existing search. Browsing facets typically are used without a text keyword, whereas refining facets usually are added after a keyword search. Incorporating browsing facets is one of the major benefits of the Phoenix library catalog. Its design goal is to offer users an experience like browsing the library shelf for new releases. The browsing facet actions accounted for the higher level of facet actions in the Phoenix catalog. Of the 13.7% facet uptake for the Phoenix catalog, 5.5% came from refining facets and 8.2% came from browsing facets. This finding suggests that facet usage may be higher when browsing facets is supported, and that browsing facets may be used more than faceted refining.

Another noticeable difference between the action distributions of the two catalogs is the portion of people who like to start their search using the advanced search feature. It seems that the Phoenix catalog users were much less likely to conduct advanced searches compared to the UNC searchers (0.02% versus 3.03%, respectively). There may be several possible explanations for the difference: including differences in users of the catalogs (university vs. public); differences in user interface design (in the Phoenix interface, the advanced search option does not stand out enough to be easily seen); and that the Phoenix interface provides more options initially, perhaps reducing the need to engage advanced search.

Initiating a search

Previous studies indicate that the beginning step or initial query is important in determining search success (Ramdeen & Hemminger, 2012; White & Marchionini, 2007). This study examines ways that people initiate their search when provided with different options. As introduced in Section 1.4, the UNC library offers four options as the entry point: *Search*, *Advanced Search*, *Browse New Titles*, and *Browse by Call Number*. As a result, 82.3% of searchers start from *Search*, which is the default start page. Approximately 3.8% searchers begin from the *Advanced Search* page. Neither *Browse New Titles* nor *Browse by Call Number* is a frequent starting place for searchers, 0.3% and 0.06%, respectively. This finding is in line with Lown's (2008) statement that most people start with a text search. It is worth noting that 13.5% of searchers in this study started with a request that was not anticipated as a starting point. For instance, 9.5% of searchers began with *ViewRecord*. They might jump directly into a record by pasting a URL address someone had recommended. They also might resume a search from where they left last time.

For the Phoenix library, the searcher usually had three choices to get started: text querying, browsing, and advanced search. The processed logs show that 68.7% of searchers started from text querying, and only 0.02% began from advanced search, significantly less than at UNC. Using the entry facets to select a type of material, accounting for 18.6%, seemed common as a way to begin a search. The breakdown of this first-step browsing also was checked: 7.2% *eBooks+digital media*, 4.8% *books*, 3.7% *movies*, 2.3% *music*, and 0.68% *magazines & newspapers*. Moreover, 1.8% of searchers began their search by choosing one of the shortcut categories displayed below the first-level categories.

From a facet-text search perspective, two of the four starting options in the UNC catalog could be seen as a faceted beginning. *Browse New Titles* and *Browse by Call Number* are the two options that give searchers two dimensions along which to slice the whole collection, and searchers can click just one of the available labels to reach a subset of the collection without making any text queries. For the Phoenix catalog, a faceted beginning is more straightforward for those searchers who use the general browsing categories or the shortcuts. Compared to the Phoenix catalog, the UNC catalog has limited support for faceted entry/browsing. Figure 4.3 shows the relative proportions of text beginnings and faceted beginnings as two possible ways to start a search. As indicated in this figure, a dominant proportion of searches begins with text queries across the two catalogs, whereas a much smaller proportion begins with facets. This finding indicates that text searching is still a primary starting point for a search, no matter whether and to what extent the interface supports a faceted beginning. This finding differs from that by English et al. (2002) that suggests that participants choose to begin their searches more frequently by browsing than searching. These differences in findings are probably due to the different tasks, different interface implementations (Flamenco, Matrix View and Tree View), different research methods (user experiment), and different underlying data (architectural image database) that the researchers used. Although not as much as text beginnings, faceted beginnings are much more prevalent for the Phoenix catalog than for the UNC catalog, probably because of the Phoenix catalog's better support of the browsing structure. The UNC catalog browse tab specifically supports browsing the LC classification scheme with the purpose of promoting subject access. But searchers rarely take advantage of this feature.

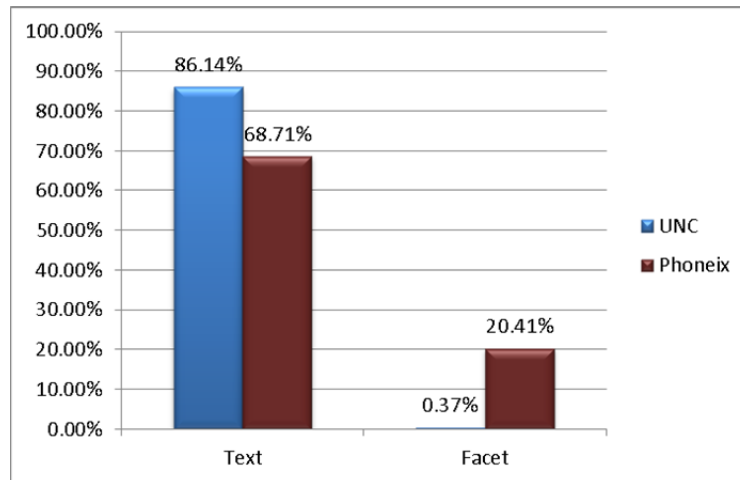


Figure 4.3 Beginning of a search: Text or facet

Clusters of users

In the previous sections, search behavior is explained at the single-action level and at the search session level, both of which concern individual behavior. Also of interest to this study, however, is the identification of search patterns that concern group behavior, such as whether search sessions naturally segregate into certain types of search groups. A cluster analysis was employed, and the distribution of actions is the characteristic on which the clusters were formed. The actions are coded at the fine grain level (Table 3.2), and the fine grain codes are grouped into common coarse code groups (Table 4.1). The coarse codes are used to classify clusters, because fine-grained codes would separate the search sessions into too many unwanted fine clusters and lose the big picture of the main groups.

Table 4.1 Coarse code of actions

coarse code	UNC action code	Phoenix action code
OptionSelection	BeginSimple	BeginSimple
	BeginAdvanced	BeginAdvanced
	BeginCallNumbers	BookSearch
	BeginNewTitles	EResourceSearch
		MovieSearch
		MagNewsSearch
		ImageSearch
TextSearch		MusicSearch
	BlankText	BlankText
	SingleTermText	SingleTermText
	MultipleTermText	MultipleTermText
	SwitchTextField	SwitchTextField
	MultipleFieldText	MultipleFieldText
	BooleanSearch	
FacetOperation	Addfacet	Browse/Refine
	RemoveFacet	RemoveFacet
	OpenFacet	ShowMoreFacet
	CloseFacet	
	ShowMoreFacet	
	AddCallNumbers	
	AddNewTitles	
ResultManipulation	NextPage	NextPage
	SortResult	SortResult
	ViewRecord	ViewRecord
	SubjectFollowup	SubjectFollowup

Note: The actions highlighted in blue are unique to the UNC catalog. The actions highlighted in brown are unique to the Phoenix catalog

The resulting clusters demonstrate the common session characteristics across the catalogs. Figure 4.4 illustrates the dendrograms produced by SAS for both catalogs. A major task of clustering is to decide the number of clusters. This study uses a semi-partial R-square as the metric for the distance between clusters. The number of the clusters is based on the first disruptive change of the semi-partial R-square value. As a result, six clusters were identified for the UNC catalog, and seven clusters were found for the Phoenix catalog. Each cluster was distinct in terms of group behavior and assigned a label (cluster name) according to its characteristics.

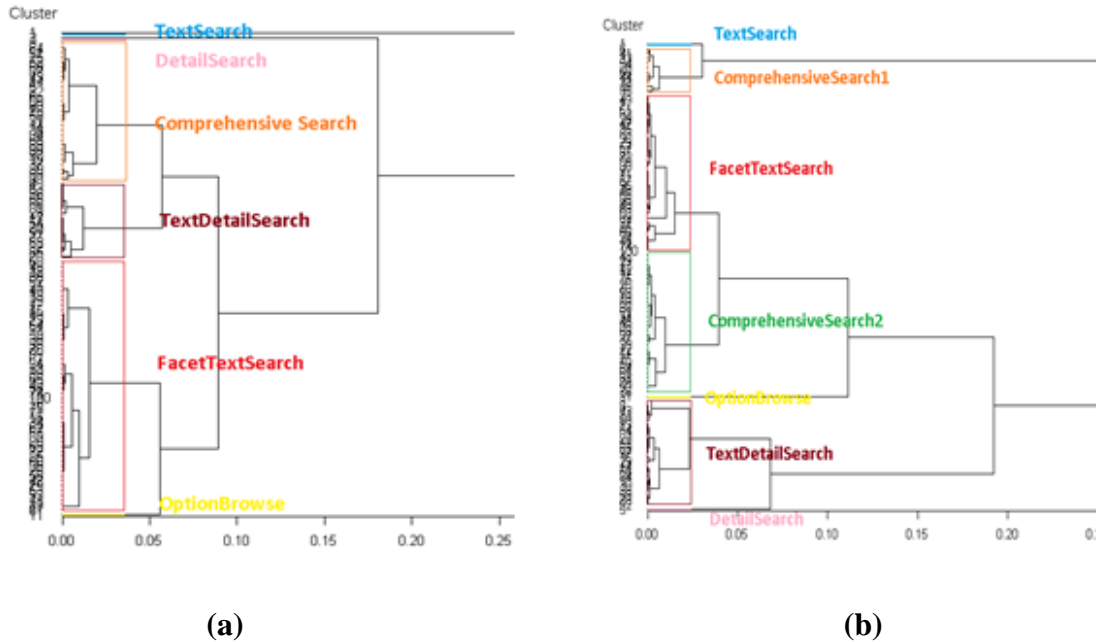


Figure 4.4 Clusters of (a) the UNC catalog searchers, and (b) the Phoenix catalog searchers

Figure 4.5 shows the distribution of the actions for each cluster. The largest cluster for the UNC catalog represents 42.6% of the sessions and is characterized by a dominant proportion (99.97%) of *TextSearch*. *TextSearch* represents traditional text querying, with simple and straightforward search strategies, and without much effort or any follow-ups of the text queries. The second largest cluster (20.0%) of the sessions, labeled *TextDetailSearch*, is characterized by a nearly even distribution of text querying and result checking. For the search sessions in this cluster, users typically enter a keyword and then click through some result items. These sessions also are common in other non-faceted library catalogs and general web search engines. The next largest cluster (18.2%) represents the search sessions where searchers utilize most of the action types, such as select an option, enter a keyword, add a facet, and click through a record. The sessions in this cluster are labeled *ComprehensiveSearch*. The next cluster (9.8%), labeled *DetailSearch*, represents the search

sessions where most actions are associated with checking the search results, such as clicking through an item, sorting the results list, and navigating through pages. The next cluster (8.3%), labeled *FacetTextSearch*, is the featured cluster of the faceted search system. It is characterized by a large percentage of *FacetOperation* (48.7%), together with some *TextSearch* actions(23.7%). This cluster is the research focus of this study. The high involvement of facets and text searches in this group suggests that facets usually are used in combination with text searches rather than used independently. This finding agrees with previous research, indicating that facets mostly are used together with text searches. For example, in the Kules, Capra, and Ryan (2010) study, the researchers found that people use facets at key points in their searches, such as just after issuing a text query or just before adding an item to their shopping cart. This finding also is biased by the system design where the facets are supposed to support refining a text search rather than browsing the collection. Finally, the smallest cluster (1.6%) of the sessions is labeled *OptionBrowse* and is where searchers might navigate across tabs without actually issuing a search.

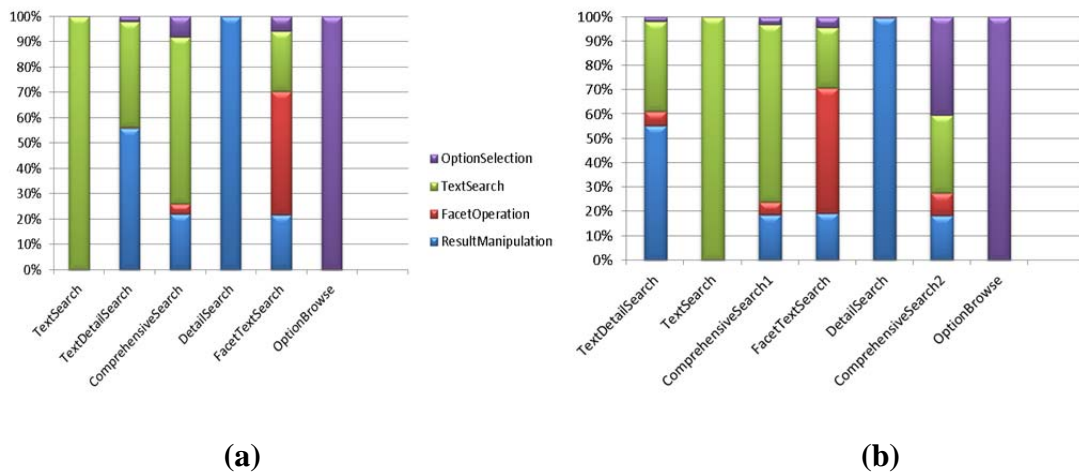


Figure 4.5 Action distribution for each cluster (a) UNC, and (b) Phoenix

The seven clusters identified for the Phoenix catalog exhibit a rough consistency with those of the UNC catalog. The slight difference between them is that the *FacetTextSearch* group has increased (from 8.3% to 11.0%) from the UNC catalog to the Phoenix catalog. *ComprehensiveSearch2* (7.1%) constitutes a new search group where searchers conduct all kinds of actions, similar to *ComprehensiveSearch* in the UNC catalog, but with a significantly higher percentage of facet operations (9.3%) and tab selections (40.3%). The tabs in the Phoenix interface are the general (first-level) browsing categories and, thus, the tab selections are in fact the facet operations. Looking across the clusters for the Phoenix catalog, as shown in Figure 4.5.b, *FacetOperation* appears in four clusters, which is two more than in the UNC catalog. This finding probably suggests a wider range of facet use for all types of searches in the Phoenix catalog as compared to the UNC catalog.

4.1.2 Facet Usage

From the catalog developers' perspective, faceted catalogs offer an intuitive way to search with almost no rules for searching. Users may increase either precision or recall simply by selecting or deselecting facet values. A theoretically motivated interface design is not necessarily the one preferred by the user (Heo & Hirtle, 2001), however. So, for users, the questions that need answers include: Do people use facets and how often do they use them? Do they use them singly or in combination? The statistics generated in the following subsections provide a descriptive picture for understanding facet usage.

Overall, faceted actions account for 8.0% of all actions for the UNC catalog and 15.8% for the Phoenix catalog. Figure 4.6 provides the breakdown of the facet actions. For the UNC catalog, the most frequent facet operation (6.4% out of 8.0%) is adding a facet, whereas for

the Phoenix catalog, *refine*, *browse*, as well as the top level browsing facets are all frequent facet operations.

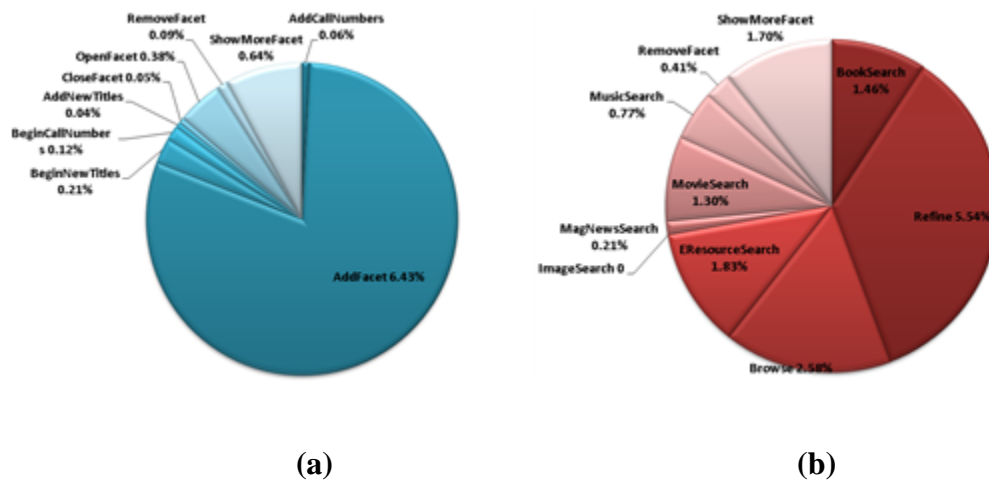


Figure 4.6 Breakdown of *FacetOperation* for (a) UNC, and (b) Phoenix

At the search session level, 12.1% of the UNC search sessions involve at least one facet action. For the Phoenix search sessions, the percentage is 41.5%, approximately 2.5 times that of the UNC sessions. For UNC search sessions, compared to text searches, however, faceted searches are supplemental and used only by a small subset of people. The striking difference between the two catalogs might be due to the Phoenix catalog's better support for faceted browsing in addition to faceted refining, as well as the better support for facets to start a search.

The processed logs also reveal the frequent facet actions. When users added facets to their queries, among the requests that include at least one facet, 58% include just a single facet, 5.3% include two facets, and 3.1% include 3 facets. Very few requests included more than four facets, probably because four facets would significantly reduce the result set, and

not much value would be gained in adding additional facets. This result is in line with Lown's (2008) finding that most searchers incorporate only one facet value into their search.

Table 4.2 summarizes the top ten frequently used single facets and facet combinations for both catalogs. For the UNC data, most of the top ten single facets relate to format and location. That is, it appears that format and location are the most popular facets. Note that neither format nor location is content-related information. Both are considered meta-information that describe the item. This finding is different from that of the Antelman et al.'s study (2006) regarding NC State University Library's popular facets. For the NC State University catalog, the top facet is *LC Classification*, followed closely by *Subject: Topic*, and then *Library (Location)* and *Format*. Different facet layouts and different subject facet implementations might explain this discrepancy. To some extent, this study's findings agree with those of Google, which has recently incorporated *format* and *location* as two of a few facets that are now available on their traditional simple and uncluttered screen. Table 4.2 lists only the top ten single facets and facet combinations. In fact, a 'long tail' list occurred also, suggesting that people use even more facets. Looking at the list for facet combinations, although most combinations are not used as frequently as the single facet, *new titles*, *DVDs*, and *CDs* often appear in the combinations, implying that facet combinations often are used for finding items for entertainment purposes.

For the Phoenix catalog, the top level browsing facets are the same as the frequently used categories. Searchers are able to take advantage of these categories on a link-rich opening page to define their search scope prior to their search. These top level facets are the entry points into the collection, such as books, movies, music, etc. Some of these may have functioned as browsing actions (like new release), but most are simply narrowing to a type of

resource. In addition to the single facets, other facet combinations, such as *Movies* and *New Arrival*, are popular. Similar to those found in the UNC catalog, most of these combinations reflect browsing for something popular for fun. This is in line with Sapiie (1995)’s study that “forty-nine percent of users browse for recreational materials” (p. 145).

Table 4.2a Popular facet list by frequency for the UNC catalog

Rank	Single Facet Value		Facet value combination	
1	Format:eBook	6.70%	Publication Year:2000 to present+New Titles:New in last week+Format:Book+Language:English	0.75%
2	Format:Book	5.78%	Media Resources Center+New in last month+Videos and DVDs	0.57%
3	Format:Online	5.13%	Format:Music+Format:SCORE	0.54%
4	Location:Davis Library	4.03%	Format:Audio+Format:AUD-CD	0.47%
5	Location:Media Resources Center	3.98%	Location:Media Resources Center+Format:Music CD	0.44%
6	Format:Videos and DVDs	3.38%	Location:Media Resources Center+Availability:Available	0.38%
7	Format:Electronic Book	3.26%	Location:Music Library+Format:Music	0.34%
8	Format:Audio	2.37%	Publication Year:2000 to present+New Titles:New in last week+Location:Davis Library+Format:Book+Language:English	0.34%
9	Language:English	2.23%	Availability:Available+Location:Davis Library	0.33%
10	Format:Journal, Magazine, or Serial	2.01%	New Titles:New in last month+Location:Media Resources Center+Format:Videos and DVDs	0.31%

Table 4.2b Popular facet list by frequency for the Phoenix catalog

Rank	Single Facet Value		Facet value combination	
1	Book	17.20%	Movies+New Arrival	1.64%
2	Movies	13.05%	Downloadable Media+Format:Downloadable eBook	1.53%
3	Downloadable Media	9.30%	Movies+Format:DVD	1.35%
4	Music	9.16%	Book+New Arrival	1.01%
5	Format:DVD	1.73%	New Arrival+Language:English+Movies+Format:DVD	0.94%
6	Format:Book	1.50%	Downloadable Media+Format:Downloadable Audio Book	0.79%
7	Format:Book on CD	1.22%	Format:Music (CD)+New Arrival+Language:English	0.58%
8	Format:Music (CD)	1.08%	New Arrival+Format:Music (CD)	0.49%
9	Magazines&Newspapers	0.87%	Book+Format:Book	0.44%
10	Fiction/Nonfiction:Non fiction	0.64%	Book+Fiction/Nonfiction:Non fiction	0.42%

4.1.3 Facets to Support Formulating or Reformulating a Search

The previous sections discuss facet usage at the single-action or search session level. The following sections discuss the role facets play in helping people formulate or reformulate their searches at the search tactical level. According to Kules (2006), the added facet feature leads to altered search tactics because it enhances the information available and the range of possible interactions.

In this study, query length, number of query submissions per search, number of viewed records, and record depth are the quantitative measures that describe search formulating or reformulating tactics. Search tactics for faceted searches and non-faceted searches are examined to see if people search differently depending on the presence of facets. Faceted searches comprise all the search sessions from the *FacetTextSearch* cluster, whereas non-faceted searches are the search sessions from the clusters other than the *FacetTextSearch* cluster. The results are summarized in Table 4.3.

Table 4.3 Quantitative measures for formulating/reformulating a search

	UNC		Phoenix	
	Non-faceted	Faceted	Non-faceted	Faceted
Query Length	3.2	2.6	3.8	2.9
Number of Query Submissions	3.5	2.9	4.7	2.7
Number of Records Viewed	0.9	0.8	1.4	1.6
Record Depth	3.6	12.1	2.9	10.4

Query length is one of the quantitative parameters used to evaluate searchers' query formulation tactics. For the UNC and Phoenix catalogs, the average query length for faceted searches has decreased, probably because with the facet feature people can begin their searches on a broad basis and rely on facets to narrow the search later. This idea suggests a

reduced effort in typing query terms. Across both faceted searches and non-faceted searches, the query length is generally short, with two to three terms. This finding is in line with results found in the literature that suggest that people tend to use short queries when they search through OPACs. The most common session length is two terms (Jansen, Spink, Blakely, & Koshman, 2007; Jones et al., 2000; Lau & Goh, 2006; Mahoui & Cunningham, 2001; Wallace, 1993).

Typically, users need to go through several iterations of formulating and reformulating queries before reaching the documents they need. Because IR is an interactive and iterative process, query formulation usually involves the initial query formulation as well as a query reformulation in which the initial query is adjusted. It is believed that query reformulation is not any easier than the initial query formulation.

In this study, for non-faceted searches, people submitted 3.5 (for UNC) or 4.7 (for Phoenix) queries on average per search, whereas for faceted searches, the number of submitted queries is 2.9 (for UNC) or 2.7 (for Phoenix). It seems that the facets improve the search efficacy by reducing the number of search iterations. This result is probably due either to the fact that the searcher found the information he/she needed quickly and thus did not need a new iteration, or that the searcher used facets to adjust his/her original query and did not need to initiate a new query.

Furthermore, many of the queries beyond the first iteration are simple deviations from the initial one. This observation agrees with White and Marchionini's finding (2007) that many further queries are simply "syntactic variants" of the initial one. The initial query appears to be important in determining search success. When checking for the initial query in this study, the average length for non-faceted searches is 3.9 terms and for faceted searches is

2.5 terms. Larger differences are observed in the initial query length than in the total query length.

The parameter, *number of records viewed*, is a tricky measurement to explain. Many searchers are found to be satisfied with a snippet of information available on the results list without actually clicking through to an item. Whether the searcher is able to find the information he/she needs just by the number of records viewed is difficult to discern. It does not seem to matter whether the number of records viewed increases or decreases.

This study's results indicate a large increase for *record depth* for faceted searches. *Record depth* is defined for non-faceted searches as the location or rank of a particular item in a results list. For faceted searches, however, record depth is the original location had the facets not been selected. As Table 4.3 shows, the depth for non-faceted searches is 3.6 (for UNC) or 2.9 (for Phoenix), which is a relatively high location on the first page (20 or 10 results per page). Basically, searchers tend to stay on the first results page without going further to the following pages. This finding is consistent with that of studies that found that most searchers do not examine more than the first page of the search results (Jansen et al., 2000). The average depth for faceted searches is 12.4 or 10.4, much deeper than for non-faceted searches. Especially for the Phoenix catalog interface, participants were able to access the items beyond the first page (10 results per page). This ability is due to the fact that facets flatten out the depth of the original results list by providing an overview of the shortcut categories. The range of items people can access is thereby greatly enhanced, which confirms the results of a previous study (Kules, 2006) that people explore more deeply when facets are available.

4.1.4 Summary

The results derived from transaction log analysis are aimed at answering the first three research questions raised in Section 1.2. The author will reiterate the three questions below:

1. What are the possible interactions people employ with faceted library catalogs, and how many of these interactions are faceted interactions?
2. Do search sessions segregate naturally into clusters according to their search behavior?
3. Do searchers change the way they formulate or reformulate their search with facets compared to without facets?

In summary, people were conducting over twenty kinds of interactions (actions) with the faceted catalogs. Of all the interactions, 8.0% for UNC and 15.8% for Phoenix were faceted interactions. Search sessions segregated into six clusters for UNC and seven clusters for Phoenix according to the session characteristics. Due to the availability of facets, searchers changed their way of formulating or reformulating their search. Specifically, they tended to issue shorter queries, submit fewer queries, and were able to reach deeper in the result list than without facets.

4.1.5 Discussion

Transaction log analysis serves as a preliminary study for providing general statistics and ideas about the ways patrons in this study interacted with the two faceted catalogs. Most of the search sessions were brief traditional text querying actions. Participants were able to use facets, but compared to text searching, faceted search was used only by a small group of people. Facet usage was higher when faceted browsing was supported in addition to faceted

refining. When both faceted browsing and refining were supported, faceted browsing was found to be used more than faceted refining. As for facet uptake frequency, meta-information and ‘administrative’ facets are the most popular facets for the UNC catalog. For the Phoenix catalog, the top-level browsing categories are the most frequently used facets. Participant searchers who used facets tended to formulate or reformulate their searches differently by issuing short queries, submitting few queries, and exploring the results list in depth.

The percentages of facet operations (8.0% and 15.8%) are lower than those found from other research that concerns faceted catalogs. For example, Antelman et al. (2006) found that 30% of the search requests to the NC State University Library involve post-search refinements obtained from the dimensions on the results page. Lown’s (2008) finding with the same library (NC State University Library) indicates that facet searching accounts for 15% to 18% of all requests, and 34% of search sessions include at least one facet search. These two studies excluded many actions when calculating the percentages, whereas in this study, every human action is taken into account. In spite of the percentage differences, this study agrees with all the previous studies that have found that faceted searches represent a smaller part of the search process than text searches. This finding is due either to the fact that searchers avoid the extra cognitive or technical burdens needed to incorporate facets if the text search does a good enough job, or that searchers are not able to understand facets well enough to use them properly.

The most important finding derived from the logs is that faceted browsing is an important activity in addition to faceted refining. This finding is demonstrated by the Phoenix catalog where faceted browsing is supported and well utilized by patrons. In the UNC library where facet operations only support narrowing a text search, facet operations account for

only 8.0% of the facet operations. In contrast, for the Phoenix catalog, the percentage of facet operations is 15.8 percent. Marchionini characterizes browsing strategies as “informal and opportunistic” (Marchionini, 1995). A good catalog browse should simulate the experience of browsing the stacks and even improve it, because the online catalog overcomes the physical constraints of a brick and mortar library and supports nonlinear browsing. Compared to the Phoenix catalog, the number of browse search sessions is much lower for the UNC catalog. Even for the Phoenix catalog, the number of true browses with no search term entered is much fewer than for text searches.

The UNC catalog implements a combination of LC classification numbers and LCSHs for the purpose of enhancing subject access. For example, a broad topic search might return thousands of results without the classification numbers or subject headings. The two together can be used to define logical clusters or post-coordinate refinement for browsing. Because refinement is based on the authority data, users are able to add or remove values in any order to assist them and can never refine to zero records (dead ends). But, according to the log data, the classification numbers and subject headings are rarely used by searchers. A gap may exist between the way a question is asked and the way it might be answered (Borgman, 1996). The Phoenix catalog implements two sets of subject headings as facets: BISAC and LCSHs. It turns out that some BISAC subject headings are listed in the top ten frequently-used facets but no LCSH is listed. It appears, therefore, that the BISAC subject headings are easier to initiate browsing than the LCSHs. Also, the BISAC subject headings are more user-friendly than LCSHs for searchers who need to find something quickly but not necessarily in depth. This might confirm the rationale for implementing BISAC subject

headings in a handful of public libraries across the United States in order to satisfy the needs of patrons who have demonstrated browsing preferences (Beezley, 2011).

For both catalogs, the popularity of facets seems to parallel their presentation order in the search interfaces. For the UNC catalog interface, the top-placed facets on the left panel are *availability*, *location* and *format*, which are also the most popular facets for the Phoenix catalog interface. The category presentation order for *books*, *movies*, and *music* also roughly parallels the frequency of the use of those categories. How much facet popularity is a result of presentation order is difficult to discern. While there is likely some biasing towards use of the top listed facets, these choices appear to be fairly consistent. The current order of presentation of the UNC catalog facets was chosen based on their usage during initial catalog use (originally other facets, including subject headings, were higher up). Changing the order of displayed facets to study the effects on facet uptake could be a focus of future studies.

The ‘least effort principle’ applies to the faceted search system. Most search sessions are very brief with only a few actions, the queries users type into the text box are usually two- or three-term words, the advanced search feature rarely is used, and most search strategies are naïve with few query modification iterations. It seems that searchers do not spend unlimited time and effort on a search. They stop if they achieve reasonably good outcomes. Possible explanations for this behavior include: 1) answers are easily found. Most library patrons have a simple and straightforward information need, and the library catalogs appear to do a sufficient job of providing relevant information quickly; 2) novice searchers perform only very simple searches (with easy answers) in library catalogs; and 3) searchers become dissatisfied quickly and may terminate the search because the facets were ill-

specified by the novice searcher. (Many users may be novices as information seekers, who use naïve search strategies. Wang, Barry, & Yang, 2003).

Limitations

Limitations of this study include the inability to determine searchers' intentions from the TLA. Log analysis is limited to some quantitative numbers and figures, which are not sufficient for understanding the whole search story about people's actions and the reasons they perform those actions. A complementary experimental study that recruits subjects to search through the two catalogs was conducted, and combining this with information-seeking behavior captured through screen logging and interviews helps to better understand the ways people search through faceted library catalogs.

Another potential limitation for session-level analysis is the identification of the session boundaries. Without applications to track when sessions begin and end, any session identification method is always an estimate. In addition, the logged data do not capture the requests cached on the local machine or proxy servers. For example, the web server is unable to record the action if a user visits a page using the *Back* or *Forward* button. Therefore, in this study, a few actions within a session were missed when analyzing the transaction logs.

Finally, cluster analysis is subjective in terms of determining the number of clusters, i.e., deriving characteristic variables, etc. The results of cluster analysis are limited to serving as an exploratory method to reveal the big picture of the whole data. In order to know the details of the search process and ways people actually use facets, other methods, such as pattern identification and automatic visualization, are used for this study.

4.2 User Experiment 1: Faceted Search vs. Non-faceted Search

Although the TLA provided an overview of search patterns and facet usage in this study, it also raised questions, such as: What is the underlying motivation of a particular action? Why does a person spend so much time on an action? Is it because he/she was thinking about what to do next, or because he/she was answering a phone call? Did the person find the items he/she wanted? What are the searcher's feelings and perceptions about faceted library catalogs? With so many questions, the author decided to conduct user experiments to complement the TLA. The main purpose of the first experiment is to compare search performance between faceted and non-faceted library catalog interfaces, as well as to understand how facets are used in the UNC catalog's faceted interface.

4.2.1 Characteristics of the Subjects

In this experiment, 24 subjects were recruited and all completed the study in its entirety. The average age of the 24 subjects was 19.4. As for gender, 11 were male and 13 were female.

Their online searching experience was assessed by the Entry Questionnaire. Overall, subjects had a mean of 8.6 years of searching experience. All subjects had at least 5 years of searching experience and one had 13 years. Subjects also were asked about their searching experience with general online library catalogs, with the UNC library catalog, and with using categories (facets) to find information. These experiences were self-rated on a 5-point scale, where 1 = very little and 5 = very much. Figure 4.7 displays the distribution of the ratings.

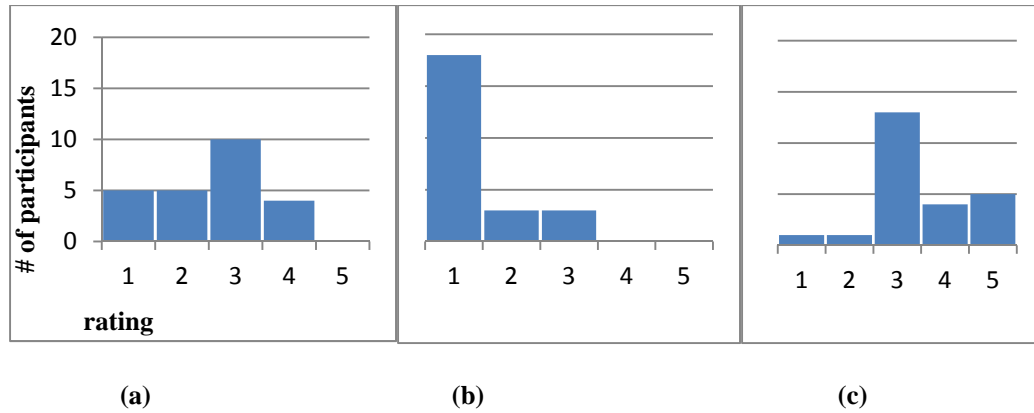


Figure 4.7 Subjects' experience with (a) online library catalogs, (b) the UNC library catalog, and (c) searching with categories

As shown in Figure 4.7, most people rated themselves at the mid-level of experience for the general online library catalogs and for using categories to find information. For the UNC library catalog, however, most people (18 out of 24) indicated that they had little experience with it. This outcome was as expected, i.e., that most searchers would not have much experience with the UNC catalog and therefore showed little possible bias towards it.

4.2.2 Search Time

In comparing search performance between the non-faceted and the faceted interfaces, search time is used as one measure of search performance. Overall, most participants could finish a search task in several minutes. The average time spent on a task was 138.8 seconds (2.3 minutes). The longest search task took 1,178 seconds (19.6 minutes), and the shortest task took 24 seconds. The search time was observed to be affected by search interfaces, task types, and task complexity.

Factors affecting search time: interface, task type, and level of task complexity

The primary research question for Experiment 1 is whether searchers could perform searches more quickly using the faceted interface than using the non-faceted search interface.

In addition to search interface, task type and task complexity are two other major factors that affect search time. Figure 4.8 presents the three comparisons of search time in terms of these three factors.

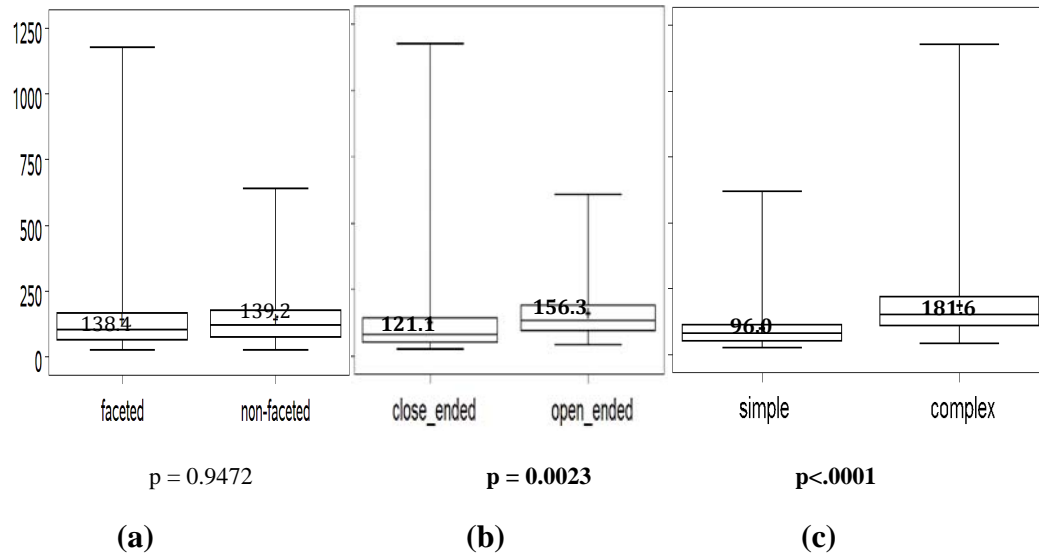


Figure 4.8 Comparison of search time results for (a) interface, (b) task type, and (c) level of task complexity

As shown in Figure 4.8, although search times are shorter with the faceted interface than with the non-faceted interface, the difference is not as significant as expected. This finding differs from most previous faceted OPAC usability research that suggests faceted search improves search speed (for example, Antelman et al., 2006). In this study, the search times of searchers using the non-faceted interface are not much slower than those for the faceted interface, probably because of the availability of the advanced search feature, which contains a full-fledged list of search limits. Another reason for the discrepancy between this study's results and those of previous research might be the fewer choices found in the non-faceted interface compared to the faceted interface where users might be overwhelmed by a large number of choices. During the interviews, several participants said they felt the facet

feature was able to get them information more quickly than with the non-faceted feature. But the quickness comparison turned out not to be significant. Rather, it appears that the participants were able to search reasonably quickly when facets were not available. The wider distribution of search time for the faceted interface compared to the non-faceted interface suggests that people might spend a very long time on a search with facets, confirming the idea that facets might provide support for investigating and exploring.

As to task type and task complexity, a significant difference is found for both. The close-ended tasks took significantly less time than the open-ended tasks. This result is not surprising because the open-ended tasks usually require more effort for exploring and investigating than the close-ended tasks. The multiple target items take longer to find as well. The simple tasks took much less time than the complex tasks. This finding, at the group level, to some extent confirms the feasibility of the pre-defined task complexity based on how fuzzy the task description is. To complement the pre-defined task complexity, user's perceived task complexity also was collected via the post-search questionnaire.

Interaction between interface and task type

Search time might also be affected by interactions between interface, task type, and task complexity. To look for such interactions, a general linear model (GLM) regression analysis was conducted. The results are displayed in Table 4.4. None of the interactions between interface, task type, and task complexity is significant, suggesting that interface, task type, and task complexity are three independent factors in affecting search time.

Table 4.4 GLM regression results for search time

factor	F value	P value
interface	0.01	0.9419
task type	11.15*	0.0009
complexity	66.49*	<.0001
interface*task type	1.62	0.2041
task type*complexity	3.32	0.0693
interface*complexity	2.96	0.0864

*Significant at the 0.05 level.

To better understand the ways that search time is affected by the three primary factors of interface, task type, and task complexity a factorial subgroup comparison was made, and the results are displayed in Table 4.5 as a pairwise comparison both vertically between interfaces or levels of complexity, and horizontally between task types. As shown in three-way comparison table, levels of complexity seems playing a significant role in affecting the search time because three out of four pairwise comparisons between levels of complexity are significant. Complex tasks take significantly longer than simple tasks to finish. Apparently, the faceted interface reduces the search time discrepancy between the two task types and speeds up the open-ended tasks. For the close-ended tasks, searchers were quicker with the non-faceted interface, whereas for the open-ended tasks, this was not necessarily the case. The benefits of using the faceted interface to accelerate searches are notable primarily for the open-ended tasks.

Table 4.5 Factorial result for search time

Task Complexity	Task Type	Close-ended	Open-ended
	Interface		
Simple	Faceted	73.8	99.3
	Non-faceted	64.0	146.8
Complex	Faceted	181.3	199.2
	Non-faceted	165.9	180.0

Note: numbers with connectors are the pairs with significant difference at the 0.05 level

Search time for individual tasks

In addition to examining search time for interface and task type at the aggregate level, search time is also examined at the individual task level to investigate the effects of interface on a particular task. Figure 4.9 displays the search time for each task for both interfaces. The two curves are roughly parallel, suggesting that search time across the two search interfaces is consistent. It appears that the faceted interface works no better than the non-faceted interface for the close-ended tasks (A1-A4 and B1-B4), but does so for the open-ended tasks (A5-A8 and B5-B8), which confirms previous results in Table 4.5. Task A7 (microfinance, open-ended) is an exception, where the non-faceted interface shows quicker search results. If search accuracy is concerned, however, the non-faceted interface is shown to be less accurate than the faceted interface (Figure 4.11). This outcome is likely due to the fact that the availability of facets supports exploring different subtopics and regions, which increases the search time and improves the search accuracy as well. The largest discrepancy (70.1 seconds) between the two interfaces is seen for Task A5 (*non-English movies*). The apparent reason might be the support of deselecting facets to eliminate unwanted information. Otherwise, on the non-faceted interface, searchers had to try a non-English language one at a time or manually scan to find a non-English movie. Either way took a significantly long time.

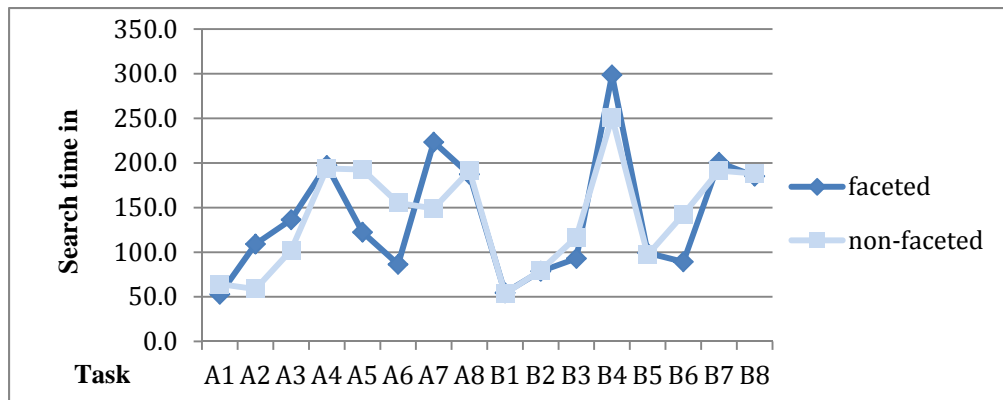


Figure 4.9 Search time comparison between the two interfaces

4.2.3 Accuracy

The second measure of search performance is search accuracy. Overall, most subjects could successfully finish most of the search tasks. Of the 384 tasks performed by the 24 users, 291 (100%) were completely correct and 10 (0.03%) were completely incorrect. The average accuracy was 91.7%.

Factors affecting search accuracy: interface, task type, and level of task complexity

This study also investigates the impact of interface, task type, and task complexity on search accuracy. Figure 4.10 presents the three comparisons in terms of the three major factors.

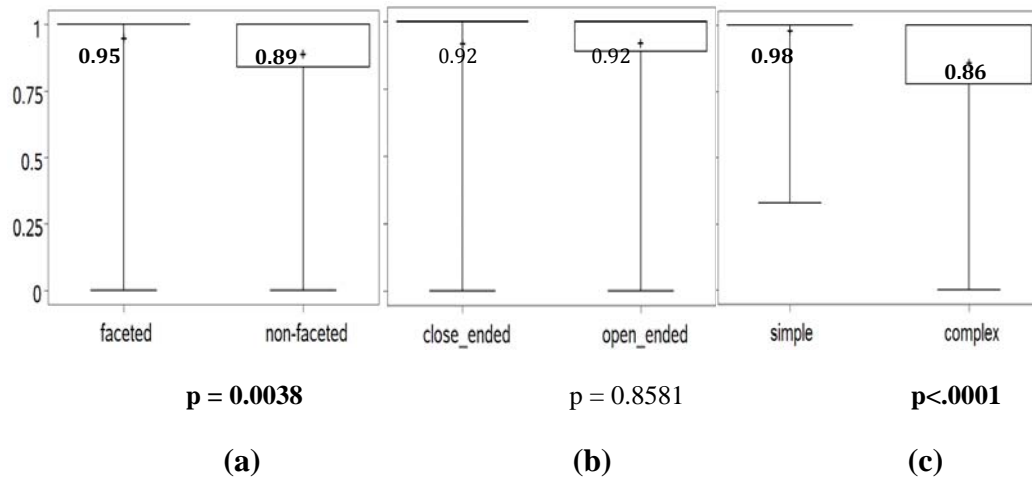


Figure 4.10 Comparison of search accuracy results for (a) interface, (b) task type, and (c) level of task complexity

As shown in Figure 4.10, the faceted interface leads to significantly greater search accuracy than the non-faceted interface. Although participants did not necessarily search more quickly with the faceted interface, they did so more accurately. This outcome is probably due to people spending some time viewing and selecting facets, which could increase the search time. The facet operations did not save search time but offered more

options for searchers to explore and exposed them to more relevant items, therefore possibly improving the search accuracy. Another significant difference is between the simple and the complex tasks. The simple tasks maintained a significantly higher degree of accuracy than the complex ones. Taking into account that the simple tasks also took less time to complete, the study is able to confirm that the study's procedure for task generation leads to well-grounded tasks. The results indicate that the task design succeeds in creating two main groups of tasks: simple and complex.

Interaction between interface and task type

GLM regression analysis was again performed to look for the interaction effects between interface and task type, and the results are displayed in Table 4.6. No statistically significant interaction is found for interface, task type, and task complexity, suggesting that these three factors are independent of each other in affecting search accuracy.

Table 4.6 GLM regression results for search accuracy

factor	F value	P value
interface	9.38*	0.0024
task type	0.04	0.8492
complexity	40.03*	<.0001
interface*task type	0.10	0.7514
task type*complexity	3.69	0.0554
interface*complexity	1.46	0.2270

*Significant at the 0.05 level

A factorial subgroup comparison was made to help understand ways that search accuracy is affected by interface, task complexity, and task type. The results are displayed in Table 4.7. Only one significant difference is evident: the open-ended task is more accurate using the faceted interface than when using the non-faceted interface. This result is also consistent with the search time results; i.e., the advantage of the faceted interface is seen

primarily for the open-ended tasks. The possible reasons for this finding include the availability of facets, which alters the range of information users can access and therefore improves search performance. With facets, searchers are able to search more quickly and more accurately during exploratory tasks.

Table 4.7 Factorial results for search accuracy

Task Complexity	Task Type	Close-ended	Open-ended
	Interface		
Simple	Faceted	99%	100%
	Non-faceted	99%	93%
Complex	Faceted	89%	91%
	Non-faceted	79%	85%

Note: numbers with connectors are the pairs with significant difference at the 0.05 level

Search accuracy for individual tasks

Search accuracy also is examined at the individual task level to understand the interface effects on a particular task. As displayed in Figure 4.11, the approximate parallel of the two curves suggests the accuracy consistency across the two interfaces. Mostly, the faceted accuracy is higher than the non-faceted, except for Task B4 (*travel to Europe*) and B7 (*Mubarak*), both of which are not selected explicitly to match facets provided by the catalog. The biggest discrepancy (31%) between the two interfaces is found for Task A4 (*PHP*), with the faceted interface performing much better than the non-faceted interface. The apparent reason for the discrepancy is that the task contains only a partial word match. Using the facet feature, searchers were able to browse the available authors' names from the authors' list after text querying with 'PHP'. One searcher also duplicated the author's name he/she saw from the author facet to conduct a new text search. Without the facet feature,

however, it was hard for searchers to specify the partial match. Some searchers reordered the results list by author, but that decision turned out not to be effective.

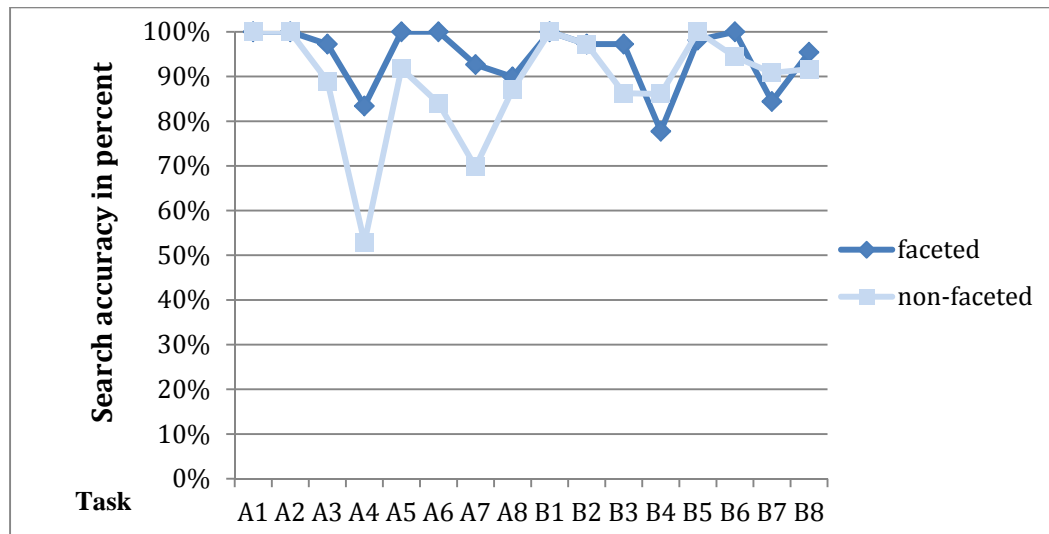


Figure 4.11 Search accuracy comparison between the two interfaces

4.2.4 Users' Perceptions

Perceptions from the post-study questionnaires

In the first experiment, users were asked to rate the two search interfaces based on the degree to which the interfaces are intuitive, enjoyable, easy, and quick. Users rated these variables using a 5-point Likert scale, where 1 = very little and 5 = very much. Figure 4.12 shows the average ratings for the two interfaces. Users rated the faceted interface significantly higher than the non-faceted interface for all four aspects.

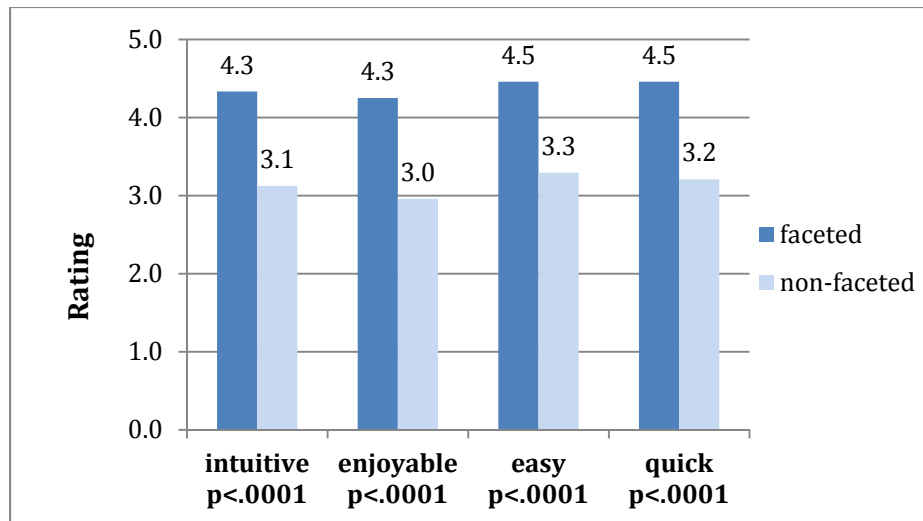


Figure 4.12 Users' ratings for the two interfaces

When asked which interface they liked better, 21 out of 24 users preferred the faceted interface, only 1 preferred the non-faceted interface, and the other 2 said they liked both equally.

Perceived complexity is examined at the individual task level as well. As shown in Figure 4.13, the perception rating for the faceted interface is higher (easier) for most tasks. However, Task A8 (*science fiction*) and B1 (*the catcher in the rye*) were perceived to be much harder with the faceted interface than with the non-faceted interface. For Task A8, the genre requirement (*science fiction*), which is not easily matched to available facets, might make people perceive the task to be difficult. However, from the search time and search accuracy figures, participants performed this task equally well with both interfaces. For Task B1, which is a well-specified task, the advanced search with many specific fields might make the searcher perceive the task to be easier than it would be using facets.

The biggest difference between the two interfaces is seen in Task A3 (*Stephen Hawking*). Probably the subject facets, such as *cosmology* and *astrophysics*, facilitated the

search process. Otherwise, having to type the topic in the search box was unpleasantly difficult for users. Task A6 (*browse movies*) is another task where the perceptions differ greatly. This outcome is not surprising because Task A6 is a less specified task and is better addressed by browsing. In other words, Task A6 is heavily dependent on the facet structure to lead the user down the correct search path. Without facets, determining which keywords might find the results was difficult for users.

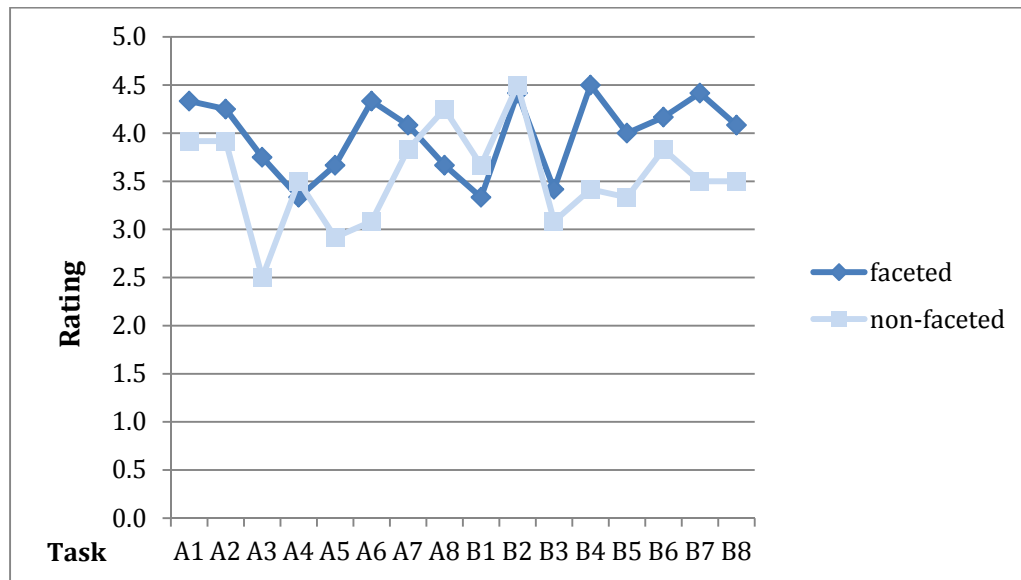


Figure 4.13 Users' ratings of task complexity (Question 1 in the post-search questionnaire)

Perceptions from the post-study interviews

During the interviews, participants were asked to name some tasks for which facets helped significantly and tasks for which facets did not help much. Task B2 (*without you*) and B3 (*Clint Eastwood*) were named most frequently as tasks where facets helped significantly. According to the searchers, the format facet *CD* and the subject facet *World War II* were especially helpful. Both of these facets are considered to be specific 'marginal' information that searchers appended to their text queries.

Tasks B4 (*travel to Europe*) and A8 (*science fiction*) were most frequently listed as examples where facets did not help much, which is confirmed by these tasks' ratings. Facets were not particularly helpful in terms of both partial match and genre specification.

All participants were asked about their perceptions of the two interfaces, and most participants explicitly expressed their preference of the faceted interface over the non-faceted interface.

It (non-faceted) was easier than the first one (faceted).

It (non-faceted) was OK. I preferred the other one (faceted) better.

Not all participants preferred the faceted interface, however. Some participants liked the non-faceted interface or (dis)liked the two equally.

It (faceted) took more effort coz I had to click more options.

It (non-faceted) was... hmm... some of them were easier, some were hard. Yes compared to the first one (faceted), the first sometimes I got too much and this one (non-faceted) was too little.

Most people liked the faceted search interface because it offers facets to help them narrow their search, generate ideas, and save the typing effort. Some participants also mentioned that some degree of learning was involved in order to become familiar with the facet feature and use it effectively.

It (faceted) is really an easier format to use than the first one (non-faceted), it was just all right there and to have selections that are for you so you could like see how specific you can get instead of like the first one (non-faceted) which was more like trial and error.

I liked the second one (faceted) better – only after I learned how to use it, in the beginning it was kind of hard because I was used to the first interface (non-faceted), maybe it was just me, the categories and searching broad was easier.

(On this non-faceted interface), you had to do a lot more manual stuff like typing in, it was sometimes better but sometimes easier to look at something than to type it in and then give you more ideas when you are not exactly sure what you are looking for, if

you want to do politics you can click politics, it gives you subjects then you can search the subjects.

When asked if they searched differently on the two search interfaces, most people said ‘yes’. Only two said they searched the same way.

I guess I would start broader on the second one (faceted) because I could use the ones on the side to narrow it down than on the first one (non-faceted) because on the first one I would try to be as specific as possible because it was more difficult to narrow it down.

(For the non-faceted search), yes I had to be more creative with how I searched things. You have to think a lot more about different ways you can type things so you get more results.

The first thing (non-faceted) is to use a good keyword and try and find as well as using the small parameters that would help, the second one (faceted) it was more just mathematical, trying to pinpoint it down.

These participant comments were coded and listed in terms of search behavior differences. Table 4.8 presents the codes and the number of times each was mentioned by the participants.

Table 4.8 Codes and the number of times each was mentioned

Non-faceted	Count	Faceted	Count
Issue more keywords	3	Rely on subject facets	2
Type everything	5	Start broader, and narrow down later	7
Start over	2	Turn to facets rather than starting over	2
Think a lot more about different ways you can type things (being creative)	1		

Participants also were asked about the features they liked and disliked about both search interfaces. Their responses were grouped into positive and negative responses, and then coded by theme. The comments fall into one of these three major categories: 1) presentation/layout, 2) search interaction, and 3) search results.

The faceted interface elicited approximately 1.5 times as many favorable comments as the non-faceted interface. The majority of the favorable comments for the faceted interface can be found under the search interaction category and concern the feature of the facets that helped participants interact with the systems.

It has the facets on the side so you can keep it narrowing down as much as you need to.

Well, it was just so much easier, it's easier to narrow it down because in the other one, you keep going back and if you are doing everything in this one, you can just keep adding.

Most of the positive comments for the non-faceted interface came from participants to whom this interface was presented first. These participants focused primarily on the advanced search feature, which helped them specify their needs in different fields.

I used the advanced search mostly because that way you could put specific information in.

The unfavorable comments for the faceted and non-faceted interfaces are similar. The most frequent unfavorable comments for the faceted interface include that the facet categories were confusing and not helpful. Two participants mentioned the choice overload the facets had imposed on them. Another participant was concerned that the facets may have excluded the very information he wanted.

Sometimes, it was difficult to find what you were looking for from these categories.

The options, it makes it confusing, I am thinking of too many things, which one should I use, there are too many choices, it can be confusing. It could be me, too many choices in my brain.

I used it to narrow down. It's so quick maybe you exclude something you want.

The most negative comments for the non-faceted interface include that it lacked the ability to modify or refine the search and the user had to start over with the search. Several participants mentioned they would prefer to click rather than type during the search process.

I just don't like how you would have to go back into the form and submit something if you didn't narrow it down the first and you need to start over.

I don't really like to type too much. When you want to add something to your search, you have to search again, restart. Enter everything again.

4.2.5 The Role Facets Play during the Search

Beyond the statistics of search performance and user perceptions about the faceted catalog, the role facets play during the search is also of interest to this research. That is, this study also seeks to understand ways that people use facets and the ways facets help people search.

Facets vs. task type and level of task complexity

In order to investigate facet usage in different situations, this study also examines the extent to which facets are used for different task types and complexity levels. The extent is represented by the percentage of facet operations (according to Table 4.1) during search sessions.

Overall, for the UNC faceted interface, facet operations account for 46.7% of all operations, which is much higher than is found from the log analysis (8.0%). This is likely because of the experimental tasks were more involved, or more suited to facets (i.e. they were not simple lookups of titles/author combinations like many everyday library catalog queries). It might also be due to the training the subjects received before search. Facets were used most frequently (61.2%) for Task B6 (*journal & magazine*) and least frequently (20.2%) for Task A2 (*sweet dreams*). The former is an under-specified query and the latter is a well-

specified one. According to Capra et al. (2007), hierarchical navigation tends to help people with less specified queries because the menu labels provide strong semantic cues. Task B6 is very broad, so users had complexity coming up with specific keywords and, therefore, used more facets to narrow their search. Keyword specification was much easier for Task A2, which is well-specified and did not require facets as much as the other tasks.

Figure 4.14 shows the breakdown of the facet operation percentages for different task types and task complexity levels. As displayed in Figure 4.14, users incorporated significantly more facets for open-ended and complex tasks, implying that facets are needed more for open-ended and complex tasks. This finding confirms that of previous research (Kules et al., 2009; Marchionini & Brunk, 2003; Tunkelang, 2009) and indicates that facets are a particularly useful technique for those tasks that need more effort to analyze, explore, and navigate.

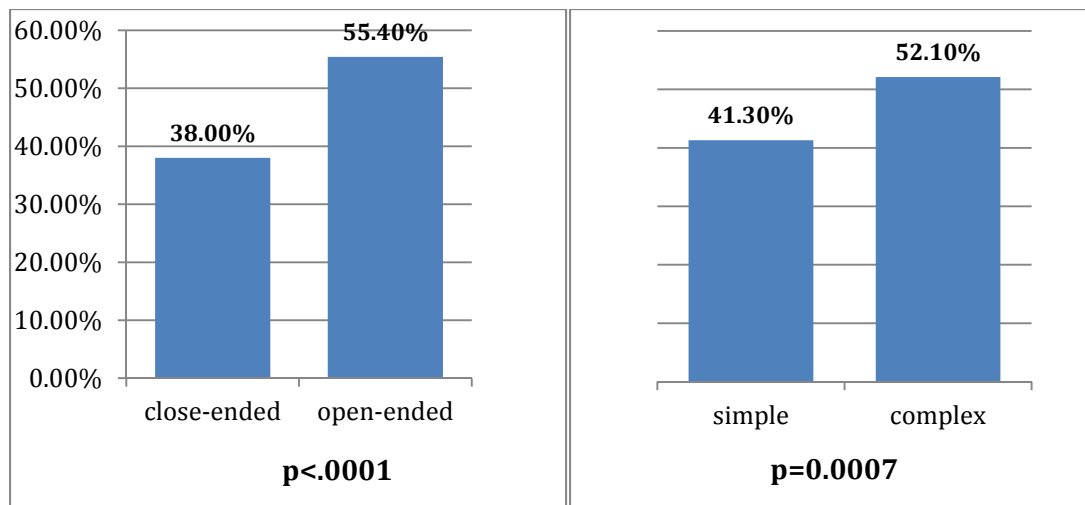
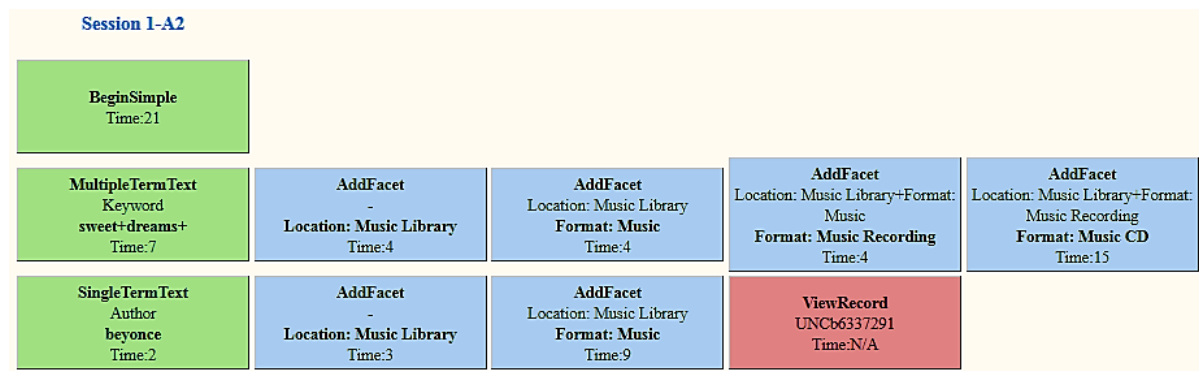


Figure 4.14 Facet usage for task type and level of task complexity

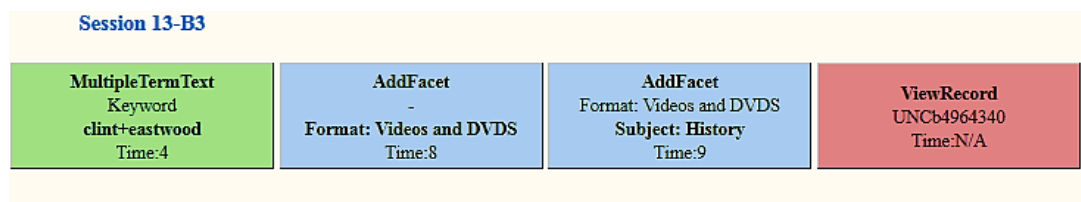
Facet use based on observation

In addition to the quantitative breakdown of facet usage, this study also seeks to understand the search process with facets in a qualitative manner. One way to understand the ways people use facets is to visualize the actions from the search logs, and scan them for patterns. Scanning these logs is an efficient way to investigate how, why, and when people tend to incorporate facets into their search activity. We developed and used the tool VUTL to study the logs. In looking through all the visualized search sessions of experiment 1, five primary roles emerged for facets to play during the search. The most common reason is to retrieve specific information to narrow the search. Such specific information is usually non-core or marginal information that remains after an initial search, for example, format and language. The searchers usually start by entering some core information, such as *title* and *author*, in the search box, and then adding facets as a way of appending some aspects to the text query. Typically, faceted operations immediately follow the text search, indicating that most searchers incorporate facets at an early stage (immediately after the initial text query).

Figure 4.15 shows two example VUTL visualizations. In the first example, the searcher performed two search iterations. In each iteration, he/she incorporated the location facet *music library* and the format facet *music* immediately after typing the title *sweet dreams* or the author *Beyoncé*. The facets used in this example are specific pieces of information that were supplemental compared to the text query. In the second example, the searcher consecutively added the format facets *videos and DVDs* and the subject facet *history* to append these facets to the movie actor *Clint Eastwood*.



(a)



(b)

Figure 4.15 Examples: Facet for appending specific information

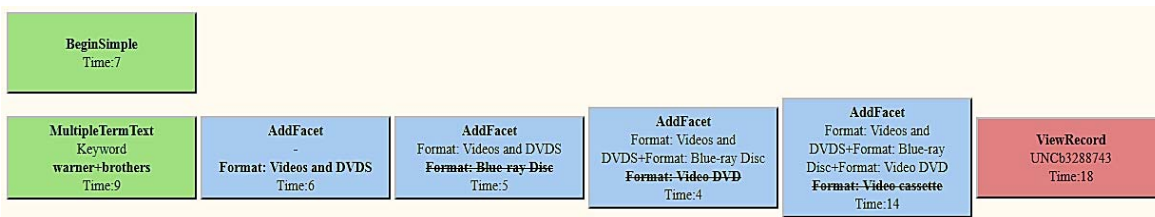
People use facets not only to obtain information they want, but also to eliminate information they do not want. For instance, as shown in Figure 4.16 (a), the searcher deselected English in order to find non-English items. In the second example, the searcher deselected *DVDs* and *blue-ray disk*. Having the excluding selection available was practical because sometimes people do not have a clear idea of the information they need but they can delete information they know they do not need. The excluding selection is especially helpful for people who prefer the bottom-up path to arrive at information by eliminating undesired items.

Session 7-A5 (Partial)



(a)

Session 15-B5 (Partial)



(b)

Figure 4.16 Examples: Facets for eliminating information

Some searchers use facets to explore the collection and to see what information is available. This search method is similar to the ‘organize by overview’ method mentioned by Kules and Shneiderman (2006). Because facets are able to flatten out information that otherwise might be buried deep into the list, some searchers take advantage of this feature to see available choice options when they do not have a particular idea of the information they want.

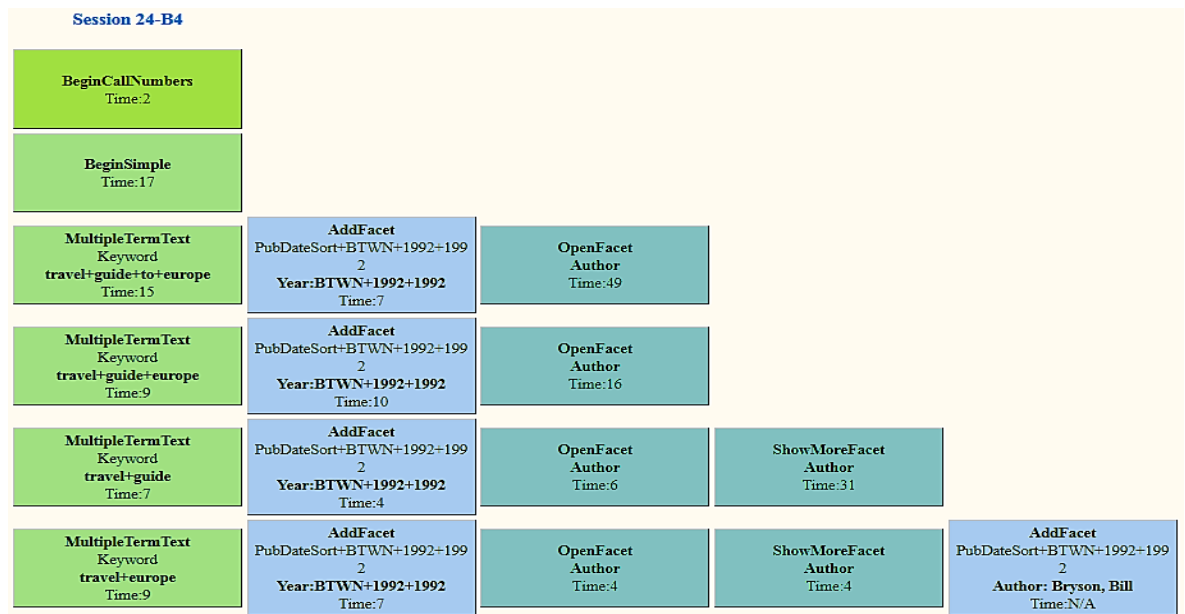
For example, a searcher only knew the book was about PHP programming and that the author’s name started with *Vas*. It was difficult for the searcher to come up with an author keyword because the author’s name was only partially known. Therefore, the searcher typed *PHP*, which pulled up an authors’ list relevant to PHP. From this list, the searcher was able to select a name starting with *Vas*. The second example is similar, where the searcher looked

through the author list to find the author Bill Bryson when he only knew that the author's initials were *B* and *B*. In this case, the searcher submitted quite a few text queries before finding *Bill Bryson*. The author facets generated by the previous queries did not have *Bill Bryson* because these queries were either too narrow or too broad. The chosen lens was not accurate enough to include Bill Bryson as one of the authors. In this case, an artificial decision was involved when designing the faceted categories about the cut-off point for a long list of authors to be presented. In addition to the artificial decision, the chosen lens, i.e., the initial keywords, were very important in determining the search success.

Session 4-A4 (partial)



(a)



(b)

Figure 4.17 Examples: Facets for exploring

Several searchers were observed to use facets as sources of terms for query reformulation, especially for some complex tasks. Searchers were also found to have similar strategies using query suggestions (Kelly et al., 2010). For simple topics, participants were able to think of a variety of queries based on the task scenario or their pre-search topic knowledge. However, for complex tasks, because of a lack of clues or pre-search knowledge, the searchers were able to use the facets generated by the initial query as new query sources for their next text search. Figure 4.18 presents an example where the participant employed a facet value as a new text query. At the third run of the text search, the author keywords *Vaswan* and *Vikram* were the words the searcher duplicated from the author facet during the second run of the text search. This example also confirms findings of previous studies (Kules et al., 2009; Ramdeen & Hemminger, 2012) that users might use facets implicitly by viewing them rather than explicitly clicking on them.

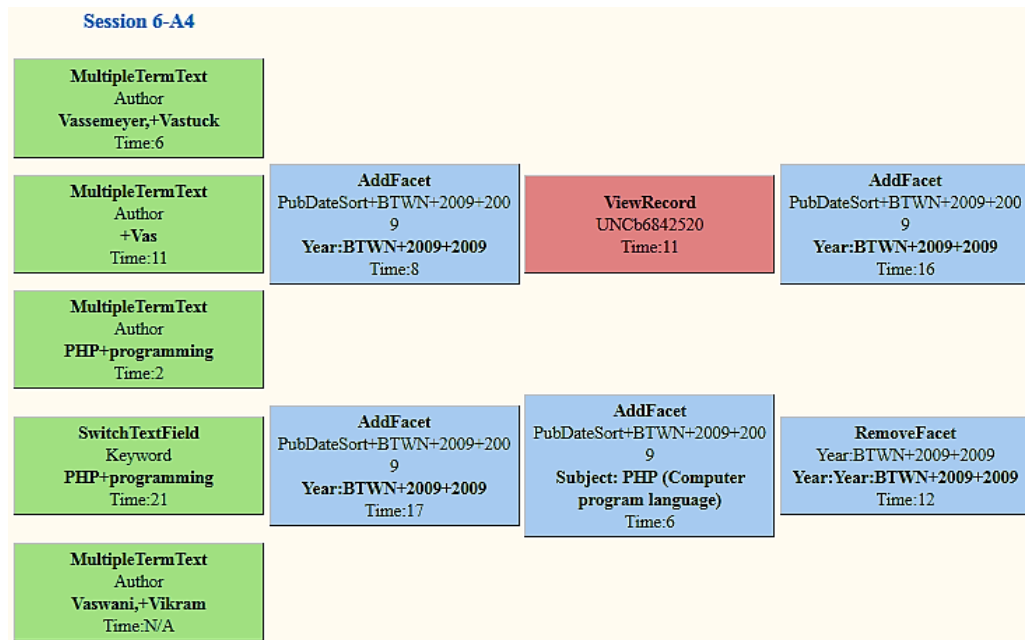


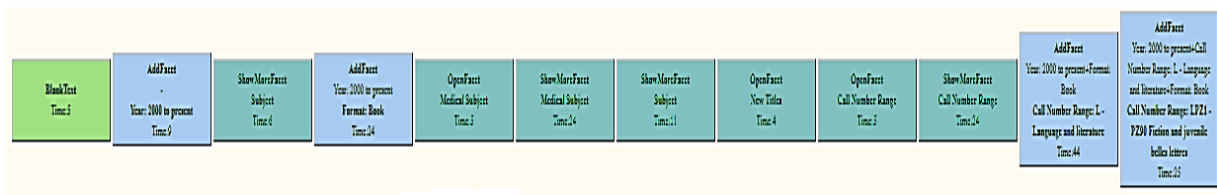
Figure 4.18 Example: Facets for reformulating text queries

Some searchers start a search by clicking on a facet instead of typing a query. Facets provide searchers a browsing feature, which is helpful when searchers have vague information needs or have a ‘cold start’. The browsing structure is even more helpful when the search scenario is broad, under-specified but well-structured, and the structure is easily mapped to the interface’s browsing structure to form a path. Figure 4.19 (a) displays a search session where the searcher is trying to find new titles held by the UNC Media Resources Center. He/she started looking through new titles, then added the location facet *Media Resources Center*, and last added the format facet *Videos and DVDs*. The searcher reached the target items through browsing without making any text queries. In Figure 4.19 (b), the searcher started with a blank text search in order to access the available facets, and then continually conducted many facet operations (*year, format, subject, medical subject, new titles, and call number range*) in an attempt to reach science fiction options. Finally, this search iteration did not work, and the searcher had to re-issue a second run. For the UNC faceted interface, issuing a blank search is an artificial and extra action of viewing the whole collection and accessing the facets. Issuing a blank search is also a starting point for browsing.



(a)

Session 11-A8 (partial)



(b)

Figure 4.19 Examples: Facets for browsing

Facet use based on the post-study interviews

Participant responses are roughly consistent with those observed from the visual logs. Most subjects remarked that they used facets similarly for different tasks.

If I typed in keywords, and the result number is too large, then I refine the search.

Like when I want to find movies by the Warner Brothers, like I excluded like the DVDs and Blue-rays, because you don't have that, I really liked about like you could exclude certain media, I felt a lot easier.

You can also probably find, get the subjects that are similar to it, that can contribute to your research get you an idea you probably hadn't thought of in doing your research.

One participant noted that he used facets because they made the search quicker.

It was more like...intuitive when I would use that one, like, you would get the results out of the second one (sweet dreams) quicker than if I was to type it in, I'm not really sure why.

Several participants said that if they had difficulty or a 'cold start' problem, they would turn to facets.

If I was having trouble finding what I needed otherwise then I would turn to that, like history of education on the side.

4.2.6 Summary

In this section, the author described the results from the first experiment which contrasted the searchers' performance and satisfaction between a non-faceted and a faceted interface. The results primarily address the Research Question #4 and #5 as below:

4. Does a faceted search improve the search performance and user satisfaction?
5. What is the role that facets play during a search?

To sum up, faceted search does not necessarily speed up the search, especially the close-ended search, but is able to improve the search accuracy. Facets are especially helpful for those open-ended and complex tasks. 21 out of 24 subjects liked the faceted interface better. Overall, people rated the faceted interface higher than the non-faceted in terms of intuitiveness, enjoyment, easiness, and quickness.

Based on the observation through the visualized logs, facets primarily played five roles in supporting search: to append specific information; to exclude unwanted information; to see what is available; to get an idea for query terms; and to solve a "cold start" problem by browsing.

4.2.7 Discussion

Interface differences in terms of search performance

While overall no significant differences in search time are found between the faceted and non-faceted interfaces, results related to the open-ended tasks are suggestive that facets help more in these situations because the facets have reduced the search time gap between the closed-ended and open-ended tasks. This finding differs slightly from that of previous studies (Antelman et al., 2006; Ramdeen & Hemminger, 2012) that suggest faceted interfaces significantly reduce the search time. This study did not establish a statistically significant

results, possibly because of differences in the interfaces (for instance the faceted interface did not include advanced search like Ramdeen and Hemminger's study (2012)), and differences in the task scenarios. Also, it was clear that there was little advantage to faceted search for simple, basic exact match lookup tasks; and possibly faceted interfaces perform slightly worse because of being more complex than needed for those tasks. During the post-study interviews, several participants mentioned that they had too much information or too many choices with facets, which might be an associated burden. The associated burden might increase the search time because searchers need to spend some time viewing the facets and choosing one of them. The associated burden is most evident when the tasks are simple and straightforward and the traditional search is able to handle them well. On the other hand, the availability of the advanced search on the non-faceted interface, where people are able to specify their need in many different fields, makes the advanced search option effective for some tasks, further reducing the search time difference compared to that of the faceted search interface.

The difference in search accuracy between interfaces was significant—it is significantly higher with facets than without them. This finding is consistent with the searchers' comments during the interviews, i.e., that facets provide them with some results that they may have otherwise missed. The difference is also somewhat explained by task type in that most of the accuracy improvement occurs for open-ended tasks, i.e., the search accuracy is dramatically different across the two interfaces for open-ended tasks. This finding once again suggests that facets are more supportive for open-ended tasks. For the close-ended tasks, most of which are straightforward fact-finding tasks, the advantage of the facets is not as significant as for the open-ended ones where searchers need to analyze,

explore, and integrate. This finding is in line with that of previous research efforts that indicate faceted searches and categorized overviews are especially effective in supporting complex information-seeking tasks, such as exploratory, open-ended tasks (Kules et al., 2009; Marchionini & Brunk, 2003; Tunkelang, 2009).

Facets for complex tasks

In addition to the open-ended tasks, facets are used also for complex tasks. Complex tasks are defined as fuzzy tasks in this study. They are the tasks for which the information need is not well-specified. The results suggest that facets are able to provide special assistance for complex tasks. For example, for Task B3 (*Clint Eastwood*), which provides a synopsis of a movie, specifying a keyword from the synopsis is difficult. But the searcher was able to select one of the subject facet labels that matched the synopsis. In this way, not only were facets helping to narrow the search, but they were also able to alter the range and the extent to which searchers might interact with the catalog.

Searchers' preferences

Most subjects preferred the faceted interface over the non-faceted interface. They described the faceted interface as “*fantastic*”, “*easier format*”, and “*get specific as much as possible*”. This preference of interface is consistent with that of previous usability research into faceted library catalogs (for example, Ramdeen & Hemminger, 2012; Antelman, Lynema, & Pace, 2006). For the study participants, the faceted search provided a natural and easily understood interface with almost no rules for searching. Any keyword or no keyword would generate results instantly. Furthermore, the use of facets seems to rationalize the two conflicted purposes of finding something quickly or finding something using complex criteria (Collins, Samples, Pennell, & Goldsmith, 2007). But not all the participants liked the

faceted interface; some searchers still preferred the non-faceted interface, because “*there was not much to figure out*” before the search, which is a comment that relates to the ‘information overload’ that facets might bring to the searchers.

Differences in search behavior

Some behavioral differences are evident across the two interfaces. Participants commented on several interesting effects that the facets had on their searching behavior. They confirmed expectations that they would change their searching behavior based on the availability of facets. This finding also agrees with Kules’ (2006) finding that the facet feature allows searchers to draw on new search tactics for reducing effort and improving outcomes. Some searches started as broad and were narrowed later to make the search iteratively closer to the target documents. This finding is similar to O’Day and Jeffries’ (1993) concept of ‘orienteering’ which is a post-query navigation. In reality, the results from the initial query were not sufficiently relevant, and participants needed the post-query interaction to reach the complete set of information they needed. Most searchers explicitly expressed their preference toward this iterative process over the non-faceted interface that either led to success or required starting over.

Several participants said they would issue fewer and shorter queries with the faceted interface compared to the non-faceted interface. This outcome is confirmed by the log analysis, especially for open-ended tasks. The facets are supposed to provide searchers with cues, similar to the notion of ‘information scent’ (Kules & Shneiderman, 2008; Pirolli & Card, 1999), which encourages searchers to click on an option rather than modify their query. On the non-faceted interface, however, the searchers had to think of different ways to enter

keywords. They had to go through a trial and error process and therefore make more runs of text search before reaching the target.

Subjects were observed to use subject facets constantly and several subjects explicitly expressed their preference for subjects, especially for some topics, such as microfinance and Egyptian politics. Searchers preferred to use the facets to explore topics rather than put the topics as keywords in the search box, because according to one participant, the subject facet reflected the catalogers' view of the hierarchy and also the catalogers' vocabulary, and thus was more likely to obtain good results.

Facets also affect the ways people think during the search. Some searchers used the numbers next to the facet value to gain an idea of how specific or how broad the next search would be. They were able to use the facets to understand the distribution of the items across categories. For example, for Task A7 (*microfinance*), when searchers refined their search using region, for example Asian or African countries, they would pay attention to the number next to the country name to decide whether they could stop searching. Other searchers used categories when they felt 'stuck' or when they had a cold start problem. Still other searchers used facets to eliminate the items they did not want.

Subject headings vs. LC classification numbers

One of the interesting implementations of the UNC faceted catalog is that it reveals two types of topic facets to improve topical searching. One is the LCSHs, and the other is the LC classification numbers. The former organizes the subject headings in a decoupled flat structure and reveals the most relevant ones according to the current text query. LCSHs update after each text search. They maintain the first four levels of LC classification hierarchy and present the relevant root and leaf categories of this four-level hierarchy to

searchers. According to Chan (2001), subject headings and classification numbers often operate in isolation from each other. The UNC faceted catalog serves as an experiment in presenting users with these two different, but complementary approaches.

During Experiment 1, subjects were observed to use the subject headings frequently. They also expressed their preference for these headings that helped them explore different topics. For example, for Task A3 (*Stephen Hawking*), they added *cosmology*, *astrophysics* and *big bang theory* from the subject heading facet to fine-tune their search after entering the author's name as the text query. In contrast, fewer people incorporated LC classification numbers during their search. Most people did not navigate the LC classification hierarchy to find *science* and then *astronomy* and last *cosmology*. This finding is different from that of Antelman et al. (2006) that LC classifications are the most heavily used facets. The reason for this discrepancy may be that the UNC catalog truncated its hierarchy at the fifth level and only kept the first four levels. Without the whole structure, searchers may find difficulty in locating specific information. Another reason is related to the well-known discussion about hierarchical and flat structures to present facet values. Several searchers explicitly expressed their dislike of the hierarchy because they risked ending up with no results after scanning through the hierarchy, and because the cataloger's way of organizing items might be different from the searcher's own view. The subject headings are just flat-structured labels assigned to current items. Compared to the vertically structured LC classification numbers, LCHSs are easier to navigate and use.

Finding ways to present facets that have a large number of values has always been a point of discussion. A flat structure and or a hierarchical structure are the two primary choices, each of which has advantages and disadvantages. In the UNC catalog, although the

LC classification has a meaningful hierarchy, it might take more effort on the part of the user to make sense of it than to find value from it. In the UNC catalog, a hard-to-understand hierarchy is probably worse than having no hierarchy at all.

Task efficacy

The 16 tasks used for the searches were intended to be of two types (close-ended and open-ended) and two complexity levels (simple and complex). Although at the group level, the open-ended tasks took longer to finish than the close-ended tasks, and the complex tasks were perceived as more complex than the simple ones, at the individual task level, these distinctions are not as clear. This phenomenon is due in part to the fact that information fuzziness might not be rigorous enough to take into account the task complexity. For example, the partial match requirement (e.g., last name starts with *Vas*) was perceived to be operationally harder than the topic requirement (e.g., the topic should be *microfinance*). The lack of a clear definition for task complexity has hindered the construction of the topics due to the lack of guidance criteria in the field. Participants varied in their interpretations of the topics, and some searchers had previous knowledge that made them perceive a task to be easy. In terms of task efficacy, the degree to which tasks depend on the interface, and to what degree they depend on individual differences, is difficult to discern. With hindsight, the tasks were controlled at the aggregate level. In addition, user-reported data also were collected to complement the pre-defined data.

False positives & negatives

As mentioned earlier, some researchers doubt the validity of using satisfaction and perception as evaluation measures for search systems. A phenomenon was described by Hildreth (2001) as *false positive* where users express satisfaction with poor search results.

During this study, not only false positives but also false negatives were noted. The searchers' comments and perceptions were not always consistent with their actions. Some searchers expressed their preference for the faceted interface but they did not perform better with the faceted interface. The tasks they described as frustrating or difficult were not necessarily completed poorly, whereas the tasks they described as easy were not necessarily completed well. For example, Tasks B4 and B7 were perceived to be easier with the faceted interface but they took a long time to complete and were not necessarily accomplished accurately.

This issue of false positives and false negatives raises concerns about the validity of the users' ratings and semi-structured interviews. The respondents may have imperfect recall. Given time to reflect on a particular action, they might try to make sense of their behavior. Their explanation for their behavior might be very different from what they actually felt at the time.

Limitations

The study aims at understanding the search behavior of average searchers. It is limited by the fact that all the participants were students from UNC-CH. Although most of these students were freshmen with little experience and bias with the UNC faceted library catalog, they all were 'digital natives' or 'Google generation'. They may not be representative of other searchers.

The study is also limited by constraints imposed on the open-ended tasks. Only three target items were required for each open-ended task in order to measure the search time. This constraint reduced the exploratory nature of the open-ended tasks and made it difficult to capture the quality of the search in terms of recall, exploration paths, search diversity, and search depth.

Although the training time was sufficient for participants to learn to use facets and practice tasks, it took time for them to reflect on the facets. At the beginning of the experiment, the participants were often still in the process of learning and adjusting. Some participants remarked that they were getting used to the facets as the search process went on. Generally, the later tasks were conducted more smoothly than the early ones.

Another limitation is that it was hard to tell which facets were used frequently through the user experiment because the artificial search tasks might have affected which facets were used. For example, availability did not factor in to the tasks. But from the logs harvested from the general public, the availability facet was found to be a commonly used category. Individual differences also affected the way the search task was approached and solved. These factors probably contributed to the lack of objective differences in the two interfaces.

The experiment also revealed some interface design bugs; for example, at some point, adding a facet value would result in a blank screen or lose the text query, both of which were unintended events in the design. But these bugs seemed minor and tolerable by most of the participants.

4.3 User Experiment 2: UNC catalog facets vs. Phoenix catalog facets

The main purpose of the second experiment is to investigate ways different facet implementations affect the ways people search, particularly the ways people use facets to search. The UNC Libraries faceted catalog and the Phoenix Public Library faceted catalog are chosen as two cases to investigate the ways people use facets to make searches.

4.3.1 Characteristics of the Subjects

Sixteen subjects were recruited and all completed the study in its entirety. The average age of the subjects was 18.3. As for gender, 7 were male and 9 were female. Overall, they had a mean of 8.1 years of searching experience. Three subjects had less than three years of searching experience and four had eleven years. Subjects were asked about their searching experiences with online library catalogs, with the UNC library catalog, and using categories (facets). These experiences were self-rated on a 5-point scale, where 1 = very little and 5 = very much. Figure 4.20 presents the distribution of the ratings. All the subjects were UNC-CH freshmen and they were assumed to have little experience with the Phoenix Public Library catalog.

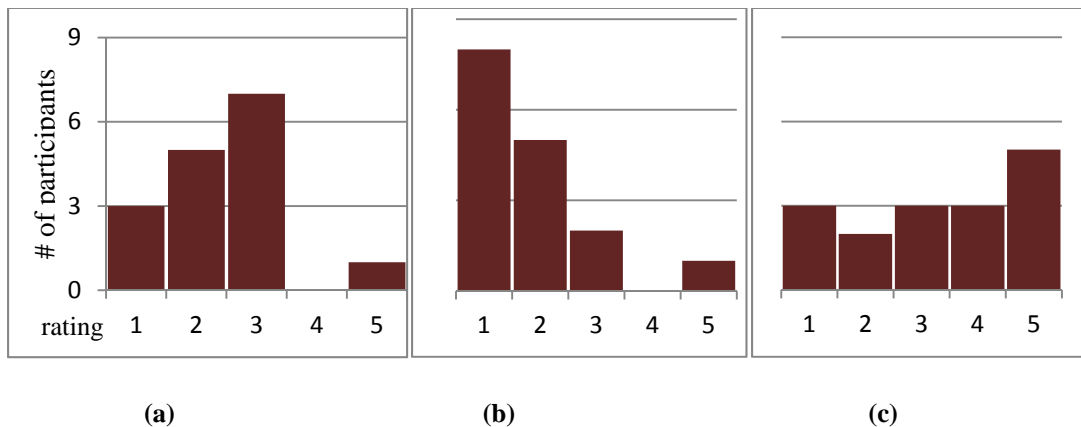


Figure 4.20 Subjects' experience with (a) online library catalogs, (b) the UNC library catalog, and (c) categories

As shown in Figure 4.20, most participants exhibited mid-level experience with the online library catalogs, but most people (13 out of 16) indicated that they had limited experience with the UNC catalog, as was expected. As to experience using categories, participants were nearly evenly distributed over the different experience levels.

4.3.2 Search Time

The primary goal of the second experiment in this study is to understand the impact of different facet implementations on the ways people interact, using quantitative methods. The focus of this experiment is not so much on a comparison of quantitative search performance characteristics, because too many confounding factors for the two library catalogs exist that might reduce the statistical power. Rather, this experiment focuses on general statistics and an overall contrast of how searches were performed.

As such, most participants finished a search task in several minutes. The average time for a task is 162.1 seconds (2.7 minutes). The longest search task took 953 seconds (15.9 minutes), and the shortest took 29 seconds.

Factors affecting search time: interface, task type, and level of task complexity

Figure 4.21 shows the search time broken down by the categories of interface, task type, and task complexity. While it does not make sense to directly compare search results between the two categories because of factors like different interfaces, different content materials, and (in real life) potentially different user audiences, we will try to understand some of the qualitative differences borne out by the quantitative measures. Overall, searchers took a significantly longer time with the Phoenix catalog than with the UNC catalog. The time distribution is more widely spread for the Phoenix catalog as well. Different factors might contribute to these differences between the two catalogs, and facets are probably one of them. According to some participants' comments, the Phoenix interface facets, especially the subject facets, are not as accessible as those for the UNC interface. In addition, the Phoenix catalog contains two sets of facets, two sets of subject headings, and more added features, which together may provide more visual complexity than is found in the UNC

catalog. The participants probably needed extra time to understand these facets and features prior to initiating their searches. As expected, participants needed more time to complete open-ended and complex tasks.

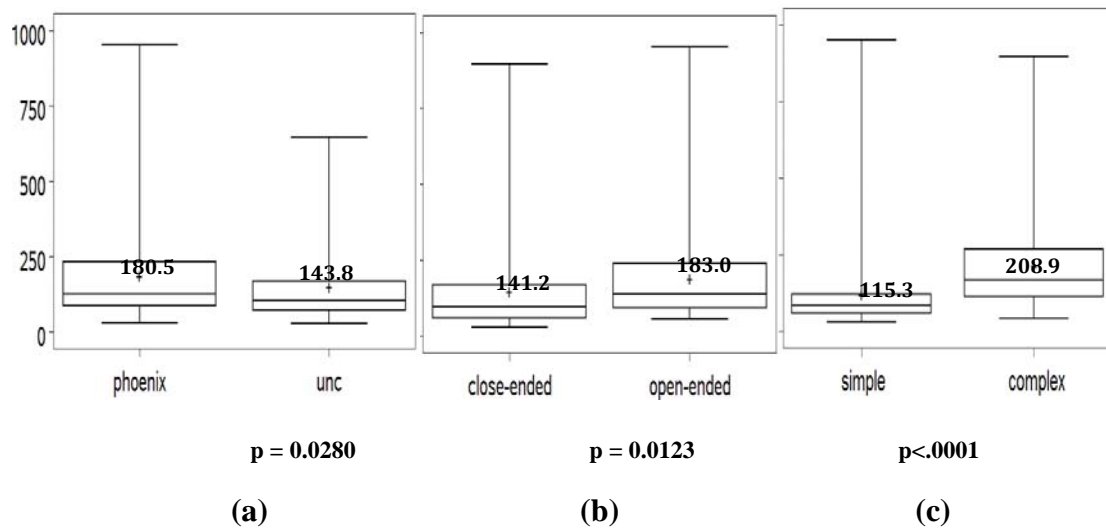


Figure 4.21 Comparison of search time results in terms of (a) interface, (b) task type, and (c) level of task complexity

Interaction between interface and task type

Like the first experiment, GLM regression analysis was performed to look for the interaction effect of interface and task type on search time. The results are displayed in Table 4.9. None of the three interactions among interface, task type, and task complexity is significant, suggesting that the three major factors are independent in affecting search time.

Table 4.9 GLM regression results for search time

factor	F value	P value
interface	5.79*	0.0168
task type	7.49*	0.0066
complexity	37.56*	<.0001
interface*task type	3.31	0.0702
interface*complexity	3.52	0.0617
task type*complexity	0.22	0.6358

*Significant at the 0.05 level

Table 4.10 presents a three-way sub-group comparison. Similar to the result displayed in Table 4.5, levels of task complexity is a significant factor in effecting the search time. The open-ended tasks take significantly longer to complete with the Phoenix interface than with the UNC interface.

Table 4.10 Factorial results for search time

Task Complexity	Task Type	Close-ended	Open-ended
	Interface		
Simple	UNC	84.8	101.8
	Phoenix	75.3	199.3
Complex	UNC	188.6	199.7
	Phoenix	216.1	231.3

Note: numbers with connectors are the pairs with significant difference at the 0.05 level

Search time for individual tasks

The Task C set (first set of tasks for experiment 2) took longer to complete with the Phoenix interface than with the UNC interface, especially for Task C4 (*PHP*) and C5 (*non-English*). For Task C4, the possible explanation might be that this topic on programming language is ‘academic’ and can be found more easily with facet labels and similar books that the academic library of UNC provides. Task C5 saw the biggest difference between

interfaces, probably due to the *non-English* requirement. The Phoenix catalog is not able to exclude from the result set content with matches a given facet value, and thus is unable to eliminate unwanted information like the UNC catalog can. Task C8 (*science fiction*) is the only task in Set C that requires less time with the Phoenix catalog than with the UNC catalog. Not surprisingly, the Phoenix BISAC genre facet that is available to searchers makes the search for science fiction quicker than with the UNC catalog.

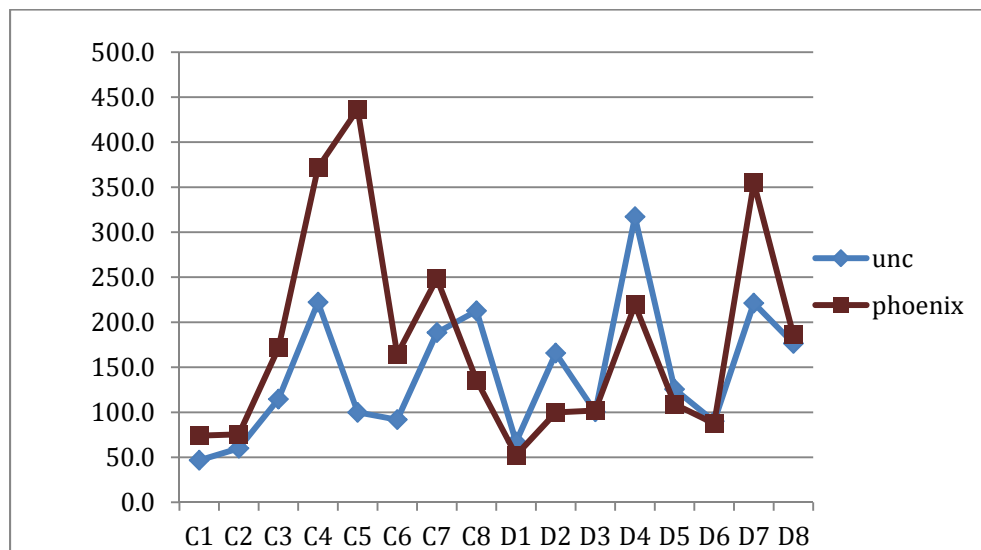


Figure 4.22 Search time comparison between the two interfaces

For most of Task Set D, using the Phoenix catalog is shown to be quicker than using the UNC catalog, especially for Task D4 (*travel to Europe*) and D2 (*without you*). For Task D4, several people were able to take advantage of the browsing structure to find the target item. They started with browsing *books*, and then picked out *travel books* under *non-fiction*, and finally found *Bill Bryson* from the author facet. This path was not anticipated during the task design but is an excellent example of ways that browsing facets can lead down a search

path. This path was observed to be a faster route than taking the traditional text search route with *travel Europe* as the keywords, which is the only possibility in the UNC catalog. Task D2, an entertainment topic, is found more quickly through the public library catalog than through the university library catalog. In addition, one more correct title (*CD*) was found in the Phoenix collection than in the UNC collection, thus increasing its chance of being found. Task D7 (*Mubarak*) is the only task in Set D that takes more time with the Phoenix interface than with the UNC interface. As mentioned before, this task was not designed to match available subject labels particularly closely. Therefore, participants relied heavily on text querying to specify the topic. Participants' actions were similar for both catalogs. Apparently, the Phoenix catalog contained fewer eligible items which were not easily retrieved by a single query submission. This made the search more time-consuming than with the UNC catalog.

4.3.3 Accuracy

Overall, most subjects successfully finished most of the search tasks. Of the 256 tasks performed by the 16 users, 190 were completely correct, and 16 were completely incorrect. The average task performance accuracy was 88.0 percent.

Factors affecting search accuracy: interface, task type, and level of task complexity

Figure 4.23 illustrates a non-significant difference in terms of search accuracy between the UNC and Phoenix catalogs. Taking into account the results in Figure 4.21, however, participants spend longer with the Phoenix interface than with the UNC interface. Although this study does not aim to provide quantitative comparisons during the experimental design phase, it attempts to counterbalance the difference between the two catalogs and intentionally selected balanced sets of tasks. Most of the selected tasks were

found to be challenging, which required effort on the part of the searchers to explore and investigate the options. The Phoenix catalog is a public library that serves the general public, which, according to one searcher's comments, makes it easy to find something quickly, but difficult to find something sophisticated. The facet design, i.e., two sets of facets and two systems of subject headings, also is a contributing factor for accuracy. As for task type, although the close-ended tasks were less time-consuming, they were completed as well as the open-ended tasks. As expected, the simple tasks maintained a higher accuracy rate than the complex tasks.

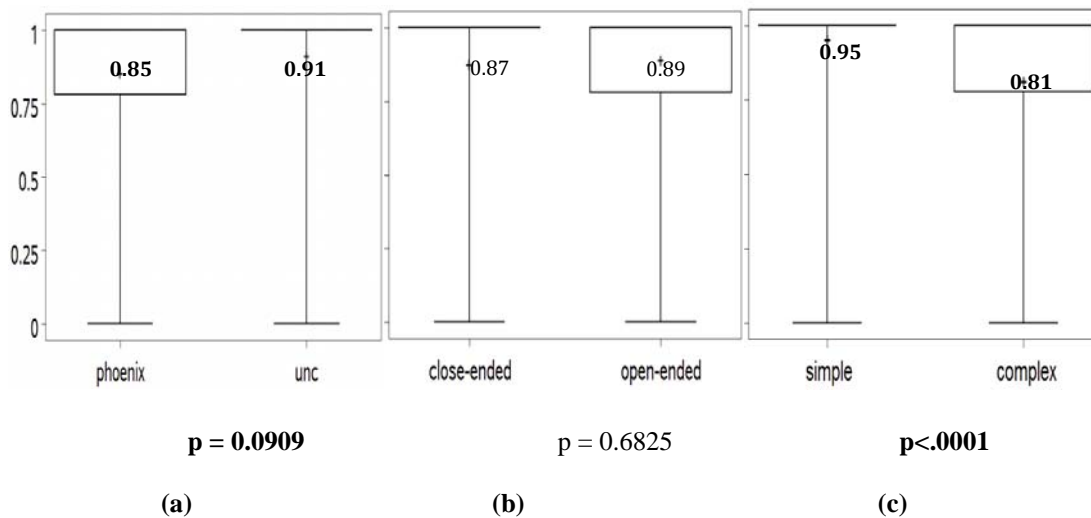


Figure 4.23 Comparison of search accuracy in terms of percentage between (a) interface, (b) task type, and (c) level of task complexity

Interaction between interface and task type

GLM regression analysis was applied to examine the interaction effects between interface and task type. Table 4.11 presents the results. The interaction between interface and complexity is significant, suggesting that the impact of interface on search accuracy is dependent on level of task complexity. Figure 4.24 indicates that the level of task complexity has a greater effect on search accuracy for the Phoenix catalog than for the UNC catalog.

Meanwhile, the interface has a greater impact on search accuracy for the complex tasks than the simple tasks. Therefore, the difference between the two interfaces is seen primarily for the complex tasks that maintained a lower accuracy with the Phoenix catalog than with the UNC catalog.

Table 4.11 GLM regression results for search accuracy

factor	F value	P value
interface	3.13	0.0783
task type	0.18	0.6683
complexity	19.60*	<.0001
interface*task type	1.63	0.2026
interface*complexity	5.17*	0.0238
task	0.06	0.8034
type*complexity		

*Significant at the 0.05 level

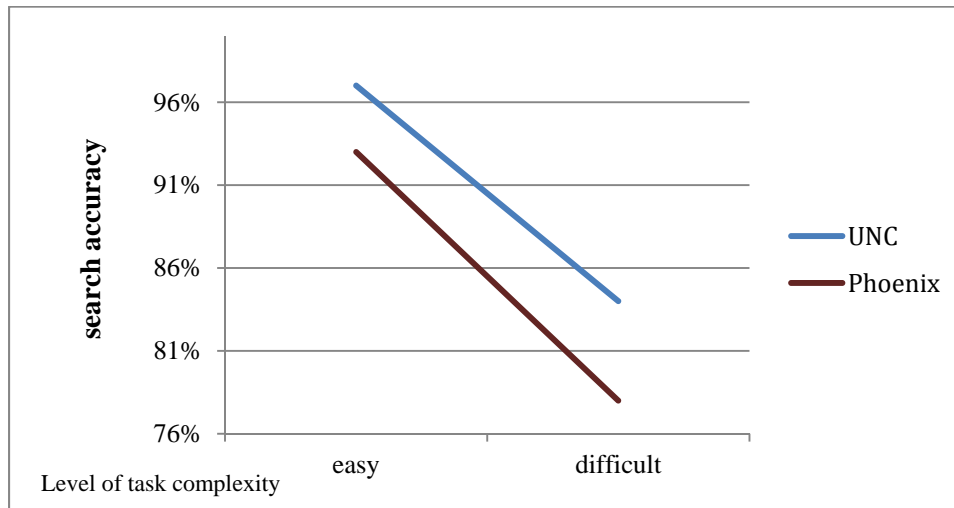


Figure 4.24 Interaction effect of interface and task complexity on search accuracy

Table 4.12 presents the search accuracy results of the three-way sub-group comparison. Similar to the result in Table 4.10 that the open-ended tasks took longer with the

Phoenix catalog than the UNC catalog, they have seen a less accuracy rate with Phoenix than with UNC.

Table 4.12 Factorial results for search accuracy

Task Complexity	Task Type		Close-ended	Open-ended
	Interface			
Simple	UNC		97%	98%
	Phoenix		99%	86%
Complex	UNC		79%	89%
	Phoenix		74%	81%

Note: numbers with connectors are the pairs with significant difference at the 0.05 level

Search accuracy for individual tasks

Figure 4.25 displays the search accuracy at the individual task level across the two interfaces. The two accuracy curves are statistically parallel, suggesting consistency across the two search interfaces. The two curves also demonstrate an inverse trend compared to the search time curves. Overall, the Phoenix catalog exhibits lower accuracy rates than the UNC catalog, except for Task C2 (*sweet dreams*) and D4 (*travel to Europe*). Task C2 is an easy topic for the Phoenix catalog. For Task D4, the accuracy is higher probably for the same reason that the search time is shorter, as mentioned above. The biggest discrepancy occurs in Task C5 (*non-English*), which is also the task that shows the biggest search time difference. Again, the reason is probably the same, that is, the Phoenix catalog offers no way to exclude unwanted items/options.

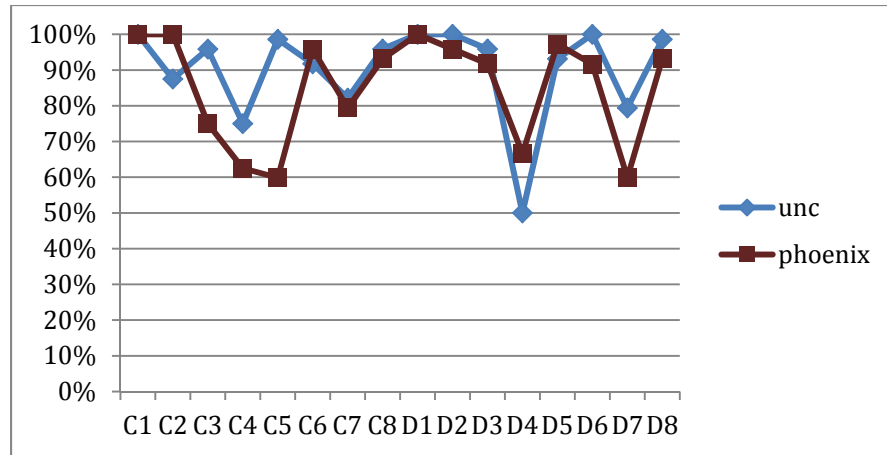


Figure 4.25 Search accuracy comparison between the two interfaces

4.3.4 Users' Perceptions

Perceptions from post-study questionnaires

For Experiment 2, Figure 4.26 shows the average ratings for intuitiveness, enjoyableness, easiness, and quickness. Users rated the UNC interface significantly higher than the Phoenix interface in every aspect except intuitiveness.

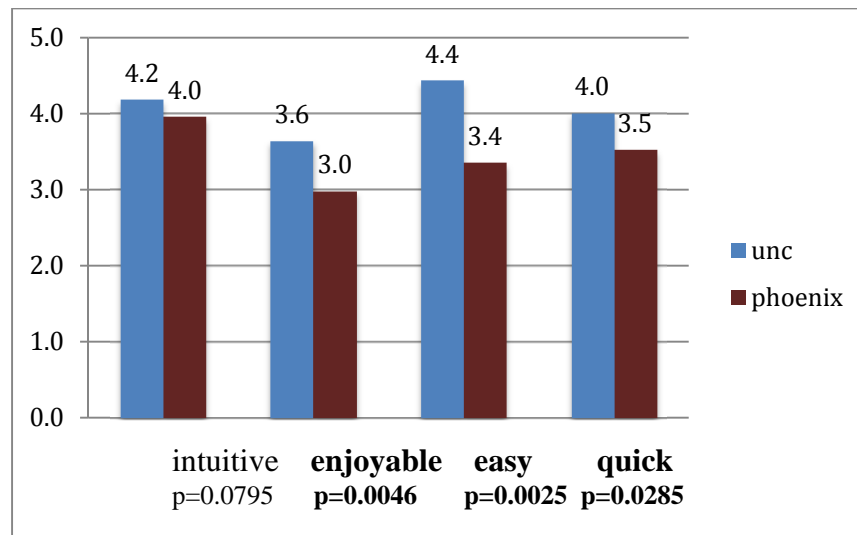


Figure 4.26 Users' ratings for the two interfaces

When asked which interface they prefer, 10 out of 16 participants preferred the UNC library catalog, 4 preferred the Phoenix search interface, and the other 2 said they liked both equally.

The four people who preferred Phoenix all were introduced to the Phoenix interface first and the UNC interface second. The results suggest that people tend to like what they see first. Getting used to the first interface is another possible explanation.

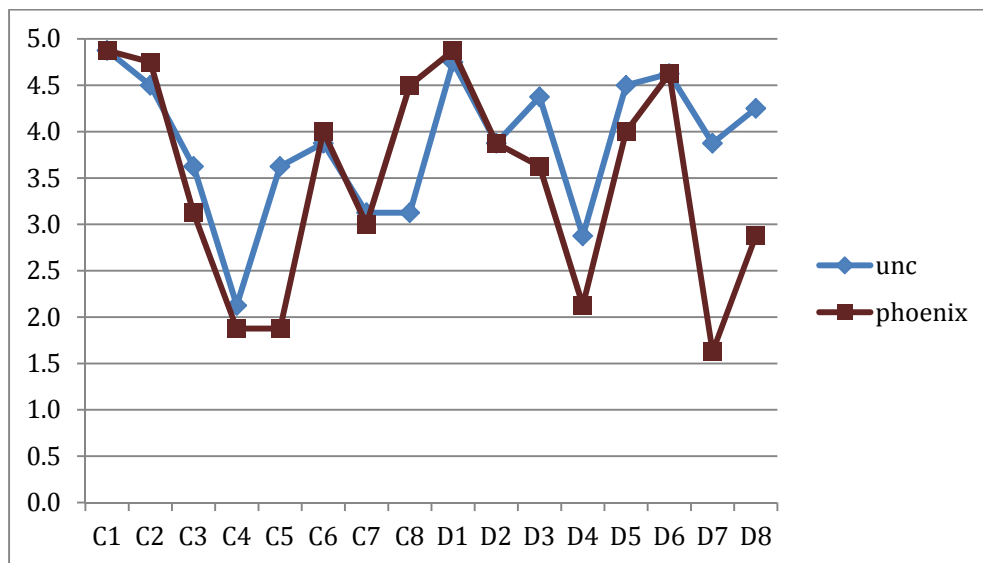


Figure 4.27 Users' ratings of task complexity (Question 1 in the post-search questionnaire)

Like the search time and search accuracy curves, the user rating curves for individual tasks are also consistent across the two search interfaces and seem to match the ups and downs of the accuracy curves. Tasks C5 (*non-English*) and D7 (*Mubarak*) show the biggest discrepancy between the two interfaces. For both Tasks C5 and D7, the reasons are probably the same as those mentioned previously.

Task C8 (*science fiction*) and D8 (*graphic novel*) were designed to be counterparts, but were in fact perceived differently by the searchers on the two interfaces. Task C8 was

perceived as easier with the Phoenix interface probably because of the genre (*science fiction*) requirement, whereas Task D8 was considered easier with the UNC catalog probably due to the subject facet (*fictitious mouse*).

Perceptions from the post-study interviews

Participants also were asked to give an example of a task where facets helped significantly and an example where facets did not help much or even hindered the search. For the UNC catalog, most people named the same tasks as in the first experiment. Tasks A3 (*stephen hawking*) and B6 (*journals & magazines*) were cited as the ones where facets helped the most. Tasks A4 (*PHP*) and A8 (*science fiction*) were listed most frequently as examples where facets did not help much.

For the Phoenix catalog, Tasks C8 (*robots*) and C6 (*adventure movies*) were mentioned most frequently as the ones where facets were the most helpful. Most people chose these two examples because they said the catalog offered the specific facets that could be applied to the tasks, such as *science fiction* under *fiction*, and *adventure* under *popular movie genre*. Browsing to find answers was convenient. Tasks D7 (*Mubarak*) and C3 (*Stephen Hawking*) were named most frequently as examples where facets did not help much. According to searchers, the facets did not hinder the search, but neither did they help to find the information they needed. Both of these tasks are topic tasks that required searchers to explore subjects in order to select appropriate ones. The available subject terms seemed not to help searchers.

All participants were asked about their preference between the two interfaces. Ten out of 16 participants liked the UNC interface better than the Phoenix interface.

(Phoenix) not as easy as the other one (UNC).

This one (Phoenix) was OK. But I like the first one (UNC) better.

A few, however, did not prefer the Phoenix interface over the UNC interface.

It was good. I like the first (Phoenix) one better. It looks better to my eyes. The second thing is I feel the categories are better. It fits what I was looking for.

It (Phoenix) was a little bit faster and quicker (compare to the UNC).

They are both very similar, it's just that, especially for students like me, it's (Phoenix) more appealing.

Most people liked the UNC interface because it offers categories that are more accessible. Several people mentioned they liked the UNC catalog's year facet, which is not available in the Phoenix catalog.

But if you are looking for something deep, this one (Phoenix) probably wasn't the one to choose between the two if had options. It's not as accessible as UNC.

Because this one (Phoenix) was hard to get subject. I mean the topic. It's easier for UNC to get specific when I wanted. Here you couldn't eliminate. This one you couldn't take the multiple things at the same time.

Participants also were asked about what they liked and disliked about both search interfaces. As in the first experiment, their responses were grouped into positive and negative, and then coded by theme. The comments fall into one of these three major categories: 1) presentation/layout, (2) search interaction, and (3) search results.

Most of the positive comments about the Phoenix interface came from participants when they were presented the Phoenix catalog first and had not seen the UNC search interface. The favorable comments were about the facets that helped them narrow the search, which in fact is a common feature for both interfaces. The unique positive opinions about the Phoenix interface fall primarily under the presentation/layout category. Several participants also mentioned that they liked the recommendation feature, for example, 'popular genres' and 'age level'. Some also liked the separation of non-fiction and fiction and the genres

under each. They said the genres were particularly helpful for the science fiction task (Task A8) and the graphic novel task (Task B8).

It (Phoenix) looks better to my eyes. One thing: the age level, fiction, non-fiction were very useful. I guess this is public library vs. a university library.

I didn't really use this, (age level), but I guess it would have been helpful if you are trying to find like an adult book or teen and I think children in some categories so I think this will be very helpful. That's like fiction and non-fiction are like that. I guess just being able to click on the logo and scroll down cause it gives you options, music movies, books, magazines, newspapers so I think it's pretty cool.

The difference like the sub-categories, like when they (interface) asked me to find this science fiction, it was really easy. Coz they (interface) had the category for the science fictions.

The number of negative comments is approximately three times greater than the number of positive comments about the Phoenix Public Library interface. The majority of the unfavorable comments concern the search interaction. Participants remarked that the facets were too fine grained, the year facet was not available, and they could not eliminate facets.

You couldn't really search for the year but you can type in. It is not clear. Like subject region, I don't know what it is about.

It like fictions or juvenile fictions, it's still fictions I feel like. They are separated too much. And you have to think about the books. It might be too much sometimes depending what you are looking for.

Because this one was hard to get subject. I mean the topic. It's easier to get specific when I wanted. Here you couldn't eliminate. This one you couldn't take the multiple things at the same time.

4.3.5 The Role Facets Play during the Search

Facets vs. task type and level of task complexity

Similar to Experiment 1, Experiment 2 focuses on the extent to which facets are used for different task types and different complexity levels. For the Phoenix catalog, the extent is represented by the percentage of facet operations (according to Table 4.1) in the search

sessions. Overall, for the Phoenix interface, facet operations account for 25.2% of all operations, which is much higher than the results (15.8%) from the log analysis for the Phoenix interface. This percentage, however, is lower than the UNC counterpart, suggesting that people use fewer facets than they do with the UNC interface in the experiment setting but use more facets in the ‘wild’ as is shown in the log analysis. Facets were used most (52.9%) for Task D6 (*kids’ music*) and used least (0%) for Task D1 (*the catcher in the rye*). The former is an under-specified browsing query and the latter is a well-specified one. Figure 4.28 shows the breakdown of the facet operation percentages for the different task types and levels of task complexity. Similar to their use of the UNC catalog, participants incorporated significantly more facets for open-ended tasks and complex tasks.

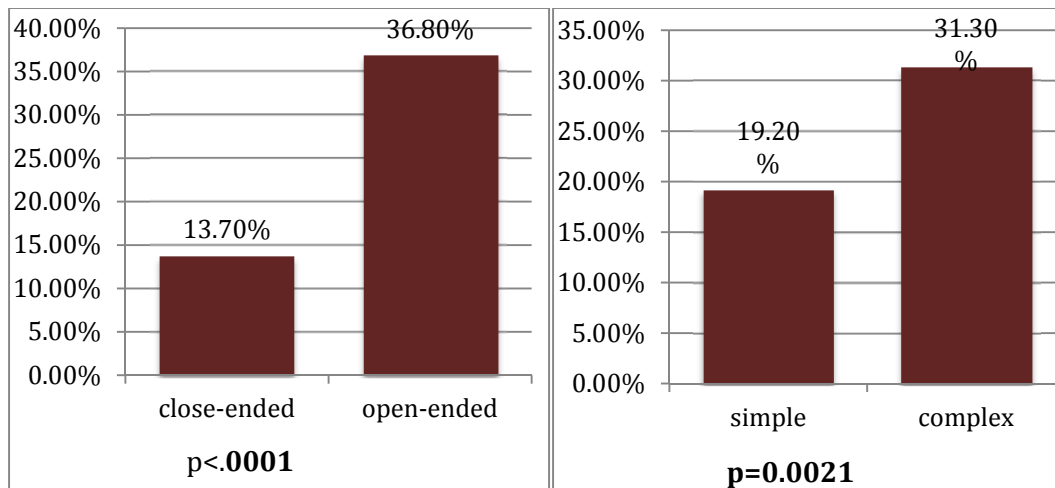


Figure 4.28 Facet usage for task type and level of task complexity (Phoenix only)

Facet use based on observation

To understand more about the roles facets play during the search and the qualitative search process beyond those statistics and numbers, the author scanned all the search sessions through the visualized logs. The qualitative search process that involves how, why, and when

people use facets is also of interest to this study. Scanning through the visualized logs, people use facets with the Phoenix interface for similar reasons as they do with the UNC interface. They append specific information after typing the core information in the search box. They deselect facets to exclude unwanted information. They use facets to explore the collection and investigate available options. They also scan facets to gain an idea of query terms. Finally, they browse facets to initiate a search if they run out of ideas or have a ‘cold start’ problem.

In spite of such commonality, compared to UNC catalog searchers, Phoenix catalog searchers utilize facets slightly differently due to the two catalogs’ different facet implementations. The most evident difference is that participants tended to browse more often in the Phoenix structure, presumably due in part to its better browsing structure, than in the UNC catalog Figure 4.29 presents two examples.

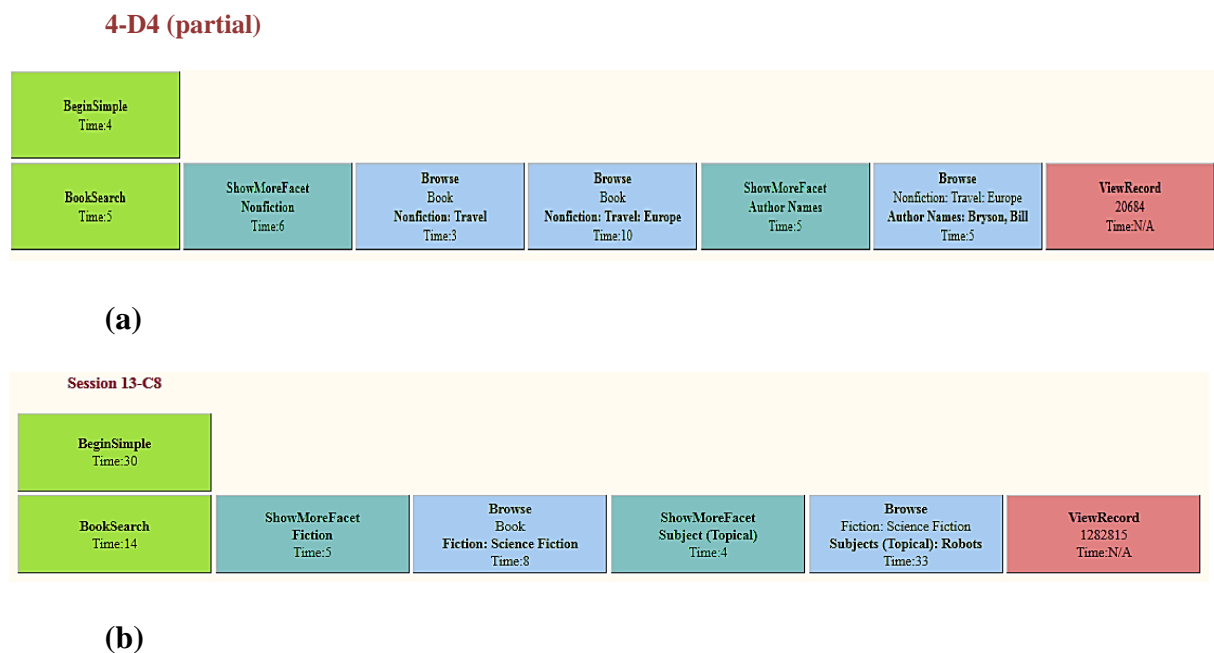


Figure 4.29 Examples: Subjects tended to browse often

As shown in Figure 4.29 (a), the searcher is trying to find a travel book on Europe and the author's initials are *B* and *B*. He/she started by browsing *nonfiction*, then found *travel* under *non-fiction*, and then chose *travel: Europe*. Finally, the searcher was able to find *Bill Bryson* in the authors' list. This example is one where the browsing structure leads the search down the correct path. Compared to the UNC searchers, who would type, as an example, *travel guide to Europe* in the search box, the Phoenix browsing structure was quicker and saved much effort for deriving keywords. In the second example, the searcher was able to narrow the search from *fiction* to *science fiction*, and finally to *robots* to reach the target result set. This path is different from that taken by the UNC searchers, most of whom started by entering *robots* in the search box.

Sometimes, however, browsing might distract searchers and increase the search time, with the searchers ending up with nothing gleaned from browsing. For example, Figure 4.30 presents a search session in which the searcher clicked *political science* under *non-fiction* to find *egyptian politics*, but this way did not work. He/she had to resubmit a new search with the keyword *mubarak*. The searcher thought he/she could browse to find the desired information. Because browsing did not work, the searcher ended up using extra time to obtain no meaningful results. Scanning the visualized logs brought up examples of cases that began as browsing but switched to text searching; however, cases of the reverse (from text searching to browsing) were negligible. This finding suggests that people would like to try browsing first if they saw this feature as a possibility. This preference is probably due to the fact that browsing requires less effort by the searcher than text querying and might return relevant and complete results.

Session 4-D7					
BeginSimple Time:45					
BookSearch Time:14	Browse Book Nonfiction: Political Science Time:17	ShowMoreFacet Subject Region Time:10	ShowMoreFacet Subject Names Time:21	ShowMoreFacet Subject (Topical) Time:15	ShowMoreFacet Nonfiction Time:18
SingleTermText Search+All mubarak Time:9	Refine Book Location: Agave Time:3	Refine Location: Agave Location: Burton Barr Time:8	Refine Location: Burton Barr Subject Region: Egypt Time:10	ViewRecord 31177 Time:8	StartBrowse Time:2

Figure 4.30 Example: Browse might increase search time

Unlike UNC catalog searchers, Phoenix catalog searchers were not able to eliminate unwanted information by excluding records matching a facet value. Typically, Phoenix catalog searchers manually selected the item they wanted one by one. For example, Figure 4.31 illustrates a search session to find non-English movies, where the searcher tried *Spanish*, *Russian*, and *Italian* one at a time. Other searchers chose ten multiple languages (*non-English*) at the same time. Although they made these selections, they commented later that they preferred the excluding selection that the UNC interface provides. This example also shows that the facet implementation affects the ways people interact with the catalog.

11-C5 (partial)

MultipleTermText Search+All paramount+pictures Time:9	ShowMoreFacet Language Time:7	Refine - Language: Spanish Time:7	ViewRecord 430022 Time:20
MultipleTermText Search+All paramount+pictures Time:12	ViewRecord 393523 Time:7		
MultipleTermText Search+All paramount+pictures Time:11	Refine Language: Spanish Language: Russian Time:11	ViewRecord 522827 Time:14	
MultipleTermText Search+All paramount+pictures Time:28	Refine Language: Russian Language: Italian Time:5	ViewRecord 881179 Time:N/A	

Figure 4.31 Example: Subjects tended to try one language at a time

When both BISAC and LCSHs were available, participants tended to use genre facets, such as *fiction* or *non-fiction* from the BISAC subject headings rather than the topic facets from the LCSHs. Figure 4.32 presents an example of people entering the topic (*mice*) in the search box and then choosing *comics and graphic novels* from *juvenile fiction*. This behavior is different from that of UNC catalog searchers, because specifying genre is not easy with the LCSHs. Genre was not present as a separate facet in the UNC catalog in the study, but the some genres did exist within the subject facet (which contained mainly subjects, but some genre categories like “poetry”). Most searchers entered the genre (like “graphic novel”) in the search box and then chose *mice* from the LCSHs.

Session 3-D8 (partial)

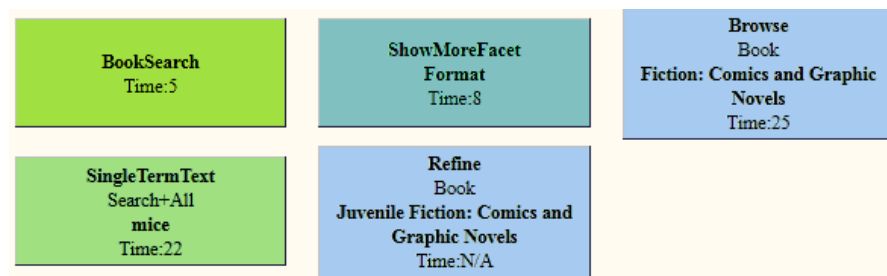


Figure 4.32 Example: Subjects tended to use BISAC subject headings for genre information

Another difference is that participants who searched using the Phoenix catalog tended to issue long queries rather than relying on facets later. Figure 4.33 provides an example where the searcher entered the title, the publisher, and even the publishing year into the search box. Participants tended to enter this much information in part because during the interview they were told they could do so, because the Phoenix search box can interpret the year information. Compared to most UNC searchers, who entered only the title in the search box and used the year facet to narrow their search later, the Phoenix searchers did not take

advantage of the facet feature, but they did save some search time. Whether people preferred this way or whether they had no choice because Phoenix does not offer the year facet is difficult to discern. According to the post-study interviews, some searchers remarked that they wished the Phoenix catalog had the year facet like the UNC catalog.

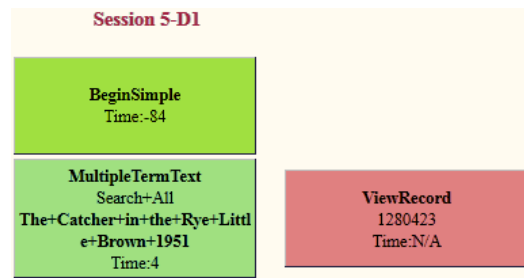


Figure 4.33 Example: Subjects tended to put the year information into the search box

Facet use based on the post-study interviews

One of the questions in the post-study interview asked searchers if they prefer searching or browsing, and the responses are split almost evenly.

I didn't like browsing because I know I was looking for something really specific.

If I didn't have much information to search. I had to use the categories to get what I want.

I definitely used it differently for different tasks. Some of them, I would start from the left side and browse to find what I am looking for, like children's books.

Most participants said they used fewer facets with the Phoenix catalog than with the UNC catalog, either because the Phoenix search box is more user-friendly in terms of being able to interpret more facets, or because the categories are not as helpful as those found in the UNC catalog.

Categories here (Phoenix)? I didn't use them as much...It was kind of easier to just type in the keywords.

It (Phoenix) wasn't as easy as UNC, I used it(UNC) more effectively because it had numbers.

I just, they didn't like there wasn't as many like I noticed on the UNC one, like there was a lot, they would separate even when there's like even two things are similar there would be two different categories. For these ones (Phoenix), they just seemed kind of more general.

Searchers also felt organized when they used the UNC catalog facets. They used words and phrases such as “*step by step*”, “*always available*”, and “*specific*”. On the other hand, they used phrases such as “*random*” “*general*”, “*too much*”, “*might be gone*” in terms of the Phoenix catalog categories.

4.3.6 Summary

The second experiment is aimed to contrast how people utilize facets in the two library environments. Both quantitative and qualitative results were reported and qualitative results were the focus the this experiment and addressed the last two research questions put forward in Section 1.2.

6. Is there any difference in terms of search performance and user satisfaction between the two library catalogs (UNC and Phoenix)?

7. Do people use facets differently in the two faceted catalogs? If so, in what ways?

Overall, there was difference in both search performance and user satisfaction due to different facet implementations and other interface or collection differences. People used facets differently with the Phoenix interface compared to the UNC interface: they tended to browse more with the convenient browsing options offered at any search point; they were unable to eliminate unwanted information because the facet selection did not allow them to; they tended to use BISAC subject headings to begin a search or specify the genre

information; and they were likely to put more in the search box because the search box was able to interpret more.

4.3.7 Discussion

Interface differences in terms of search performance

This study finds that searchers performed differently with the UNC interface and the Phoenix interface. When performance is broken down by task type, the performance difference is seen largely in the open-ended tasks. The research design tried to counterbalance the numerous confounding factors, so the most likely reason for this finding is the difference in facet design between the two catalog interfaces. This result confirms searchers' comments during the interview, i.e., that Phoenix facets "*were not as accessible as UNC*", "*random*", "*general*", and "*always be away*". Some helpful features are unavailable in the Phoenix interface, such as the numbers that indicate how many items belong to a value and the excluding selection that are both found in the UNC interface. In addition, the Phoenix interface seems to reveal an intractable number of facets to searchers, which burdened and confused the participants. For example, the Phoenix interface shows eight facets for topic refinement: *fiction*, *non-fiction*, *juvenile fiction*, *juvenile non-fiction*, *subject names*, *subject region*, *subject era*, and *subjects (topical)*. Most searchers felt confused by these overly separated facets. However, for some tasks that needed genre specification or a browsing path, such as Tasks D4 (*travel to Europe*) and A8 (*science fiction*), the Phoenix interface performs better than the UNC interface. In addition, some searchers favored the Phoenix interface's commercial bookstore flavor and some sophisticated features (e.g., the slide show of book covers and recommendation feature at the top of the results list), which are visually appealing. For the open-ended tasks, most of which need exploration and investigation, the

Phoenix facets, especially the subject facets, did not help users much. This outcome confirms the searchers' comments that the Phoenix interface is easy for finding something for fun but difficult for finding in-depth information.

Facet usage

In this experiment, facet usage was much higher than in the log analysis, which brought up the interesting comparison between the naturalistic setting and the experimental setting. Although the study design tried to make the tasks simulate real-life problems, to some extent these tasks were artificially tailored for the functionality of the facet feature, as conceptualized by the author. This fact probably greatly boosted facet usage. In addition, most of these designed tasks were very challenging and might differ from the information needs of most of the general public. The presence of complex or challenging tasks might have increased the facet uptake as well. Whether experiments are suitable for usability studies is always a matter of ongoing debate. Some researchers indicate the inadequacy of such experiments compared to naturalistic and longitudinal research designs (Tang, 2005; Yuan, 1997).

In this experiment, an interesting finding is that the facet usage is lower with the Phoenix catalog than with the UNC catalog, but the opposite is true for the log analysis. The experimental results are consistent with the searchers' comments that they prefer typing with the Phoenix catalog but they felt the Phoenix facets are not as accessible as those of the UNC catalog. This finding is contrary to the findings derived from log analysis, where the Phoenix facet usage is almost double the UNC facet usage. Assigned tasks, underlying information needs, and searchers' context are different in the two settings. For example, the log data show that the most frequently used facets are the general browsing categories, such as *books* and

movies. But these facets were hardly ever used to complete the experiments' tasks. In a real-life context, many people have very simple and straightforward information needs. They might want to find something quick and fun for entertainment and, therefore, might be easily satisfied simply browsing the available categories without going further.

Browsing

During the experiment, many searchers started their search by facet browse. Some participants found the target items through the browsing path and some ended up switching to text querying. Browsing is a subtle searching activity (Ingwersen & Wormell, 1989) in that during browsing, people scan information items, and each new information scent that is gathered may provide new ideas. For example, for Task D4 (*travel to Europe*), until the searcher clicked *travel* under *non-fiction* he/she had no way to select the region *Europe*. This behavior is similar to what Bates refers to as *berry-picking*, meaning the incremental collection of pieces of information.

However, searchers risked not finding anything after they spent a significant amount of time browsing the hierarchy. That is, browsing might lead searchers astray. Therefore, the situations and task type that searchers choose to start browsing reflect decisions that the searchers must make themselves. A good interface design can lead searchers to the right decisions by organizing the catalog collection in such a way as to be meaningful to searchers and that offers a browsing hierarchy that is easy-to-predict for searchers. For the current catalogs, most searchers based these decisions on their feelings and experience. Usually, if the task is well-structured but under-specified, and is composed of several aspects but lacking specific information to type in as an entry, then searchers tend to start their search with

browsing. For example, for Task D6 (*kid's music*), searchers were able to find three music CDs for children that should be among the new arrivals in the library.

To some extent, searchers were puzzled by browsing and searching during the initial stages of the search. Sometimes, they would specify keywords for a task better solved by browsing, and sometimes they would start browsing for a task suitable for query specification. But, most participants were able to learn from the searching process quickly and switched to the better way. They were rarely observed switching back and forth between browsing and text querying. More searches initiated as browsing but developed into text querying rather than the other way around.

Facet features

For most participants, the Phoenix catalog facets were not as accessible as the UNC facets in assisting them complete the tasks. Based on the author's observations, the reasons for this finding might be: 1) the Phoenix catalog facets are separated too much (for example, there are eight subject facets), which is likely to confuse searchers; 2) some facet names and assignments are difficult to understand, such as the difference between *Subject Names* and *Subject (Topical)* and why *United States* appears in both *Subject Names* and *Subject Region* if *microfinance* is entered in the search box; 3) the Phoenix catalog facets do not have numbers that indicate how many items belong to a particular facet value (without this feature, searchers experience less control over actions they may take and knowing in advance the effect of selecting a particular facet); and 4) some facets disappear unexpectedly after the search parameter updates. For example, *Subject (Topical)* is only allowed to be used once. After the searcher clicks on it, the category disappears in the next stage. Without the

accumulative refining ability, the subject facet is less helpful in narrowing the search. This user experiment suggests that even a lightweight facet feature may be beneficial.

BISAC vs. LCSHs

The Phoenix Public Library has licensed the BISAC subject headings for browsing its holdings. BISAC subject headings are assumed to be more intuitive and user-friendly than LCSHs. For example, BISAC subject headings are used for the signage for the various sections in Barnes & Noble bookstores (Schallan, 2007). However, the Phoenix catalog still maintains the LCSHs for the subject facets for searchers to narrow their topics. Showing the two systems of subject headings on one interface proved to be an interesting combination for this research, and participants were able to distinguish between these two subject headings and used the two differently on demand. For example, if they needed to filter the genre, they would turn to BISAC for *fiction* or *non-fiction*; whereas if they needed to explore topics such as *cosmology* and *small business*, they would use LCSHs. Some people complained that they were confused by so many subject categories. They could perform much better if they could organize the same type of values under one facet.

People are able to adapt

The experimental results suggest that people are able to adapt to different library catalogs through the process of learning and adjusting. The behavioral differences across the two faceted catalogs result from the effects of the catalog implementation. For example, participants would select categories that the opening page offered, so they used the trial and error method one item at one time when an excluding selection was not available, and they entered the year information in the search box if no other way to refine the publication year was available. This adaptability is similar to the finding by Capra et al. (2007) that suggests

that interaction style is affected by information architecture. It is also similar to Kules' (2006) conclusion that searchers are able to draw on new tactics as well as revise old ones without much increase in effort. Although training is a way to assist searchers in becoming familiar with the system, most searchers also demonstrate the ability to self-learn and self-adapt.

Limitations

In addition to the limitations that are applicable to Experiment 1, some other limitations apply for Experiment 2. First and foremost, the Phoenix search interface was not a standalone fixed configuration setup just for the experiment like the UNC interface was. Users used the same live interface as patrons of the Phoenix library. As a result, during the peak usage hours, the library catalog could sometimes encounter a slight delay, but it was not generally noticeable to the participants. The UNC system also suffered occasional delays when its server was overloaded.

Without any manipulation of the interface, participants had to be told explicitly not to use the advanced search feature in order to match the UNC faceted catalog that does not offer the advanced search feature. Beyond this directive, many other confounding factors, such as the underlying collection, the way the search box interprets queries, and the indexing rules, are of interest to this research. The confounding factors greatly affected the power of quantitative comparisons. Thus, the study was constrained to focus on the qualitative contrasts surrounding the search process, search stories, and searchers' perceptions of the two search interfaces. Another possible problem with using the public version is that the search interface could be improved upon over time, and sudden interface changes could occur at any given time. In order to be aware of whether changes to the interface could occur during the

period of the experimental studies, the author contacted the Phoenix Public Library catalog developers in advance to ensure that an interface change was not scheduled during the experiment period. However, the collection was updated from time to time, and some new items were added during the experiment. Fortunately, this change did not significantly affect the search results, or the ways the participants searched for information.

The study also is limited because the search tasks were not perfectly fair for the two different types of library. When designing the search tasks, an attempt was made to balance the task topics and facet selections between the two catalogs to mitigate their differences. For example, an ‘academic’ topic such as *PHP programming* would be counterbalanced by a ‘daily life’ topic such as *comics and graphic novels*. In reality, however, although most tasks were reasonably fair at the aggregate level, some individual tasks, such as *PHP programming* and *egyptian politics* were considered to be more difficult with the Phoenix catalog than *graphic novels* or *popular music* were with the UNC catalog. Balancing the topic complexity absolutely was a difficult task to accomplish due to the catalog differences.

During the experiment 2, some bugs and flaws of the facet and/or catalog implementations were found. For example, some facet values appeared twice under a particular facet group because of typos or special symbols, such as *Africa* and *Africa.*, both of which appear in *Subject Region* just because of the period after *Africa* in the second instance of the word. Another example of a flaw is that the book collection does not contain all books. Searchers would be able to find many more books if he/she could search the overall collection rather than clicking *books* and searching only the book collection. These minor bugs and flaws might have jeopardized the searchers’ trust in the catalogs.

5. Conclusions and Discussion

5.1 Conclusions

This dissertation seeks to understand whether faceted search improves the interactions between searchers and library catalogs and to understand ways that facets are used in different library environments. Interactions under investigation include possible search actions, search performance, and user satisfaction. Two faceted catalogs from two libraries, the University of North Carolina at Chapel Hill (UNC-CH) Library and the Phoenix Public Library, are chosen as two examples of different facet implementations.

The single most significant finding of this dissertation is that facets improve the interactions between searchers and catalogs under some specific situations. Faceted searches demonstrate statistically significant improvements in search accuracy for complex and open-ended tasks. Additionally, facets are used statistically more frequently in complex and open-ended tasks than in simple and close-ended tasks. For less-specified tasks or for tasks where the information need is vague, facets offer searchers options for browsing the collection and allow searchers to overcome the problem of a ‘cold-start’.

Facets also change the way people formulate or reformulate their search. With facets, users tend to issue shorter queries, submit fewer queries, and scan deeper along the result list than without facets. Most people use facets to narrow their search. They use facets to append specific information, to exclude undesired documents, to see the information that might be available, to avoid or delete the need for new text queries, and to browse. People utilize

facets differently based on different facet implementations. That is, they browse more if browsing is well supported, adopt trial and error strategies if the facets cannot be logically combined, and they enter information into the search box that is not supported by facet refinement.

Most searchers are able to understand facets without much training and learning. Faceted search is supplemental compared to text search, which is the dominant mode of searching with most catalogs. Browsing with facets significantly increases their overall use, beyond using facets just for refining (narrowing) search results. Searchers do not necessarily search more quickly with the faceted catalog, but they do so more accurately and with greater satisfaction than with a non-faceted catalog.

Due to the differences in their facet implementations, the UNC and Phoenix catalog facets each exhibit respective advantages and limitations. Compared to the UNC facets, the Phoenix facets are not as accessible for narrowing the search, but they offer searchers a better way for browsing the collection.

Overall, the results paint a detailed picture of the ways people utilize facets and the ways facets benefit searchers. The results of this research have some practical implications for librarians and catalog designers in the form of user interface design guidelines, which take into consideration constraints, capabilities, features, tradeoffs, domain knowledge and human factors.

5.2 Benefits of Facets

As a result of the two study experiments, this research finds that most participants preferred faceted search over non-faceted search and appreciated this feature in helping them narrow their search. Faceted catalogs offer an intuitive library catalog interface for users with

almost no rules for searching. Any term or no term entered will generate instant results. To some extent, faceted catalogs reconcile two seemingly conflicting purposes of searching: to find target information quickly and to find information on a specific topic using complex criteria. In addition to narrowing the search, facets can potentially provide users with an overview of the information collection as well as provide them with easy entry points into different parts of the collection.

Although the study was not able to establish a statistically significant difference in search time between the non-faceted and the faceted interfaces, there was a difference in search accuracy. This result is probably due to the fact that facets are able to reveal more relevant items to searchers and therefore enhance the library catalog as a discovery tool. The literature suggests that most searchers do not examine more than the first page of search results. For a typical list, searchers may only scan ten to twenty results, thereby limiting the possibility of finding valuable information. Facets extend the typical range in which people generally interact with catalogs and present the results in flattened categories, thereby helping people to find deeply buried valuable information.

Based on the actions of the study participants and resulting data, this study shows that searchers use facets more for open-ended and complex tasks than for close-ended and simple tasks because facets provide post-search refinements to explore the information space. This aspect of facets particularly benefits searchers who are familiar with web search engines and who do not plan an elaborate search process before initiating a search, but prefer simply to enter a few terms (the initial result set) and refine their search ‘accumulatively’ (Novotny, 2004).

Another major benefit of facets is the incorporation of browsing capabilities into the text search. The experimental results from this study (i.e., results from Experiment 1 and Experiment 2) indicate that searchers do take advantage of a browsing structure to target the items they need. But browsing does not always save the searchers time and effort. In this study, some participants initiated their search by browsing but later switched to text querying. This behavior contains the associated risk that users may end up with no useful results at all after spending much time navigating along the hierarchy. That is, browsing may increase the search time and lead searchers down incorrect search paths.

The results of this study suggest that most people are able to use subject headings to explore different topic areas. Subject headings and classification numbers have been notoriously difficult to understand by average users. One of the benefits of faceted library catalogs is that they expose both the subject headings and classification numbers to searchers. Because of the virtual, nonlinear digital environment, faceted library catalogs are also able to offer multiple sets of subject headings to searchers. In this study, participants preferred subject headings over classification numbers because subject headings have a flat and flexible structure where searchers do not have to figure out the meaning of the hierarchy. Both the BISAC subject headings and the LCSHs offer respective benefits. The study results suggest that most people choose to use BISAC subject headings for genre information and LCSHs for topic information.

5.3 Limitations of Facets

Facets also may bring some problems to the interface, such as choice overload, visual complexity, and information overload. Not all of the participants liked the faceted catalog; some still preferred traditional search because they thought it was straightforward and they

did not need to know much about the topic prior to the search. In fact, for every feature added to the interface, the searcher experiences an associated burden. Therefore, faceted catalogs need to ensure that any facet (or burden) is worthwhile to the searcher, because offering searchers options that they cannot use effectively is pointless. The experimental results from this study reveal that, for some tasks, facets do not help searchers, but actually burden them instead.

Furthermore, this study finds that current facet implementations are not perfect. Facets may be separated either too much or too little, rarely used but nonetheless placed in a prominent position and potentially helpful but missing altogether from a relevant list of facets. For some facets, the presentation technique needs improvement and the category names are not sufficiently intuitive. Some facets even contain typos and bugs. Some study participants remarked that sometimes they were confused by facets, and would have preferred to make more effort pursuing the search path rather than being confused.

Sometimes, facets can lead to either too little or too much information. For some tasks in the study experiments, the facets were not helpful because the dimensions did not match what the searchers needed to cluster the information space. This phenomenon is understandable, because in theory, the ways to classify information space are unlimited, so therefore, the number of possible facets is limitless. However, in practice and due to the space limit, only a limited number of facets are presented and, therefore, the limited number of facets cannot necessarily satisfy all the clustering needs. In other cases, facets can lead the search in a direction the searcher does not want to go or to information the searcher does not need. Some participants remarked that they felt “unsafe” using facets because facets narrowed the search from thousands of results to several results too quickly. During this

filtering process, some valuable information might be excluded. Therefore, the extent to which facets can filter the search needs special attention.

Another limitation of facets is that, for most faceted systems, users are able to refine a query, but not expand a query. For example, if the user enters ‘American history’ in the search box, it is easier for him/her to narrow the query in terms of year, author, and/or region. However, the user is not able to ‘zoom out’ to a larger view about world history. Expanding a search is particularly useful and common in exploratory searches where the initial chosen lens need to be adjusted.

Other limitations of facets are inherent in the system. For example, faceted systems do not support facets of facets. Facets themselves cannot be assigned facets. Therefore, faceted search does not allow set-based relationship browsing. So, exploring a collection of books based on the facets that apply to its authors is impossible. This limitation creates logical complexities and problems for faceted search because a title can have multiple authors. The set of titles authored by someone from Ireland who is affiliated with the University of Dublin, for example, is not the same set as the intersection of titles by someone from Ireland and the set of people affiliated with the University of Dublin, because the title may have two authors, each of whom satisfies one of the criteria. To solve this problem, semantic web browsing is needed, allowing users to perform faceted searches on multiple entity types at one time and to perform joint operations between entity types as part of a query.

The last but not least limitation of faceted search in terms of query formulation lies in the limited number of dimensions and available values. Compared to other query elicitation techniques, such as query expansion, term suggestion, and relevance feedback, faceted

metadata provide users a rather narrow avenue for formulating a query. Users are forced to abide by the fixed number of dimensions and dimension values to refine their queries.

5.4 Search Context: Laboratory vs. Naturalistic Setting

Both TLA and laboratory experiments were conducted as part of this research to investigate ways that people search using faceted library catalogs. This study compares user behavior in both a naturalistic setting and a laboratory setting. As expected, the study participants demonstrated diverse and different behaviors in the two settings. Overall, compared to the natural setting, in the laboratory setting the participants spent a longer amount of time, conducted more operations, and submitted more queries in one search session, they tried some advanced search techniques, and they were more persistent in obtaining the information they needed. The facet usage also saw a large difference, for example, 8.0% in the natural setting and 46.7% in the laboratory for the UNC faceted library catalog, which is a 484% increase. Although the artificial search scenarios greatly affected the percentages, context may also contribute to the differences. The study participants were more likely to be motivated to search in a laboratory environment than in a natural setting when their performance was recorded and they were paid. Also, most of the information needs from the general public tend to be simple and easily satisfied with just a few actions. Their behavior is based on their simple information need.

The debate on whether to use a naturalistic setting or an experimental setting to investigate human behavior is an ongoing one. Each setting has its own advantages and limitations. This dissertation combines the two settings to take advantage of the benefits of each and to balance the disadvantages of each with complementary factors whereby one setting's advantages offset the disadvantages of the other setting.

5.5 Individual Differences

The participants' search behaviors demonstrated individual differences. This finding is found from both the log analysis and the user experiments. For the same task and using the same catalog, participants chose different ways of searching. Even across different tasks, a personal search pattern could easily be identified. For example, some participants were 'query' people who tended to modify their queries until they reached a satisfactory result set. This query modifying behavior could be seen across many tasks. Some participants were 'record' people who rarely adjusted their original query but rather continued to scan and check records until they found the target items. This pattern could be identified across different tasks and different interfaces as well. These findings agree with previous studies (for example, Davies, 1998) that both information searching and handling behavior are very personal. The findings also suggest that individual differences do matter, and that personal approaches to modeling would be more effective than general-purpose approaches.

5.6 Information Barriers in Library Catalogs

The information barriers in traditional library catalogs observed by Borgman (1996) are the "gap between the way a question is asked and ways it might be answered." Therefore, matching or entry vocabularies address the general problem of reconciling a user's query with the vocabulary presented in the catalog. Although faceted search reveals some authority data to searchers and addresses some information asymmetry between the information collection and the information need (as shown in Figure 5.1), its exposure of the index vocabulary to the user in the subject facet is limited to controlled vocabulary derived from the bibliographic records. Relevant records may not be retrieved because of a mismatch

between the vocabulary of the users and that of the bibliographic records, or because bibliographic record vocabulary is missing from the facets.

Research (Antelman et al., 2006) shows that users' vocabulary is large and diverse – that is, users rarely choose the same term to describe the same concept – and that users' vocabulary also is inflexible – that is, users are unable to repair searches using synonyms. Without the ability to stem or handle synonyms, users are not able to employ faceted search sufficiently to overcome such information barriers.

Another essential reason for the existence of information barriers lies in the presentation of the collection. Library catalogs, unlike web search engines, do not allow a search of the entire collection, but rather a search for the surrogates of the collection (MARC records). Any catalog with a slick appearance and fantastic facet design, but that misses the underlying artificial and inflexible surrogates that usually contain many typos, will not see a drastic improvement in user-catalog interaction.

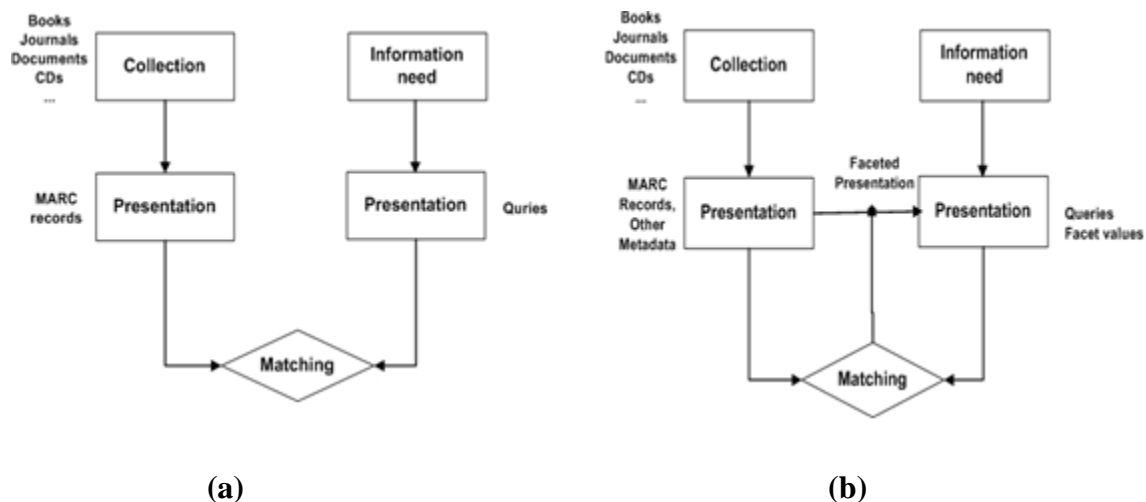


Figure 5.1 Before (a) and after (b) adding facets to library catalogs

5.7 Design Principles for Faceted Library Interfaces

The results of this research can be used to propose or refine a set of design guidelines for faceted library catalogs. Such guidelines are intended to inform librarians and library IT staff about ways to make the catalogs effective in helping people find the information they need. User interface design guidelines take into consideration constraints, capabilities, features, tradeoffs, domain knowledge and human factors. Through best practices, they provide practical advice to OPAC designers. The proposed principles are suggested to create guidelines that:

- Incorporate browsing facets
- Add/remove facets selectively
- Support including and excluding by facets
- Provide a flat vs. hierarchical structure
- Provide popular vs. long-tail data
- Consolidate the same types of facet values
- Support 'AND', 'OR' and 'NOT' selections
- Incorporate predictable schema

Incorporate browsing facets

This study indicates that people are able to take advantage of browsing facets, and that browsing facets boosts the facet uptake. Future faceted OPACs could incorporate faceted browsing structures to accommodate searchers' browsing behavior. The depth and breadth of the hierarchy should be considered carefully to avoid any confusion or burden to searchers. Structures that are either too deep or too wide will cause usability issues. Arranging facet values into a meaningful hierarchy is also important because sometimes searchers require more effort to make sense of a browsing structure than to find value from it.

Add/remove facets selectively

Due to space limitations and computational costs, facets must be chosen selectively for placement on the search interface. More importantly, a large number of facets can confuse searchers. From the log analysis conducted as part of this research, some participants rarely used some facets, such as the author facet or the MeSH facet. So, some facets should simply be removed if they are found not to be useful. On the other hand, some facets, such as the genre facet, should be added for their added value and usefulness.

Provide a flat vs. hierarchical structure

Determining possible ways to present facets that have a large number of values is a matter of ongoing debate. A flat structure and a hierarchical structure are the two primary choices. In a flat structure, facet values are presented one by one, according to some ranking criterion. Due to the screen limit, the top ranked values are displayed by default, with the remaining ones in a ‘see more’ option. Flat data are criticized for lacking a well-organized structure to lead users to the information they need. Presented with a long list, the participants in this study had to scan through the list one entry at a time in order to choose one. Presenting the users with only the top posted labels might also risk hiding the long-tail information that could be valuable.

In contrast to a flat structure, a hierarchical structure is proposed. A hierarchical structure offers a good way to organize the subject values. However, the depth and the width of the hierarchy must be considered carefully to avoid any confusion or burden to users. Facets are to help users, not to distract them with an impenetrable hierarchy (Tunkelang, 2009). The findings of this study suggest that, unless the hierarchy makes perfect sense to searchers, a flat structure should be used to present the facet values.

Provide popular vs. long-tail data

Both the UNC and Phoenix catalogs display facets with a large number of values by ‘cutting off’ a long list and showing only the top values. The underlying assumption is that the top posted values are more helpful to searchers than deeply buried ones. This assumption is somewhat problematic, however, because sometimes the long-tail data are actually valuable to searchers. Therefore, future catalogs should not only consider the popular values, but also provide a way for searchers to access the deeply buried long-tail data.

Consolidate the same types of facet values

Although the definition of *facet* is not as rigorous as the classic faceted classification that organizes a domain into mutually exclusive and collectively exhaustive dimensions, during the user experiments in this study, participants experienced confusion when *topical* and *name* subjects were separated, and *fiction* and *juvenile fiction* were split. Therefore, facets of the same type of value should be analyzed to determine whether they should be restructured and consolidated into one facet.

Support ‘AND’, ‘OR’, and ‘NOT’ selections

This study demonstrates that the user selects one value per facet, but people actually need multiple selections. When multiple selections were made available in this study, most participants were able to take advantage of them. So far, the logical relationships of queries supported by most faceted search systems are quite simple: an ‘or’ relationship among facet values and an ‘and’ relationship among facets. However, what if the user wants an ‘and’ among facet values as well as an ‘or’ among facets? The ‘not’ relationship supported by the UNC catalog proved helpful to users as well. Ideally, future faceted catalogs should be able to support complex logical relationships among facets as much as SQL can.

Incorporate predictable schema

The study participants were found to incorporate facets at an early stage of their searches. Therefore, showing facets before searchers have seen any search results has the potential to quicken their search, but it can also lead them down the incorrect path because the searchers are not able to predict the effect of choosing these facets. This phenomenon is similar to the idea that Beaulieu and Jones (1998) refer to as ‘functional visibility’ in the context of query expansion. They suggest that searchers must be aware of the options that are available at any stage, and also must be aware of the effect of these options. For example, the numbers next to facet labels are one type of predictable scheme. In addition, a preview of facet values, perhaps appearing by mouse over the facet value, could be potentially helpful for searchers to assess the facet values.

5.8 Limitations

In addition to the limitations mentioned in the discussion section at the end of each result section, such as disadvantages of transaction log analysis, population bias, artificial nature of search tasks, internet delay, below the author will talk about the overall major limitations of this study.

The major limitation lies in the fact that there were many confounding factors when we tried to compare or contrast the UNC and Phoenix catalogs. Due to different audience, different collections, different local needs, different catalogs, and different interfaces, it is hard to make any direct comparison meaningful. However, through the results from the two library catalogs, we are able to ‘qualitatively’ better understand how people used facets in different library environments. The understanding also helps minimize these confounding factors through experiment manipulations and randomizations.

In this dissertation research, because of the constraint (each task requires and only requires 3 answers) imposed on the open-ended tasks, search time is a reasonable measurement to show how efficient the search is while keeping the accuracy easily comparable to each other. Using search time as a measurement of search performance, however, is not perfect, especially for the open-ended tasks. Some researchers (for example, Capra & Marchionini, 2008) suggest that time might not be a suitable measure for exploratory tasks. Completing an exploratory task quickly may suggest that a search system does not provide support for investigating and exploring. This finding is backed up by the Kammerer, Narin, Pirolli, and Chi' study (2009) results that suggest that the participants who used the MrTaggy interface spent more time and produced better reports than participants who used other interfaces. Time, in this case, is a positive measure for the system. Future research might consider removing the constraint on the open-ended tasks and investigating the subtle relationship between search time and search performance.

It is challenge to generalize the results of this study to some other faceted library catalogs. An evaluation of one library's faceted catalog does not substitute for another due to the size and scope of local collections, cataloging practices, and metadata (Fagan, 2010). It is even more challenge to apply the results of this research to special libraries and information centers which have 'special' collections, patrons, and services. Facets might be more supportive for narrowing search because special libraries are likely to have well-defined metadata that is helpful to search by; on the other hand, might be less likely to initially browse like with the Phoenix Public Library.

5.9 Future Work

Results of this research suggest an additional research agenda that involves both new data collection efforts and the re-use of the existing datasets. The existing data are incredibly detailed, user-focused, and in large quantity. The author plans to conduct a detailed study on ways that people search based on individual task level. Search path analysis and sequential event analysis are proposed to better understand ways people handle different tasks using different catalogs. Deep manipulations of faceted interfaces are also needed to investigate searchers' behavior even further. For example, the facet presentation order, facet value display, and browsing structure (for example, make the UNC catalog appearance look like that of the Phoenix catalog, but without changing the underlying collection), are all features that can be manipulated and tested. The author also plans to conduct more in-depth studies based on the visualized logs with the help of VUTL, for example, the state (action) transition comparison, pattern identification, and analysis of dwell time on an action.

The author's future research will continue to move along the path it has been moving thus far, but will involve more search situations and varied contexts. In particular, the author plans to extend her research scope from the library environment to the general web search environment, and narrow the scope from generic libraries to special libraries, to gain a better understanding of the ways people use facets in different contexts.

The author also plans to focus more intentionally on the evaluation aspects of IR user studies, such as task complexity control. Self-report complexity is one the of the primary data types on which researchers must rely when studying human search behavior. Despite a large body of research on self-report measures and method bias in other disciplines, few serious

treatments of this topic in the IR field currently exist. Understanding, documenting, and addressing these issues is fundamental to information systems.

Another project the author has planned is the design and usability testing of a mobile faceted search interface. Mobile devices, such as smart phones and tablet computers, are becoming increasingly popular as information access tools. Until recently, most mobile interfaces maintained only a few basic and essential functions of desktop-based interfaces. Small screens, awkward input devices, and relatively low bandwidths are all challenges for information-rich applications such as faceted search. By investigating ways that mobile search interfaces differ from those of desktop search interfaces, including ways that mobile queries differ from general web queries, and developing specialized techniques for specifying queries, incorporating facet refinement, and viewing retrieval results, the author would like to develop a mobile faceted search interface that would make searching on mobile platforms feel more like desktop-based web searches.

In conclusion, this research has contributed to a better understanding of how facets can be used to support the interactions between searchers and library catalogs. The research findings have practical implications for librarians and catalog designers to design future faceted library catalogs. The research results also suggest future research agendas that can further address these issues.

Appendix A. Consent Form

University of North Carolina-Chapel Hill
Consent to Participate in a Research Study
Adult Participants
Social Behavioral Form

IRB Study # 11-0726
Consent Form Version Date: 09/13/2011

Title of Study: Beyond Text Querying and Ranked List: How People Use Facets in Faceted Library Catalogs
Principal Investigator: Xi Niu
UNC-Chapel Hill Department: School of Information and Library Science
UNC-Chapel Hill Phone number: (919) 381-7759
Email Address: xiniu@email.unc.edu

Faculty Advisor: Dr. Brad Hemminger
UNC-Chapel Hill Department: School of Information and Library Science
UNC-Chapel Hill Phone number: (919) 966-2998
Email Address: bmh@ils.unc.edu

Study Contact telephone number: (919) 381-7759
Study Contact email: xiniu@email.unc.edu

What are some general things you should know about research studies?
You are being asked to take part in a research study. To join the study is voluntary.
You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study. There also may be risks to being in research studies.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study.

You will be given a copy of this consent form. You should ask the researchers named above, any questions you have about this study at any time.

What is the purpose of this study?
The purpose of this study is to compare two user interfaces of the UNC library catalog system.

You are being asked to be in the study because you are a student at UNC and may use these interfaces in your academic work.

How many people will take part in this study?

If you decide to be in this study, you will be one of approximately 48 people in this research study.

How long will your part in this study last?

We expect the total duration of your participation to be about 2 hours.

What will happen if you take part in the study?

Users will be introduced to two different interfaces of the library catalog and then asked to do searches using both interfaces. During the testing users will rate the effectiveness of the interfaces used. At the conclusion of the session they will be asked questions about their experience to help interpret the effectiveness of the two interfaces.

What are the possible benefits from being in this study?

Research is designed to benefit society by gaining new knowledge. You may also expect to benefit by participating in this study by increasing your skills using library catalogs.

What are the possible risks or discomforts involved from being in this study?

There are no known risks to participating in this study.

How will your privacy be protected?

Subject privacy and confidentiality will be maintained in several ways. The principal investigator will do the recruiting and arrange times by email. All participants will be able to select the time of their participation.

The user test results will be coded with unique identifiers. Linkage between the unique identifiers and the volunteers' names will be retained separately from the user test results to facilitate follow-up questions for clarification. The identifiers, test results, and screen/audio recordings will be destroyed after the data analysis is complete.

Participants will not be identified in any report or publication about this study. Although every effort will be made to keep research records private, there may be times when federal or state law requires the disclosure of such records, including personal information. This is very unlikely, but if disclosure is ever required, UNC-Chapel Hill will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies for purposes such as quality control or safety.

We would like to use screen capture software to record what users were doing with the interfaces. Users will also be audiotaped during the interview for transcription and analysis. The screen and audio recording will be kept until analysis is complete and then it will be destroyed. Until that time, the recordings will be stored in a secure place. Users' name will not be stored with either the screen recording or the audio recording.

Check the line that best matches your choice:
____ OK to record me during the study
____ Not OK to record me during the study

Will you receive anything for being in this study?

You will be receiving compensation of \$30 for taking part in this study. At any point in time you may withdraw from the study; if you do you will receive a prorated portion of the \$30 stipend.

Will it cost you anything to be in this study?

There will be no costs for being in the study

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

Title of Study: Beyond Text Querying and Ranked List: How People Use Facets in Faceted Library Catalogs

Principal Investigator: Xi Niu

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

Signature of Research Participant

Date

Printed Name of Research Participant

Signature of Person Obtaining Consent

Date

Printed Name of Person Obtaining Consent

Appendix B. Questionnaires

ENTRY QUESTIONNAIRE

1. What's your age? ____
2. What's your gender? A. Male B. Female
3. Overall, for how many years have you been searching online? ____
4. How much **experience** do you have using online library catalog search interfaces?

Very little

Very much

1 2 3 4 5

5. How much **experience** do you have using the UNC Library catalog search interface in particular?

Very little

Very much

1 2 3 4 5

6. How much **experience** do you have using categories to narrow search? For example, how much experience do you have using systems like Amazon (where you can limit Books to Children's Books, and further limit to 9-12 year old appropriate books) and eBay (limit cloth to Men's and further to Men's Swimwear)?

Very little

Very much

1 2 3 4 5

POST-SEARCH QUESTIONNAIRE

1. How **easy** was it to find the relevant item(s) for the topic?

very difficult very easy

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

2. How **satisfied** are you with the search result?

very unsatisfied very satisfied

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

3. How **enjoyable** was the search process?

very frustrating very enjoyable

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

POST-INTERFACE QUESTIONNAIRE

	Strongly Disagree Strongly Agree				
	1	2	3	4	5
The interface was intuitive .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interface makes the searching process enjoyable .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interface makes it easy to find relevant items.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This interface makes it quick to find relevant items.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Appendix C. Interview Questions

PART I CLOSE-ENDED QUESTIONS

1. Overall, which of the two library catalog interfaces you tried did you like the best?

- ☐ the interface with the facet controls in the left column
- ☐ the interface without the facet controls in the left column
- ☐ I liked them the same

PART II OPEN-ENDED QUESTIONS

Faceted UNC Interface

1. Specifically, what are the things you like and don't like for the interfaces?
2. Can you tell me for what reasons or when you would use the categories? Can you give me an example?
3. Did you notice any similarity or difference in how you used the categories each time? Can you describe an example?
- 4a. Can you describe an example where the categories helped the search?
- 4b. Can you describe an example where the categories hindered the search?
5. Are there any things about the interfaces that you found confusing?
6. Do you have any suggestions for changes of both interfaces?

Non Faceted UNC Interface

1. Specifically, what are the things you like and don't like for the interfaces?
- 2a. What is your best experience with the search interface (the search interface helped you successfully find the target quickly, easily, and correctly)
- 2b. What is your worst experience with the search interface (the search interface was frustrating to use, didn't help you find results you wanted, even hindered your search)
3. Are there any things about the interfaces that you found confusing?
4. Do you have any suggestions for changes of both interfaces?

Phoenix Interface

1. Specifically, what are the things you like and don't like for the interfaces?
2. Can you tell me for what reasons or when you would use the categories? Can you give me an example?
3. Did you notice any similarity or difference in how you used the categories each time? Can you describe an example?
- 4a. Can you describe an example where the categories helped the search?
- 4b. Can you describe an example where the categories hindered the search?
5. Are there any things about the interfaces that you found confusing?
6. Do you have any suggestions for changes of both interfaces?

General

1. Which interface do you like better? Why?
2. Did the (different) categories change the way you searched? Can you describe an example? Why?
3. Is there anything else you'd like to talk about?

Appendix D. Sources of Questionnaire and Interview Questions

I. Kelly et al. (2010)

Exit Questionnaire (partial)

	Strongly Disagree Strongly Agree				
	1	2	3	4	5
The system was easy to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to find relevant documents with the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall, the system was effective in helping me complete the search tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Post-Search questionnaire (partial)

How difficult was it to find relevant documents for the topic?

Very Difficult				Very Easy
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How satisfied are you with your search?

Very Unsatisfied				Very Satisfied
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How satisfied are you with the system's performance?

Very Unsatisfied				Very Satisfied
O	O	O	O	O

II. Capra et al. (2007)

Presession Questionnaire (patial)

5. Overall, for how many years have you been searching online? _

9. How much experience have you had searching on computerized library catalogs either locally (e.g., you library) or remotely (e.g., Library of Congress)

No experience		Some experience		A great deal of experience
O	O	O	O	O

13. How much experience do you have using categories to narrow searches? For example, how much have you used categories on systems such as Amazon (i.e. Books -> Children's Books -> Ages 9-12) and eBay (i.e. Clothing -> Men's -> Sweaters)?

No experience		Some experience		A great deal of experience
O	O	O	O	O

III. Ramdeen & Hemminger (2012)

Exit questions

1. So tell me about this experience? How did the interface(s) work for you?
2. Specifically, what were the things you liked and didn't like about the Metadata based search interface?
3. Are there any features or system behaviors that you found confusing?
4. Specifically, what were the things you liked and didn't like about the Text based Faceted Browse based search interface?
5. Are there any features or system behaviors that you found confusing?
6. Is there anything else you'd like to talk about?

Appendix E. Tasks

Set A

Type	Close-ended	Open-ended
Complexity		
Simple (3 exact clues)	1. You want to find the classic piece of “To kill a mockingbird”, published by “HarperCollinsPublishers” in 1995.	5. You want to find three movies that were originally released in a language other than English. The movies should be published by Paramount Pictures.
Simple (3 exact clues)	2. You just heard a pop song at a restaurant. Somebody told you the song was “Sweet Dreams” by Beyoncé. You want to find if the library has a music CD containing this song.	6. You want to watch some recently released movies; you are to find three movies that have been received in last month by the UNC "Media Resources Center".
Complex (1 exact and 2 fuzzy clues)	3. A friend recommended a book to you. But you can't remember the exact title. You know the author is Stephen Hawking and it's about advanced series in astrophysics and cosmology, particularly big bang theory.	7. You are taking Economics class. For your term paper you want to find three resources (any type) that cover the topic of microfinance as related to small business in either Asian or African countries.
Complex (1 exact and 2 fuzzy clues)	4. You are developing an online questionnaire and need some entry-level guidance on PHP programming. A professor recommended a book to you, but all you can remember is the author's last name started with "Vas", like Vasse Meyer, Vastuck. And that the book was published in 2009.	8. You want to find three science fiction books that have stories involving robots published after 2000.

Set B

Type Complexity	Close-ended	Open-ended
Simple (3 exact clues)	1. You want to find the classic piece of “The catcher in the rye”, published by “Little, Brown” in 1951.	5. You want to find three movies by Warner Brothers. The movies should be in any media other than DVD or BlueRay disk, because you do not have a DVD player.
Simple (3 exact clues)	2. You just heard a pop song at a café. Somebody told you the song was “Without You” by Mariah Carey. You want to find if the library has a music CD containing this song.	6. You want to find three journals, magazines, or serials that the Information & Library Science Library has recently received in last three months.
Complex (1 exact and 2 fuzzy clues)	3. Your dad is a war film lover. He asked you if you could help him find the movie about a group of people raising the American flag at the battle of Iwo Jima in World War II. The movie director is Clint Eastwood.	7. You have heard many news stories recently about Egypt. You would like to know some more background information in order to better understand the news. You want to find three titles about Egyptian politics, particularly Hosni Mubarak’s influence on Egyptian politics. They should be either an e-book or an online resource.
Complex (1 exact and 2 fuzzy clues)	4. Next month you are taking a trip to Europe. To get ready for your trip you want to read a travel guide to Europe. You remember you have read a good one. The author's first and last names both begin with the letter "B". You also remember it was published in 1992. Try to find if the library catalog has the book.	8. You want to find three comics or graphic novels containing stories about fictitious mouse characters. They should be published after 2005.

Set C

Type Complexity	Close-ended	Open-ended
Simple (3 exact clues)	1. You want to find the classic piece of “To kill a mockingbird”, published by “HarperCollinsPublishers” in 1995.	5. You want to find three movies that were originally released in a language other than English. The movies should be published by Paramount Pictures.
Simple (3 exact clues)	2. You just heard a pop song at a restaurant. Somebody told you the song was “Sweet Dreams” by Beyoncé. You want to find if the library has a music CD containing this song.	6. You want to find three adventure movies, that are very popular at the library. Choose three that are currently on the "most borrowed list" maintained by the library.
Complex (1 exact and 2 fuzzy clues)	3. A friend recommended a book to you. But you can't remember the exact title. You know the author is Stephen Hawking and it's about advanced series in astrophysics and cosmology, particularly big bang theory.	7. You are taking Economics class. For your term paper you want to find three resources (any type) that cover the topic of microfinance as related to small business in either Asian or African countries.
Complex (1 exact and 2 fuzzy clues)	4. You are developing an online questionnaire and need some entry-level guidance on PHP programming. A professor recommended a book to you, but all you can remember is the author's last name started with "Vas", like Vassemeyer, Vastuck. And that the book was published in 2009.	8. You want to find three science fiction books that have stories involving robots published after 2000.

Set D

Type Complexity	Close-ended	Open-ended
Simple (3 exact clues)	1. You want to find the classic piece of “The catcher in the rye”, published by “Little, Brown” in 1951.	5. You want to find three movies by Warner Brothers. The movies should be in any media other than DVD or BlueRay disk, because you do not have a DVD player.
Simple (3 exact clues)	2. You just heard a pop song at a café. Somebody told you the song was “Without You” by Mariah Carey. You want to find if the library has a music CD containing this song.	6. You are babysitting three kids age less than 10. You heard that children are better behaved when there is background music played. Your babysitting job is stressful, and you want to find three music CDs for kids. You also want them to be among the new arrivals in the library.
Complex (1 exact and 2 fuzzy clues)	3. Your dad is a war film lover. He asked you if you could help him find the movie about a group of people raising the American flag at the battle of Iwo Jima in World War II. The movie director is Clint Eastwood.	7. You have heard many news stories recently about Egypt. You would like to know some more background information in order to better understand the news. You want to find three titles about Egyptian politics, particularly Hosni Mubarak’s influence on Egyptian politics. They should be in either Agave or Burton Barr branch.
Complex (1 exact and 2 fuzzy clues)	4. Next month you are taking a trip to Europe. To get ready for your trip you want to read a travel guide to Europe. You remember you have read a good one. The author's first and last names both begin with the letter "B". You also remember it was published in 1992. Try to find if the library catalog has the book.	8. You want to find three comics or graphic novels containing stories about fictitious mouse characters. They should be published after 2005.

Appendix F. Subject Training Scripts and Tasks

i. UNC faceted interface

Interactive Task 1:

You vaguely remember a comedy/science fiction film from the early 1980s about a group of people who repossess cars for a living. You're pretty sure the movie stars Emilio Estevez. Find out whether the library has a copy of the DVD.

Points: actor can be treated as an author; subject facet, year facet; guess by repossess

Interactive Task 2:

You want to find three biographies of American first ladies that were published in 2010.

Points: use an item to find another item; take a suggestion; broader queries

Independent Task 1:

You want to find the title of "Physics of semiconductor devices", published by "Wiley-Interscience" in 1969.

Points: put everything but year in the same box, no need to do advanced search.

Independent Task 2:

You are at the Library Service Center. You want to find three titles on History of Education under the broader topic "Education". The titles should be held by the Library Service Center and written in either Turkish or Serbian.

Points: browse by call numbers; OR selection; make sure the item is the right one

Independent Task 3:

Find three new (got in last month) popular music CDs, but not from Music Library.

Points: new titles; subject facet; NOT selection

ii. UNC non-faceted interface

Interactive Task 1:

You heard a book mentioned on the radio that you'd like to read. All you remember is that the book was published in 2005 and has the word "Germ" in the title. It's about the advancement of human civilizations throughout history. Locate the book.

Points: subject in the keyword field; broader queries

Interactive Task 2:

A friend was raving about novels by Pynchon he just finished. You also want to read them.

Find Pynchon's three distinct novels published by Penguin Press.

Points: Publish needs broadening; how to make sure it's novels

Independent Task 1:

You want to find an e-book titled "The Story of an African Farm" published by Chadwyck-Healey.

Points: title field; format subject; all put title and publisher together in the search box, anywhere; manually scan

Independent Task 2:

Find three movies on women's rights that were released in the 1990s.

Points: subject in the keyword field

Independent Task 3:

You happen to read an interesting chapter of a book. The chapter is titled "Can I Believe the Bible?" by a well-known Christian writer Josh McDowell. You want to follow this up and find the whole book.

Points: partial content; either way works (simple or advanced)

iii. Phoenix

Interactive Task 1:

You heard a book mentioned on the radio that you'd like to read. All you remember is that the book was published in 2005 and has the word "Germ" in the title. It's about the advancement of human civilizations throughout history. Locate the book.

Points: subject in the keyword field; broader queries; year in the search box

Interactive Task 2:

A friend was raving about novels by Pynchon he just finished. You also want to read them. Find Pynchon's three distinct novels published by Penguin Press.

Points: Publish needs broadening; how to make sure it's novels

Independent Task 1:

You want to find the classic book "Pride and Prejudice" published by "Oxford University Press" in 1999.

Points: all put together in the search box

Independent Task 2:

You want to find three readings for your spare time. Find three Pulitzer Prize winners of non-fictions about wars.

Points: browse

Independent Task3:

You happen to read an interesting chapter of a book. The chapter is titled "Is it true? : is it believable?" by a well-known Christian writer Josh McDowell. You want to follow this up and find the whole book.

Points: partial content

Appendix G. Email Exchange with Subjects

Recruiting Email

Subject: Freshman: Earn \$30 for performing searches using library catalogs

Body:

Dear Freshmen,

Do you like challenges? Do you like searching for answers to questions?
Would you like to participate in a study comparing two interfaces to the
UNC library catalog?

This research study is investigating the differences between two library
catalog interfaces. If you choose to participate in this study, you
will meet with a researcher individually and be asked to perform
searches using the two interfaces, and then answer a few questions. The
whole study takes about 2 hours.

Participants who complete the study
will receive \$30
as compensation for their time.

This is an IRB approved study (IRB#: 11-0726, approved date: 4/28/2011).
Oversight for this study is provided by the UNC Chapel Hill
Institutional Review Board (IRB) for Social and Behavioral Research. If
you
have questions or concerns about this study please contact the IRB at
919-966-3113 or by email at irb_questions@unc.edu.

For more information or to take part in this study please contact Xi Niu
at xiniu@live.unc.edu, or call 919-381-7759. Participants should be at
least 18 years of age, and should be Freshmen at UNC (Native English
Speaker is preferred).

Sincerely,
Xi Niu, Brad Hemminger

Email Reminder

Subject: Reminder for participating in our experiment

Body:

Dear [STUDY PARTICIPANT],

I am writing to remind you that you are scheduled to participate in a research study about library catalogs on TOMORROW at [INSERT TIME].

Please come to the entrance to Davis Library on time. My cell phone number is 919-381-7759. Could you please tell me yours?

Your participation will take approximately 2 hours. You will receive \$30.00 for participating in the study.

We rely on the generosity of persons such as yourself for conducting research and are very appreciative of your willingness to help. Choosing or declining to participate in this study will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research; it is purely voluntary.

Best,
Xi Niu

Appendix H. Receipts

IRB Study #11-0726

I acknowledge receipt of [\$30] for participating in this research study.

Signature of Research Participant Date

Printed Name of Research Participant

Signature of Researcher Date

Printed Name of Researcher

Principal Investigator

Xi Niu
CB #3360
School of Information & Library Science, UNC-CH
Chapel Hill NC, 27599-3360
xiniu@live.unc.edu

References

- Antelman, K., Lynema, E., & Pace, A. K. (2006). Toward a twenty-first century library catalog. *Information Technology and Libraries*, 25(3), 128-138.
- Babu, B. R., & O'Brien, A. (2000). Web OPAC interfaces: An overview. *The Electronic Library*, 18(5), 316-330.
- Bast, H., & Weber, I. (2006). When you're lost for words: Faceted search with auto completion. In *Proceedings of ACM Special Interest Group on Information Retrieval (SIGIR 2006)*, 31-35.
- Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424.
- Beaulieu, M. & Jones, S. (1998). Interactive searching and interface issues in the Okapi best match probabilistic retrieval system. *Interacting with computers*, 10(3), 237-248.
- Beezley, S. A. (2011) Dare to be Deweyless: An Evaluation of BISAC-based Headings in Public Libraries. Master's Paper. University of North Carolina at Chapel Hill.
- Belkin, N. J. (1980). Anomalous States of Knowledge as a Basis for Information Retrieval. *The Canadian Journal of Information Science*, 5, 133-143.
- Belkin, N. J. (2000). Helping people find what they don't know. *Communication of ACM (CACM)*, 43(8), 58-61.
- Belkin, N., Cool, C., Kelly, D., Lee, H. J., Muresan, G., Tang, M.C., et al. (2003). In Query length in interactive information retrieval. *Proceedings from SIGIR 2003: The 26th Annual International ACM Conference on Research and Development in Information Retrieval*, 205-221. Toronto, Canada.
- Ben-Yitzhak, O., Golbandi, N., Har'El, N., Lempel, R., Neumann, A., Ofek-Koifman, S., ... Yogev, S. (2008). Beyond basic faceted search, *Proceedings from WSDM 2008: The ACM International Conference on Web Search and Data Mining*. Stanford, CA.
- Bilal, D. (2002). Children's use of the Yahoooligans! Web search engine III. Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, 53(13), 1170-1183.
- Borgman, C.L. (1986a). The user's mental model of an information retrieval system: An experiment on a prototype online catalog. *International Journal of Man-Machine Studies*, 24(1), 47-64.
- Borgman, C. L. (1986b). Why are online catalogs hard to use? lessons learned from information-retrieval studies. *Journal of the American Society for Information Science*, 37(6), 387-400.

- Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science*, 47(7), 493-503.
- Borgman, C. L., Hirsh, S. G., Walter, V. A., & Gallagher, A. L. (1995). Children's searching behavior on browsing and keyword online catalogs: The science library catalog project. *Journal of the American Society for Information Science*, 46(9), 663-684.
- Breeding, M. (2007). Introduction [next-generation catalogs] *Library Technology Reports* 43(4), 5-14.
- Capra, R. G., & Marchionini, G. (2008). The relation browser tool for faceted exploratory search. Proceedings from JCDL '08: The 8th ACM/IEEE-CS Joint Conference on Digital libraries. Pittsburgh, PA.
- Capra, R. G., Marchionini, G., Oh, J. S., Stutzman, F., & Zhang, Y. (2007). Effects of Structure and Interaction Style on Distinct Search Tasks. Proceedings from JCDL '07: The 8th ACM/IEEE-CS Joint Conference on Digital libraries. Vancouver, Canada
- Chan, L. M. (2001). *Exploiting LCSH, LCC, and DDC to Retrieve Networked Resources: Issues and Challenges*. Washington, D.C.: Library of Congress.
- Chau, M., Fang, X., & Sheng, O. R. L. (2005). Analysis of the Query Logs of a Web Site Search Engine. *Journal of the American Society for Information Science and Technology*, 56(13), 1363-1376.
- Chen, H. M., & Cooper, M. D. (2001). Using clustering techniques to detect usage patterns in a web-based information system. *Journal of the American Society for Information Science and Technology* 52(11), 888-904.
- Chen, H. M., & Cooper, M. D. (2002). Stochastic modeling of usage patterns in a web-based information system. *Journal of the American Society for Information Science and Technology*, 53(7), 536-548.
- Chen, H., Yim, T., Fye, D., & Schatz, B. (1995). Automatic thesaurus generation for an electronic community system. *Journal of the American Society for Information Science and Technology*, 46(3), 175-193.
- Cochrane, P. A., & Markey, K. (1983). Catalog use studies--since the introduction of online interactive catalogs: Impact on design for subject access. *Library and Information Science Research*, 5(4), 337-363.
- Collins, M., Samples, J., Pennell, C., & Goldsmith, D. (2007). Magnifying the ILS with Endeca. *The Serials Librarian*, 51(3/4), 75-100.
- Connaway, L., Budd, J., & Kohtanek, T. (1995). An investigation of the use of an online catalog: User characteristics and transaction log analysis. *Library Resources & Technical Services*, 39(2), 142-152.

- Cutrell, E., Robbins, D. C., Dumais, S. T., & Sarin, R. (2006). Fast, flexible filtering with Phlat-Personal search and organization made easy. Proceedings from CHI 2006: *Conference on Human Factors in Computing Systems*. Montreal, Quebec, Canada.
- Dalrymple, P.W., & Zweizig, D.L. (1992). Users' experience of information retrieval systems: An exploration of the relationship between search experience and affective measures. *Library and Information Science Research*, 14, 167–181.
- Davies, M. (1998). Impact of information developments on the information handling techniques of research scientists: Implications of selected results. *New Review of Information Networking*, 11, 53–70.
- Doan, K., Plaisant, C., Shneiderman, B., & Bruns, T. (1997). Query previews for networked information systems: A case study with NASA environmental data. *SIGMOD Record*, 26, 75-81.
- Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. C. (2003). Stuff I've seen: a system for personal information retrieval and re-use. Proceedings from SIGIR 2003: *The 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 72-79. Toronto, Canada.
- Eastman, C. M., & Jansen, B. J. (2003). Coverage, relevance, and ranking: The impact of query operators on web search engine results. *ACM Transactions on Information Systems (TOIS)*, 21(4), 383-411.
- Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, 45(3), 171–212.
- English, J., Hearst, M., Sinha, R., Swearingen, K., & Lee, K. (2002). Flexible search and navigation using faceted metadata. Technical Report. University of California at Berkeley.
- Ensor, P. (1992). User characteristics of keyword searching in an OPAC. *College and Research Libraries*, 53(1), 72-80.
- Fagan, J. C. (2010). *Usability Studies of Faceted Browsing: A Literature Review*. *Information Technology and Libraries*, 29(2), 58-66.
- Frei, H.P., & Jauslin, J.F., (1984) Two-dimensional representation of information retrieval services. In: *H.J. Dietschmann, Editor, Representation and exchange of knowledge as a basis of information processes*, North Holland, New York (1984), 383–396.
- Greenberg, J. (2003). Metadata and the World Wide Web. *Encyclopedia of Library and Information Science*, 1876-1888.
- Goker, A., & He, D. (2000). Analysing Web Search Logs to Determine Session Boundaries for User-Oriented Learning. *Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems*, 319-322.

- Grossman, T., & Fitzmaurice, G. (2010). ToolClips: An investigation of contextual video assistance for functionality understanding. Proceedings from CHI 2010: *The 28th International Conference on Human factors in Computing Systems*, 1515–1524. Atlanta, GA
- Hancock-Beaulieu, M. (1990). Evaluating the impact of an online library catalogue on subject searching behavior at the catalogue and at the shelves. *Journal of Documentation*, 46(3) 18-338.
- Harman, D. (1988). Towards interactive query expansion. Proceedings from SIGIR 1988: *The 11th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 321-331. Grenoble, France.
- Hearst, M. A. (2000). Next generation web search: Setting our sites. *Bulletin of the Technical Committee on Data Engineering*, 23(3), 38-48.
- Hearst, M. A. (2006). Clustering versus faceted categories for information exploration. *Communications of the ACM*, 49(4), 59-61.
- Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K., & Yee, P. (2002). Finding the Flow in Web Site Search, *Communications of the ACM*, 45 (9), 42-49.
- Hert, C. A., & Marchionini, G. (1997). Seeking Statistical Information In Federal Websites: Users, Tasks, Strategies, and Design Recommendations. *Final Report to the Bureau of Labor Statistics*. <http://ils.unc.edu/~march/blsreport/blsmain.htm>
- Hemminger, B. M., Lu, D., Vaughan, K., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 58(14), 2205-2225.
- Heo, M. & Hirtle, S. C. (2001). An empirical comparison of visualization tools to assist information retrieval on the Web. *Journal of the American Society for Information Science and Technology*, 52(8), 666-675.
- Hildreth, C. R. (1991). Advancing toward the E3 OPAC: the imperative path. In *Proceedings of Think Tank on the Present and Future of the Online Catalog*, 17-38.
- Hildreth, C. R. (2001). Accounting for users' inflated assessments of on-line catalog search performance and usefulness: an experimental study. *Information Research*, 6(2). Available at: <http://InformationR.net/ir/6-2/paper101.html>
- Hutchinson, H., Bederson, B.B., Druin, A., (2007) Supporting Elementary-Age Children's Searching and Browsing: Design and Evaluation Using the International Children's Digital Library, *Journal of the American Society for Information Science and Technology*, 58 (11), 1618-1630.
- Hunter, R. N. (1991). Successes and failures of patrons searching the online catalog at a large academic library: A transaction log analysis. *RQ*, 30(3), 395-402.

- Ingwersen, P., and Wormell, I. (1989). Modern indexing and retrieval techniques matching different types of information needs. In: *Koskiala, S., and Launo, R. (Eds.) Information, Knowledge, Evolution* (pp. 79–90). London, UK: North-Holland.
- Janosky, B., Smith, P., & Hildreth, C. (1986). Online library catalog systems: An analysis of user errors. *International Journal of Man-Machine Studies*, 25(5), 573-592.
- Jansen, B. J., & Pooch, U. (2001). A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology*, 52(3), 235-246.
- Jansen, B. J., Spink, A., Blakely, C., & Koshman, S. (2007). Defining a session on web search engines. *Journal of the American Society for Information Science and Technology*, 58(6), 862-871.
- Jansen, B. J., Spink, A., & Saracevic, T. (2000). A study of users queries on the Web. *Information Processing and Management*, 36(2) (Special Issue: Web Research & IR), 207–227.
- Järvelin, K. & Ingwersen, P. (2004). Information seeking research needs extension towards tasks and technology. *Information Research*, 10(1), 212. Available at <http://InformationR.net/ir/10-1/paper212.html>
- Jones, S., Cunningham, S. J., McNab, R., & Boddie, S. (2000). A transaction log analysis of a digital library. *International Journal on Digital Libraries*, 3(2), 152-169.
- Jones, W. P. (2007). *Keeping found things found: The study and practice of personal information management*. San Francisco, CA: Morgan Kaufmann.
- Käki, M. (2005). Findex: search result categories help users when document ranking fails. Proceedings from CHI 2005: *The 23rd International Conference on Human Factors in Computing Systems*, 140. Portland, Oregon.
- Kammerer, Y., Narin, R., Pirolli, P., & Chi, E. (2009). Signpost from the masses: Learning effects in an exploratory social tag search browser. Proceedings from CHI 2009: *The 27th International Conference on Human Factors in Computing Systems*, 625–634. Boston, MA.
- Kelly, D., Cushing, A., Dostert, M., Niu, X., & Gyllstrom, K. (2010). Effects of Popularity and Quality on the Usage of Query Suggestions during Information Search, Proceedings from CHI 2010: *The 28th ACM Conference on Human Factors in Computing Systems*. Atlanta, GA
- Kim, K. (2001). Information-seeking on the Web: Effects of user and task variables. *Library & Information Science Research*, 23(3), 233-255.
- Kwasnik, B. H. (1992). A descriptive study of the functional components of browsing. Proceedings from: *Engineering for Human-Computer Interaction: The IFIP TC2/WG2.7*

Working Conference on Engineering for Human-Computer Interaction, 191-203. Ellivuori, Finland

Knutson, G. (1991). Subject enhancement: Report on an experiment. *College and Research Libraries*, 52(1), 65-79.

Kules, B. (2006). Supporting exploratory web search with meaningful and stable categorized overviews. Dissertation. University of Maryland at College Park.

Kules, B., Capra, R., Banta, M., & Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? Proceedings from JCDL 2009: *The Joint International Conference on Digital libraries*, 313-322. Austin, TX.

Kules, B., Capra, R., & Ryan, J. (2010). Transaction log analysis of user actions in a faceted library catalog interface. Proceedings from HCIR 2010: *The 4th Workshop on Human-Computer Interaction and Information Retrieval*. New Brunswick, NJ.

Kules, B., & Capra, R. (2011). Influence of Training and Stage of Search on Gaze Behavior in a Library Catalog Faceted Search Interface. *Journal of the American Society for Information Science and Technology*. 63(1), 114-138.

Kules, B., & Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing and Management*, 44(2), 463-484.

Lown, C. (2008). A transaction log analysis of NCSU's faceted navigation OPAC. Master's Paper. University of North Carolina at Chapel Hill.

La Barre, K. (2007). The heritage of early FC in document reference retrieval systems, *Library History*, 23(2), 129-149.

Large, A., & Beheshti, J. (1997). OPACs: A research review. *Library and Information Science Research*, 19(2), 111-133.

Lau, E. P., & Goh, D. H. L. (2006). In search of query patterns: A case study of a university OPAC. *Information Processing and Management*, 42(5), 1316-1329.

Lewis, D. W. (1987). Research on the use of online catalogs and its implications for library practice. *Journal of Academic Librarianship*, 13(3), 152-157.

Lown, C. (2008). A transaction log analysis of NCSU's faceted navigation OPAC. Master's Paper. University of North Carolina at Chapel Hill.

Luther, J. (2003). Trumping google? metasearching's promise. *Library Journal*, 128(16), 36-40.

Magennis, M., & Rijsbergen, C.J. (1997). The potential and actual effectiveness of interactive query expansion. Proceedings from SIGIR 1997: *The 20th Annual*

International ACM SIGIR Conference on Research and Development in Information Retrieval, 324-332. Philadelphia, PA

- Mahoui, M., & Cunningham, S. J. (2001). Search behavior in a research-oriented digital library. *Lecture Notes in Computer Science*, 2163, 13-24.
- Mat-Hassan, M., & Levene, M. (2005). Associating Search and Navigation Behavior through Log Analysis. *Journal of the American Society for Information Science and Technology*, 56(9), 913-934.
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. New York: Cambridge University Press.
- Marchionini, G. (2002). Co-Evolution of User and Organizational Interfaces: A Longitudinal Case Study Of WWW Dissemination of National Statistics. *Journal of the American Society for Information Science and Technology*, 53(14), 1192-1209.
- Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4), 41-46.
- Marchionini, G., & Brunk, B. (2003). Towards a general relation browser: a GUI for information architects. *Journal of Digital Information*, 4(1).
- Muramatsu, J., & Pratt, W. (2001). Transparent Queries: investigation users' mental models of search engines. Proceedings from SIGIR 2001: *The 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 217-224. New Orleans, LA.
- Nahl, D. (1997). Information counseling inventory of affective and cognitive reactions while learning the internet. *Internet Reference Services Quarterly*, 2(2-3), 11-33.
- Niu, X., Lown, C., & Hemminger, B. M. (2009). Log based analysis of how faceted and text based search interact in a library catalog interface. Proceedings from HCIR 2009: *The 3rd Workshop on Human-Computer Interaction and Information Retrieval*. Washington, DC.
- Noerr, P. L., & Noerr, K. T. B. (1985). Browse and navigate: An advance in database access methods. *Information Processing and Management*, 21(3), 205-213.
- Novotny, E. (2004). I Don't Think I Click: A Protocol Analysis Study of Use of a Library Online Catalog in the Internet Age. *College & Research Libraries*, 65(6), 525-37.
- O'Brien, A. (1990). Relevance as an aid to evaluation in OPACs. *Journal of Information Science*, 16, 265-271.
- O'Day, V., & Jeffries, R. (1993). Orienteering in an information landscape: how information seekers get from here to there. Proceedings from CHI 1993: *The ACM SIGCHI*

- Conference on Human Factors in Computing Systems*, 438–445. Amsterdam, The Netherlands.
- Olson, T. A. (2007). Utility of a faceted catalog for scholarly research. *Library Hi Tech*, 25(4), 550-561.
- Peters, T. A. (1989). When smart people fail: An analysis of the transaction log of an online public access catalog. *Journal of Academic Librarianship*, 15(5), 267-273.
- Peters, T. A. (1993). The history and development of transaction log analysis. *Library Hi Tech*, 11, 41-41.
- Pirolli, P. L., & Card, S. K. (1999). Information foraging. *Psychological Review*, 106(4), 643-675.
- Ramdeen S., & Hemminger B.M. (2012). A tale of two interfaces: How Facets affect the library catalog search experience. *Journal of the American Society for Information Science*, (published online 2011). Print 2012. DOI: 10.1002/asi.21689.
- Riewe, . (2008). Survey of Open Source Integrated Library Systems. Master's Paper. San Jose State University.
- Rocchio, J. (1971). *Relevance feedback in information retrieval*. In G.Salton (Ed.), *The smart retrieval system—Experiments in automatic document processing* (pp. 313–323). Englewood Cliffs, NJ: Prentice-Hall.
- Sapiie, J. (1995). Reader interest classification: The user friendly schemes. *Cataloging & Classification Quarterly*, 19(3), 143-155.
- Schallan, J. (2007). New search and browse at Phoenix Public Library. Public email from [Publib]. Retrieved on Feb, 23, 2012.
- Sharit, J., Hernández, M.A., Czaja, S.J., & Pirolli, P. (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, 3:1–3:25.
- Sheble, L., & Wildemuth, B. (2009). Chapter 18. Transaction logs. In Wildemuth, B., *Applications of social research methods to questions in information and library science* (pp.166-177). Santa Barbara, CA :Libraries Unlimited.
- Silverstein, C., Henzinger, M., Marais, H., Morics, M. (1999). Analysis of a Very Large Web Search Engine Query Log. *SIGIR Forum*, 33(1), 6-12.
- Sit, R. A. (1998). Online library catalog search performance by older adult users. *Library and Information Science Research*, 20(2), 115-131.
- Soergel, D. (1999). The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science*, 50(12), 1119-1120.

- Solomon, P. (1993). Children's information retrieval behavior: A case analysis of an OPAC. *Journal of American Society for Information Science and Technology*, 44(5), 245-264.
- Spink, A., Wolfram, D. & Jansen, B.J., & Saracevic, T. (2001). Searching the Web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 226-234.
- Tang, M. C. (2005). Browsing and Searching in a Faceted Information Space: A Longitudinal Study of PubMed Users' Interactions with a Novel Display Tool, PhD Dissertation. Rutgers University.
- Taylor, A. G. (1992). *Introduction to cataloging and classification*. Colorado: Libraries Unlimited.
- Taylor, A. G. (2006). *Introduction to cataloging and classification*. Westport, CT: Libraries Unlimited.
- Tolle, J. E., & Hah, S. (1985). Online search patterns: NLM CATLINE database. *Journal of the American Society for Information Science and Technology*, 36(2), 82-93.
- Tunkelang, D. (2009). *Faceted search*. San Rafael, CA: Morgan & Claypool Publishers.
- Yee, K. P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. Proceedings from CHI 2003: *The 21st Conference on Human Factors in Computing Systems*, 401-408. Fort Lauderdale, FL.
- Yee, M. M. (1991). System design and cataloging meet the user: User interfaces to online public access catalogs. Retrieved March 21, 2010, from <http://www.Escholarship.org/uc/item/2rp099x6>.
- Young, M., & Yu, H. (2004). The impact of web search engines on subject searching in OPAC. *Information Technology and Libraries*, 23(4), 168-180.
- Yuan, W. (1997). End-user searching behaviors in information retrieval: a longitudinal study. *Journal of the American Society for Information Science*, 48(3), 218-234.
- Wallace, P. M. (1993). How do patrons search the online catalog when no one. *RQ*, 33(2), 239-252.
- Wang, P., Berry, M.W., & Yang, Y. (2003). Mining longitudinal Web queries: Trends and patterns. *Journal of the American Society for Information Science and Technology*, 54(8), 743-758.
- Warren, P. (2000). Why they still cannot use their library catalogues. In *Proceedings of Informing Science Conference*, 19-22.

- Wildemuth, B., & Cao, L. (2009). Chapter 12. Experimental Studies. In Wildemuth, B., *Applications of social research methods to questions in information and library science* (pp.105-115). Santa Barbara, CA :Libraries Unlimited.
- Wildemuth, B., & Moore, M. E. (1995). End-user search behaviors and their relationship to search effectiveness. *Bulletin of the Medical Library Association*, 83(3): 294–304.
- White, M. D., & Iivonen, M. (2001). Questions as a factor in web search strategy. *Information Processing & Management*, 37(5), 721–740.
- White, M.D., Iivonen, M., & Shneiderman, B. (1994). Dynamic queries for visual information seeking. *IEEE Software* 11(6), 70–77.
- White, R. W., & Drucker, S. M. (2007). Investigating behavioral variability in web search. *Proceedings from WWW 2007: The 16th Annual World Wide Web Conference*, 21–30. Banff, Alberta, Canada.
- White, R. W., & Marchionini, G. (2007)Examining the effectiveness of real-time query expansion. *Information Processing & Management* 43(3), 685–704.
- White, R. W., & Roth, R. A. (2009). *Exploratory search: Beyond the query-response paradigm*. San Rafael, CA: Morgan & Claypool Publishers.
- Wolfram, D., Wang, P., & Zhang, J. (2009). Identifying web search session patterns using cluster analysis: A comparison of three search environments. *Journal of the American Society for Information Science and Technology*, 60(5), 896-910.
- Zhang, J., & Marchionini, G. (2005) “Evaluation and evolution of a browse and search interface: Relation browser,” In *Proceedings of the National Conference on Digital Government Research*, 179–188.