Usability Studies of Faceted Browsing:
A Literature Review
Jody Condit Fagan

Faceted browsing is a common feature of new library catalog interfaces. But to what extent does it improve user performance in searching within today's library catalog systems? This article reviews the literature for user studies involving faceted browsing and user studies of "next-generation" library catalogs that incorporate faceted browsing. Both the results and the methods of these studies are analyzed by asking, What do we currently know about faceted browsing? How can we design better studies of faceted browsing in library catalogs? The article proposes methodological considerations for practicing librarians and provides examples of goals, tasks, and measurements for user studies of faceted browsing in library catalogs.

Many libraries are now investigating possible new interfaces to their library catalogs. Sometimes called "next-generation library catalogs" or "discovery tools," these new interfaces are often separate from existing integrated library systems. They seek to provide an improved experience for library patrons by offering a more modern look and feel, new features, and the potential to retrieve results from other major library systems such as article databases.

One interesting feature these interfaces offer is called "faceted browsing." Hearst defines facets as a "a set of meaningful labels organized in such a way as to reflect the concepts relevant to a domain."[1] LaBarre defines facets as representing "the categories, properties, attributes, characteristics, relations, functions or concepts that are central to the set of documents or entities being organized and which are of particular interest to the user group."[2] Faceted browsing offers the user relevant subcategories by which they can see an overview of results, then narrow their list. In library catalog interfaces, facets usually include authors, subjects, and formats, but may include any field that can be logically created from the MARC record (see figure 1 for an example).

Using facets to structure information is not new to librarians and information scientists. As early as 1955, the Classification Research Group stated a desire to see faceted classification as the basis for all information retrieval.[3] In 1960, Ranganathan introduced facet analysis to our profession.[4] Librarians like metadata because they

know its power, and facets can showcase metadata in new interfaces.

According to McGuinness, facets perform several functions in an interface:

- vocabulary control
- site navigation and support
- overview provision and expectation setting
- browsing support
- searching support
- disambiguation support[5]

These functions offer several potential advantages to the user: The functions use category systems that are coherent and complete, they are predictable, they show previews of where to go next, they show how to return to previous states, they suggest logical alternatives, and they help the user avoid empty result sets as searches are narrowed.[6] Disadvantages include the fact that categories of interest must be known in advance, important trends may not be shown, category structures may need to be built by hand, and automated assignment is only partly successful.[7] Library catalog records, of course, already supply "categories of interest" and a category structure.

Information science research has shown benefits to users from faceted search interfaces. But do these benefits hold true for systems as complex as library catalogs? This paper presents an extensive review of both information science and library literature related to faceted browsing.

Method

To find articles in the library and information science literature related to faceted browsing, the author searched the Association for Computing Machinery (ACM) Digital Library, Scopus, and Library and Information Science and Technology Abstracts (LISTA) databases. In Scopus and the ACM Digital Library, the most successful searches included the following:

- (facet* or cluster*) and (usability or user stud*)
- facet* and usability

In LISTA, the most successful searches included combining product names such as "aquabrowser" with "usability." The search "catalog and usability" was also used. The author also searched Google and the Next Generation Catalogs for Libraries (NGC4LIB) electronic discussion list in an attempt to find unpublished studies.

Search terms initially included the concept of "clustering"; however, this was quickly shown to be a clearly defined, separate topic. According to Hearst, "Clustering refers to the grouping of items according to some measure

Jody Condit Fagan (faganjc@jmu.edu) is Content Interfaces Coordinator, James Madison University Library, Harrisonburg, Virginia.

Showing 1 - 20 of 1346 Results for **interpersonal communication**		Sort Relevance ⌄
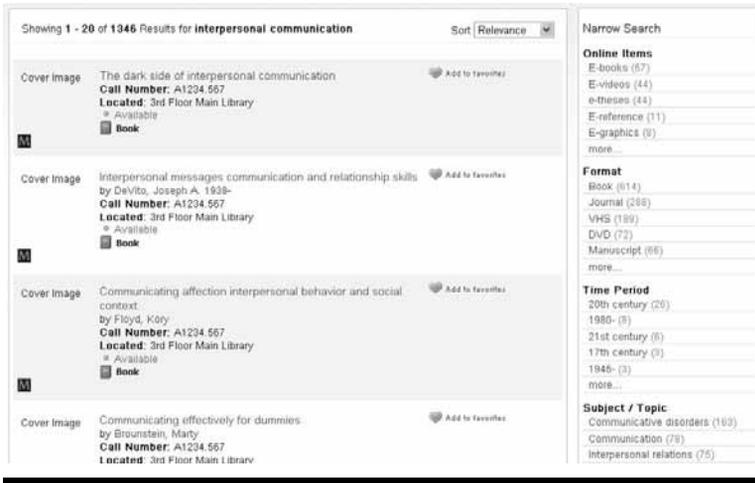
Figure 1. Faceted results from JMU's VuFind implementation

of similarity . . . typically computed using associations and commonalities among features where features are typically words and phrases."[8] Using library catalog keywords to generate word clouds would be an example of clustering, as opposed to using subject headings to group items. Clustering has some advantages according to Hearst. It is fully automated, it is easily applied to any text collection, it can reveal unexpected or new trends, and it can clarify or sharpen vague queries. Disadvantages to clustering include possible imperfections in the clustering algorithm, similar items not always being grouped into one cluster, a lack of predictability, conflating many dimensions, difficulty labeling groups, and counterintuitive subhierarchies.[9] In user studies comparing clustering with facets, Pratt, Hearst, and Fagan showed that users find clustering difficult to interpret and prefer a predictable organization of category hierarchies.[10]

▋ Results

The author grouped the literature into two categories: user studies of faceted browsing and user studies of library catalog interfaces that include faceted browsing as a feature. Generally speaking, the information science literature consisted of empirical studies of interfaces created by the researchers. In some cases, the researchers' intent was to create and refine an interface intended for actual use; in others, the researchers created the interface only for the purposes of studying a specific aspect of user behavior. In the library literature, the studies found were generally qualitative usability studies of specific library catalog interface products. Libraries had either implemented a new product, or they were thinking about

doing so and performed a user study to inform their decision.

Results: empirical studies of faceted browsing

The following summaries present selected empirical research studies that had significant findings related to faceted browsing or interesting methods for such studies. It is not an exhaustive list.

Pratt, Hearst, and Fagan questioned whether faceted results were better than clustering or relevancy-ranked results.[11] They studied fifteen breast-cancer patients and families. Every subject used three tools: a faceted interface, a tool that clustered the search results, and a tool that ranked the search results according to relevance criteria. The subjects were given three simple queries related to breast cancer (e.g., "What are the ways to prevent breast cancer?"), asked to list answers to these before beginning, and to answer the same queries after using all the tools. In this study, subjects completed two timed tasks. First, subjects found as many answers as possible to the question in four minutes. Second, the researchers measured the time subjects took to find answers to two specific questions (e.g., "Can diet be used in the prevention of breast cancer?") that related to the original, general query. For the first task, when the subjects used the faceted interface, they found more answers than they did with the other two tools. The mean number of answers found using the faceted interface was 7.80, for the cluster tool it was 4.53, and for the ranking tool it was 5.60. This difference was significant ($p<0.05$).[12] For the second task, the researchers found no significant difference between the tools when comparing time on task. The researchers gave the subjects a user-satisfaction questionnaire at the end of the study. On thirteen of the fourteen quantitative questions, satisfaction scores for the faceted interface were much higher than they were for either the ranking tool or the cluster tool. This difference was statistically significant ($p < 0.05$). All fifteen users also affirmed that the faceted interface made sense, was helpful, was useful, and had clear labels, and said they would use the faceted interface again for another search.

Yee et al. studied the use of faceted metadata for image searching, and browsing using an interface they developed called Flamenco.[13] They collected data from thirty-two participants who were regular users of the Internet, searching for information either every day or a few times a week. Their subjects performed four tasks (two structured and two unstructured) on each of two interfaces. An example of an unstructured task from their study was "search for images of interest." An example of a structured task was to gather materials for an art history

essay on a topic given by the researchers and to complete four related subtasks. The researchers designed the structured task so they knew exactly how many relevant results were in the system. They also gave a satisfaction survey. More participants were able to retrieve all relevant results with the faceted interface than with the baseline interface. During the structured tasks, participants received empty results with the baseline interface more than three times as often as with the faceted interface.[14] The researchers found that participants constructed queries from multiple facets in the unstructured tasks 19 percent of the time and in the structured tasks 45 percent of the time.[15] When given a post–test survey, participants identified the faceted interface as easier to use, more flexible, interesting, enjoyable, simple, and easy to browse. They also rated it as slightly more "overwhelming." When asked to choose between the two, twenty-nine participants chose the faceted interface, compared with two who chose the baseline ($N = 31$). Thirty-one of the thirty-two participants said the faceted interface helped them learn more, and twenty-eight of them said it would be more useful for their usual tasks.[16] The researchers concluded that even though their faceted interface was much slower than the other, it was strongly preferred by most study participants: "These results indicate that a category-based approach is a successful way to provide access to image collections."[17]

In a related usability study on the Flamenco interface, English et al. compared two image browsing interfaces in a nineteen-participant study.[18] After an initial search, the "Matrix View" interface showed a left column with facets, with the images in the result set placed in the main area of the screen. From this intermediary screen, the user could select multiple terms from facets in any order and have the items grouped under any facet. The "SingleTree" interface listed subcategories of the currently selected term at the top, with query previews underneath. The user could then only drill down to subcategories of the current category, and could not select terms from more than one facet. The researchers found that a majority of participants preferred the "power" and "flexibility" of Matrix to the simplicity of SingleTree. They found it easier to refine and expand searches, shift between searches, and troubleshoot research problems. They did prefer SingleTree for locating a specific image, but Matrix was preferred for browsing and exploring. Participants started over only 0.2 percent of the time for the Matrix compared to 4.5 percent for SingleTree.[19] Yet the faceted interface, Matrix, was not "better" at everything. For specific image searching, participants found the correct image only 22.0 percent of the time in Matrix compared to 66.0 percent in SingleTree.[20] Also, in Matrix, some participants drilled down in the wrong hierarchy with wrong assumptions. One interesting finding was that in both interfaces, more participants chose to begin by browsing (12.7 percent) than by searching (5.0 percent).[21]

Uddin and Janacek asked nineteen users (staff and students at the Asian Institute of Technology) to use a website search engine with both a traditional results list and a faceted results list.[22] Tasks were as follows: (1) look for scholarship information for a masters program, (2) look for staff recruitment information, and (3) look for research and associated faculty member information within your interested area.[23] They found that users were faster when using the faceted system, significantly so for two of the three tasks. Success in finding relevant results was higher with the faceted system. In the post–study questionnaire, participants rated the faceted system more highly, including significantly higher ratings for flexibility, interest, understanding of information content, and more search results relevancy. Participants rated the most useful features to be the capability to switch from one facet to another, preview the result set, combine facets, and navigate via breadcrumbs.

Capra et al. compared three interfaces in use by the Bureau of Labor Statistics website, using a between-subjects study with twenty-eight people and a within-subjects study with twelve people.[24] Each set of participants performed three kinds of searches: simple lookup, complex lookup, and exploratory. The researchers used an interesting strategy to help control the variables in their study:

> Because the BLS website is a highly specialized corpus devoted to economic data in the United States organized across very specific time periods (e.g., monthly releases of price or employment data), we decided to include the US as a geographic facet and a month or year as a temporal facet to provide context for all search tasks in our study. Thus, the simple lookup tasks were constructed around a single economic facet but also included the spatial and temporal facets to provide context for the searchers. The complex lookup tasks involve additional facets including genre (e.g. press release) and/or region.[25]

Capra et al. found that users preferred the familiarity afforded by the traditional website interface (hyperlinks + keyword search) but listed the facets on the two experimental interfaces as their best features. The researchers concluded, "If there is a predominant model of the information space, a well designed hierarchical organization might be preferred."[26]

Zhang and Marchionini analyzed results from fifteen undergraduate and graduate students in a usability study of an interface that used facets to categorize results (Relation Browser ++).[27] There were three types of tasks:

- Type 1: Simple look-up task (three tasks such as "check if the movie titled *The Matrix* is in the library movie collection").
- Type 2: Data exploration and analysis tasks (six tasks

that require users to understand and make sense of the information collection: "In which decade did Steven Spielberg direct the most movies?").

- Type 3: (one free exploration task: "find five favorite videos without any time constraints").

The tasks assigned for the two interfaces were different but comparable. For type 2 tasks, Zhang and Marchionini found that performance differences between the two interfaces were all statistically significant at the .05 level.[28] No participants got wrong answers for any but one of the tasks using the faceted interface. With regard to satisfaction, on the exploratory tasks the researchers found statistically significant differences favoring the faceted interface on all three of the satisfaction questions. Participants found the faceted interface not as aesthetically appealing nor as intuitive to use as the basic interface. Two participants were confused by the constant changing and updating of the faceted interface.

The above studies are examples of empirical investigations of experimental interfaces. Hearst recently concluded that facets are a "proven technique for supporting exploration and discovery" and summarized areas for further research in this area, such as applying facets to large "subject-oriented category systems," facets on mobile interfaces, adding smart features like "autocomplete" to facets, allowing keyword search terms to affect order of facets, and visualizations of facets.[29] In the following section, user studies of next-generation library catalog interfaces will be presented.

Results: library literature

Understandably, most studies by practicing librarians focus on products their libraries are considering for eventual use. These studies all use real library catalog records, usually the entire catalog's database. In most cases, these studies were not focused on investigating faceted browsing per se, but on the usability of the overall interface. In general, these studies used fewer participants than the information science studies above, followed less rigorous methods, and were not subjected to statistical tests. Nevertheless, they provide many insights into the user experience with the extremely complex datasets underneath next-generation library catalog interfaces that feature faceted browsing. In this review article, only results specifically relating to faceted browsing will be presented.

Sadeh described a series of usability studies performed at the University of Minnesota (UM), a Primo development partner.[30] Primo is the next-generation library catalog product sold by Ex Libris. The author also received additional information from the Usability Services lab at UM via e-mail. Three studies were conducted in August 2006, January 2007, and October 2007. Eight users from various disciplines participated in each

of the first two studies: The first study comprised one faculty member, five graduate students, and two undergraduate students; the second comprised two faculty members, four graduate students, and two undergraduate students. The third study did not report results related to faceted browsing and is not discussed here. The first study had seven scenarios; the second study had nine. The scenarios were complex: for example, one scenario began, "You want to borrow Shakespeare's play, *The Tempest*, from the library," but contained the following subtasks as well:

1. Find *The Tempest*.
2. Find multiple editions of this item.
3. Find a recent version.
4. See if at least one of the editions is available in the library.
5. What is the call number of the book?
6. You'd like to print the details of this edition of the book so you can refer to it later.

Participants found the interface friendly, easy to use, and easy to learn. All the participants reported that faceted browsing was useful as a means of narrowing down the result lists, and they considered this tool one of the differentiating features between Primo and their library OPAC or other interfaces. Facets were clear, intuitive, and useful to all participants, including opening the "more" section.[31] One specific result from the tests was that "online resources" and "available" limiters were moved from a separate location to the right with all other facets.[32]

In a study of Aquabrowser by Olson, twelve subjects—all graduate students in the humanities—participated in a comparative test in which they looked for additional sources for their dissertation.[33] Aquabrowser was created by MediaLab but is distributed by Serials Solutions in North America. This study also had three pilot subjects. No relevance judgments were made by the researchers. Nine of the twelve subjects found relevant materials by using Aquabrowser that they had not found before.[34] Olson's subjects understood facets as a refinement tool (narrowing) and had a clear idea of which facets were useful and not useful for them. They gave overwhelmingly positive comments. Only two felt the faceted interface was not an improvement. Some participants wanted to limit to multiple languages or dates, and a few were confused about the location of facets in multiple places, for example, "music" under both format and topic.

A team at Yale University, led by Bauer, recently conducted two tests on pilot VuFind installations: a subject-based presentation of e-books for the Cushing/Whitney Medical Library and a pilot test of VuFind using undergraduate students with a sample of 400,000 records from the library system.[35] VuFind is open-source software developed at Villanova University (http://vufind.org).

The team drew test questions from user search logs in their current library system. Some questions targeted specific problems, such as incomplete spellings and incomplete title information. Bauer notes that some problems uncovered in the study may relate to the peculiarities of the Yale implementation.

The medical library study contained eight participants—a mix of medical and nursing students. Facets, reported Bauer, "worked well in several instances, although some participants did not think they were noticeable on the right side of the page."[36] The prompt for the faceted task in this study came after the user had done a search: "What if you wanted to look at a particular subset, say 'xxx' (determine by looking at the facets)."[37] Half of the participants used facets, half used "search within" to narrow the topic by adding keywords. Sixty-two percent of the participants were successful at this task.

The undergraduate study asked five participants faced with a results list, "What would you do now if you only wanted to see material written by John Adams?"[38] On this task, only one of the five was successful, even though the author's name was on the screen. Bauer noted that in general, "the use of the topic facet to narrow the search was not understood by most participants. . . . Even when participants tried to use topic facets the length of the list and extraneous topics rendered them less than useful."[39]

The five undergraduates were also asked, "Could you find books in this set of results that are about health and illness in the United States population, or control of communicable diseases during the era of the depression?"[40] Again, only one of the five was successful. Bauer notes that "the overly broad search results made this difficult for participants. Again, topic facets were difficult to navigate and not particularly useful to this search."[41] Bauer's team noted that when the search was configured to return more hits, "topic facets become a confusingly large set of unrelated items. These imprecise search results, combined with poor topic facet sets, seemed to result in confusion for test participants."[42] Participants were not aware that topics represented subsets, although learning occurred because the "narrow" header was helpful to some participants.[43] Other results found by Bauer's team were that participants were intrigued by facets, navigation tools are needed so that patrons may reorder large sets of topic facets, format and era facets were useful to participants, and call-number facets were not used by anyone.

Antelman, Pace, and Lynema studied North Carolina State University's (NCSU) next-generation library catalog, which is driven by software from Endeca.[44] Their study used ten undergraduate students in a between-subjects design where five used the Endeca catalog and five used the library's traditional catalog. The researchers noted that their participants may have been experienced with the library's old catalog, as log data shows most NCSU users enter one or two terms, which was not true of study participants. The researchers measured task success, duration, and difficulty, but did not measure user satisfaction. Their study consisted of four known-item tasks and six topic-searching tasks. The topic-searching tasks were geared toward the use of facets, for example, "Can you show me how would you find the most recently published book about nuclear energy policy in the United States?"[45]

All five participants using Endeca understood the idea of facets, and three used them. Students tried to limit their searches at the outset rather than search and then refine results. An interesting finding was that use of the facets did not directly follow the order in which facets were listed. The most heavily used facet was Library of Congress Classification (LCC), followed closely by topic, and then library, format, author, and genre.[46] Results showed a significantly shorter average task duration for Endeca catalog users for most tasks.[47] The researchers noted that none of the students understood that the LCC facet represented call-number ranges, but all of the students understood that these facets "could be used to learn about a topic from different aspects—science, medicine, education."[48]

The authors could find no published studies relating to the use of facets in some next-generation library catalogs, including Encore and WorldCat Local. Although the University of Washington did publish results of a WorldCat Local usability study in a recent issue of *Library Technology Reports*, results from the second round of testing, which included an investigation of facets, were not yet ready.[49]

Discussion

Summary of empirical evidence related to faceted browsing

Empirical studies in the information science literature support many positive findings related to faceted browsing and build a solid case for including facets in search interfaces:

- Facets are useful for creating navigation structures.[50]
- Faceted categorization greatly facilitates efficient retrieval in database searching.[51]
- Facets help avoid dead ends.[52]
- Users are faster when using a faceted system.[53]
- Success in finding relevant results is higher with a faceted system.[54]
- Users find more results with a faceted system.[55]
- Users also seem to like facets, although they do not always immediately have a positive reaction.
- Users prefer search results organized into predictable, multidimensional hierarchies.[56]
- Participants' satisfaction is higher with a faceted system.[57]

- Users are more confident with a faceted system.[58]
- Users may prefer the familiarity afforded by traditional website interface (hyperlinks + keyword search).[59]
- Initial reactions to the faceted interface may be cautious, seeing it as different or unfamiliar.[60]

Users interact with specific characteristics of faceted interfaces, and they go beyond just one click with facets when it is permitted. English et al. found that 7 percent of their participants expanded facets by removing a term, and that facets were used more than "keyword search within": 27.6 percent versus 9 percent.[61] Yee et al. found that participants construct queries from multiple facets 19 percent of the time in unstructured tasks; in structured tasks they do so 45 percent of the time.[62]

The above studies did not use library catalogs; in most cases they used an experimental interface with record sets that were much smaller and less complicated than in a complete library collection. Domains included websites, information from one website, image collections, video collections, and a journal article collection.

Summary of practical user studies related to faceted browsing

This review also included studies from practicing librarians at live library implementations. These studies generally had smaller numbers of users, were more likely to focus on the entire interface rather than a few features, and chose more widely divergent methods. Studies were usually linked to a specific product, and results varied widely between systems and studies. For this reason it is difficult to assemble a bulleted summary as with the previous section. The variety of results from these studies indicate that when faceted browsing is applied to a real-life situation, implementation details can greatly affect user performance and user preference.

Some, like LaBarre, are skeptical about whether facets are appropriate for library information. Descriptions of library materials, says LaBarre, include analyses of intellectual content that go beyond the descriptive terms assigned to commercial items such as a laptop:

> Now is the time to question the assumptions that are embedded in these commercial systems that were primarily designed to provide access to concrete items through descriptions in order to enhance profit.[63]

It is clear that an evaluation of commercial interfaces or experimental interfaces does not substitute for an OPAC evaluation. Yet it is a challenge for libraries to find expertise and resources to conduct user studies. The systems they want to test are large and complex. Collaborating with other libraries has its own challenges: An evaluation of one product's faceted system for a library catalog does not substitute for another, the size and scope of local collections may greatly affect results, and cataloging practices and metadata will affect results. Still, it is important for practicing librarians to determine if new features such as facets truly improve the user's experience.

Methodological best practices

After reading numerous empirical research studies (some of which critique their own methods) and library case studies, some suggestions for designing better studies of facets in library catalogs emerged.

Designing the study
- Consider reusing protocols from previous studies. This provides not only a tested method but also a possible point of comparison.
- Define clear goals for each study and focus on specific research questions. It's tempting to just throw the user into the interface and see what happens, but this makes it difficult, if not impossible, to analyze the results in a useful way. For example, one of Zhang and Marchionini's hypotheses specifically describes what rich interaction would look like: "Typing in keywords and clicking visual bars to filter results would be used frequently and interchangeably by the users to finish complex search tasks, especially when large numbers of results are returned."[64]
- Develop the study for one type of user. Olson's focus on graduate students in the dissertation process allowed the researchers to control for variables such as interest of and knowledge about the subject.
- Pilot test the study with a student worker or colleague to iron out potential wrinkles.
- Let users explore the system for a short time and possibly complete one highly structured task to help the user become used to the test environment, interface, and facilitator.[65] Unless you are truly interested in the very first experience users have with a system, the first use of a system is an artificial case.

Designing tasks
- Make sure user performance on each task is measurable. Will you measure the time spent on a task? If "success" is important, define what that would look like. For example, English et al. defined success for one of their tasks as when "the participant indicated (within the allotted time) that he/she had reached an appropriate set of images/specific image in the collection."[66]
- Establish benchmarks for comparison. One can test for significant differences between interfaces, one can test for differences between research subjects and an expert user, and one can simply measure against

expectations or against previous iterations of the same study. For example, "75 percent of users completed the task within five minutes." Zhang and Marchionini measured error rates, another possible benchmark.[67]

- Consider looking at your existing OPAC logs for zero-results searches or other issues that might inspire interesting questions.
- Target tasks to avoid distracters. For example, if your catalog has a glut of government documents, consider running the test with a limit set to exclude them unless you are specifically interested in their impact. For example, Capra et al. decided to include the United States as a geographic facet and a month or year as a temporal facet to provide context for all search tasks in their study.[68]
- For some tasks, give the subjects simple queries (e.g., "What are the ways to prevent breast cancer?") as opposed to asking the subjects to come up with their own topic. This can help control for the potential challenges of formulating one's own research question on the spot. As librarians know, formulating a good research question is its own challenge.
- If you are using any timed tasks, consider how the nature of your tasks could affect the result. For example, Pratt, Hearst, and Fagan noted that the time that it took subjects to read and understand abstracts most heavily influenced the time for them to find an answer.[69] English et al. found that the system's processing time influenced their results.[70]
- Consider the implications of your local implementation carefully when designing your study. At Yale, the team chose to point their VuFind instance at just 400,000 of their records, drew questions from problems users were having (as shown in log files), and targeted questions to these problems.[71]

Who to study?
- Try to study a larger set of users. It is better to create a short test with many users than a long test with a few users. Nielsen suggests that twenty users is sufficient.[72] Consider collaborating with another library if necessary.
- If you test a small number, such as the typical four to eight users for a usability test, be sure you emphasize that your results are not generalizable.
- Use subjects who are already interested in the subject domain: for example, Pratt, Hearst, and Fagan used breast cancer patients,[73] and Olson used graduate students currently writing their dissertations.[74]
- Consider focusing on advanced or scholarly users. La Barre suggests that undergraduates may be over-studied.[75]
- For comparative studies, consider having both between-subjects and within-subjects designs.[76]
    - A between-subjects design involves creating two groups of participants, each of which tests a different system.
    - A within-subjects design has one group of participants test both systems.

It is hoped that if libraries use the suggestions above when designing future experiments, results across studies will be more comparable and useful.

Designing user studies of faceted browsing
After examining both empirical research studies and case studies by practicing librarians, a key difference seems to be the specificity of research questions and designing tasks and measurements to test specific hypotheses. While describing a full user-study protocol for investigating faceted browsing in a library catalog is beyond the scope of this article, reviewing the literature and the study methods it describes provided insights into how hypotheses, tasks, and measurements could be written to provide more reliable and comparable evidence related to faceted browsing in library catalog systems.

For example, one research question could surround the format facet: "Compared with our current interface, does our new faceted interface improve the user's ability to find different formats of materials?" Hypotheses could include the following:

1. Users will be *more accurate* when identifying the formats of items from their result set when using the faceted interface than when using the traditional interface.
2. Users will be able to identify formats of items *more quickly* with the faceted interface than with the traditional interface.

Looking at these hypotheses, here is a prompt and some example tasks the participants would be asked to perform: "We will be asking you to find a variety of formats of materials. When we say formats of materials, we mean books, journal articles, videos, etc."

- Task 1: Please use interface A to search on "interpersonal communication." Look at your results set. Please list as many different formats of material as you can.
- Task 2: How many items of each format are there?
- Task 3: Please use interface B to search on "family communication." What formats of materials do you see in your results set?
- Task 4: How many items of each format are there?"

We would choose the topics "interpersonal communication" and "family communication" because our local catalog has many material types for these topics and because these topics would be understood by most of our students. We would choose different topics to

help minimize learning effects. To further address this, we would plan to have half our users start first with the traditional interface and half to start first with the faceted interface. This way we can test for differences resulting from learning.

The above tasks would allow us to measure several pieces of evidence to support or reject our hypotheses. For tasks 1 and 3, we would measure the number of formats correctly identified by users compared with the number found by an expert searcher. For tasks 2 and 4, we would compare the number of items correctly identified with the total items found in each category by an expert searcher. We could also time the user to determine which interface helped them work more quickly. In addition to measuring the number of formats identified and the number of items identified in each format, we would be able to measure the time it takes users to identify the number of formats and the number of items in each format. To measure user satisfaction, we would ask participants to complete the System Usability Scale (SUS) after each interface and, at the very end of the study, complete a questionnaire comparing the two interfaces.

Even just selecting the format facet, we would have plenty to investigate. Other hypotheses and tasks could be developed for other facet types, such as time period or publication date, or facets related to the responsible parties, such as author or director:

> **Hypothesis**: Users can find more materials written in a certain time period using the faceted interface.
> **Task**: Find ten items of any type (books, journals, movies) written in the 1950s that you think would have information about television advertising.

> **Hypothesis**: Users can find movies directed by a specific person more quickly using the faceted interface.
> **Task**: In the next two minutes, find as many movies as you can that were directed by Orson Welles.

For the first task above, an expert searcher could complete the same task, and their time could be used as a point of comparison. For the second, the total number of movies in the library catalog that were directed by Welles is an objective quantity. In both cases, one could compare the user's performance on the two interfaces.

▌ Conclusion

Reviewing user studies about faceted browsing revealed empirical evidence that faceted browsing improves user performance. Yet this evidence does not necessarily point directly to user success in faceted library catalogs, which have much more complex databases than those used in these experimental studies. Previous case-study investigations of library catalog interfaces with facets have proven inconclusive. By choosing more specific research questions, tasks, and measurements for user studies, libraries may be able to design more objective studies and compare results more effectively.

References

1. Marti A. Hearst, "Clustering versus Faceted Categories for Information Exploration," *Communications of the ACM* 49, no. 4 (2006): 60.

2. Kathryn La Barre, "Faceted Navigation and Browsing Features in New OPACS: Robust Support for Scholarly Information Seeking?" *Knowledge Organization* 34, no. 2 (2007): 82.

3. Vanda Broughton, "The Need for Faceted Classification as the Basis of All Methods of Information Retrieval," *Aslib Proceedings* 58, no. 1/2 (2006): 49–71.

4. S. R. Ranganathan, *Colon Classification Basic Classification*, 6th ed. (New York: Asia, 1960).

5. Deborah L. McGuinness, "Ontologies Come of Age," in *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, ed. Dieter Fensel et al. (Cambridge, Mass.: MIT Pr., 2003): 179–84.

6. Hearst, "Clustering versus Faceted Categories," 60.

7. Ibid., 61.

8. Ibid., 59.

9. Ibid.. 60.

10. Wanda Pratt, Marti A. Hearst, and Lawrence M. Fagan, "A Knowledge-Based Approach to Organizing Retrieved Documents," *Proceedings of the Sixteenth National Conference on Artificial Intelligence, July 18–22, 1999, Orlando, Florida* (Menlo Park, Calif.: AAAI Pr., 1999): 80–85.

11. Ibid.

12. Ibid., 5.

13. Ka-Ping Yee et al., "Faceted Metadata for Image Search and Browsing," 2003, http://flamenco.berkeley.edu/papers/flamenco-chi03.pdf (accessed Oct. 6, 2008).

14. Ibid., 6.

15. Ibid., 7.

16. Ibid.

17. Ibid., 8.

18. Jennifer English et al., "Flexible Search and Navigation," 2002, http://flamenco.berkeley.edu/papers/flamenco02.pdf (accessed Apr. 22, 2010).

19. Ibid., 7.

20. Ibid., 6.

21. Ibid., 7.

22. Mohammed Nasir Uddin and Paul Janecek, "Performance and Usability Testing of Multidimensional Taxonomy in Web Site Search and Navigation," *Performance Measurement and Metrics* 8, no. 1 (2007): 18–33.

23. Ibid., 25.

24. Robert Capra et al., "Effects of Structure and Interaction Style on Distinct Search Tasks," *Proceedings of the 7th ACM-IEEE-CS Joint Conference on Digital Libraries* (New York: ACM, 2007): 442–51.

25. Ibid., 446.

26. Ibid., 450.

27. Junliang Zhang and Gary Marchionini, *Evaluation and Evolution of a Browse and Search Interface: Relation Browser++* (Atlanta, Ga.: Digital Government Society of North America, 2005): 179–88.

28. Ibid., 183.

29. Marti A. Hearst, "UIs for Faceted Navigation: Recent Advances and Remaining Open Problems," 2008, http://people.ischool.berkeley.edu/~hearst/papers/hcir08.pdf (accessed Apr. 27, 2010).

30. Tamar Sadeh, "User Experience in the Library: A Case Study," *New Library World* 109, no. 1/2 (Jan. 2008): 7–24.

31. Ibid., 22.

32. Jerilyn Veldof, e-mail from University of Minnesota Usability Services Lab, 2008.

33. Tod A. Olson, "Utility of a Faceted Catalog for Scholarly Research," *Library Hi Tech* 25, no. 4 (2007): 550–61.

34. Ibid., 555.

35. Kathleen Bauer, "Yale University Library VuFind Test—Undergraduates," May 20, 2008, http://www.library.yale.edu/usability/studies/summary_undergraduate.doc (accessed Apr. 27, 2010); Kathleen Bauer and Alice Peterson-Hart, "Usability Test of VuFind as a Subject-Based Display of Ebooks," Aug. 21, 2008, http://www.library.yale.edu/usability/studies/summary_medical.doc (accessed Apr. 27, 2010).

36. Bauer and Peterson-Hart, "Usability Test of VuFind as a Subject-Based Display of Ebooks," 1.

37. Ibid., 2.

38. Ibid., 3.

39. Ibid.

40. Ibid., 4.

41. Ibid.

42. Ibid., 5.

43. Ibid., 8.

44. Kristin Antelman, Andrew K. Pace, and Emily Lynema, "Toward a Twenty-First Century Library Catalog," *Information Technology & Libraries* 25, no. 3 (2006): 128–39.

45. Ibid., 139.

46. Ibid., 133.

47. Ibid., 135.

48. Ibid., 136.

49. Jennifer L. Ward, Steve Shadle, and Pam Mofield, "User Experience, Feedback, and Testing," *Library Technology Reports* 44, no. 6 (Aug. 2008): 22.

50. English et al., "Flexible Search and Navigation."

51. Peter Ingwersen and Irene Wormell, "Ranganathan in the Perspective of Advanced Information Retrieval," *Libri* 42 (1992): 184–201; Winfried Godert, "Facet Classification in Online Retrieval," *International Classification* 18, no. 2 (1991): 98–109.; W. Godert, "Klassifikationssysteme und Online-Katalog [Classification Systems and the Online Catalogue]," *Zeitschrift für Bibliothekswesen und Bibliographie* 34, no. 3 (1987): 185–95.

52. Yee et al., "Faceted Metadata for Image Search and Browsing"; English et al., "Flexible Search and Navigation."

53. Uddin and Janecek, "Performance and Usability Testing"; Zhang and Marchionini, *Evaluation and Evolution*; Hao Chen and Susan Dumais, *Bringing Order to the Web: Automatically Categorizing Search Results* (New York: ACM, 2000): 145–52.

54. Uddin and Janecek, "Performance and Usability Testing."

55. Ibid.; Pratt, Hearst, and Fagan, "A Knowledge-Based Approach"; Hsinchun Chen et al., "Internet Browsing and Searching: User Evaluations of Category Map and Concept Space Techniques," *Journal of the American Society for Information Science* 49, no. 7 (1998): 582–603.

56. Vanda Broughton, "The Need for Faceted Classification as the Basis of All Methods of Information Retrieval," *Aslib Proceedings* 58, no. 1/2 (2006): 49–71; Pratt, Hearst, and Fagan, "A Knowledge-Based Approach," 80–85.; Chen et al., "Internet Browsing and Searching," 582–603; Yee et al., "Faceted Metadata for Image Search and Browsing"; English et al., "Flexible Search and Navigation using Faceted Metadata."

57. Uddin and Janecek, "Performance and Usability Testing"; Zhang and Marchionini, *Evaluation and Evolution*; Hideo Joho and Joemon M. Jose, *Slicing and Dicing the Information Space Using Local Contexts* (New York: ACM, 2006): 66–74.; Yee et al., "Faceted Metadata for Image Search and Browsing."

58. Yee et al., "Faceted Metadata for Image Search and Browsing"; Chen and Dumais, *Bringing Order to the Web*.

59. Capra et al., "Effects of Structure and Interaction Style."

60. Yee et al., "Faceted Metadata for Image Search and Browsing"; Capra et al., "Effects of Structure and Interaction Style"; Zhang and Marchionini, *Evaluation and Evolution*.

61. English et al., "Flexible Search and Navigation," 7.

62. Yee et al., "Faceted Metadata for Image Search and Browsing," 7.

63. La Barre, "Faceted Navigation and Browsing," 85.

64. Zhang and Marchionini, *Evaluation and Evolution*, 183.

65. English et al., "Flexible Search and Navigation."

66. Ibid., 6.

67. Zhang and Marchionini, *Evaluation and Evolution*.

68. Capra et al., "Effects of Structure and Interaction Style."

69. Pratt, Hearst, and Fagan, "A Knowledge-Based Approach."

70. English et al., "Flexible Search and Navigation."

71. Bauer, "Yale University Library VuFind Test—Undergraduates."

72. Jakob Nielsen, "Quantitative Studies: How Many Users to Test?" online posting, Alertbox, June 26, 2006 http://www.useit.com/alertbox/quantitative_testing.html (accessed Apr. 7, 2010).

73. Pratt, Hearst, and Fagan, "A Knowledge-Based Approach."

74. Tod A. Olson used graduate students currently writing their dissertations. Olson, "Utility of a Faceted Catalog for Scholarly Research," *Library Hi Tech* 25, no. 4 (2007): 550–61.

75. La Barre, "Faceted Navigation and Browsing."

76. Capra et al., "Effects of Structure and Interaction Style."