One Size Doesn't Fit All: Tailoring Discovery through User Testing

Citation to Publisher Version: Deodato, Joseph, Gambrell, Khalilah & Frierson, Eric. (2016). One Size Doesn't Fit All: Tailoring Discovery through User Testing. In K. Varnum (Eds.), *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content* (183-199). Chicago: ALA Editions.

Citation to *this* Version: Deodato, Joseph, Gambrell, Khalilah & Frierson, Eric. (2016). One Size Doesn't Fit All: Tailoring Discovery through User Testing. In K. Varnum (Eds.), *Exploring Discovery: The Front Door to Your Library's Licensed and Digitized Content* (183-199). Chicago: ALA Editions. Retrieved from doi:10.7282/T34T6MH2.

CHAPTER FOURTEEN

ONE SIZE DOESN'T FIT ALL

Tailoring Discovery through User Testing

JOSEPH DEODATO, KHALILAH GAMBRELL, AND ERIC FRIERSON

Implementing a web-scale discovery service can be a complex and challenging task, but the process doesn't simply end once your service goes live. In order to ensure that your new service is meeting the needs of its users, your implementation plan should include iterative user testing. Different library users have different needs, making it nearly impossible for any discovery product to offer a one-size-fits-all solution. Fortunately, most of these products offer a variety of configuration and customization options to help libraries tailor the experience to the needs of their users.

User testing is an effective and affordable method for libraries to configure their discovery service based on their users' needs and expectations. The following chapter offers practical guidance on how to apply insights from user testing to the customization of web-scale discovery services. However, customization alone only goes so far. To address the full range of challenges uncovered in usability testing also requires close collaboration between libraries and vendors to help develop services and solutions that become a part of future web-scale discovery service offerings. Accordingly, the authors recommend a collaborative three-pronged approach to improving web-scale discovery that includes product enhancement, local customization, and user education.

BACKGROUND

Like other academic libraries, the Rutgers University Libraries (RUL) turned to web-scale discovery services in the hopes of improving access to its collections. For years, users had expressed dissatisfaction with the libraries' search experience, which they found to be confusing, cumbersome, and fragmented. Web-scale discovery promised the ability to deliver a vast amount of high-quality, scholarly content from a single, easy-to-use interface. However, past disappointments with federated search had led some librarians to approach discovery with a bit more caution. RUL undertook a rigorous nine-month evaluation of the leading products that included extensive product and user research, an exhaustive list of product requirements, detailed correspondence with vendors and customers, and an extended period of product trials and testing.[1] At the conclusion of its proceedings, the evaluation team was certain of only one thing: none of the products on the market seemed truly capable of fully satisfying all of the libraries' needs. The linchpin of the team's analysis was a weighted evaluation rubric that was used to score each product in terms of coverage, performance, functionality, and technical architecture. Although one product scored highest overall, the final tally was remarkably close. The team found that each product had its own particular strengths and weaknesses and none had truly stood out from the pack. Library users, who participated in comparative usability testing of the three final candidates, seemed to reach an analogous conclusion. Although users had greater success and more favorable impressions of one product over the others, they struggled with similar problems in all three interfaces. While all agreed that these products made finding library materials easier, they also made it clear that there was ample room for improvement.

Ultimately, the evaluation team recommended that Rutgers acquire EBSCO Discovery Service (EDS). EDS had received the team's highest overall rating and was the favorite in user testing in addition to offering a full range of customization options. However, the evaluation team's recommendation came with two important caveats. First, the team recommended that additional usability testing be conducted to determine how to customize EDS for the best user experience. Second, the team suggested that Rutgers seek to work directly with EBSCO on testing so that findings could be shared among all stakeholders.

Coincidentally, the EBSCO User Research group had recently shifted from being an internal working group to a team that extends its expertise and services to customers. User Research has played a critical role in EBSCO's product development life cycle by using various user research methods to understand users' research workflows and in turn ensure that EBSCO products satisfy users' needs and expectations. The group has collaborated with several institutions to understand students' research workflows. Collaborations have been in the form of data analysis, surveys, contextual inquiries, and usability testing. Mostly

comprised of librarians with a background in user research, EBSCO User Research was eager to work with Rutgers to observe how students used EDS.

User testing is an important tool for improving library services, but its value is limited when divorced from the ability to control or influence the design of that service. Library-vendor collaboration is critical for improving the user experience of any licensed product or resource. Librarians gain a knowledgeable and committed user experience (UX) partner and vendors get direct feedback from real library users about their product. Furthermore, usability issues cannot be addressed by local customization alone. Some issues can only be solved through product enhancement while others require user education. For this reason, it's important that user testing involve all key stakeholders, including those responsible for developing, customizing, and teaching the library's web-scale discovery service. The following case study is intended to serve as an instructive example of how libraries and vendors might collaborate in user testing to address UX issues holistically using the three-pronged approach of product enhancement, local customization, and user education.

LITERATURE REVIEW

As the adoption of web-scale discovery services continues to rise, a growing body of literature has developed devoted to usability testing of these products. Some studies examine only one product while others offer comparative evaluations of multiple vendor offerings. Regardless of which product is the subject of evaluation, findings have tended to be remarkably consistent. The following is a summary of some common themes found within the usability literature on web-scale discovery services.

One commonly cited problem with web-scale discovery services is that users rarely have a clear understanding of scope and coverage. Studies suggest that students frequently have trouble determining what types of content are searchable within their library's discovery tool. For many users, these tools appear as a black box whose contents and inner workings remain a mystery. As Fagan et al. point out, librarians, themselves, do not always know what titles are covered or what level of indexing is offered by their discovery service provider because this information is not always readily available.[2] Majors, who conducted a comparative analysis of the leading commercial services, concluded that none of these products "offer any transparency about what is being searched and/or indexed, and user behavior reflected a trial and error methodology of figuring out what user tasks were actually supported."[3] In a search log analysis conducted at Montana State University, Meadow and Meadow found that users enter all types of unsupported queries into the library's discovery tool, including the titles of specific databases or journals, the names of frequently used services like

course reserves and interlibrary loan, and even the URLs of popular websites like Facebook.[4] As Teague-Rector and Ghaphery point out, a "larger, more visible search box may lead users to falsely assume that the library search functions like an online search engine."[5] Libraries need to make clear the purpose and scope of their discovery tool and ensure that resources and collections which fall outside of that scope are still discoverable. Lown, Sierra, and Boyer recommend that librarians think carefully about how discovery tools are presented on their website and alongside other valuable resources like the catalog and specialized databases.[6] Unlike online search engines, the goal of a library website is not merely to provide useful results but also to promote an understanding of the content, services, and tools that the library offers. Library search interfaces should help users differentiate between the kinds of resources available to them and guide them to those that will best serve their need.

Another frequently mentioned area of concern is search query construction. Although it may not come as a surprise, it turns out that most library users are not expert searchers. Studies suggest that the average user tends to rely on basic keyword searching, rarely modifies default settings or fields, and almost never uses Boolean logic or advanced search operators. This finding is not specific to discovery tools per se and is equally true for other library resources. This behavior is typically attributed to users' habituation to popular search engines where sophisticated ranking algorithms don't require precise queries and tend to be far more forgiving of error. For example, Asher, Duke, and Wilson's comparative study of EDS, Summon, Google Scholar, and conventional library resources found that students "treated almost every search box like a Google search box" using simple keyword searches in 81.5 percent (679/829) of the searches observed.[7] Of the 824 queries studied in Meadow and Meadow's search log analysis, only 27 (3.3 percent) utilized Boolean operators.[8] Foster and MacDonald note that even users who have attended one or two library instruction sessions tend to perform only the simplest of searches.[9]

One could argue that discovery tools are in fact designed to accommodate this sort of behavior. Users are expected to begin with broad keyword searches and progressively refine their results through the use of post-search facets and limiters. However, findings are somewhat mixed on whether or not users take full advantage of these functions. For example, Cassidy et al. assert that the majority of participants in their EDS study did not notice or attempt to use any facets or limiters in their searching.[10] In contrast, Fagan et al. found that their EDS users made frequent use of pre- and post-search refinements to limit searches. In separate studies of Primo, both Nichols and Comeaux noted that the use of facets and limiters increased as testing progressed. These findings suggest that users' awareness and use of search refinements may increase as they learn the system.

The lack of sophisticated user search strategies combined with broad, undefined content coverage produce large, heterogeneous result sets that require careful and attentive evaluation. However, many studies suggest that users often have difficulty interpreting results and distinguishing different types of information. While testing Summon, Gross and Sheridan found that users often had trouble distinguishing between a journal and a newspaper article or between a book and a book review.[11] These observations suggest that the blending of different types of information under one roof may pose serious challenges for novice users who are ill equipped to tell them apart and generally "see all information as being of equal value."[12] Indeed, researchers have shown that many users pursue evaluation superficially by skim-reading titles and abstracts and rarely investigating beyond the first page of results. In the comparative usability study conducted by Asher, Duke, and Wilson, 92 percent (598/649) of the resources used by students were found on the first page of search results.[13] The authors suggest that students attempt to compensate for their lack of evaluation skills by relying on the search engine to determine the quality of resources for them. Even when they know what characteristics to look for, they spend little time actually doing so. Instead, they prefer to put all of their faith in the ranking algorithm and expect the most useful and relevant results to be returned on the first page. As a result, users typically begin a new search rather than refine their current query or navigate to the next page of results.

The usability study conducted at Rutgers follows methods similar to those used in other studies and shares many of the same findings about user behavior. Its uniqueness lies primarily in the collaborative nature of the project, which involved the participation of both library and vendor representatives, as well as the holistic approach to resolving usability problems through a combination of product enhancement, local customization, and user education.

METHODS

Twelve Rutgers students and faculty members were recruited to participate in moderated remote usability testing. Participants were asked to use EDS to complete a series of tasks while sharing their thoughts aloud and having their verbal responses and on-screen activity recorded. Video transcripts of the sessions were analyzed to identify and address usability issues.

Testing was conducted using UserTesting (www.usertesting.com), a commercial online usability testing service. UserTesting offers a full-service user research platform that allows customers to set up usability tests and receive videos of users interacting with their website. Clients can test with their own users or take advantage of UserTesting's on-demand user panel. The researcher

dashboard provides useful metrics such as time on task and level of difficulty as well as tools to annotate videos and create highlight reels.

Remote testing was chosen for its ease, convenience, and flexibility. Testing remotely made it easier to reach a wider range of users dispersed across multiple campuses. Users were able to participate at a place and time of their own choosing as long as they had Internet access and a microphone. Remote testing is also arguably less obtrusive since it allows users to be observed in the comfort of their own environment rather than a lab or conference room. Finally, this method made it possible for Rutgers and EBSCO researchers to collaborate virtually without necessitating the time or expense of interstate travel.

Nielsen has famously argued that the majority of usability problems can be uncovered with just five users.[14] However, for the purposes of this study, the investigators wanted to obtain a sample that was large enough to represent the diversity of users that the libraries' discovery service is intended to serve. Sauro has suggested a formula for determining sample size based on the chance of detecting a problem during testing and its probability of occurring.[15]

Sample size = log(1-Chance of Detecting) / log(1-Probability of Occurring)

Using this formula, the investigators decided to recruit 12 participants in order to obtain an 85% chance of identifying any problems that impact 15% or more of users.

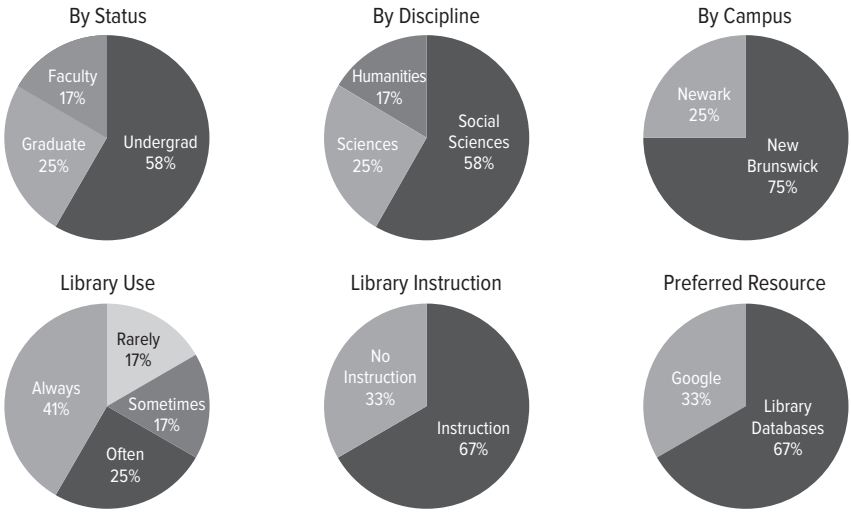
12 = log(1-.85) / log(1-.15)

Participants were recruited through flyers posted in all campus libraries as well as announcements on the libraries' website, Facebook, and Twitter pages. Users interested in participating in the study were asked to contact the investigators via e-mail and complete a brief online pre-test questionnaire. The questionnaire was used to collect demographic information and screen candidates to ensure the selection of a diverse sample representing different campuses, disciplines, and levels of education. The selected candidates were sent a URL with instructions on how to complete the test. Each participant received a $25 Amazon gift card courtesy of EBSCO.

Quota sampling was used to ensure a diverse group of participants. See figure 14.1 for a statistical overview of the sample. Since recruitment ads were posted exclusively within the libraries' physical and virtual spaces, it is highly likely that some amount of selection bias occurred. Generally speaking, the majority of volunteers appeared to be active library users who were well acquainted with library resources and services. Eight out of 12 users (67 percent)

FIGURE 14.1

Participant Overview



- said they "always" or "often" use library resources in their academic coursework or research
- had previously attended a library instruction class
- identified the library catalog or a library database as the resource they use most often when conducting research
- rated their ability to find information using library resources to be "good" to "excellent"

Participants were asked to complete five tasks designed to determine how well EDS supports core user needs related to finding, identifying, selecting, obtaining, and using information. Each task was communicated in the form of a hypothetical scenario to provide the participant with an interpretive context for completing the task. In order to test different aspects of the user interface, scenarios typically included two parts. The first part usually required locating a piece of information while the second involved doing something with it. For example, a participant might be asked to find an article on a topic as well as capture its citation for a bibliography. See figure 14.2 for a list of tasks used in testing.

Each session lasted approximately thirty minutes and was moderated by a UserTesting staff member. At the conclusion of the test, participants were asked to complete a brief online questionnaire asking them about their experience and satisfaction with EDS.

FIGURE 14.2

Tasks

Task	Scenario
1	You are writing a research paper that argues that social networking sites like Facebook are a threat to privacy. Find two scholarly articles that support your argument and e-mail them to yourself.
2	You are working on a group project about wearable technology. Find one article on this topic and decide how to share it with the other members of your group.
3	Find one news article about U.S. foreign debt published in the *New York Times* during the last three years. How would you capture citation information for this article to use in your bibliography?
4	Someone in your department recommended you read an article called "On Magic Realism in Film" by Fredric Jameson. Search for this article and explain how you would obtain a copy of it.
5	You are interested in the latest research on the use of artificial intelligence in perinatal medicine. Find one conference paper on this topic. From the results page, explain how you would broaden your search to find additional results from other sources.

RESULTS

Overall, EDS performed reasonably well, with users achieving an 80–100 percent success rate in three of the five tasks. Most users had little difficulty fulfilling the general requirements of these tasks, but few spent any significant time evaluating the results that were returned. Records were typically selected after a cursory glance and sometimes without a particularly compelling rationale other than the fact that their search terms were mentioned in the title. For example, although task 1 required users to find peer-reviewed articles, some selected articles from newspapers and magazines. It is not clear whether users would have taken more care in evaluation if they were selecting sources for a real assignment or project. Nor is it obvious that participants would have behaved any differently had they been searching a different resource.

Success rates were notably lower for tasks 3 and 5. Task 3 required users to construct a complex search query limited by date and publication. Many had trouble selecting the appropriate field options or limiters that would have yielded the desired result. Task 5 asked users to find one conference paper on perinatal medicine and then, from the results page, attempt to broaden their search to include results from additional sources. The goal here was to see if users would notice and understand how to use EDS's built-in federated search function to integrate results from databases that are not included in the central index. However, none of the users seemed to notice the invitation in the right column to "Expand Your Search" by selecting these additional resources. Instead, all users attempted to broaden their search by altering or omitting

a few of their original search terms. In designing the test, the investigators sought to avoid using leading language that might influence user responses, but it's possible that this part of the task might have been too vague. Figure 14.3 shows the average success rate and completion time for each task along with the standard deviation.

In the post-test questionnaire, the majority of participants indicated that they found EDS to be a helpful research tool and said that they would use it again. In their comments, users praised EDS for its simplicity, efficiency, and ease of use. They particularly appreciated the time-saving benefits of searching across multiple library resources. As one participant noted, EDS "made it easier to find articles and cut down my search time." Echoing similar sentiments, another said, "I liked the fact that you were able to get a lot of results all at once." Participants also liked the ability to refine searches using faceted browsing and noted the variety of useful ways to narrow results by subject, language, or date of publication.

However, much of this praise was tempered by reservations. As one participant put it: "It is an effort in the direction of making the university library website the go-to site for research . . . but it has some way to go." In particular, several participants experienced difficulty constructing searches and obtaining relevant

FIGURE 14.3

Task success rates

Task	Success		Time		Observations
	Mean	SD	Mean	SD	
1	80%	20%	6:03	3:35	The majority of users completed this task without difficulty, although some had trouble distinguishing between scholarly and non-scholarly sources.
2	100%	0%	2:46	0:42	All users completed this task without difficulty, but several commented that they were not aware or had never used the product's share functions.
3	60%	40%	5:45	2:20	A significant number of users had difficulty formulating an effective search query and/or limiting their results by publication.
4	80%	50%	3:14	1:04	Although most users were eventually successful, many experienced difficulty locating the given article, even when the exact title of the work was entered.
5	50%	10%	4:30	1:45	Most users experienced little difficulty with the first part of this task, but all failed on the second part requiring them to use federated search (which may not have been entirely clear in the instructions).

results. During testing, users often revised their searches multiple times in order to find the requested information. The following comments illustrate some of the frustration users felt when search results did not match their expectations:

> "I did not like the search engine itself and had some trouble finding certain articles when typing in specific authors."

> "It was difficult to figure out what to type in the search engine to help narrow down my search for certain articles."

> "I am most likely not seeing the best results on the first search. You have to modify your search about 2–3 times to find what you're looking for."

Several participants also commented on the design of the user interface. Many found it to be too cluttered, busy, or confusing. This seemed especially applicable to the search results screen. The following comments illustrate users' concerns with and difficulty navigating the interface.

> "The interface was not absolutely clear, but probably I just have to spend more time working with this system."

> "There is a lot going on . . . the tools on the side take away from the main part of the page . . . I didn't need any of it and I'm not even sure what some of the buttons mean."

> "I would have preferred the UI to be more easy and intuitive to use with fewer clicks."

> "The UI still needs to be simplified . . . Make the UI less cluttered visually."

DISCUSSION

This study generated a number of valuable insights about user search behavior that carry implications for those involved in developing, managing, and teaching web-scale discovery services. Based on these observations, the investigators formulated a series of recommendations aimed at improving the user's experience with EDS. Each recommendation was classified into one of three categories based on whether it required product enhancement, local customization, or user education. Figure 14.4 provides a summary of the team's observations and recommendations.

FIGURE 14.4

Observations and recommendations

1	Observation	Most users rely on basic keyword or natural language searching more often than Boolean or field searching.
	Recommendation(s)	[User Education] Improve user education on search query construction including keyword selection, field searching, search operators, facets, and limiters. [Customization] Change the default search mode from Boolean/Phrase to Find All My Search Terms, which seems to be more in keeping with how they expect search to function.
2	Observation	Most users find the interface to be too cluttered, particularly the right column of the search results screen.
	Recommendation(s)	[Customization] Remove the right column to simplify the search results screen and find alternative ways of integrating this content into the interface. Replace federated search with a database recommender system and incorporate database and research guide recommendations into a placard that appears at the bottom of the results screen. Replace static chat widget with a dynamic slide-out tab.
3	Observation	Most users have difficulty searching for known items even when the exact title of the article is entered.
	Recommendation(s)	[Product Enhancement] Improve searching and relevance ranking for known items. One possibility may be to boost the ranking of items where the user's search terms are found in proximity. [Update: Since testing, EBSCO Product Management has made significant relevancy ranking improvements to known item searching.]
4	Observation	Most users had difficulty interpreting terms such as "journal" and "peer reviewed."
	Recommendation(s)	[Product Enhancement] Change "Journal Title/Source" in advanced search field dropdown menu to "Publication." [Customization] Change "Journal Name" limiter in basic and advanced search options to "Publication." [Customization] Change "Peer Reviewed" limiter in basic and advanced search options and results list to "Scholarly/Peer Reviewed."
5	Observation	Most users did not navigate beyond the first page of results.
	Recommendation(s)	[Customization] Increase the number of search results per page from 10 to 20. [Product Enhancement] Replace results list pagination with infinite scrolling.
6	Observation	Some users could not accurately distinguish between different content formats.
	Recommendation(s)	[Product Enhancement] Improve icons and display of item formats. Consider including tool-tip format definitions. [User Education] Improve user education about content formats.

Figure 14.4 continued on next page.

FIGURE 14.4

Observations and recommendations (continued from previous page)

7	Observation	Some users did not notice when advanced search limiters from a previous search were being applied to their current search.
	Recommendation(s)	[Product Enhancement] Make applied limiters more prominent. Make advanced search limiters and database facets easier to remove from a search. [Update: Since testing, advanced search limiters are now displayed in and removable from breadbox. Database facet still is not.]
8	Observation	Some users did not notice system error messages notifying them of spelling errors or the need to enter search terms.
	Recommendation(s)	[Product Enhancement] Make system error messages more prominent by, for example, adding a border and/or background color.
9	Observation	Most users had little trouble identifying and using output tools such as e-mail, share, and export although some noted that they do not often use these options.
	Recommendation(s)	[User education] Promote and explain utility and function of output tools such as e-mail, share, and export.

As documented in other studies, users tended to rely on basic keyword or natural language searching more often than Boolean or field searching. They often entered several search terms on a single line without any search operators or fields. These simple queries worked well enough when the bar for precision was low, as in the first two tasks involving broad topical searches. However, users did not fare so well on the next two tasks where they had to find articles from specific publications or by specific authors. For example, when asked to find an article on U.S. foreign debt published in the *New York Times,* the most commonly used query was a keyword search for *u.s. foreign debt new york times.* A similar search strategy was employed when users were asked to find a copy of Fredric Jameson's article "On Magic Realism in Film." Several users' initial query was a variation of *on magic realism in film fredric jameson.* Users were often confused as to why these queries, which seemed both logical and precise to them, did not produce the results they expected. However, the default search mode in EDS is "Boolean/Phrase," which requires search terms to be found in proximity. Therefore, keyword searches combining authors, subjects, and publications in a single undifferentiated query string tend to fail because these terms are not typically found in proximity. Of course, both of the above queries work flawlessly in Google and it seems that users expected EDS to function in much the same way. To make EDS conform more to user expectations, the investigators recommended changing the default search mode to "Find All My Search Terms." This suggestion would remove the proximity requirement and automatically join all search terms in a Boolean AND statement. With this option set as the default, the initial search conducted by users given this task (the title followed by the author's

name as keywords, e.g., *on magic realism in film fredric jameson*) returns the desired article as the second result. The investigators also recommended that librarians devote more instruction and tutorials to keyword selection and search query construction as a means to improve overall user efficacy with search systems.

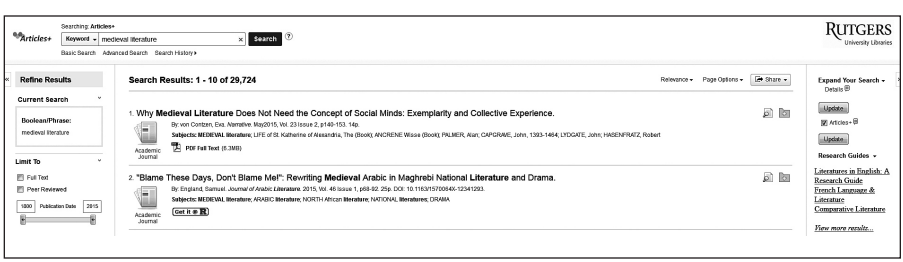
The study also found that users had difficulty searching for known items even when the exact title of an article was entered. Taking the example of the Fredric Jameson article again, many entered the exact title as a keyword search without quotation marks. However, this search strategy tended to be unsuccessful since the article does not appear until result #13 on the second page and most users never navigated beyond the first page of results. Incidentally, the ranking of the sought article actually dropped to #18 even after some users switch from a keyword to a title search. Unfortunately, none of the users utilized quotation marks to increase the precision of their query. A keyword search with quotation marks returned the article as result #4 while a title search with quotations bumped it to #1. Accordingly, the investigators recommended that EDS product management explore ways to improve searching for known items. Since testing, several enhancements have been implemented for known item searching including the following: (1) relevancy increase applied to exact title matches in library catalog, (2) autocomplete has been improved to support common misspellings, popular searches, and current topics, and (3.) a publication title placard displays above the results and allows a user to see the exact title match for a book or journal that the library has access.

To counter users' reluctance to probe beyond the first page of results, the investigators recommended customizing EDS to increase the number of results per page and for product management to consider infinite scrolling as a future enhancement.

Many study participants found the EDS results list to be cluttered. Therefore, the investigators recommended minimizing cognitive load by removing the right column of the search results screen and redistributing the content that previously occupied this space. The right column of the EDS search results screen functions as a placeholder for local add-on content. At Rutgers, the right column (see figure 14.5) was used to display EDS federated search option, a library chat widget,

FIGURE 14.5

Right-hand column display before usability testing



and recommended research guides. However, few participants seemed to use this content and many actually found it distracting. To reduce visual clutter on the interface, content from the right column was transitioned to EDS placards. Placards are call-out boxes that present the user with information related to his or her search and can be placed at the top, bottom, or somewhere in the middle of the results set. In Rutgers's case, the Springshare "Recommended LibGuides" right-column app was replaced with a placard that appears at the bottom of the results list when there are relevant research guides available (see figure 14.6). Rutgers is also considering known-search placards hard-coded to respond to frequent searches that are not intended for the discovery index (e.g., "JSTOR," "parking," "hours") with appropriate information or links.

The investigators found that users were either unaware or unsure of how to use EDS's federated search option. Task 5 asked participants to broaden the results of a preceding search to include results from additional sources. However, none of the users seemed to notice the "Expand Your Search" option that would allow them to incorporate results from external databases. Although the objective of the task may have been too vague for some, it seems fair to conclude that users would not choose this option without explicit instruction to do so. This finding is consistent with several other EDS studies. Williams and Foster, for example, found that users struggled to complete a task requiring them to use federated search without prompting from the test moderator.[16] Cassidy et al. noted that even after the purpose of this tool was explained to users, the majority were still unable to use it successfully.[17] Fagan et al. found that less than 1 percent of all EDS sessions over the course of a semester included any interaction with federated search.[18] With all of this in mind, the investigators recommended retiring federated search and replacing it with a database recommender system that guides users to specialized resources on their topic. It was suggested that these database recommendations be combined with the existing research guide recommendations into a placard that would appear near the bottom of the search results screen. This way, if the user scrolls through the first page of EDS results without finding anything of relevance, they are prompted to consider consulting a specialized database or research guide related to their topic.

FIGURE 14.6

Recommended research guides placard

Lastly, the investigators recommended using on-demand controls or modal windows to deliver other content such as the chat reference service so as to minimize use of screen real estate when this content is not needed (see figure 14.7).

Users had difficulty interpreting certain labels, especially when those labels were based on library jargon. For example, when asked to find an article published in the *New York Times,* many users struggled to find the appropriate field or limiter that would allow them to specify this. Although all users eventually made their way to the advanced search screen, most failed or were reluctant to use the limiter labeled "Journal" for this task. As one user rightfully pointed out, the *New York Times* is a newspaper, not a journal, so using this option didn't seem to make sense. If this wasn't problematic enough, the investigators soon noticed that the labels for this characteristic are not consistent throughout the interface. The advanced search limiter uses the term "Journal," but the field menu uses "Journal Title/Source" and the facets use "Publication." Accordingly, the investigators recommended that labels in all three locations be standardized to avoid confusion. "Publication" was recommended as the standard label because it was deemed more user-friendly than "Journal" or "Periodical."

Users also frequently overlooked important system messages during testing. This finding is consistent with Cassidy et al., who noted that users did not always notice EDS's autocorrect suggestions.[19] Similarly, messages indicating spelling errors or the need to input search terms often went unnoticed, even when these errors produced no results. Investigators recommended an EDS product enhancement to make these messages more prominent by, for example, adding borders or background colors.

Finally, as previous studies have suggested, users had notable difficulty interpreting results and distinguishing formats. For example, although task 1 required users to identify two scholarly articles, several participants selected non-scholarly sources. It is unclear whether these users truly did not understand the difference or were merely careless in their evaluation. For example, when one of these users was asked to explain their rationale, she simply noted that the titles of the selected articles contained the word she searched. Interestingly, only one user explicitly used the peer-review limiter to complete this task, which led the

FIGURE 14.7

Right-hand column after usability testing

investigators to recommend changing the label in order to clarify its purpose. Task 5 required users to identify a conference paper and, while all users were successfully able to use the source type limiter to find the requested format, two undergraduate participants admitted that they did not know what a conference paper was.

Recommendations resulting from this study led to further customization of EDS at Rutgers University as well as recent enhancements to the product as a whole. While these changes have yet to undergo another round of user testing, they represent a critical part of iterative testing and improvement by both the library's discovery service administrator as well as the discovery vendor.

CONCLUSION

Web-scale discovery services, while a welcome improvement to the library search environment, are not a panacea. There is still ample room for improvement through product development and local customization. However, the findings from this study and observations at other institutions suggest that library instruction must also play a prominent role in improving users' experiences with discovery services. In particular, users need guidance in developing effective search strategies, distinguishing between different formats, and critically evaluating the quality of sources.

Discovery services often come with out-of-the-box configurations that may not be perfectly suited for every library. However, with the ability to change settings and apply custom code to the user interface, the opportunity for improvement and innovation is high. Librarians should develop a plan of iterative testing and collaborate with vendors in order to address the weaknesses and challenges in discovery systems. While customization options can solve many user experience problems, advancing the state of the art will require vendor commitment to improving their products as well as instructing users on how to get the most out of them.

Notes

1. Joseph Deodato, "Evaluating Web-Scale Discovery Services: A Step-by-Step Guide," *Information Technology and Libraries* 34, no. 2 (2015): 19–75. doi: 10.6017/ital.v34i2.5745.
2. Jody Condit Fagan et al., "Usability Test Results for a Discovery Tool in an Academic Library," *Information Technology and Libraries* 31, no. 1 (2012): 83–112. doi: 10.6017/ital.v31i1.1855.
3. Rice Majors, "Comparative User Experiences of Next-Generation Catalogue Interfaces," *Library Trends* 61, no. 1 (2012): 186–207. doi:10.1353/lib.2012.0029.

4. Kelly Meadow and James Meadow, "Search Query Quality and Web-Scale Discovery: A Qualitative and Quantitative Analysis," *College & Undergraduate Libraries* 19, nos. 2–4 (2012): 163–175. doi: 10.1080/10691316.2012.693434.
5. Susan Teague-Rector and James Ghaphery, "Designing Search: Effective Search Interfaces for Academic Library Websites," *Journal of Web Librarianship* 2, no. 4 (2008): 479–492. doi: 10.1080/19322900802473944.
6. Cory Lown, Tito Sierra, and Josh Boyer, "How Users Search the Library from a Single Search Box," *College & Research Libraries* 74, no. 3: (2013): 227–41. doi: 10.5860/crl-321.
7. Andrew D. Asher, Lynda M. Duke, and Suzanne Wilson, "Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources," *College & Research Libraries* 74, no. 5 (2013): 464–88. doi: 10.5860/crl-374.
8. Meadow and Meadow, "Search Query Quality and Web-Scale Discovery."
9. Anita K. Foster and Jean B. MacDonald, "A Tale of Two Discoveries: Comparing the Usability of Summon and EBSCO Discovery Service," *Journal of Web Librarianship* 7, no. 1 (2013): 1–19. doi: 10.1080/19322909.2013.757936.
10. Erin Dorris Cassidy et al., "Student Searching with EBSCO Discovery: A Usability Study," *Journal of Electronic Resources Librarianship* 26, no. 1 (2014): 17–35. doi: 10.1080/1941126X.2014.877331.
11. Julia Gross and Lutie Sheridan, "Web Scale Discovery: The User Experience," *New Library World* 112, no. 5 (2011): 236–47. doi: 10.1108/03074801111136275.
12. Ibid.
13. Asher, Duke, and Wilson, "Paths of Discovery."
14. Jacob Nielsen, "Why You Only Need to Test with 5 Users," *Nielsen Norman Group,* March 19, 2000, www.nngroup.com/articles/why-you-only-need-to-test-with-5-users.
15. Jeff Sauro, "How to Find the Right Sample Size for a Usability Test," *MeasuringU,* December 7, 2011, www.measuringu.com/blog/sample-size-problems.php.
16. Sarah C. Williams and Anita K. Foster, "Promise Fulfilled? An EBSCO Discovery Service Usability Study," *Journal of Web Librarianship* 5, no. 3 (2011): 179–98. doi: 10.1080/19322909.2011.597590.
17. Cassidy et al., "Student Searching with EBSCO Discovery."
18. Fagan et al., "Usability Test Results for a Discovery Tool in an Academic Library."
19. Cassidy et al., "Student Searching with EBSCO Discovery."