Seek and You May Find: Successful Search in Online Finding Aid Systems

Morgan G. Daniels and Elizabeth Yakel

Abstract

This article reports on a series of retrieval experiments using online finding aid systems in two archival institutions to identify heuristics for successful search in archival systems. The project analyzes the success rates of search strategies of forty-three participants and finds that highly successful searchers possess prior archival experience and use both Boolean searches and the browser's page find functionality. Successful searchers also utilize a broader range of search strategies and query reformulation techniques. Many searchers had difficulty choosing appropriate search terms and understanding archival terminology. The results of this study have implications for archival instruction, federated search, and the design of interfaces for online finding aids.

As archives put more finding aids and digital representations online, the number of people potentially accessing archives and manuscripts increases dramatically. Online users of an archives can enter through its Web page or through a search engine. Online users bring a wide range of expectations and some are perhaps confused by finding themselves in the middle of a finding aid. Designing systems to facilitate resource discovery by this ever-broadening archival audience is difficult.[1]

[1] For evidence of the increasing audiences the Web has brought to archives, see the Council of State Archivists, *State of State Archives: A Status Report on State Archives and Records Management Programs in the United States* (2007), available at http://www.statearchivists.org/reports/2007-ARMreport/StateARMs-2006rpt-final.pdf, accessed 7 February 2010. See also Amanda Hill, "Serving the Invisible Researcher: Meeting the Needs of Online Users," *Journal of the Society of Archivists* 25, no. 2 (2004): 139–48, available at http://www.informaworld.com/smpp/content~content=a713996502&db=all, accessed 14 May 2010.

Library retrieval studies propose a series of heuristics for search strategies and query reformulation. *Search strategy* refers to an overall approach to looking for information, including choice of databases, working from broader to narrower searches, and citation tracing. *Query reformulation* refers to the changes made to a search—such as use of synonyms, Boolean logic, or system-designated delimiters—to retrieve a set of results different from those obtained by an initial query. Whether these heuristics apply to archives, or how they apply, has not been tested.

Archival repositories describe their materials differently than do libraries, but they often use similar search interfaces. Users unfamiliar with primary sources but who have used online library catalogs may not realize the differences between library catalogs and cataloging practices and those of online finding aid systems. Similarly, when using archival finding aid systems, users may expect to encounter the kind of search structure, capabilities, and results they encounter with Web search engines. By understanding the search techniques that users bring to online archival collections, we can offer recommendations for increasing the ease of use of search systems through improved interface design and for the most advantageous search strategies for researchers to pursue when using these systems. This study explores the following questions:

- How do people search archival collections online?
- What are the characteristics of a successful search?
- Are there optimal search strategies for accessing information about primary sources in online finding aid systems?
- What role does query reformulation play in a successful search of online finding aids?

Literature Review

While there are numerous studies on search tactics for online library catalogs, library databases, and the Web, investigations of online archival systems are more concerned with usability issues. At the same time, many studies look at the general information-seeking patterns of historians and others who use archives, but not specifically in an online context. For example, Wendy Duff and Catherine Johnson identify four distinct information-seeking activities: 1) orienting oneself to archives, finding aids, sources, or a collection; 2) seeking known material; 3) building contextual knowledge; and 4) identifying relevant material.[2] However, they do not describe what constitutes identification in detail. In a complementary study, Margaret Dalton and Laurie Charnigo report that study participants ranked finding aids and library catalogs first and third among a long list of sources for finding primary source materials, but we do not know the

---

[2] Wendy M. Duff and Catherine A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (October 2002): 472–96.

search strategies they employed to locate the library catalogs and finding aids nor the heuristics they used to search within them.[3]

Only a few studies examine users' search strategies for archival materials. Charles Cole studied the behavior of researchers searching through primary sources, rather than searching finding aids or other representations. Although the name collection technique he identifies would also be a viable approach to online systems, his study is limited to the search activity of PhD students, therefore generalizations cannot be made to the behavior of other types of users.[4] Susan Hamburger's study is an interesting corollary to Cole's as it examines methods of resource discovery. She surveyed a cross-section of researchers— undergraduate and graduate students and other researchers (*n*=131)—from 6 research libraries, concluding that most online search strategies rely on proper names and that users are often dissatisfied with their results.[5] Much of the literature in library and information studies concerning searching by historians and others who might use archives addresses their search strategies for secondary sources, rather than for archival materials.[6]

Over the last decade and a half, studies of online finding aids have increased and are of 2 types. The first type does not involve users directly but instead raises important issues about the interfaces and functionalities of finding aid systems based on content analyses of interfaces. The second group is comprised of actual usability studies, which can be further grouped into two genres: 1) independent usability tests, and 2) tests by developers to evaluate and improve their sites. All of these studies focus on the user interface and how to design interfaces that are more intuitive for users of archival finding aids. Although our study employs usability testing techniques, we investigate how researchers search for archival materials in existing finding aids systems and the implications of these search techniques for archival systems.

Three content analyses of interfaces for online finding aids reveal significant problems in their design and use of terminology. Jihyun Kim notes important differences in these interfaces, such as varied use of data elements, labeling, and browsing categories. She also identifies limited search functions in many systems.[7] Xiaomu Zhou judges search features to be poor in many archival

---

[3] Margaret Steig Dalton and Laurie Charnigo, "Historians and Their Information Sources," *College and Research Libraries* 65, no. 5 (September 2004): 407.

[4] Charles Cole, "Name Collection by Ph.D. History Students: Inducing Expertise," *Journal of the American Society of Information Science*. 51, no. 5 (2000): 444–55.

[5] Susan Hamburger, "How Researchers Search for Manuscript and Archival Collections," *Journal of Archival Organization* 2, nos. 1–2 (2004): 79–102.

[6] A typical example is Donald Owen Case, "The Collection and Use of Information by Some American Historians: A Study of Motives and Methods," *The Library Quarterly* 61, no. 1 (January 1991): 61–82.

[7] Jihyun Kim, "EAD Encoding and Display: A Content Analysis," *Journal of Archival Organization* 2, no. 3 (2004): 41–55.

interfaces; in some cases, entirely absent.[8] Rita Czeck compares subject terms in archival MARC records and online finding aids; she finds little consistency and argues the importance of full text searching in all sections of the finding aid.[9] This combination of inconsistent terminology and varying functionality does not bode well for users.

Studies involving users of online finding aid systems corroborate and broaden the findings on interface design and terminology in these content analyses. Although there are a number of these studies, their diverse methodologies and findings make generalizations difficult. The earliest study of online finding aid systems was Burt Altman and John Nemmers's focus group, which identified both usability and archival terminology issues in online finding aids.[10] Other findings confirm their claims and extend the list of usability issues. Yakel identifies problems with the hierarchical structure of finding aids and users' lack of familiarity with the parts of a finding aid and its terminology, which makes navigation between the left and right frames of online finding aids confusing.[11] Christopher Prom focuses on the differences between novice and expert users of several online access systems. All of his participants did well on the more straightforward interfaces. Confirming earlier studies, Prom also finds that the novices were unfamiliar with archival terminology and the hierarchical presentation of the finding aids. Participants also made substantial use of the browser's "find in page" (Ctrl+F) function. Almost half of the 35 participants in 4 of Prom's 5 sites regularly used "find in page."[12] Wendy Scheir focuses exclusively on novices and her findings echo the earlier studies on terminology and structure. However, Scheir did observe that participants experienced a learning curve during the experiment. Participants reported greater confidence and more ease in using the system as the study progressed.[13] Dawne Howard focuses on the positioning of the frame with the navigation bar/container information. She finds that the position of the navigation bar vis-à-vis the actual content of the finding aid has little effect on navigability.[14]

[8] Xiaomu Zhou, "Examining Search Functions of EAD Finding Aids Web Sites," *Journal of Archival Organization* 4, nos. 3–4 (2006): 99–118.

[9] Rita L. H. Czeck, "Archival MARC Records and Finding Aids in the Context of End-User Subject Access to Archival Collections," *American Archivist* 61 (Fall 1998): 426–39.

[10] Burt Altman and John R. Nemmers, "The Usability of On-line Archival Resources: The Polaris Project Finding Aid," *American Archivist* 64 (Winter 2001): 121–31.

[11] Elizabeth Yakel, "EAD: Are Finding Aids Boundary Spanners or Barriers for Users?," *Journal of Archival Organization* 2, nos. 1–2 (2004): 63–77.

[12] Christopher J. Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2 (Spring 2004): 258.

[13] Wendy Scheir, "First Entry: Report on a Qualitative Exploratory Study of Novice User Experience with Online Finding Aids," *Journal of Archival Organization* 3, no. 4 (2005): 49–85.

[14] Dawne E. Howard, *The Finding Aid Container List Optimization Survey: Recommendations for Web Usability*, master's paper for the master's degree in library science at the University of North Carolina (2006): 18, available at http://ils.unc.edu/MSpapers/3223.pdf, accessed 1 September 2009.

In addition to such peer-reviewed usability studies, major consortia, such as the Online Archive of California (OAC) and the Northwest Digital Archives (NWDA), have tested the usability of online finding aids. The most active is the Online Archive of California, which completed 4 rounds of usability testing between 2001 and 2009, both summative, to improve existing interfaces, and formative, to design new interfaces.[15] As a result, the OAC improved search functions and display. The latest round of usability testing led to an entire redesign of the interface for the online finding aids, which was released in June 2009.[16]

NWDA has also done several rounds of usability testing, finding that study participants wanted more detail, more image content, and less text.[17] These results are contradictory, to say the least, as many of the subjects claimed that they desired sufficient detail

> to get started on their research, which meant enough to see if the collection is likely to contain relevant information and details about what is in each box. However, while detail was expected, many still said they would rather not have to read long blocks of text.[18]

---

[15] *Online Archive of California, OAC Usability Test Summary* (October 2001), http://www.cdlib.org/inside/ assess/evaluation_activities/oac_usabilitytest2001_summary.pdf, accessed 31 March 2010; *Online Archive of California, OAC Usability and Survey Results Summary* (July 2002), http://www.cdlib.org/ inside/assess/evaluation_activities/OACUsabilitySummaryJuly2002.pdf, accessed 31 March 2010; Jane Lee, preparer, *Online Archive of California, OAC First Round Usability Test Findings: OAC Redesign Project* (11 September 2008), available at http://www.cdlib.org/inside/assess/evaluation_activities/ docs/2008/oac_usability_aug2008.pdf, accessed 31 March 2010; Jane Lee, preparer, *Online Archive of California, OAC Second Round Usability Test Findings OAC Redesign Project* (23 June 2009), available at http://www.cdlib.org/inside/assess/evaluation_activities/docs/2009/oac_usability_april2009.pdf, accessed 31 March 2010.

[16] See http://www.oac.cdlib.org/, accessed 3 June 2010.

[17] Tiah Edmunson-Morton, UTWG chair, comp., *Northwest Digital Archives, Executive Summary: Usability Testing Round 4* (12 March 2008), available at http://orbiscascade.org/index/cms-filesystem-action?file=nwda/reports/nwda_utwg_ut4_report_20080312.pdf, accessed 31 March 2010; *Northwest Digital Archives, Usability Design Working Group, Executive Summary Usability Test #5—Search Functionality* (6 November 2008), available at http://orbiscascade.org/index/cms-filesystem-action?file=nwda/reports/udwg_ut5_executive_summary_20081106.pdf, accessed 31 March 2010; Jodi Allison-Bunnell, NWDA program manager, comp., *Northwest Digital Archives, Survey of Digitizing Initiatives: A Report on Current and Desired Digital Initiatives Among Orbis Cascade Alliance and Northwest Digital Archives Members* (20 November 2007), available at http://orbiscascade.org/index/cms-filesys-tem-action?file=nwda/files/di survey report 1107.doc, accessed 31 March 2010; Jodi Allison-Bunnell, NWDA Program Manager, comp., *Northwest Digital Archives, Researcher Survey Narrative Report*, 23 January 2008, available at http://orbiscascade.org/index/cms-filesystem-action?file=nwda/files/researcher_ survey_narrative_20080123.pdf, accessed 31 March 2010; Jodi Allison-Bunnell NWDA program manager, comp., *Northwest Digital Archives, Digital Initiative Planning: Institutional Needs Survey*, 25 October 2008, available at http://orbiscascade.org/index/cms-filesystem-action?file=nwda/files/institutional_ needs_survey_report_20081025.pdf, accessed 31 March 2010.

[18] Edmunson-Morton, *Northwest Digital Archives, Executive Summary: Usability Testing Round 4*, 1.

In later usability testing focused primarily on the search interface, NWDA found that participants incorrectly identified the browse options as search delimiters.[19]

These studies of online finding aids share several common findings. First, users are confused by archival terminology and practice, particularly the organization of information in the finding aid. Second, this confusion leads to problems navigating through online finding aids and the use of the "find in page" work-around to deal with the large amounts of text. Finally, real differences appear to exist between novices and more experienced users, which affect the use of online finding aid systems. We will elaborate on these themes in this article as we examine search tactics and strategies as well as query reformulation and how these do or do not lead to success.

There have been few retrieval studies in archives. Tim Hutchinson examines the retrieval effectiveness of different methods of searching online findings aids: 1) searching entire finding aids; 2) searching introductory material to finding aids; 3) searching introductory material to finding aids enhanced by controlled vocabulary terms; and 4) searching collection-level catalog records. He finds that searching entire finding aids improves recall but decreases precision, while searching different sections of the finding aid increases precision, giving users more targeted search results.[20] Fernanda Ribeiro's retrieval study demonstrates how controlled vocabulary could be used to bridge linguistic problems in finding aids and in the underlying collections.[21]

Studies of information retrieval behaviors in libraries and on the Web provide further insights into the search strategies of novice and expert searchers, and provide a useful point of comparison for this study. Karen Drabenstott examines the search strategies of nondomain experts (14 undergraduate students) using a university library website and compares them to strategies employed by domain experts. Experts tend to use advanced search techniques, such as footnote chasing, scanning issues in a journal run, and author searching, when looking for information within their field.[22] Drabenstott observed the behaviors of undergraduate students conducting searches, primarily in their major area of study. Only 5 students used one or more strategies characteristic of domain experts. Overall, these techniques made up a small part of their

[19] *Northwest Digital Archives, Usability Design Working Group, Executive Summary Usability Test #5—Search Functionality.*

[20] Timothy Hutchinson, "Strategies for Searching Online Finding Aids: A Retrieval Experiment," *Archivaria* 44 (Fall 1997): 72–101.

[21] Fernanda Ribeiro, "Subject Indexing and Authority Control in Archives: The Need for Subject Indexing in Archives and for an Indexing Policy Using Controlled Vocabulary," *Journal of the Society of Archivists* 17, no. 1 (1996): 27–54.

[22] Karen M. Drabenstott "Do Nondomain Experts Enlist the Strategies of Domain Experts?," *Journal of the American Society for Information Science and Technology* 54, no. 9 (2003): 836–54.

information retrieval strategies. Instead, the students predominantly employed subject searching, which novices often use in a domain because they are unfamiliar with the relevant authors and journals in a field and have difficulty evaluating the quality of the sources they find. Undergraduates searched subjects in 106 queries; by comparison, they used domain-expert strategies in only 20 queries.[23] Furthermore, their subject searches were frequently unsuccessful: keywords entered yielded no hits, they received error messages from the system, or, upon browsing a list of subject headings, they were not able to select one appropriate to their inquiry.

Query reformulation—modifying a question to change the search results—is another search strategy dealt with at length in the library and Web search literature. Soo Young Rieh and Hong (Iris) Xie authored a recent study focusing on query reformulation on the Web. Using search logs from 313 search sessions on the Web search engine Excite, Rieh and Xie examined query reformulations and find that overlapping term meanings from one search iteration to the next, rather than narrowing or broadening a search, characterized most of them. Synonym generation accounted for only 4.9% of the changes to the content of a query. Searchers most often changed the format of a query by altering the format of terms (exchanging acronyms for proper names, for example) (37.5%) or correcting errors (36.2%) rather than by using Boolean operators (26.3%).[24]

These authors identify a number of search strategies unavailable to researchers using archival systems. For example, many of the techniques experts employ, such as citation tracing, journal investigation, and author search, are either not applicable to archives or are not present in the same way in archives as they are in libraries, requiring even domain experts to change developed search routines. However, other strategies reported in the Drabenstott and Rieh and Xie studies are useful for considering searches in archives: broader to narrower search, synonym generation, and the use of Boolean operators. These aspects of query reformulation may have greater significance for archives. In addition, the name collection strategy noted by Cole is appropriate for primary source research, but may not be useful in other contexts. Given the difference between archival and other types of searches, we were interested in understanding how researchers adapt their existing search methods to archival systems. Thus, in this study we explore this question through an analysis of successful search techniques.

---

[23] Drabenstott, "Do Nondomain Experts Enlist the Strategies of Domain Experts?," 849.

[24] Soo Young Rieh and Hong (Iris) Xie, "Analysis of Multiple Query Reformulations on the Web: The Interactive Information Retrieval Context," *Information Processing and Management* 42, no. 3 (2006): 751–68.

M e t h o d o l o g y

To better understand the ways people search for primary sources, this series of experiments tested users' abilities to search and retrieve records from 2 online archival databases: the Bentley Historical Library's Online Finding Aids system at the University of Michigan and the federated statewide search tool at the Online Archive of New Mexico (OANM). (See Figures 1 and 2.) The interfaces and functionalities of these two systems provided the requisite diversity to examine participants' search and retrieval behavior under different conditions.[25]

Several major differences between the 2 systems merit discussion. First, the Online Archives of New Mexico is designed for federated search among a number of repositories, but the Bentley Historical Library site provides access to information for a single repository. The 2 sites had different back-end systems and substantially different search functions. For example, the Bentley system allowed for basic and advanced search options and Boolean operations; the OANM search interface had fewer delimiters ("this phrase," "all of these words," and "any of these words"). Records in these systems had different types of digital objects attached to them; accordingly, a Bentley task involved locating a digital textual document, while in OANM we asked participants to find a photograph. The digital objects were located at different levels of their respective hierarchies, allowing us to see if users could figure out how to drill down into the system to find them. Both of these search engines have changed since we conducted the experiments in 2005.

The study consists of three parts: 1) a survey, 2) the actual retrieval experiments, and 3) a post-test interview. The entire protocol lasted approximately 1 hour, although the time for each participant varied. The shortest amount of time to complete the experiments was 36 minutes and 15 seconds; the longest was 93 minutes and 18 seconds. The average time to complete the group of search tasks was 54 minutes and 13 seconds. The experiments were conducted in an office that had been converted into a laboratory.

The survey (see Appendix A) gathered basic demographic information such as age and academic status, library and archival experience (including frequency of use of online systems), and self-rated technology expertise. Project personnel developed the library and archival experience questions, which had been used in previous studies. These questions were intended to test whether familiarity with libraries and archives had any effect on understanding archival records or searching effectiveness. We borrowed the scales for the technology

[25] See http://bentley.umich.edu/ and http://oanm.unm.edu/, accessed 3 June 2010.
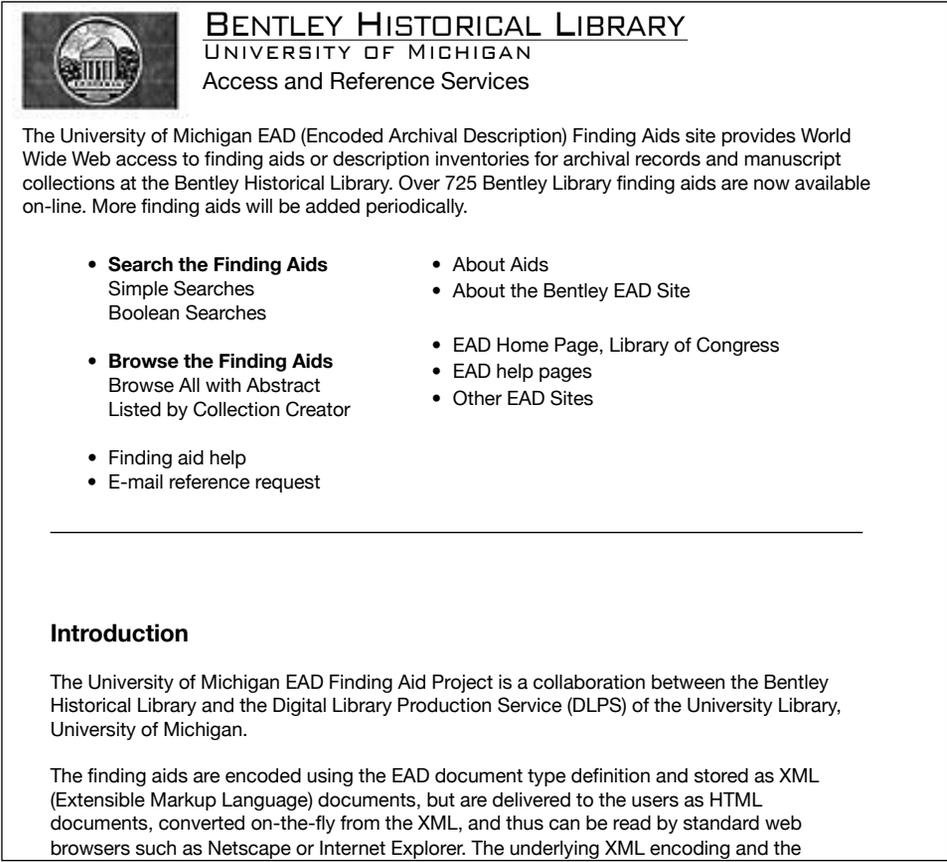
BENTLEY HISTORICAL LIBRARY
UNIVERSITY OF MICHIGAN
Access and Reference Services

The University of Michigan EAD (Encoded Archival Description) Finding Aids site provides World Wide Web access to finding aids or description inventories for archival records and manuscript collections at the Bentley Historical Library. Over 725 Bentley Library finding aids are now available on-line. More finding aids will be added periodically.

- **Search the Finding Aids**
  Simple Searches
  Boolean Searches

- **Browse the Finding Aids**
  Browse All with Abstract
  Listed by Collection Creator

- Finding aid help
- E-mail reference request

- About Aids
- About the Bentley EAD Site

- EAD Home Page, Library of Congress
- EAD help pages
- Other EAD Sites

Introduction

The University of Michigan EAD Finding Aid Project is a collaboration between the Bentley Historical Library and the Digital Library Production Service (DLPS) of the University Library, University of Michigan.

The finding aids are encoded using the EAD document type definition and stored as XML (Extensible Markup Language) documents, but are delivered to the users as HTML documents, converted on-the-fly from the XML, and thus can be read by standard web browsers such as Netscape or Internet Explorer. The underlying XML encoding and the

FIGURE 1. The Bentley Historical Library interface as it appeared at the time of the study in 2005.

self-assessment from the Flashlight Current Student Inventory, a validated questionnaire widely used in colleges and universities.[26]

The actual retrieval testing consisted of 5 tasks specially designed for each system, so that participants completed 10 tasks each. The tasks were constructed so that participants would have to use and understand several generic elements in online finding aid systems, such as the relationship between the left side navigational frame and the right side content frame, and the archival terminology found in collection descriptions. The tasks tested participants' use of search delimiters and Boolean operators, and their generation of search terms. Participants performed typical tasks that an archivist would want done using the system, such as filling out a call slip. Participants were alternately asked to complete recall (find everything about…) and precision (find a known item)

---

[26] S. C. Ehrmann, and R. E. Zúñiga, *The Flashlight Evaluation Handbook: Including the Flashlight Current Student Inventory,* version 1.0 (Washington, D.C.: Teaching, Learning, and Technology Group, American Association for Higher Education, 1997).
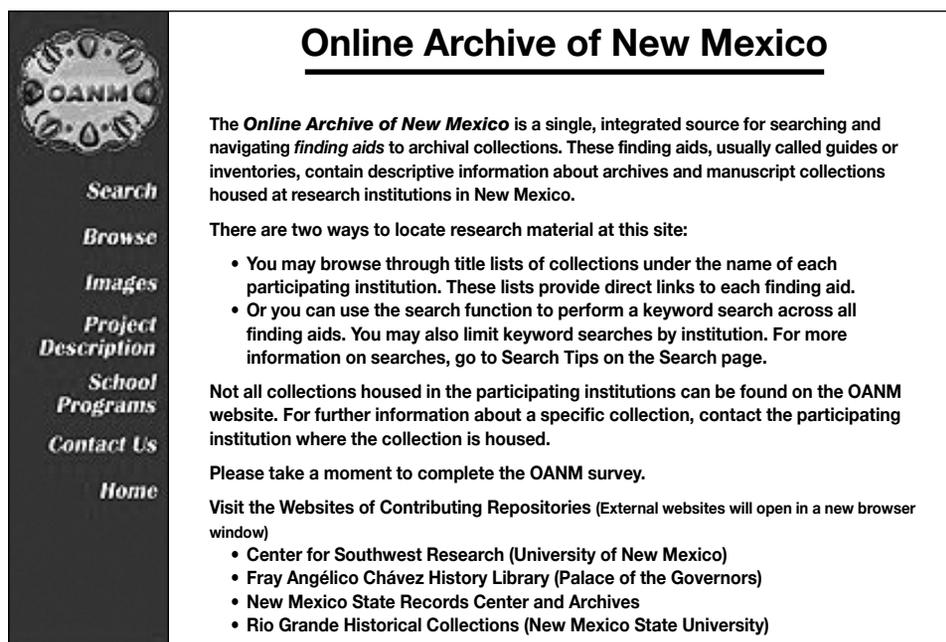
FIGURE 2. The Online Archive of New Mexico interface as it appeared at the time of the study in 2005.

searches throughout the tasks. These two types of searches mirror the kinds of search tasks that frequently take place in archives: 1) recall searches are used to find everything available on a certain topic, and 2) known-item searches are used to find something more specific, such as property records or photographs. Additionally, as part of the task development process, we identified correct answers so we could compare participants' results. Prior to the actual experiment, we pilot tested each of the tasks.

Table 1 gives the text of each of the 10 tasks, the system in which each task was conducted, and the archival knowledge tested by each. Tasks that tested knowledge of archival terminology required participants to identify series in a collection, as in task 8. Tasks 3 and 7 tested knowledge of archival organization, by asking participants to identify the box and folder numbers in which documents might be found. These tasks also tested the subjects' ability to distinguish between contextual (biographical or administrative history) and content (scope and content note) information. Tasks 2 and 9 required an understanding of search results particular to archives; both yielded results in which the correct response was not immediately identifiable but required digging through the relevant finding aid. Finally, a number of tasks tested knowledge of the structure of a finding aid.

Participants were presented with tasks similar to those encountered in library catalogs (e.g., finding materials on a subject), as well as tasks unique to archives (e.g., locating materials on people who are neither named creators nor subjects; managing and searching large blocks of text). While we wondered whether participants applied library search techniques to archival finding aids, we were particularly interested in their improvisation of search tactics specific to archives.

One limitation of the study was that tasks were administered in the same order for each participant. We did this for several reasons. During pilot testing we saw no order effects or improvement during the session as a result of increasing familiarity with the system. The tasks remained hard for participants to complete. Second, we wanted to begin with the easiest task on each system to introduce it to participants.

For this part of the experiment, data were captured using Camtasia to record keystrokes and audio of the accompanying "think-aloud" dialogue, as well as through videotape of the sessions to assess body language and other contextual information. Although we asked participants to think aloud during the actual search tests, doing so generated only some information about their rationale and strategies. Therefore, we conducted a post-test interview (see Appendix B) asking about the strategies that participants used to answer specific questions as well as their reactions to the search interfaces. During this interview, we often reset the computer screen and referred to specific tasks and actions the participant had taken.

We recruited 5 different types of participants for this experiment: academic historians (5), graduate students (8), genealogists (9), undergraduate students (11), and reference librarians (10). Participants were recruited through several means: posters placed in various humanities departments and the school of education on a university campus, announcements made to genealogical groups, and an email announcement sent to librarians at several universities. In total, we recruited 43 participants. They were predominantly (80%) female ranging in age from 19 to 71 years old and averaging 37. Their educational levels ranged from high school to PhD. However, none of these demographic factors proved to be important determinants of search success.[27] Participants with these characteristics (age, educational level, and sex) were evenly distributed among the 3 search success groupings discussed below. Table 2 shows the distribution.

---

[27] We did chi-square tests to determine whether the distribution of participants in the groupings differed by gender and highest degree attained. Neither of these proved to be significant: Gender: $X^2(2, n = 43) = 0.36$, $p = .83$; Highest degree: $X^2 (8, n = 43) = 7.25$, $p = .51$. Since there were cells with fewer than 5 observations, we also performed exact tests (Gender: $p = .9$; Highest degree: $p = .3$ both using the Freeman-Halton extension of the Fishers exact test). A one-way ANOVA was used to test for difference in the mean age among members of the 3 success groupings. Age also did not differ significantly across the 3 groups, $F (25, 17) = 1.014$, $p = .499$.

Table 1. Tasks Assigned to Study Participants

Task	Site	Type	Archival Knowledge Tested				
			Archival Terminology	Archival Organization	Distinguish Between Context and Content	Understand Search Results	Structure of a Finding Aid
1. Find collections concerning the Black Panther Party • List collections and box numbers of all relevant materials	BHL	Recall			X		X
2. Find a digital copy of a speech given by a former U-M president • Find the URL for the digital document • List series in the digital documents section	BHL	Precision	X	X		X	
3. Locate the Ferry Family Papers and locate specific correspondence • Name correct collection and box number for correspondence	BHL	Precision		X			X
4. Find collections concerning rock and roll in Detroit • Determine if the archives has collections on this topic • List all collections containing materials	BHL	Recall			X		X
5. Find former Michigan governor Blanchard's papers • Find the "blue folders" subseries; explain its arrangement • Fill out a call slip for materials on prescription drugs for seniors	BHL	Precision	X	X			X
6. Find Kit Carson's Oath of Allegiance to the U.S. • Find the date of Kit Carson's Oath of Allegiance to the U.S. • Fill out a call slip to retrieve this item	OANM	Precision		X	X	X	X
7. Locate Antonio Terrazas's wedding photo • Find the URL for the digital photo • Give the collection name, box, and folder for the print copy of the photo	OANM	Precision		X			
8. Find the Yrissari Family papers • Find this collection and list the series within it	OANM	Precision	X	X			X
9. Locate records of the sale of land involving Joseph Barton • Give the collection's call number • Find the location of a specific deed within the collection • List the lot and block number for the deed	OANM	Precision				X	X
10. Find collections concerning slavery • Find the repository with the largest number of collections on this topic • List which collections have 5 or more references to actual archival materials	OANM	Recall		X	X	X	X

Table 2. Distribution of the Sex and Age of Participants by Success Grouping

Success Grouping	Sex			Age	
	Female	Male	Total	Mean	n
Low: Below 25th percentile	8 72.7%	3 27.3%	11 100%	38.00	11
Medium	17 81.0%	4 19.0%	21 100%	36.33	21
High: Above 75th percentile	9 81.8%	2 18.2%	11 100%	37.00	11
Total	34 79.10%	9 20.90%	43 100%	36.93	43

Findings

Our research questions were
- What are the characteristics of successful search?
- How do people search archival collections online?
- Are there particularly advantageous search strategies for accessing information about primary sources in online finding aid systems?

To address these questions, we have divided our findings into 3 sections. First, we discuss our metric for search success and explore the characteristics of successful searchers. Second, we discuss overall search patterns to identify behaviors across all participants and tasks. Finally, we provide a more detailed analysis of 3 types of tasks that involve locating information in a finding aid: 1) navigating the hierarchy, 2) finding a digital photograph attached to a finding aid, and 3) total recall tasks requiring participants to identify all collections related to a particular subject.

Search Success

When developing the tasks, we determined what constituted the correct answer or answers (if the task had several sections) for each one. For the purposes of the search success metric, correctness meant completing all parts of a task correctly. While the tasks varied in complexity, all participants had an equal opportunity to complete them all. We developed a metric for search success based on participants' overall performance by calculating the percentage of correct responses from the total number of tasks attempted by each individual. The percentage correct ranged from 13% to 89%, with a mean percentage of 57.9%. We then divided the participants into 3 percentiles based on a

breakdown of the participants: the most and least successful groups contained 11 participants each, while the remaining 21 participants fell into the middle group of searchers who had average success. The top and bottom quartiles (75% correct and above, 45% correct and below) were then isolated as highly successful and relatively unsuccessful searchers, respectively.

While these participants were recruited based on their group affiliations (historians, graduate students, genealogists, undergraduates, and librarians) and presumed varying levels of familiarity with online finding aid systems, these factors did not prove important in determining search success. Highly successful and unsuccessful searchers were found in almost every group. Table 3 illustrates that genealogists and graduate and undergraduate students tended to fall in the middle range of search success, while historian and librarian searchers tended to be at either end of the success range. The librarians' domination of the high success group is noteworthy and will be explored in more depth. In spite of these apparent differences, the distribution does not exceed what might be randomly expected (p =.19 Freeman-Halton extension of the Fisher exact test).

Table 3. Search Success by Success Grouping

| Success Grouping | Occupation | | | | | |
	Historian (n = 5)	Genealogist (n = 9)	Librarian (n = 10)	Graduate Student (n = 8)	Undergraduate Student (n = 11)	Group Totals
Low: Below 25th percentile	2 / 18.2%	3 / 27.3%	2 / 18.2%	1 / 9.1%	3 / 27.3%	11 / 100.0%
Medium	1 / 4.8%	6 / 28.6%	3 / 14.3%	6 / 28.6%	5 / 23.8%	21 / 100.0%
High: Above 75th percentile	2 / 18.2%	0 / .0%	5 / 45.5%	1 / 9.1%	3 / 27.3%	11 / 100.0%
Total	5 / 11.6%	9 / 20.9%	10 / 23.3%	8 / 18.6%	11 / 25.6%	43 / 100.0%

(n = 43)

After participants completed 4 of the tasks, we asked them to rate their confidence in the correctness of their answers on a scale from 1 to 10. As Table 4 shows, average confidence was not closely aligned with search success. While confidence was higher, on average, for the most successful searchers, it was lowest for the middle range of searchers. Confidence did not vary significantly across the three groups, F (26, 16) = 1.71, p = .13. Patricia Maughan's 2001 review of information literacy and confidence assessments of undergraduate students shows that self-assessment of skill is not often aligned with actual proficiency, and most students rate their ability much higher than their test

performance results.[28] Our participants' confidence ratings were also inconsistent with their search success.

Average task completion time was also not consistent with search success: the least successful group tended to complete tasks most quickly, while the middle group took the longest, on average. As has been documented in other studies, searchers frequently trade off speed and accuracy.[29] Some searchers chose to double check their answers before considering their search complete. For example, one librarian successfully completed a search and then decided to see what results she would get with different search terms (Participant 2L). A one-way ANOVA was used to test the difference in task completion time among the 3 groups. Task completion time did not vary significantly, $F(37, 4) = 4.427$, $p = .08$.

Table 4. Average Confidence and Task Completion Time by Success Grouping

Success Grouping	n	Range of Scores	Average Percent Correct	Average Confidence	Average Task Completion Time
Low: Below 25th percentile	11	0%–45% correct	34%	6.4	5:49
Medium	21	46%–74% correct	59%	6.0	6:25
High: Above 75th percentile	11	75%–100% correct	80%	7.8	6:00
Total	43		58%	6.6	6:10

($n = 43$)

The large number of highly successful librarian searchers raises the question: How much does prior use of libraries and archives influence search success? Table 5 shows that 81.9% of searchers with a high degree of success accessed libraries "a few times a week" or more in the previous year, while Table 6 shows that 63.7% of this group used online library catalogs that frequently. However, neither of these results was statistically significant. The success groupings did not differ significantly by use of libraries in the past year, $\chi^2(6, n = 43) = 6.94$, $p = .33$, nor by frequency of library catalog use, $\chi^2(6, n = 41) = 5.78$, $p = .45$.[30]

---

[28] Patricia Davitt Maughan, "Assessing Information Literacy among Undergraduates: A Discussion of the Literature and the University of California-Berkeley Assessment Experience," *College and Research Libraries* 62 (2001): 71–85.

[29] Mika Käki and Anne Aula, "Findex: Improving Search Result Use through Automatic Filtering Categories," *Interacting with Computers* 17, no. 2 (2005): 187–206.

[30] Since multiple cells contained less than 5 observations, we went on to do exact tests (Frequency of library use: *p=.23*, Frequency of online library catalog use: *p=.5*, both using the Freeman-Halton extension of the Fisher exact test).

Table 5. Use of Libraries in the Past Year

Success Grouping	In the past year, how often did you access libraries either in person or remotely? (*n* = 43)				
	Almost every day	A few times a week	Two or three times a month	A few times a year	Totals
Low: Below 25th percentile	3 / 27.3%	4 / 36.4%	2 / 18.2%	2 / 18.2%	11 / 100.0%
Medium	3 / 14.3%	6 / 28.6%	9 / 42.9%	3 / 14.3%	21 / 100.0%
High: Above 75th percentile	5 / 45.5%	4 / 36.4%	2 / 18.2%	0 / .0%	11 / 100.0%
Total	11 / 25.6%	14 / 32.6%	13 / 30.2%	5 / 11.6%	43 / 100.0%

(*n* = 43)

Table 6. Use of Online Library Catalogs in the Past Year

Success Grouping	In the past year, how often did you access an online library catalog? (*n* = 41)				
	Almost every day	A few times a week	Two or three times a month	A few times a year	Totals
Low: Below 25th percentile	3 / 30.0%	2 / 20.0%	3 / 30.0%	2 / 20.0%	10 / 100.0%
Medium	3 / 15.0%	6 / 30.0%	6 / 30.0%	5 / 25.0%	20 / 100.0%
High: Above 75th percentile	5 / 45.5%	2 / 18.2%	4 / 36.4%	0 / .0%	11 / 100.0%
Total	11 / 26.8%	10 / 24.4%	13 / 31.7%	7 / 17.1%	41 / 100.0%

(*n* = 41)

We hypothesized that those who reported using finding aids in the past, particularly online finding aids, would have greater search success. As Table 7 shows, prior use of a finding aid was highest among the most successful users (91%); still this was not a statistically significant correlation χ^2 (2, n = 43) = 5.78, *p* = .27 (Freeman-Halton extension of the Fisher exact test *p* = .28). In fact, as seen in Table 8, the smallest percentage of people in any group who had previously used finding aids via computer was in the most successful group (60%). This group also had the largest percentage of in-person users of finding aids (70%). It may be that in-person use of hard-copy finding aids created more familiarity among participants and a greater ability to use their online counterparts.

Table 7. Prior Use of Finding Aids

	Have you ever used a finding aid? (n = 43)		
Success Grouping	Yes	No	Total
Low: Below 25th percentile	7/ 63.6%	4 / 36.4%	11 / 100.0%
Medium	14 / 66.7%	7 / 33.3%	21 / 100%
High: Above 75th percentile	10 / 90.9%	1 / 9.1%	11 / 100%
Total	31 / 72.1%	12 / 27.9%	43 / 100%

(*n* = 43)

Table 8. Method of Prior Finding Aid Usage

	How have you accessed finding aids? (n = 31)					
Success Grouping	A. In person, visiting archival repositories	B. Remotely, via computer on computer network or Internet	C. In person and remotely via computer	(A+C) Any in-person use of finding aids	(B+C) Any remote use of finding aids	Totals
Low: Below 25th percentile	0 / 0.0%	3 / 42.9%	4 / 57.1%	4 / 57.1%	7 / 100.0%	7 / 100.0%
Medium	2 / 14.3%	7 / 50.0%	5 / 35.7%	7 / 50.0%	12 / 85.7%	14 / 100.0%
High: Above 75th percentile	4 / 40.0%	3 / 30.0%	3 / 30.0%	7 / 70.0%	6 / 60.0%	10 / 100.0%
Total	6 / 19.4%	13 / 41.9%	12 / 38.7%	18 / 58.1%	25 / 80.6%	31 / 100.0%

(*n* = 31)   The question "How have you accessed finding aids?" was only posed to those who answered yes, they had used a finding aid, so only 31 people responded to this question.

O v e r a l l   S e a r c h i n g   P a t t e r n s

In this section, we will first discuss 4 common characteristics of participants' searches: number of searches per task, use of Boolean operators, navigating through large amounts of text (particularly using Ctrl+F), and coordinating the left and right frames in finding aids. The first two of these characteristics pertain to searching in library and archival access systems and are often cited in the library literature as important dimensions of a search. The latter two aspects pertain particularly to searching in online finding aids and are identified in archival usability studies as impacting researchers' use of online finding aids. Because finding aids are quite voluminous compared to library catalog records,

using left and right frames and scrolling or searching through text are salient to the study of searching in archival access systems. After this general discussion of the overall characteristics of the search, we will provide a detailed look at several of the tasks and discuss search tactics in context.

While individual searchers employed a wide range in the number of searches (up to 13 on an individual question), the average number of searches per question across all participants was 2.79 (see Table 9). The average number of searches per participant varied the most in the highly successful group, which included those with the highest (5.56) and lowest (1.25) average number of searches across all participants. We used a one-way ANOVA to test for differences in the average number of searches across all tasks among the 3 success groupings. The average number of searches did not differ significantly across the three groups, $F(28, 14) = .711$, $p = .786$.

Table 9. Average Number of Searches for Each Success Grouping

Success Grouping	Mean	Median
Low: Below 25th percentile	2.85	2.67
Medium	2.80	2.50
High: Above 75th percentile	2.73	2.40
Total	2.79	2.50

($n = 43$)

Search techniques also varied a great deal between participants. On the Bentley site, which offered a Boolean search drop-down option, almost half (46.42%) of all searches were constructed using Boolean connectors. (See Figure 3.) Six of the searchers did not use the Boolean search option at all, but they were evenly distributed among the 3 success groupings (2 each in the low, middle, and high success groups). The most successful group used the highest percentage of Boolean searches, but the difference in frequency of using Boolean operators between the most successful and the other groups was just around 5%. We used a one-way ANOVA to test for differences in the use of Boolean operators among the 3 success groupings. Use of Boolean operators did not differ significantly across the three groups, $F(2, 40) = .09$, $p = .91$. The average use of Boolean searches on the Bentley system (46.42%) was much higher than that observed by Rieh and Xie (26.3%), perhaps because Rieh and Xie's data were drawn from a Web search engine, where use of Boolean operators may be less common overall.[31]

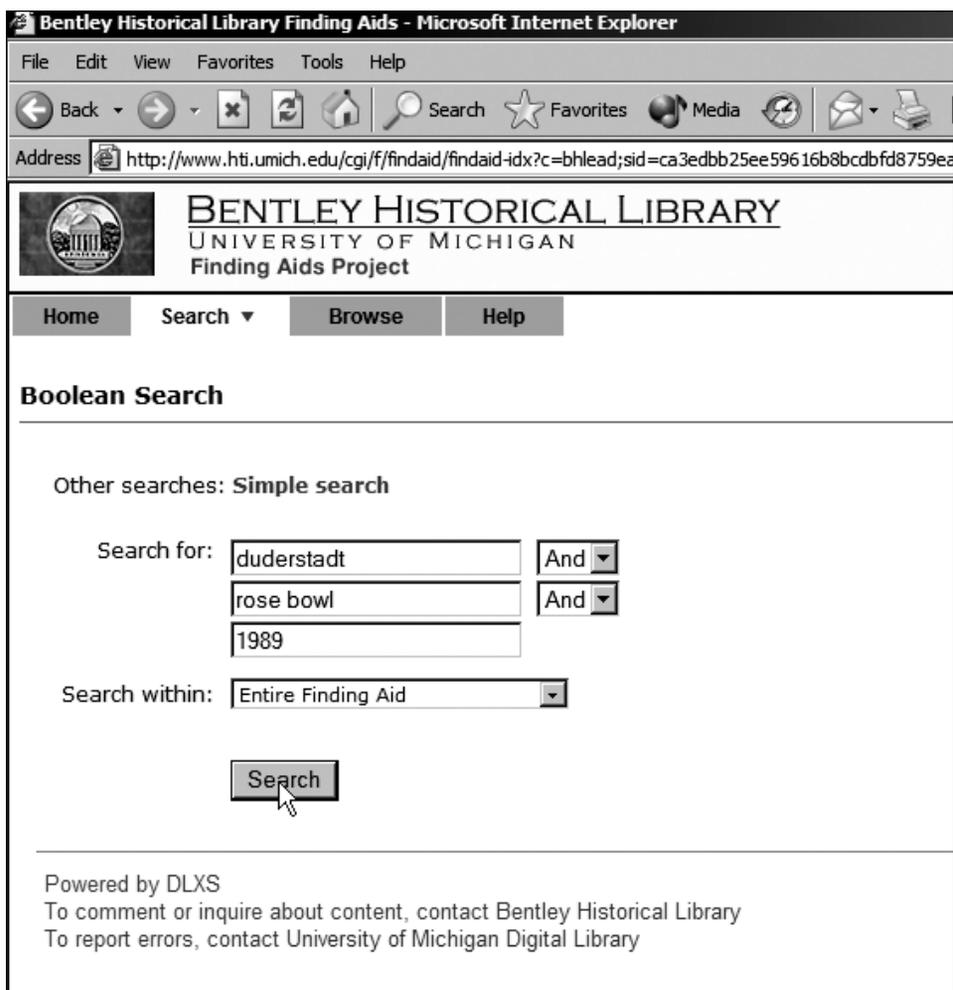[31] Rieh and Xie, "Analysis of Multiple Query Reformulations on the Web," 758.

F I G U R E 3 . **Boolean search option on the Bentley Historical Library interface as it appeared at the time of the study in 2005 (Participant 2L).**

A more striking difference among the 3 groups can be seen in the participants' use of the browser's function—the keyboard combination Ctrl+F—to search the text of a Web page (see Table 10). While the participants' use of the browser's search function (on both the Bentley and OANM sites) was much lower as a percentage of all searches than was their use of Boolean search operators on the Bentley site, it was dramatically higher among more successful searchers. Among the successful searchers, 21.64% used the Ctrl+F function, while this number dropped to 8.38% for the middle group and 4.18% for the group with the lowest success rate. Use of the browser's search function should be viewed in the appropriate context, however. Participants generally used this

search function after they had completed a search and wanted to see where their search terms appeared on a page, to determine the relevance of their results. We used a one-way ANOVA to test for differences in the use of the Ctrl+F function among the 3 success groupings. Use of the Ctrl+F function differed significantly across the 3 groups, $F(2, 40) = 3.95$, $p = .03$.

Table 10. Average Use of Boolean Search Capabilities (Bentley Site) and Average Use of Browser Search Capabilities (Both Sites)

		Average use of Boolean search capabilities (Bentley site)		Average use of browser search capabilities (both sites)	
Success Grouping	*n*	Mean	Std. Deviation	Mean	Std. Deviation
Low: Below 25th percentile	11	45.27%	.35733	4.18%	.10137
Medium	21	45.00%	.33333	8.38%	.14576
High: Above 75th percentile	11	50.27%	.34797	21.64%	.20805
Total	43	46.42%	.33565	10.70%	.16548

(*n* = 43)

N a v i g a t i o n   t h r o u g h   F i n d i n g   A i d s

One common feature in many online finding aid systems is a navigation bar on the left side of the screen that controls the content in the frame on the right side. The left side of the screen generally lists sections of the finding aid. The left side navigation bar can also include search features to assist resource discovery within a finding aid. When this experiment took place, the Bentley site had a tool in the left navigation bar called "Search Within This Text," which provided a handy option for searching quickly within a finding aid. (See Figure 4.) Unfortunately, most of the participants did not notice this navigation option, or, if they did, they did not understand its significance. They used the Ctrl+F search feature in the browser instead. Only 7 individuals used "Search Within This Text," and 6 of them used it only once each over all 5 of the Bentley tasks (the other person used it twice). These 7 searchers were evenly distributed in the 3 success categories. In contrast, 13 searchers used the browser's search feature (Ctrl+F) on the Bentley site for a total of 59 times over the course of the 5 tasks. However, we cannot claim that the participants who chose "Search Within This Text" did not know about the browser search function: 3 of them used both when searching the Bentley. While they used "Search Within This Text" once each, they used the browser search function 2, 6, and 10 times respectively. Other participants used neither the browser nor the archives' features for finding keywords on a page, instead scanning through multiple screens of finding aids until they found the relevant sections. The terms people used with the

Ctrl+F feature or for which they scanned were largely those we provided in the task. There was little synonym generation.
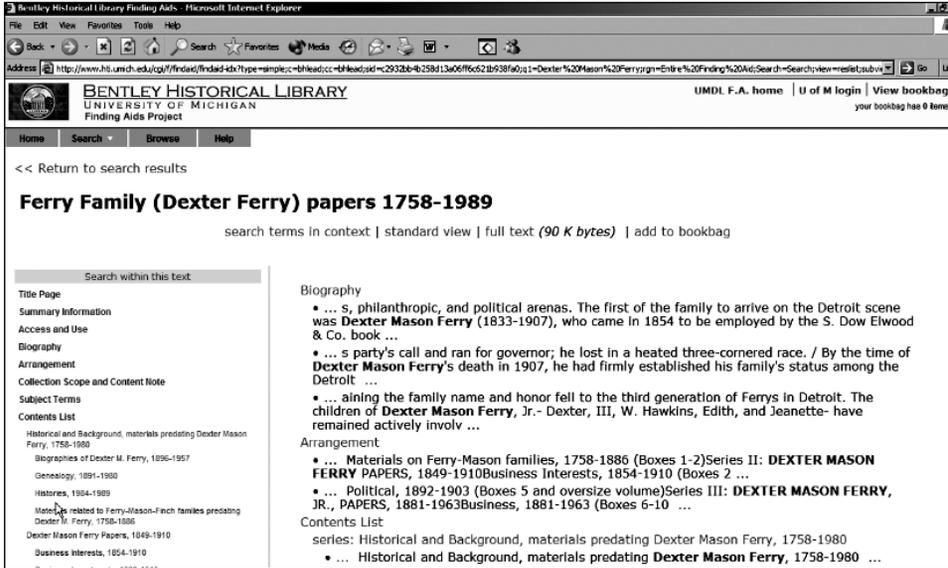
F I G U R E   4 .   **Search within this text on the Bentley Historical Library interface as it appeared at the time of the study in 2005 (Participant 5E).**

Finding information in the finding aid: Kit Carson oath of allegiance

A number of the participants struggled to use the left navigation menu to locate search terms within a finding aid successfully. Understanding where results were likely to be found requires some knowledge of the structure of a finding aid. One task that required this understanding asked participants to locate the Oath of Allegiance to the United States taken by frontiersman Kit Carson on the Online Archive of New Mexico site. We hypothesized that this was one of the simpler tasks we assigned. Thirty-seven of the 42 participants to undertake this task were able to find the correct collection and determine the date on which Carson took the oath, for an 88% success rate. Success with this task tracked logically within the 3 success groupings, with the highly successful searchers 100% correct, the middle group 90.5% correct, and the less successful group 70% correct. More successful searchers also used a smaller average number of searches, as seen in Table 11.

Table 11. Success on Task: Kit Carson Oath of Allegiance

Kit Carson Oath of Allegiance date task				
Success Grouping	Incorrect	Correct	Total	Average number of searches
Low: Below 25th percentile	3 / 30.0%	7 / 70.0%	10 / 100.0%	3.20
Medium	2 / 9.5%	19 / 90.5%	21 / 100%	2.14
High: Above 75th percentile	0 / .0%	11 / 100.0%	11 / 100%	2.09
Total	5 / 11.9%	37 / 88.1%	42 / 100%	2.38

(*n* = 42)

All of the searchers used "Kit Carson" as their primary search term, sometimes adding the phrase "Oath of Allegiance" and, in a few cases, the words "date" or "when." Entering "Kit Carson" as a default search brought up 46 results, which required searchers to scroll down the page to find and select the Kit Carson papers. At that point, participants selected different options (e.g., contents list, scope and content) on the left side navigation bar to locate the Oath of Allegiance, a task that required some familiarity with the structure of a finding aid.

Searchers who took longer to locate the necessary date or did not find it at all fell into 2 groups: those who felt they had retrieved too many results and those who were unable to navigate the finding aid structure. Some searchers decided that 46 hits were too many and immediately reformulated their queries rather than analyzing the results. These individuals frequently ended up trying several other searches before repeating their first search terms and scanning the same 46 results later. Others used one of the search delimiters only to discover that their searches had become too specific. Participant 1E, for example, searched "Kit Carson, Oath of Allegiance" in the scope and content field only, retrieving no results. When she broadened her search to "any of these words" within the scope and content field, she received 1,286 finding aids. She used several more searches to fine tune her results, finally achieving success when she simplified her search using the term "Kit Carson" and the delimiter "this phrase within entire document," which gave her 41 results out of which she selected the Kit Carson papers. As this was the first task using the Online Archive of New Mexico site, experimentation with search delimiters and browsing helped searchers familiarize themselves with the site. As one undergraduate student reported, "I found it under Contents List by randomly clicking" [Participant 15U].

Drilling down the hierarchy: Antonio Terrazas wedding photograph

Finding a digital object linked to a finding aid was a bit trickier for the test participants: only half were successful when we asked them to find a digital copy of a wedding photograph. Of the 40 searchers to attempt this task, 32 began with the name of the groom as their sole search term. The other 8 searchers used the groom's name along with the date given in the question and/or the words "wedding," "marriage," and "photograph" as search terms, a technique that several other searchers adopted as the task progressed. A number of searchers added more and more terms sequentially to narrow the search. Some were dead ends for the searchers, leading them to attempt a number of different combinations of terms and to scan through the results. In this case, synonym generation was not a particularly useful strategy for query reformulation.

One successful searcher (Participant 4L), a librarian whose overall search success placed her in the average success category, began her search with the man's name and the wedding year. Upon retrieving no results, she remarked, "It didn't treat the number very well" and searched again substituting "wedding" for the year. This yielded 3 results. She noted that the papers of 3 families were returned, but none of them had the groom's last name, so she checked the date ranges for a fit. At this point, she noticed 48 hits for her search terms in the papers of one family and selected that collection. Once inside the finding aid, she viewed the scope and content note, then the container list sections. She read carefully through the finding aid, looking for "online stuff" multiple times in several sections. When she used Ctrl+F to find the groom's name, she eventually found the photographs section of the finding aid and then located the wedding photo.

The Antonio Terrazas wedding photograph task was difficult for several reasons. First, the photograph is part of the bride's family's papers, and since her maiden name was not provided in the task, searchers did not automatically realize they had found the correct collection. When their search results did not list any collections with the groom's name in their titles, they assumed that the search had failed. For this reason, this task illustrates some of the limits and difficulties of name collection as a search strategy. Second, after selecting the correct finding aid, a great deal of scrolling, searching (using CTRL+F), or scanning within the finding aid was required to find the embedded image. This also required careful selection among the search results, as in the example of the librarian for whom a date search failed, but matching the dates of the collections with the date of the wedding was more successful. Table 12 shows the results for this task.

Table 12. Success on Task: Terrazas Wedding Photograph

Success Grouping	Wedding photo URL task		
	Incorrect	Correct	Total
Low: Below 25th percentile	10 / 100.0%	0 / .0%	10 / 100.0%
Medium	9 / 45.0%	11 / 55.0%	20 / 100%
High: Above 75th percentile	1 / 10.0%	9 / 90.0%	10 / 100%
Total	20 / 50.0%	20 / 50.0%	40 / 100%

(*n* = 40)

Total recall: find everything on rock and roll in Detroit and slavery

The recall searches asked participants to identify all collections held by the repository containing materials on a given topic. We then asked them to rate their confidence in having found everything on that topic. As with the other experimental tasks, participants selected search terms mirroring the wording of the question; however, these tasks elicited the most synonym generation from participants. When asked "You are researching rock and roll in Detroit. Does the Bentley Library have any collections on this topic?," the majority of respondents searched the Bentley interface using the terms "rock and roll" and "Detroit." A number of other terms emerged in searches as well, including "music," "Motown," and "counterculture." Some searchers used their knowledge of the topic to search for the live music venues "Grande Ballroom" and "Blind Pig," while others searched for specific acts such as "MC5" and "Commander Cody." For this task, the most efficient means of identifying all collections was to use the Library of Congress Subject Term "Rock music."

A notable technique employed by several searchers was to create search terms constructed in a style similar to true controlled vocabulary terms (e.g., Library of Congress Subject Headings). Examples include "Detroit and entertainment," "music history," "rock and roll history," and "music in Detroit"; however, only 4 individuals used the subject search delimiter in the search interface. One librarian (Participant 5L) sought out subject terms used by the Bentley to enhance her search. After finding a useful collection, she clicked on "subject terms" in the navigation menu and selected one of them, noting how it was used in the finding aid. Although the Bentley interface encourages visitors to search using subject terms, this librarian was one of the few study participants to employ an actual controlled vocabulary term when using the subject search option. She had already answered the question at this point in her search and seemed to be

satisfying her curiosity about the organization of the finding aid system rather than searching for more results.

Table 13 shows that both the medium and high success groups of searchers completed this task successfully, with slightly better results from the former group. The middle success group also used the largest number of searches in this task, on average. Table 13 also provides the average number of distinct phrases searched for each group, a measure that conveys the number of times searchers changed their search terms, reformulating the content of their query. Because many searchers used the same terms multiple times, trying out variations including quotation marks to specify a particular phrase, Boolean operators, or search delimiters that indicate the parts of a finding aid to search, this measure is a useful way to determine synonym generation. The difference between the average number of searches and the average number of phrases searched indicates that many participants in all 3 groups tried the same search terms repeatedly, using different operators and delimiters. On average, those who tried a greater number of phrases in their searching had more success. Participants' infrequent use of controlled access terms perhaps relates to a general misunderstanding of how controlled vocabulary links collections together.

Table 13. Success on Task: Rock and Roll in Detroit

Success Grouping	Rock and roll in Detroit task				
	Incorrect	Correct	Total	Average number of searches	Average number of phrases searched
Low: Below 25th percentile	3 / 27.3%	8 / 72.2%	11 / 100.0%	2.91	2.55
Medium	0 / 0.0%	20 / 100.0%	20 / 100.0%	3.35	2.80
High: Above 75th percentile	1 / 10.0%	9 / 90.0%	10 / 100.0%	3.20	2.80
Total	4 / 9.8%	37 / 90.2%	41 / 100.0%	3.20	2.73

(n = 41)

We designed a similar task using the Online Archive of New Mexico, asking participants to determine which repository holds the most collections on the topic of slavery and which collections on this topic contain 5 or more references to archival materials. This varied from the "total recall" task in two ways: first, the task became a cross-repository or federated search; and, second, by asking participants to identify the repository with the most collections, this task simulated some of the decision making involved in selecting a repository to contact or visit to research a topic. (See Figure 5.) Thirty-four participants completed this task; the others ran out of time for the experiment. All but 2 of

the searchers used the word "slavery" as their sole search term with the exception of an undergraduate who searched for "history on slavery" and a historian who searched for "indent*," using the wildcard symbol to retrieve variations on the word "indenture."

A search on "slavery" retrieved 16 collection titles. Many of the participants had difficulty interpreting their search results in terms of the question. For example, one genealogist said, "None of the listed repositories relate to slavery," showing that he either did not understand the nature of the task or had misunderstood the relationship between repositories and collections (Participant 5GE). Only 23 (67.6%) of the participants to undertake this task correctly determined which repository holds the most collections on this topic. One undergraduate selected the repository with the largest collections (most linear feet) instead of the largest number of collections (Participant 13U), while another selected the collection with the most search results rather than the repository with the most collections (Participant 1U). This question requires an understanding of several relationships: finding aids are representations of collections, a number of which are held by a given repository. Many of the participants evidently did not understand these fundamental relationships as they interpreted their results for this question.

In Table 14, we show that both the medium and high success groups did well on this task. The high success group used a greater number of searches and had a higher degree of confidence in their results. This higher confidence may be cumulative, based on searchers' success at earlier tasks.

Table 14. Success on Task: Repositories with Collections Related to Slavery

Success Grouping	Finding collections related to slavery task				
	Incorrect	Correct	Total	Average number of searches	Average confidence
Low: Below 25th percentile	5 / 62.5%	3 / 37.5%	8 / 100.0%	2.50	5.71
Medium	4 / 23.5%	13 / 76.5%	17 / 100.0%	2.29	5.88
High: Above 75th percentile	2 / 22.2%	7 / 77.8%	9 / 100.0%	3.63	6.67
Total	11 / 32.4%	23 / 67/6%	34 / 100.0%	2.67	6.06

(*n* = 34)

While about two-thirds of the participants found the repository with the most collections on the topic of slavery, none of them correctly named both collections with 5 or more references to archival materials on slavery, although 18 (52.9%) were able to correctly name 1 of the 2 relevant collections. This may

have been due to an inability to read the finding aids correctly and to distinguish between information presented in the biographical/administrative history section and that in the scope and content notes and/or contents listing.

Discussion

Although none of our statistical tests assessing personal characteristics gave us any insight into predicting search success, we did note some interesting search behaviors in this study. Participants employed traditional library search strategies and also adapted to the unique online finding aid search environment, innovating their search techniques. We observed search strategies such as narrowing a search, synonym generation, and use of Boolean operators. We also saw participants innovating to search large amounts of text using the browser CTRL+F function and identifying ways to select from large sets of online finding aids presented in search results. Query reformulation was also common among participants, as they reran searches and changed search terms, delimiters, or Boolean operators to expand or refine their results. Our searchers averaged 2.79 queries per task, comparable to the average library search of 2 to 4 queries.[32]

Throughout the tests, and across the majority of participants, 3 themes continuously emerged as barriers to search success: difficulties in selecting search terms, problems with navigation through the finding aids, and unfamiliarity with the archival terminology and practice embedded in online archival finding aids.

The Kit Carson Oath of Allegiance task shows how narrowing a search can lead to success. Those who searched for "Kit Carson" received an unwieldy set of results. Participants who used search delimiters created a manageable list of results from which to choose. Other participants attempted to narrow the search using search terms alone and this was often disasterous as the system returned all collections containing any of the terms provided. The library search literature confirms that users generally stick with a search system's default values.[33] In our study, many people wanted to narrow their searches but did not understand how to use the system to do this.

Interface issues challenge researchers navigating through finding aids, as does understanding the nature of archival collections. For example, the difficulties of researchers who fail to distinguish between content and context surfaced in the Terrazas photograph query. Most of the participants had little understanding of provenance or the context in which collections might be assembled. More experienced archival researchers might realize that a family's papers

---

[32] Karen Markey, "Twenty-Five Years of End-User Searching, Part 1: Research Findings," *Journal of the American Society for Information Science and Technology* 58, no. 8 (2007): 1072.

[33] Markey, "Twenty-Five Years of End-User Searching, Part 1: Research Findings," 1077.

could include materials related to individuals outside the family, but newcomers to archives might not. As archival institutions increasingly digitize photographs and other images, they struggle to provide access to them while maintaining a digital context reflecting the original order of the collection, information that may assist users in finding and understanding digital material. The creation of Calisphere by the California Digital Library as an accessible interface for K–12 teachers and students is a good example of experimentation and rethinking the relationship between images and online finding aids.[34]

The recall searches on rock and roll in Detroit and slavery suggested the use of synonym generation, which many of the participants attempted. They also used search delimiters more frequently in these tasks. As noted above, the rock and roll task generated a number of search terms including names of local music venues or groups. Participants had difficulty generating new terms for the search on collections dealing with slavery. An added problem in the slavery task was that several participants did not seem to understand the distinction between repositories and collections. In addition, search terms were not easy to scan within the results retrieved. Both tasks required a certain amount of expertise on the searcher's part to determine the relevance of the results. For instance, finding a search term in the history rather than the scope and content section of a finding aid may indicate that the term is relevant for understanding the context of a collection but is not represented in the contents of the collection itself. Therefore, occurrence of a term in a finding aid is an inexact indicator of relevance. Ultimately, searchers must determine relevance for themselves. These tasks mirror a common goal of visitors to archival websites: to determine if the amount and type of information on their research topic at a given repository warrants a visit. While, in the end, only the researcher can determine this, the manner in which the finding aid system displays results may affect the ease with which that decision can be made.

Study participants efficiently identified keywords in the tasks they were given and used them as search terms, but when they retrieved insufficient results with those terms, they often had trouble generating new terms that led to successful searches. It is no wonder, then, that the precise known-item searches were more successful than the recall searches, which ask users to find everything on a given topic. Known-item searches provide somewhat reliable terms on which to search. Recall searching often benefits from synonym generation, identification of controlled access terms, and knowledge of the topic, which many of our study participants did not have. It also helps a great deal if searchers are able to differentiate between information in the administrative or biographical history, the scope and content, and the contents listing sections.

[34] Jane Lee, "Calisphere UI Testing Findings and Recommendations, Chico High School, 14–15 September 2005," available at http://www.cdlib.org/inside/assess/evaluation_activities/docs/2005/calisphereChicoReport_sept2005.pdf, accessed 24 September 2009.

FIGURE 5. Scanning search results in the Online Archive of New Mexico interface as it appeared at the time of the study in 2005 (Participant 7U).

As reported in the library search literature, the experts in our study did better on searches, although not statistically better. Those with higher levels of experience using libraries, online library catalogs, and archival finding aids had greater search success than the novices. This finding aligns with earlier studies, although these also did not demonstrate statistical significance. In our study, many of the experts were professional reference librarians. Of particular note, however, was the preponderance of successful searchers of online finding aids who had used paper-based finding aids. This raises a question about what prior experience with paper-based finding aids gives users, such as some type of mental model or archival expertise, that is transferrable to an online context.

Conclusion

While usability testing is important to create transparent and accessible interfaces, the data models that underlie online finding aid systems influence

search behavior. For example, in this study, the difference in success for all users between recall and known-item searches suggests that online interfaces for archives are currently more amenable to name-based than to subject-based searching. Although several searchers made use of the subject headings provided by the interfaces, most did not. To assist with recall searches, online finding aid systems might be modified to suggest related terms to searchers, or to provide guidance on making use of subject headings. While the archival profession has been sensitive to interface design, it has focused less on exactly what user behaviors are required to perform efficient and successful searches. This study attempts to begin this conversation.

While many of the library-based heuristics for successful searching translate to archives (Boolean searching and synonym generation), several do not (citation searching and journal runs). Several characteristics of online finding aids also pose special problems for searchers, such as large blocks of text and hierarchical presentation. To address effective searching of these and other elements unique to online finding aids, archivists need to develop our own set of heuristics for best searching practices to help our users navigate archival access systems. The successful searchers in this study provide a glimpse into both the difficulties of searching and the innovations that searchers use to navigate archival finding aids. Archivists must now codify these and other strategies into finding aid systems to help searchers locate archival materials.

**A p p e n d i x   A :   P r e t e s t   Q u e s t i o n n a i r e — A s s e s s i n g   A c c e s s   a n d
A c c e s s i b i l i t y   o f   I n t e r f a c e s   f o r   P r i m a r y   S o u r c e s**

A.   Library and Archives Use

   1. General use of libraries

     1. How do you usually access *libraries*? (Select one)
- [ ] In person, visiting the buildings
- [ ] Remotely, via computer on computer network or Internet
- [ ] Remotely, by telephone
- [ ] Other (please specify) _____

     2. ***In the past year***, how often did you access libraries either in person or remotely?
- [ ] Almost every day
- [ ] A few times a week
- [ ] Two or three times a month
- [ ] A few times a year
- [ ] Never

   1.1 Online library catalog use

     3. Have you ever used an ***online library catalog***?
- [ ] Yes (If yes, go to 4)
- [ ] No (If no, skip to 6)

     4. ***In the past year,*** how often did you access an online library catalog?
- [ ] Almost every day
- [ ] A few times a week
- [ ] Two or three times a month
- [ ] A few times a year
- [ ] Never

     5. ***In a typical week***, how many times do you access an online library catalog?
- [ ] Once
- [ ] 2–3 times
- [ ] 4–5 times
- [ ] More than 5 times
- [ ] Never

   1.1.1 General use of archives

     6. Have you ever used ***archives*** for your assignment or for your own research?
- [ ] Yes (If yes, go to 7)
- [ ] No (If no, skip to 10

565

7. ***In the past year,*** how often did you use archives?
- ☐ Almost every day
- ☐ A few times a week
- ☐ Two or three times a month
- ☐ A few times a year
- ☐ Never

8. ***How many separate assignments or research projects*** have you ever done using archives?

[ ]

9. ***How many archival repositories*** have you ever done research in?

[ ]

10. Have you ever used a ***finding aid***? Finding aids mean indexes or guides to collections held by archives and manuscript repositories, libraries, and museums. Finding aids provide detailed descriptions of the collections.
- ☐ Yes (If yes, go to 11)
- ☐ No (Thank you, please go on to section B)

11. How have you accessed finding aids?
- ☐ In person, visiting archival repositories
- ☐ Remotely, via computer on computer network or Internet
- ☐ Other (please specify) _____

B.   Information Technology Use

12. **Rate your ability** to do each of the following:
(Circle the appropriate number:
1 = no knowledge/ability; to **5 = expert user**)

Send and receive voice mail	1	2	3	4	5
Create a word processed document on a computer	1	2	3	4	5
Program a VCR	1	2	3	4	5
Use a video camera	1	2	3	4	5
Use a spreadsheet or database program on a computer	1	2	3	4	5
Send and receive e-mail	1	2	3	4	5
Search for information on the Internet/World Wide Web	1	2	3	4	5
Program a computer using a programming language (such as C, C++, Java)	1	2	3	4	5
Program a computer using a database language (such as ColdFusion or Oracle.)	1	2	3	4	5
Create or edit a World Wide Web site (using programs as html)	1	2	3	4	5
Electronically sending and receiving files by way of the computer (as an email attachment or FTP)	1	2	3	4	5

13. What type of **computer configuration** do you use?
(check all that apply)

	Apple / Macintosh	Windows	DOS	Unix	Other (please specify)	Not Applicable
At home						
At school or work						
In a library, computer lab, community center						

1.2.1.1 C. Demographic Information

14. What **degrees do you hold**? (Indicate all that apply)
- ☐ Bachelor's   ☐ Master's
- ☐ Ph.D.   ☐ Certificate
- ☐ J.D.   ☐ M.D.
- ☐ High School Diploma or Equivalency
- ☐ Other (Please specify) _____

15. What was your **major / concentration**? _____

16. Are you currently **pursuing a degree**?
- ☐ Yes (If yes, go to 17)
- ☐ No (If no, skip to 21)

17. What **degree are you pursuing**?
- ☐ Bachelor's   ☐ Master's
- ☐ Ph.D.   ☐ Certificate
- ☐ J.D.   ☐ M.D.
- ☐ Other (Please specify) _____

18. Have you declared a **major area** of study?
- ☐ Yes   What is it? _____
- ☐ No

19. **When** do you expect to earn this degree?
Semester _____   Year _____

20. Are you attending school:
- ☐ Full-time
- ☐ Part-time

21. **Age**_____

22. **Sex:**
- ☐ Female
- ☐ Male

Thank you!

Appendix B: Searching On-line Finding Aids Post-Test Interview

Bentley:

1. Do you prefer the outline view or the full-text view?
2. Why did you select (browse or search) to answer x question?
3. How did you find the navigation bar on the left?
4. In relation to x question, talk about your search strategy.
5. Would you please identify any unfamiliar terminology you encountered?
6. What did you think of the site's interface?
7. What improvements would you suggest for the interface?
8. Keyword in context: Do you know what this means?
9. Keyword in context: Did you like this feature?
10. Go over search options.

    Does the person understand what they are searching when it says search within, make them name the different parts.

Online Archive of New Mexico:

1. In relation to x question, talk about your search strategy.
2. Would you please identify any unfamiliar terminology you encountered?
3. What did you think of the site's interface?
4. What improvements would you suggest for the interface?
5. Go over search options.

    Does the person understand what all the limits mean?

    Does the person understand what the "Using all of these words/ This phrase, Any of these words" means?
6. Would you like to see a full-text view?
7. Would you like search terms to be highlighted?
8. Did you think about using the "find" function in Internet Explorer?