Christopher J. Prom
Assistant University Archivist
University of Illinois Archives
19 Library
1408 W. Gregory Drive
Urbana, IL 61801

(217) 333-0798
fax (217) 333-2868
prom@uiuc.edu

USER INTERACTIONS WITH ELECTRONIC FINDING AIDS IN A CONTROLLED SETTING

****

Abstract: This article presents findings from a study conducted to measure and describe
interactions with on-line archival finding aids. The study measured responses for participants
interacting with four collection-level search systems and four individual finding aids. Both
statistical and qualitative findings are provided and correlated to demographic data such as
archival experience and self-reported computer expertise.

The study provides insights regarding how users navigate archival descriptive
information and how archivists and digital librarians might design interfaces that facilitate
efficient search strategies. Experienced archival users and novices utilize different methods of
searching for archival information. Interface elements such as alphabetical lists, page-top tables
of contents, and single-page search options enabled both expert and novice users to conduct
more efficient searches than did alternate designs.

Biographical Note: Christopher J. Prom is Assistant University Archivist and Assistant
Professor of Library Administration at the University of Illinois, where he is responsible for
managing digital projects, supervising archival processing and records scheduling, providing
reference service, and managing the American Library Association Archives. His main research
interests lie in the areas of information seeking and the usability of archival electronic resources.
He is currently Chair of SAA's Description Section, a member of the Standards Committee and
the Technical Subcommittee on Descriptive Standards. Chris is also the recipient with Ellen
Swain of a 2003-04 NHPRC Archival Research Fellowship. He holds a Ph.D. in History from
the University of Illinois.

USER INTERACTIONS WITH ELECTRONIC FINDING AIDS IN A CONTROLLED SETTING [1]

As William Maher has suggested, archives may be seen as a special example of chaos theory.[2] The representation of archival materials is inherently complex, and researchers' successes in locating materials sometimes seems to show a high dependence on "strange attractors" or clusters of information that do not at first appear to be logically connected.

Yet over the past two decades, archivists have made much progress in dispersing information regarding the mass of documents, photographs, magnetic tapes, electronic files, artifacts and other materials that they have appraised as being worthy of archival management. In simple terms, archivists have moved finding aids from bookshelves and file cabinets to the Internet. The widespread implementation of MARC-AMC and EAD have revolutionized the presentation of archival information, although many archivists have also moved information on-line without applying these standards.

Archivists' digital activities beg a simple question: How do users interact with on-line finding aids? Because archival descriptive standards do not mandate how finding aids should be presented, user interactions with finding aids deserve close analysis on their own terms. In addition, research regarding user behaviors will help the archival community present descriptive data in an easier-to-use format. The research reported here offers some insight into these issues by describing and measuring the ways in which users interact with electronic finding aids in a controlled setting. Based on this information, the study suggests some specific steps which archivists and digital librarians might take to design more efficient interfaces, interfaces that meet user expectations and facilitate efficient search strategies.

It may not seem immediately apparent that such a study is needed. For the past twenty

years, the archival community has focused much attention on issues surrounding the topics of archival users and archival description. For example, SAA's Task Force on Goals and Priorities issued a call in 1986 for systematic study regarding user understanding of archival description.[3] However, only a handful of such articles exist 18 years later.

A number of important and influential studies have been conducted on the general topic of use; recent articles by Elizabeth Yakel and Kristina Southwell provide critical overviews of this literature.[4] The literature offers assistance in planning and conducting user studies,[5] probes the ways in which reference archivists can provide efficient reference services,[6] and analyzes how users seek archival information in the broad context of their research goals.[7] However, none focus specifically on how users navigate finding aids or use descriptive information, either in an paper or an electronic environment. A prominent theme in the literature about users has been the call for more user studies analyzing hard data.

The literature regarding description shows a similar paucity of analysis regarding user interactions with finding aids. One could argue with justice that the archival profession's supreme achievement over the past 20 years has been the development and application of descriptive standards and tools. The archival journals have marked this development with a large number of articles providing contextual analyses, theoretical overviews, and implementation advice. Particularly important have been the articles published in two special issues of *The American Archivist* and a double issue of *Journal of Internet Cataloging* devoted to EAD. Studies evaluating the effectiveness of descriptive standards have been one feature of this literature.[8] However, most have evaluated the description from the archivist's point of view. A few studies did focus specifically on user reactions, but only a few used a systematic

methodology, and while the findings from these studies are relevant, they apply mainly to

MARC records and on-line catalogs.[9] More recently, Altman and Nemmers provided cogent

advice regarding the usability of on-line resources, but their article was not explicitly research-

based.[10]

We therefore know relatively little about how users actually interact with the descriptive

records that archivists prepare. With rare exceptions, research about users studies and research

about description occupy separate spheres. Anne Gilliland-Swetland has noted that archivists

have often taken a materials-centric (as opposed to user-centric) approach to description, and

that consequently, the finding aids do a very poor job of meeting user needs.[11] Along with

Dennis Meissner, she has argued that rethinking the display of descriptive information will lead

to more effective user interactions, and that reformatting should be based on an analysis of

specific user needs.[12] What is missing is a detailed understanding of how users actually interact

with and navigate current interfaces.[13] At a minimum, such information would provide

hypotheses regarding user search tactics and serve as a starting point for future research and

interface redesign.

Purpose and Goals

As a small step toward this goal, the research reported here provides baseline data and

conclusions from a user interaction study conducted in the summer of 2003 at the University of

Illinois Archives. It sheds light on two discrete but related topics: the nature of participants'

navigational strategies through finding aids and the efficiency with which they searched alternate

designs. It does not purport to explain the totality of archival users' information-seeking

behaviors, although it does contrast the ways in which self-reported archival experts, self-

reported computer experts, and novices make use of finding aids to accomplish goal-driven tasks.   The study tests the hypothesis that experts and novicesemploy different search strategies and reach different search results, and it draws some preliminary conclusions regarding which finding aid features help or hinder search efficiency.

The target population included a mix of people likely to be expert and inexpert users: undergraduate students, past archives users, graduate students, and archivists. Grappling with differences between participant groups is not a trivial matter and requires some explanation before describing the precise methodology employed in this study.

Regarding the issue of expertise, information scientists have sometimes drawn a distinction between domain experts and systems experts, arguing that from an information retrieval point of view, both have advantages over inexpert users.[14] However, it is often difficult to tell exactly who is a domain or a systems expert in the archival domain.[15] As a way around this problem, this study employs the concept of archival expertise in place of the domain/systems dichotomy.   Yakel and Torres have described this concept as "archival intelligence." [16] Archival expertise/intelligence is gained by using archives and conducting historical research, either as an archivist or as an experienced user.

Many potential users of on-line finding aids of course have no archival expertise, but that does not mean they should be considered complete novices in using on-line finding aids. Many people using finding aids have no particular knowledge of how archives are organized, but they are highly skilled at computer search techniques. This study accordingly compares computer expertise with archival expertise. Participants were identified as archival experts, computer search experts, or novices (in some cases, participants were noted as both archival experts and

computer experts).   By analyzing differences in how these groups interacted with finding aids, it is possible to suggest which designs facilitate efficient searches.

Methodology

Finding aid systems available through repository web sites are intended to facilitate the discovery of collections, files, or items in which a user might be interested.[17]   To test how experts and novices navigate through finding aids, this study chose to focus on two search behaviors: looking for collections and looking for folders. Studying such simplistic behaviors may seem overly constraining. After all, real users engage in much more diffuse and complex searches. But the narrow focus provided some distinct advantages. First, it allowed for internal controls to limit bias. More importantly, it provided the opportunity to develop working hypotheses regarding the fundamental navigational strategies that finding aid users are likely to employ in their real searches, hypotheses which can be tested in subsequent research. Finally, the focused nature of the study highlights some specific design features that either help or hinder search efficiency.

Questionnaire, Tasks, and Interview—The study included elements of three techniques common to user studies: surveying, observing, and interviewing.[18] Respondents completed a short questionnaire, searched for specific pieces of information on nine web sites, and when feasible participated in a short audio-taped interview.   All respondents completed the same nine searches.

The study was administered via a website. Subjects participated either on-site (under observation) or off-site (remotely). They were solicited by placing flyers around campus and sending email to individuals who had used the University of Illinois Archives in the two months

prior to the study. Notices were also sent to listservs for students in the University of Illinois history and library science departments. A notice went to the Archives and Archivists Listserv, ensuring a response from practicing archivists, who are heavy finding aid users. The sample was drawn from people likely to include a mix of both expert and inexpert archival users: undergraduate and graduate students, campus faculty, and archivists.

The questionnaire gathered basic demographic data and allowed users to self-report their past experiences with archives, research libraries, and computers. The questions and raw response data are provided in the appendix, and a reference version of the questionnaire and study may be viewed at http://web.library.uiuc.edu/ahx/survey/usab-test/.

After completing the survey, respondents worked through nine controlled search tasks using a standard web browser. For each task, they read a brief set of instructions which directed them to search for a specific piece of information, either the identifying number for a set of personal papers or the container number for a folder concerning a named organization. After each participant read the instructions, the system loaded an HTML frameset in the browser. The main (top) frame included a website through which the user searched and browsed. Users could use any links or search forms on the page (or even leave the original site completely).   The bottom frame included a input box into which the user entered the requested data or clicked "I give up" if unable to complete the task.   Figure one shows the instructions page for task two.

**Figure one: Instructions Page for Task Two**
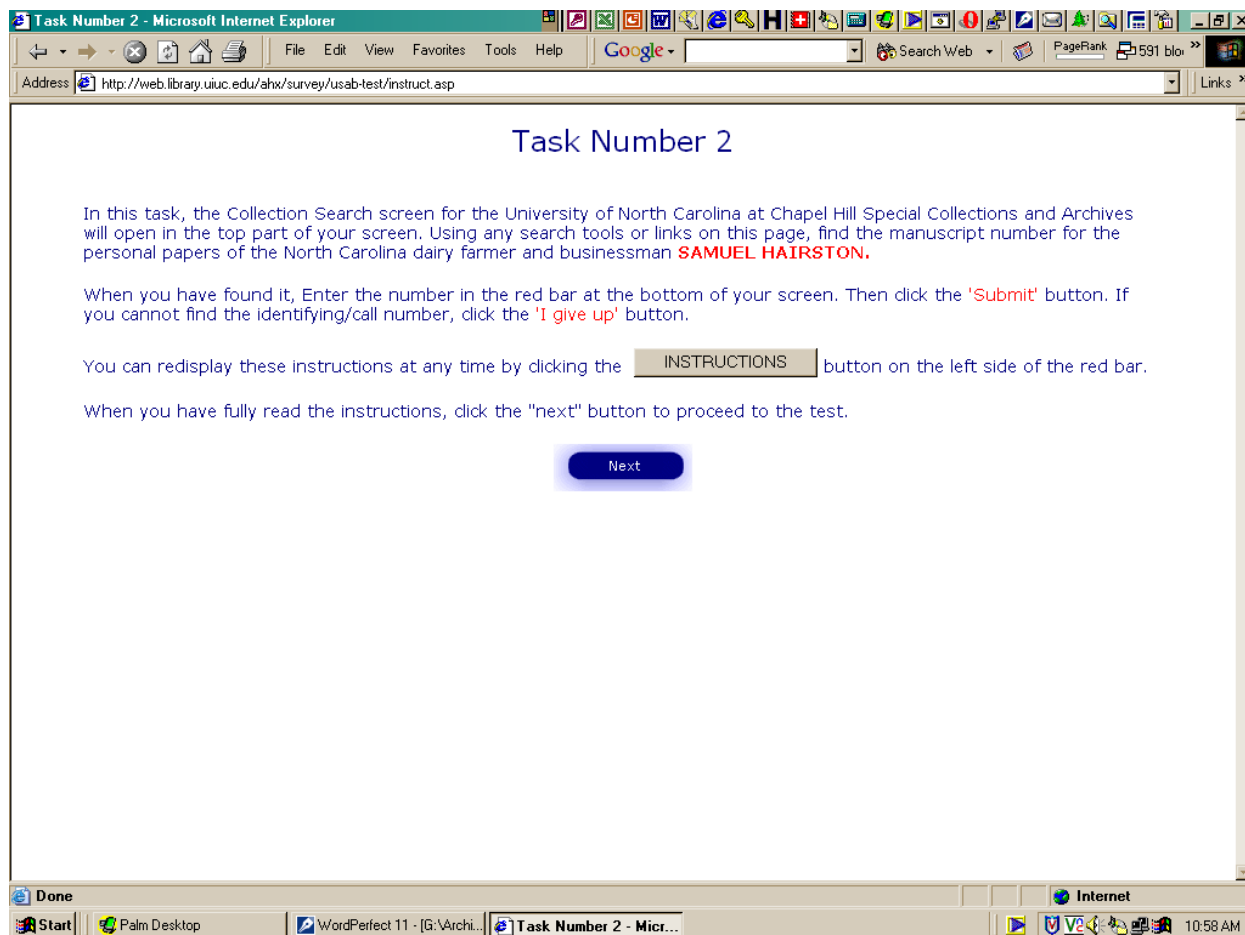
The first task was simply a "dry run" to show users how the system worked, and its results were not calculated in the data reported here. The remaining eight tasks were of two types. For tasks two through five, respondents were asked to find the collection ID number for the personal papers of a prominent individual.[19] The main search screen for a repository was loaded in the top frame of the web browser, and participants used any tools on the loaded page to find a catalog record, finding aid, or web page relating to the collection. Finding such a page, they entered the collection ID number in the bottom frame. Table one summarizes the attributes of the four interfaces, and figure two provides the initial views.

Table 1: Interfaces Tested in Collection Search

| Interface	Number of Search Options	Institution and Search options provided	Underlying architecture used
2	11	**University of North Carolina Manuscripts Department**: Hyperlink to Google site search and to nine lists of finding aids or subject-based research guides. OPAC-based quick search option on navigation bar for parent library	HTML, with links to OPAC, Google site search, and to HTML or EAD-based archival finding aids
3	2	**Online Archive of California (OAC)** Repository Page for University of California at Berkeley, Bancroft Library: "finding aid search" and a alphabetical list of hyperlinks to finding aids, Search options can be limited via menu.	EAD finding aids with underlying DLXS/XPAT architecture. Fielded searching option.
4	1	**Princeton University**, Public Policy Papers: Alphabetical list of finding aids; A-Z links near top of page. Top navigation bar for parent library.	HTML.
5	4	**University of Illinois Archives**: Left-hand search bar with phrase search, browse by subject, and browse by campus unit. Included extensive instructions.	Relational database and Active Server Pages scripting. Used archival record group/subgroup/series classifications.

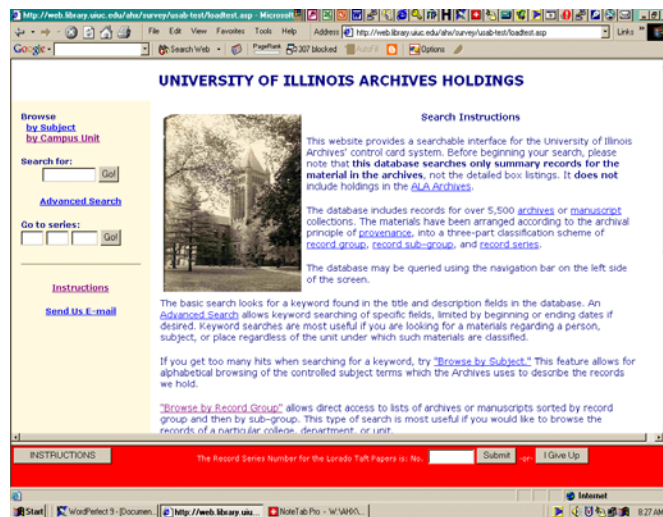# Figure 2: Interfaces Tested in Collection Search
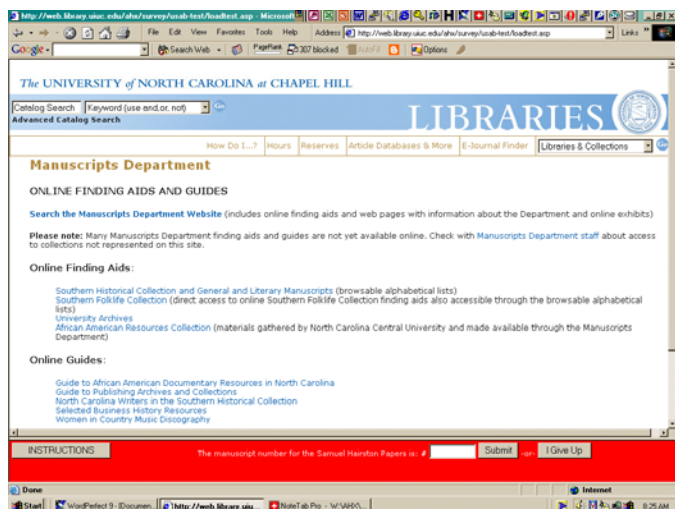
Interface 2: UNC
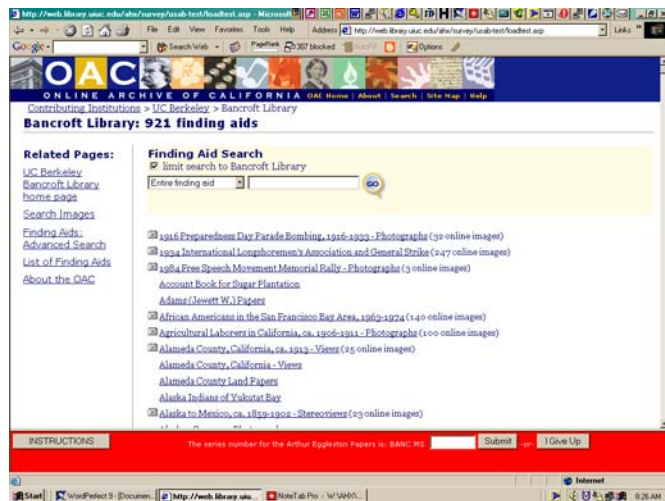
Interface 3: OAC
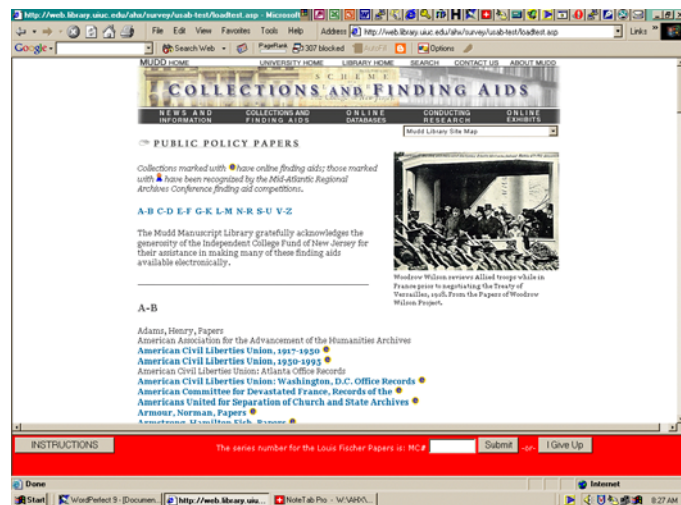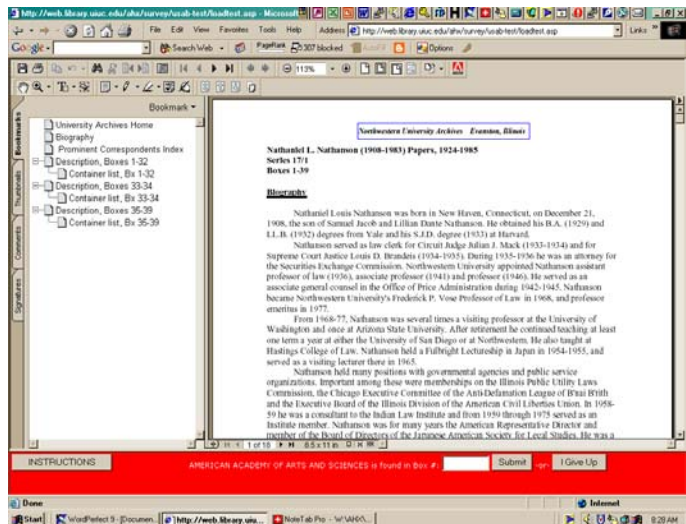

Interface 4: Princeton

Interface 5: Illinois

In tasks six through nine, respondents were asked to search an individual finding aid for a folder of material relating to a topic. The search topic was provided on the instructions page and during the search session was included in the bottom frame to serve as a reminder. The complexity of the search was controlled in a similar manner as with the collection searches.[20] Table two summarizes the characteristics of the four finding aids; the initial search screens are shown in figure three.

Table 2: Interfaces Tested in Folder Search

| Interface	Format	Finding Aid Description	Search Options Available
6	Portable Document Format (PDF)	**Northwestern University Archives**, finding aid for Nathaniel Nathanson Papers: Replicates printed finding aid. Included biographical statement, description of sub-series, and index.	Table of contents in left pane of Acrobat reader, with links to particular box numbers
7	Non-searchable EAD	**University of Houston**, finding aid for Nikki van Hightower Papers in Texas Archival Resources Online (**TARO**) EAD Cookbook-based stylesheet used to produce HTML. Entire finding aid on one page.	Left hand search bar with hyperlinks to "descriptive summary," "biographical note," "scope and contents," "restrictions," "administrative information," and "detailed description."
8	HTML	**Princeton University**, Public Policy Papers, finding aid for William Colby Papers. HTML with photograph at the top. Arrangement section includes links to summary paragraphs describing sub-series in finding aid.	Table of contents at top, with hyperlinks to "introduction," "Biography," "Arrangement," "Description," "Folder List," "Timeline" bookmarks in body of finding aid.
9	Searchable EAD	**Yale Archives and Manuscripts**, finding Aid for George Rosen Papers. Dynamically produced HTML from EAD source file. Two windows load via javascript pop-up, with main body of finding aid as HTML in larger right window; navigation bar in left pane	Left window with links to "Collection overview," "Correspondence," "American Public Health," "Association Files," "Subject Files," "Writings," "XXXX University Files," "Personal Papers and Photographs," and "Oversize." Also includes a search box

Figure 3: Interfaces Tested in Folder Search

Interface 6: Northwestern
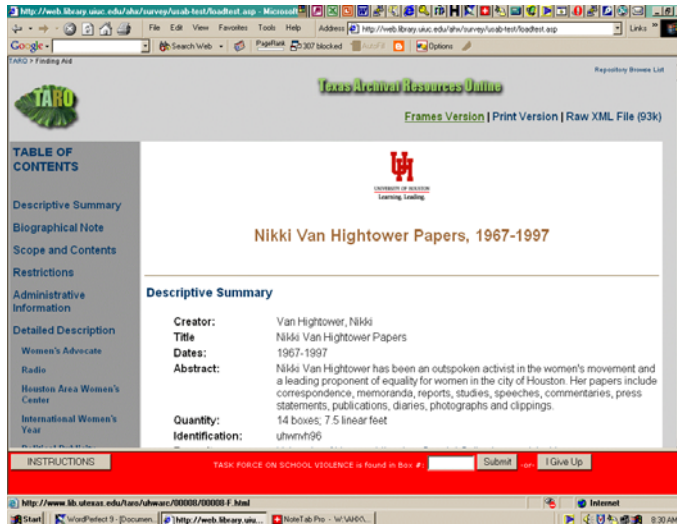
Interface 7: TARO/Univ. of Houston
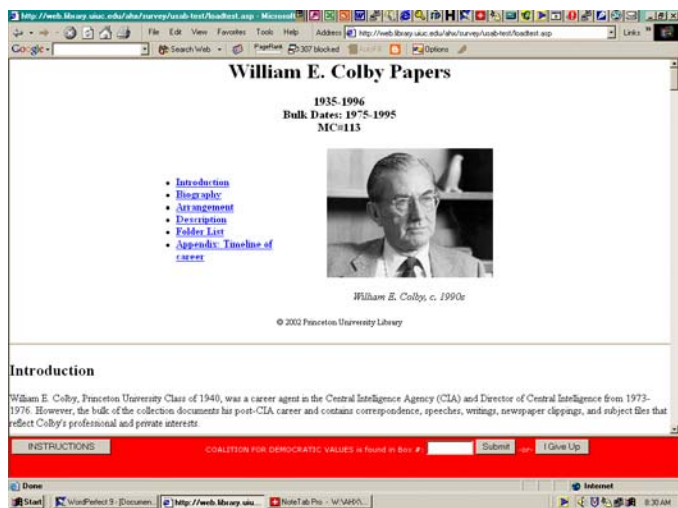
Interface 8: Princeton
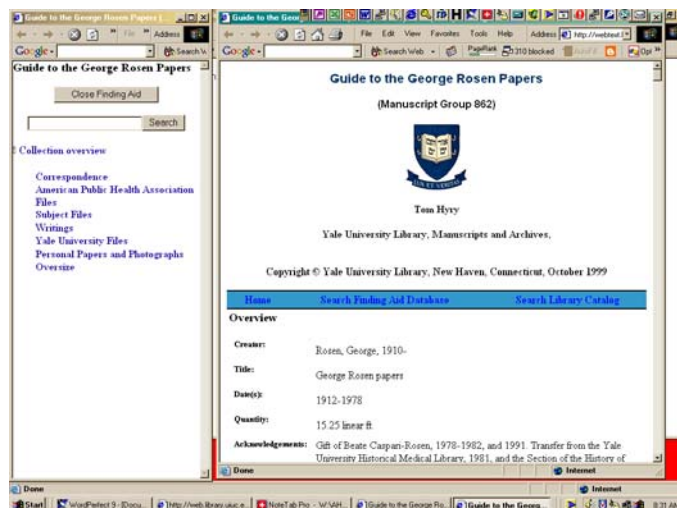
Interface 9: Yale

The on and off-site participants worked through the same tasks; a word is needed regarding the data collected for each group. For both groups, webscripts automatically recorded questionnaire responses to a database. Search results and times needed to complete or abandon each task were similarly recorded.

In addition, a great deal of qualitative information exists for on-site participants. The project director or his assistant observed the search sessions and took notes on a coding form, recording search patterns, mouse movements, and keystrokes. Distinct emotional reactions and audible comments were noted, and after the session respondents were interviewed using a set of loosely-structured questions. The project director or his assistant asked participants to provide some background on their use past use of archives and strategies in locating archival material relevant to their needs. They were encouraged to explain their thought process while searching and to suggest improvements to the interfaces.

Although it ultimately proved impossible to gather a similar amount or quality information for those participating off-site, some efforts were made.[21] Off-site participants were provided the opportunity to comment on each task via webforms. These comments offers some supplementary data to the more detailed on-site qualitative information.

Data Coding and Analysis—After participant sessions had ended, the raw data were coded, correlated, and analyzed, with more precedence being placed on the on-site results. For on-site respondents, users' preferred search strategies (as determined from observation coding forms) were recorded in the database.   Descriptive statistics were calculated for each task and t-tests were completed to determine whether differences would likely be replicated among the entire population of finding aid users. In addition, the search times were correlated to status as

an archival expert, computer expert or novice. Information from the questionnaire responses was used to classify participants.[22]   The descriptive statistics, t-scores, and correlated data were examined for patterns, and the observers' notes and transcribed interview tapes were consulted for qualitative supporting information.

Results

  Respondent characteristics—Of the 89 individuals who fully completed the study, 35 did so on-site under observation, 54 off-site and unobserved. The sample comprises a relatively balanced age group, with 48 users under age 36 and 40 users over.   Approximately two-thirds of both the on and off-site respondents were female.

  The demographic characteristics of the sample reveal heavy weighting toward graduate students, faculty, or staff members, with only 18 (20%) identifying themselves as members of the public and 11 (12%) as undergraduate students. Closer examination reveals that the targeted populations of archival experts, computer experts, and novices were well-represented within the sample. Forty six participants were classified as self-reported archival experts (9 of these participated on-site; 26% of on-site sample), 45 as self-reported computer experts (11 on site; 31%), and 22 as novices (16 on site, 46%), using the criteria discussed above. Twenty-four (one on site) reported themselves as both archival experts and computer experts.

  Twenty-seven (77%) of the on-site participants were either graduate or undergraduate students, and they included both expert and inexpert searchers. The sample of on-site users also included five members of the campus administrative staff, an important user community for most university archives. The on-site participants included twelve history students or historians (34%) and nine library school students or librarians (26%). Fourteen (40%) had other

affiliations, including psychology, electrical engineering, agricultural economics, kinesiology, and campus administration. No genealogists were included.

Other questionnaire responses support the conclusion that the sample included many people with archival and/or computer expertise. Sixty-four members of the entire sample (72%) claimed to have used archives at least once in the past year. Similarly, 67 (75%) stated that they have used electronic archival finding aids.

Interviews completed with the on-site participants provide some insight into participants' past experiences with archives and their expectations for on-line archival access.   Most of the graduate students claimed to be familiar with archival research and with on-line searching for archival materials. They tended to see finding aids as an essential but limited search tool. One history graduate student "[couldn't] say with any confidence that I have adequately dipped into an archive." As part of her training she had received no instruction in how to search for archival materials and thought her past success with archives was "really only because of the kindness of the archivist." Another dissertating graduate student described her main strategy for locating collections: "Often it's listed in publications somewhere. Somebody will mention they've found this information and such and such location." These behaviors (backward chaining, footnote chasing, reliance on personal connections) have been well described in previous studies.[23] Another graduate student noted:

> "At this point in my research, as I'm still very junior, I generally look at other monographs or surveys to see what sources they have used and where they are. If I get pointed in the right direction in the archive I usually will contact the archive librarian or the curator of the historical society and ask them how I should do it [find additional information]."

For users such as these, the finding aid serves a very specific purpose. As one noted, "when they

are very detailed. . . it can let me focus in on certain dates that I know I am particularly interested in. It might also describe the content of certain correspondents." Another respondent was very happy with one institution's printed finding aids "for everything that was in the collection. So I could sit there with the finding aid and know what I've got in each box. Exactly what folder I needed. . . . I rarely misfired." But she was also skeptical that on-line finding aids would be complete, and strongly believed she needed to view a printed finding aid or speak to an archivist about the collection before using it in person.

Novice researchers illustrated less understanding of the complexities of seeking archival information. They were much more trusting of electronic data. One thought a finding aid must be "somebody who's helping you find something whereas a search engine you're finding it yourself." Another "something that not everybody has. Probably a java applet that someone puts on a website or a paper index of an archive." One undergraduate who had used a university archives exhibited a expectation for immediate search results, nothing that "if you're looking for any resources you usually look at a quick search button. For any library resource." Another student believed that any hyperlinks in finding aids would lead to scanned copies of documents, since "actual documents online . . .would be most helpful. If they were there, you could click on them, possibly download them, possibly print them, so you could have them and use them." Campus administrators showed a similar need for direct access to information; one noted she wanted little "snippets" for inclusion in campus promotional materials. She said she would be unlikely to search a finding aid herself (although I know from past experience that she has.)

Participants' comments illustrate some of the tensions archivists face in attempting to provide information for both highly skilled and novice users. Experienced users want full access

to all descriptive information but are skeptical that they will find it on-line. Inexperienced

archival researchers have little understanding of what archives are and many assume that

everything — including the actual archives — are available through the website. One can see

these differences reflected in an analysis of the search results.

Task Results—Respondents searched each interface in a controlled setting for known

items, either a collection ID or a folder title. For clarity, the results for each set of interfaces will

be discussed separately, but a general point must be made first. As shown in figure four, the

average search times for inexpert and expert users varied significantly, with inexpert users taking

about 30 seconds longer on average to complete tasks. It is also interesting to note that computer

experts searched the finding aids about as quickly on average as archival experts, implying that

both factors play a significant role in efficient navigation through electronic finding aids.

Figure 4: Search Times for Expert and Inexpert Users in Seconds
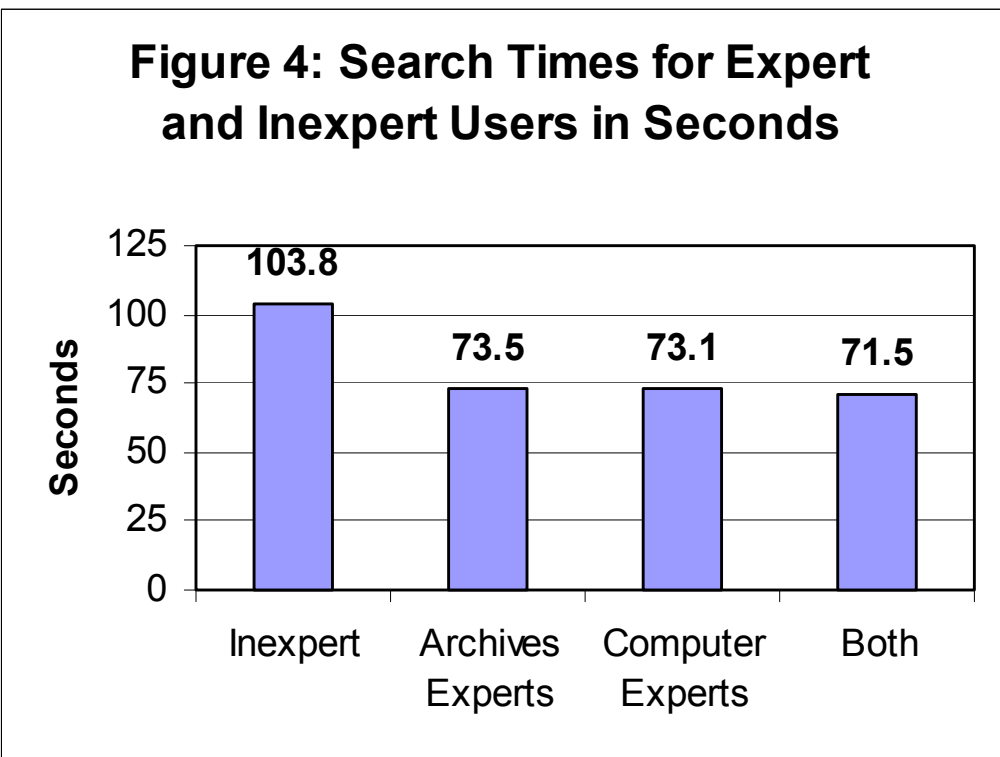
Collection Level Searches— For the collection level searches (tasks 2-5), participants

were provided the title of a known collection of personal papers and asked to find the collection

identifier. Table three provides summary data for four evaluative criteria, sorted by interface:

How many people found the collection record they were seeking? How many seconds were

required for the average successful search?   How variable were the search times? How many

steps (i.e., hyperlink clicks, text entries) did users take before completing or abandoning the

task?

Table 3: Collection Level Search Summary

	Off Site Users	On Site Users	All Users
Interface 2: 11 options — UNC			
Percent correct	85%	89%	87%
Mean time in sec.	86.5	95.7	90.2
Standard deviation in sec.	61.6	48.1	56.4
Mean number of steps	N/A	5.7	N/A
Interface 3: two options — OAC			
Percent correct	98%	97%	98%
Mean time in sec.	78.8	105.0	90.0
Standard deviation in sec.	77.1	73.1	76.1
Mean number of steps	N/A	5.3	N/A
Interface 4: one option — Princeton			
Percent correct	93%	86%*	90%*
Mean time in sec.	37.2	58.3	45.1
Standard deviation in sec.	28.9	70.9	49.9
Mean number of steps	N/A	3.1	N/A
Interface 5: three options — Illinois			
Percent correct	89%	86%	88%
Mean time in sec.	74.8	76.0	75.3
Standard deviation in sec.	50.6	63.8	55.7
Mean number of steps	N/A	5.6	N/A

*The website was unavailable at the time four on-site completed the task,
lowering the nominal success rate; results are included here with qualification.

Users most often found the requested information in interface three (OAC), although the

overall success rates were high for each task.[24] Based solely on overall success rates and time

efficiency, respectively, the OAC and Princeton designs were most efficiently searched for

known collections of personal papers. The Princeton design in particular was searched very

efficiently; on-site participants used two fewer steps, and the cited success rate of 86% would

likely have been higher, but problems with Princeton's webserver prevented at least four

participants from completing the task through no fault on their part. For those who successfully

found the requested information, the average search time in the Princeton interface was half that

of UNC's and OAC's and 3/5th of that of UIUC's. How representative are these results, and

what do they mean?

For each of the comparisons provided in table four, confidence levels indicate the

likelihood that the search time differences would be repeated across the entire population of

finding aid users.[25] The bolded lines show as strong likelihood that the Princeton interface will

be searched more quickly by the population of users conducting searches for known collections.

It provided a simple list of collections with A-Z hyperlinks at the top of page, allowing users to

jump quickly to the content of most interest. OAC provided a similar list (albeit without the A-Z

hyperlinks) in addition to a search box. OAC's search feature would of course be very useful in

an uncontrolled setting, but for this test, it appears to be related to increased search times.

Table 4: T-test Results for Collection Searches

Interfaces Compared	T-score	Level of significance	Confidence Level
2 vs 3	.03	.489	51%
2 vs 4	**5.32**	**<.001**	**>99.9%**
2 vs 5	1.67	.0489	95.1%
3 vs 4	**4.54**	**<.001**	**>99.9%**
3 vs 5	1.43	.078	92.2%
4 vs 5	**3.59**	**<.001**	**>99.9%**

Participants searched more slowly and/or with less success through the interfaces that had

the most search options or paths. The UNC site provided 11 options, including a link to a library

OPAC and a Google site search, but it posted the highest mean search times. While non-expert

participants searched the OAC interface much more slowly than experts, experts and non-experts

alike seemed to encounter some problems with the UIUC and UNC interfaces, as discussed in

more detail below.

Correlating average search times by participant type provides some unexpected results.

As noted above, participants searched the OAC interface with a very high success rate. But as

shown in table five, inexpert participants searched the OAC interface in a disproportionately slow

fashion, not only relative to experts completing the same task, but also to their own times in the

Princeton, UIUC, and UNC interfaces. This suggests that the OAC interface may be more suited

to the navigational practices of experts than archival or computer novices. Other statistical

evidence supports this conclusion. The OAC site exhibited a high variability in search times,

illustrated in a high standard deviation.   While quite a few respondents worked through the OAC

site very quickly, the response times tend to tail upward more quickly than in other interfaces.

This interface apparently caused some confusion for a significant minority of users, particularly

novices. (Similar points can be made regarding the UNC interface, but not as strongly.)

**Table 5: Average Search Times in Seconds by Task
and Status: Collection Level***

	2 (UNC)	3 (OAC)	4 (Princeton)	5 (UIUC)
All Participants	87.8	93.2	52.9	81.2
Inexpert	**112.8**	**124.5**	73.7	96.4
Archives Expert	81.5	73.3	43.7	78.8
Computer Expert	78.4	81.0	41.3	86.8
Both	81.1	60.8	32.4	101.0

*Includes times for successful and unsuccessful searches

These hypotheses can be refined and tested by analyzing on-site participants' observed

search behaviors and interview comments. Watching participants work with archival interfaces and soliciting their opinions illustrates different navigational techniques used by novices and experts and helps explain why certain designs were more efficiently utilized.

The Princeton and OAC designs seem to have aligned well with participants' stated preferences. In the questionnaire, 26 of the on-site respondents (74%) claimed that when browsing through an individual website they preferred to scroll through the website and click links. Only nine claimed to prefer search boxes or forms.   Figure five shows the observed search styles coded for the collection-level interfaces (two through five).

Figure 5: Observed Search Behaviors: Collection Search
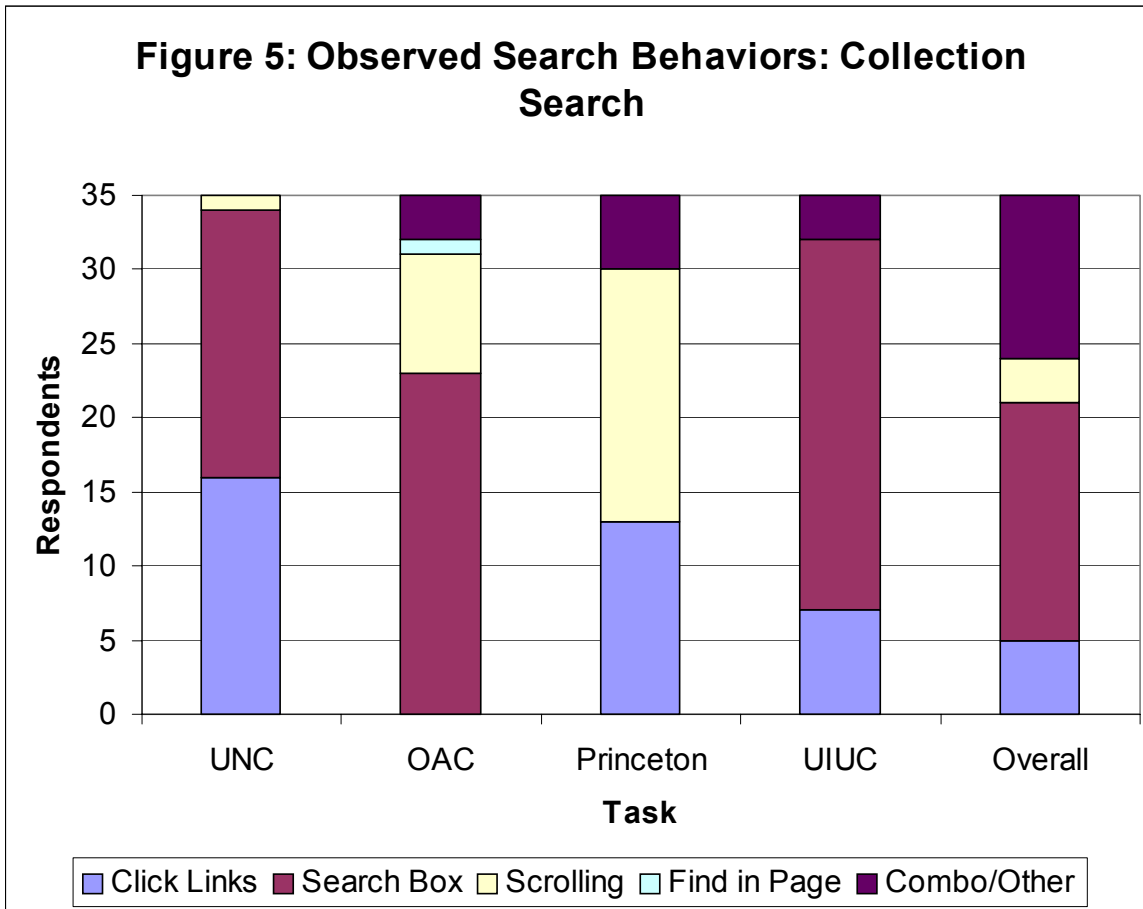
However, participant behaviors as observed and coded do not reflect their stated preferences for scrolling and clicking. Search behaviors appear to be highly conditioned by the existence of certain design features on a website. For example, most users used a search box when one was available, whatever they might have stated in a questionnaire.[26]

It appears likely that some web site designs may encourage inefficient search habits because they do not provide ready access to browse functions. Ironically, the only interface without a search feature was searched more quickly on average than the other three. As one participant noted in commenting on the UIUC interface, "I used the search function since it was so prominent. I would assume Taft's papers would be the first hit, but they were not. I definitely prefer a browse if I know what I am looking for."[27] Other participants commented that searching was "easy [when the list is] broken into alpha."[28] Another thought it was "much easier . . . than searching in a catalog."[29] Many participants specifically noted in their submitted comments or interviews that such alphabetical lists are easy to use.   If a list could not be found on a given page, some participants would begin to look for one, scrolling to the bottom of the page if necessary. These findings would seem to reinforce those of a recent study of genealogists, which highlighted the prominence they place on name searching.[30] Name browsing has been shown to be important to other archival experts, such as history students.[31] Since these behaviors are prominent and effective among archival experts, interfaces for archival finding aids should encourage them.

Some novices use browsing or search boxes as effectively as experts do. However, many others tended to follow dead ends or irrelevant links.

This can be demonstrated most readily by comparing expert and novice search sessions in

the OAC interfaces. One expert searcher entered the search term in the search box and found the

collection listed in the hit list. After hovering over the link for a few seconds, he clicked the link,

entered the full finding aid and quickly found the collection identifier. Other experts scrolled

down to the alphabetical list and found the collection in a matter of seconds. (Some novices used

similar strategies, whether by accident or design.) The OAC interface was efficiently used by

experts because the main part of the page provided two options that matched different search

styles, but it did not confuse people with too many choices.

On the other hand, many novices got lost in options not relevant to the task at hand (e.g.

those found in sidebars). This was true for all interfaces, but especcially for the OAC searches.

One novice's steps in the OAC page were recorded by the project assistant as follows:

- • Scanned interface for 20 seconds
- • Clicked back button to the task instructions pane.
- • Clicked forward to the main interface, rescanned the page
- • Clicked "UC Berkeley Bancroft Home Page" under "Related Links"
- • On Bancroft Homepage, clicked "Collections" link
- • Scanned left hand table of contents on collections page
- • Clicked "Reference and Access" link on top navigation bar.
- • Clicked link labeled "Bancroft Library Finding Aids"—Led to list of Berkeley repositories on old OAC (Dynaweb) site
- • Clicked "University Archives"—led to list of finding aids.
- • Clicked "Bancroft Library Finding Aids"
- • Enters "Arther D. Eggleston" [sic] in search box—no hits.
- • Tries several other seaches and links in Dynaweb interface
- • Gives up on search.

This participant moved very slowly throughout the task; his overall search times was over

five minutes before he gave up. Many novice participants exhibited similar confusion and false

starts while searching the UNC pages.   The initial search screen provided many options,

including a link to a Google site search, a quick link to the parent library's OPAC (which contains

catalog records for manuscript collections), and several other options. As a result, many hesitated

before searching and never seemed confident they were on the right track. Those who clicked the

OPAC search link often became particularly confused. One participant's search included the

following sequence of steps:

- read page for a few seconds
- hovered cursor over link to on-line finding aid
- scrolled down to bottom of page
- scrolled back up
- hovered over link labeled "Search Manuscripts Department Website"
- clicked the link, began to enter search term into Google site search
- abandoned the search
- clicked browser's back button, returned to main page
- clicked another link which led unexpectedly to an alphabetical list
- scrolled down the list
- clicked a hyperlink labeled with name of collection being searched for
- found the requested information.

This participant's overall search time was 90 seconds, slightly above average.   Many novice

users did not know where to begin searching. As one person noted for the Illinois search screen

(another search with multiple options), "I felt impatient about reading through all the text on the

first page and wondered which link to follow. It seems to me that users just want the key to the

finding aids as quickly as possible, so whatever gets them there fastest and with the fewest clicks

is best."[32] Several participants pointed out that advanced search options will likely be appreciated

only by a very small number of experienced users. In at least several cases, pull down menu

options led to bad search results when participants accidentally changed the search to an

irrelevant field such as "call no." or "date."

Other participants ran into problems if search engines did not use expected conventions.

For example, the UIUC interface did not divide phrases into separate terms joined Boolean "and"

operators (like Google does). Several participants were surprised when they entered the

collection name but received no hits, forcing them to redo their searches several times. The

results screen for UIUC also caused some confusion since it sorted hits by university department (reflecting the archival idea of provenance), but many participants expected the results to be relevance ranked. When searches failed in this interface, some users blamed bad indexing, but the problems they experienced could most likely be fixed by changing the search algorithm and display order.

Results from the collection-level searches illustrate that novices and experts use the same basic navigational strategies of looking for lists or using search boxes in attempting to find collection records, but that experts are much more subtle in their approach, leading to greater serach efficiency. Both archival and computer experts seem to have a good understanding of the conventions under which searching and browsing collection descriptions take place. However, novices are easily confused or led astray by extraneous options, so any interface which does not provide easy access to collection lists and a simple, one or two-path search option will likely cause confusion for users searching for a named collection.

Folder Level Searches–Results of the folder level searches, on the other hand, illustrate fundamental differences between how novices, archival experts, and systems experts navigate on-line finding aids to find information at the folder level. For the four folder-level tasks, participants were provided the name of an organization and asked to find the box number in which information regarding that organization was located. For each search, a single finding aid was loaded into the browser. Table six provides summary data, using similar evaluative criteria as provided for the collection level searches.

Table 6: Folder Level Search Summary

	Off Site Users	On Site Users	All Users
Interface 6: PDF — Northwestern			
Percent correct	91%	77%	85%
Mean time in sec.	89.1	128.4	103.1
Standard deviation in sec	70.7	104.28	85.6
Mean number of steps	N/A	9.6	N/A
Interface 7: Left nav. bar — TARO, EAD Cookbook			
Percent correct	96%	83%	91%
Mean time in sec.	57.5	73.8	63.4
Standard deviation in sec.	39.9	52.1	45.1
Mean number of steps	N/A	5.1	N/A
Interface 8: Top TOC — Princeton			
Percent correct	100%	74%*	90%*
Mean time in sec.	45.0	60.0	49.9
Standard deviation in sec.	26.8	35.9	30.7
Mean number of steps	N/A	4.0	N/A
Interface 9: Searchable EAD w/ pop ups — Yale			
Percent correct	78%	89%	82%
Mean time in sec.	97.0	75.4	87.8
Standard deviation in sec.	67.9	39.0	58.1
Mean number of steps	N/A	4.7	N/A

*The website was unavailable at the time four on-site completed the task,
lowering the nominal success rate; results are included here with qualification.

 As with the collection-level search, participants were generally very successful in finding

the information requested, although overall success rates were slightly less, in a range of 82 to

91% as compared to 87 to 98%. However those participating on-site had much less success

completing the folder-level searches. For example only 77% of those on-site found the

information requested in task six; similar differences occurred in the other searches. The mean

search times and number of steps needed to complete a search in the Princeton finding aid were

also significantly lower than for the other searches. The results in table six show that the

difference in mean search times between interface eight (Princeton) and each of the others was

statistically significant. Across the entire population of users, it is very likely that people will

navigate through the Princeton interface faster than the alternate designs provided in this test, at

least when attempting to locate information on a known topic as asked to do in this study.

Table 6: T-test Results for Folder Searches

Interfaces Compared	T-score	Level of significance	Confidence Level
6 vs 7	3.60	<.001	>99.9%
6 vs 8	**5.11**	**<.001**	**>99.9%**
6 vs 9	1.28	.102	89.8%
7 vs 8	**2.22**	**.014**	**98.6%**
7 vs 9	2.90	.002	99.8%
8 vs. 9	**4.98**	**<.001**	**>99.9%**

By contrast, participants searched the PDF finding aid (interface six, Northwestern) with the least success and speed. The overall success rate was still high, but mean search times were much higher than for other finding aids. Search times were also more variable, and more steps were needed by the average user. Figure five plots the successful search times for on-site respondents on a line graph.

Figure 5: Successful Search Times For On-Site Participants—Folder Level
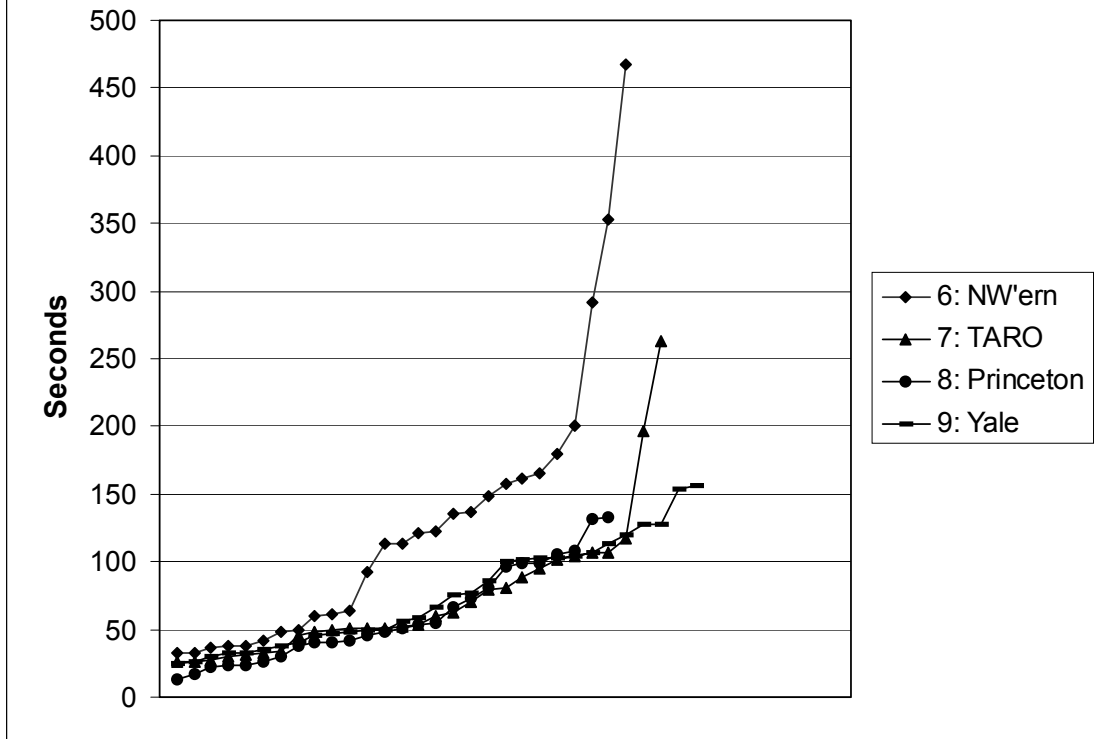
Eleven of the 27 people who successfully searched interface the PDF finding aid did so in one minute or less. (Close examination shows these were mostly expert users.) However, times shot up precipitously for the remaining individuals. Based solely on time efficiency, more users experienced problems with the PDF finding aid, and the problems they experienced were often substantial. Yale's searchable finding aid bears close analysis as well. Participants had the lowest overall success rate and the second highest search times while searching it. These differences were highly likely to be reflected in the entire population of finding aid users.

By contrast, interfaces seven (TARO) and eight (Princeton) were searched much more successfully and efficiently on average. Both included a table of contents with hyperlinks to specific sections of the finding aid. In the TARO finding aid, the table of contents was provided in a left-hand table occupying approximately 20% of the screen. This design feature is common outside the archival realm, and has become very popular for finding aids as well. The Princeton finding aid, which used a table of contents at the top of the page, was searched 13.5 seconds more quickly than the design using the left-hand navigation bar. (While the t-scores indicate that a similar difference would likely be found across the whole population of finding aid users, such small time differences are not necessarily significant in terms of the functionality of the interface.)

Whether time efficiency is deemed significant depends on the extent to which search speed is an appropriate metric for measuring finding aid effectiveness. Based solely on the quantitative results, it would appear that participants made more efficient use of the TARO and (especially) Princeton designs. Clearly, other factors such as participant background and their observed behaviors in working in with the finding aid must also be examined. They help identify

key design elements which facilitate effective (or ineffective) navigational strategies.

As shown in the bold entries in table seven, comparisons between expert and inexpert users did not show significant time differences *within* either the Princeton or Yale finding aids (although obviously when comparing *between* the two finding aids, the Princeton finding aid was searched more quickly). The design elements included in these finding aids seem to have equally benefitted (or hindered) both experts and novices. By contrast the TARO and especially the Northwestern finding aids show a significant gap between expert and inexpert search times, suggesting that some specific elements in these finding aids might account for the difference. In the Princeton finding aid, novices's search times appear to have decreased to the point where they are comparable with experts'. In the Yale finding aid, expert times appear to have increased to the point where they are comparable to novices'.. This suggests that fundamental differences in interface design can dramatically affect the ways in which (and the efficiency with which participants use archival description..

Table 7: Average Search Times in Seconds by Task and Status: Folder Level*

	6 (NW'rn)	7 (TARO)	8 (Princeton)	9 (Yale)
All Participants	126.1	72.3	49.5	88.0
Inexpert	187.5	90.6	**58.7**	**86.5**
Archives Expert	106.0	68.4	**48.2**	**88.9**
Computer Expert	102.0	60.7	**45.4**	**89.5**
Both	98.6	59.8	47.7	91.0

*Includes successful and unsuccessful searches

To test this hypothesis, participants' navigational styles, interview responses, and submitted comments were examined. Figure six shows participants' dominant search strategies for each interface.

Figure 6: Observed Search Behaviors: Folder Search
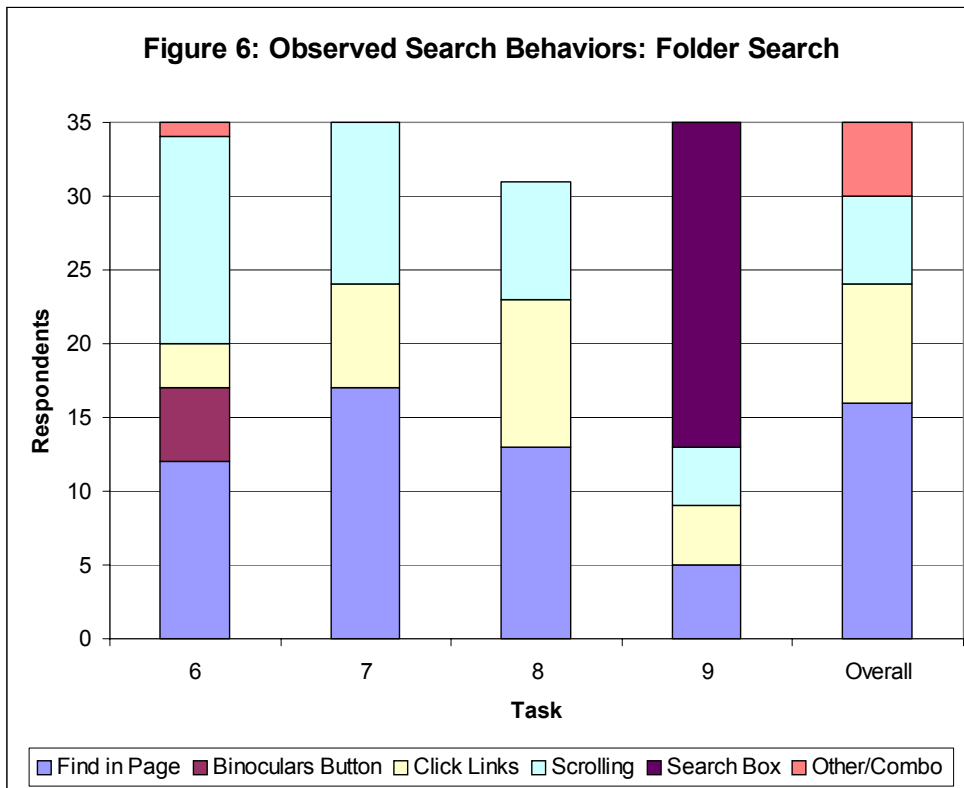
Participants generally chose one of two strategies: scrolling through the finding aid or using the browser's find-in-page (CTRL-F) function.[33] Many participants appear to be familiar with the find-in-page function, indicating that they can effectively search finding aids even when no search capability is provided on the website itself.[34] One finding aid offered a query box to search the finding aid (Yale). When this option was available, 22 of 34 participants (63%) attempted to use it, suggesting that many users of finding aids will make use of a search function when it is limited to a single collection or finding aid.[35] As discussed in more detail below, Yale's search option often failed for technical reasons, explaining its high average search time.

Many of the computer experts (which included both archival experts and novices) attempted to find the search term using tools built into the browser, taking little if any time to try to understand the organization (or even name) of the collection they were searching. A typical participant following this strategy clicked a few links or scrolled a few lines into the finding aid, then used the find-in-page function to identify the box in which it which relevant materials were located. Some novices were apparently not aware of the find-in-page function or did not think it would be helpful for the types of searches they were asked to do in the study. One had the habit of leaving the finding aid altogether to look for the requested information. During the interview, he noted that he looks immediately for a search box on each page he visits and that if one is not available, he leaves the site or goes to a higher level. During the search of the TARO finding aid, the project assistant recorded the following steps in his search:

- read top right-hand headings
- scrolled down a few lines
- scrolled back up
- clicked small TARO link at top
- Entered truncated search term in query box on main TARO page
- read links returned by search

- clicked link to Lawrence Pope Collection [i.e. the wrong collection]
- executed CTRL-F search in Pope finding aid; search failed
- clicked back button to return to search results screen.
- Gave up

Rather than searching the finding aid relevant to his search, he left the site and became lost. While this search may seem atypical, it illustrates the ways in which novice searchers and computer experts rely on computerized tools to search for them.

Some archival experts, particularly graduate students fairly early along in their training, used similar strategies, i.e. they relied on search boxes and CTRL-F. However, the project director and his assistant observed very different sets of search behaviors for many of the nine archival experts who participated on-site, especially those who during their interviews described extensive archival experience.

These participants browsed through the collection as a whole, often reading the scope and content note or the narrative series descriptions. Based on information gathered during this initial scan, they clicked links and scrolled through the collection, making educated guesses to visually locate the term searched for.

For example, one advanced graduate student habitually scanned the initial screen quickly and scroll down to the container list. If she did not find the search term immediately, she returned to the top of the finding aid and clicked a series heading that looked likely to contain the information. (Unlike many other uses, she did not use the search box in the Yale finding aid, instead clicking the link to the "Subject File," where she found the folder title only 20 seconds into her search session.)

Another archival expert followed this sequence of steps upon entering a task:

- Scanned the initial page

- hovered mouse over links on left hand table of contents
- clicked link labeled "scope and contents,"
- skimmed the scope and content note
- clicked back button to return to top
- hovered over link labeled "Detailed Description,"
- clicked link to first subseries within collection
- read short summary of subseries
- scrolled through folder list under the summary
- scrolled into next the subseries and read its summary
- scrolled down and found the folder heading he was seeking

Similar behaviors were common among other experienced archival users observed during the study. (Unfortunately, since no archivists participated under observation, it is impossible to say whether they used a similar search strategy.) During their search sessions, archival experts wondered aloud "How is this file organized" or "I wonder what category it would be in."[36] These participants noted during the interview that they liked to get an overview of the entire collection before searching it for a specific item. (Even one participant who has conducted archival research for over 50 years but never used an electronic finding aid used this search strategy when working on-line.) When asked to explain his actions, another noted:

> I was trying to figure out how this was organized and where I would go to look for something because this left hand side where it has "detailed description" and it has some highlighted topics. I tried to figure out if those were the only highlighted topics or whether there was going to be more somewhere else.[37]

The use of these tactics by archival experts strongly implies that experienced researchers spend a significant amount of time making use of descriptive context and inventories. In one sense, this affirms archival descriptive practice, if not the display features typically associated with it. It corroborates the maxim that archivists should provide complete context for archival materials, whether the medium of delivery is electronic or paper.

In addition to the specific behaviors noted above, many users commented during their

interviews on the specific design features they liked or disliked in using finding aids. Since the comments submitted by expert and novice users were similar, they are included here as examples of common behaviors and preferences which should be taken into account when designing interfaces for individual finding aids.

Participants indicated that the use of unexpected file formats impeded efficient finding aid usage, even if they were eventually able to find the information they are seeking. This was clearly the case with the PDF finding aid. PDF format replicates a printed finding aid. People searched it more slowly and with less success (on average) than the other formats, and they tended to dislike it. After noting blandly that "PDF's are a problem," one summarized his objections to working with the files:

> When I see a PDF to me that's a snapshot of something printed. . . . When I'm looking for functionality in a document 99% html files are better. PDF's are bigger and more bulky, if I'm God forbid working from a modem, they take forever to download. If I want to open more than one window they take forever. PDF's are not good.[38]

Another blamed problems working with the format on himself, noting that he "didn't see a strategy for finding [something], and it may be that I don't understand archives or what I was looking for." Although sentiments from these inexperienced users would seem to indicate that archivists should not use PDF files to deliver finding aid text, they may have a useful role to play. PDF allows a high degree of format control and may also be useful as an alternate presentation to HTML, so that users can print copies of the finding aid to browse off line or bring to the archives.

Institutions should avoid complex search options within individual finding aids. Many participants encountered problems working with the Yale interface, which was delivered in two windows. The larger right-hand window contained the full finding aid text; the left hand window was a navigation bar with a search box and links to parts of components of the collection

description. It invoked functions on the main window. When these functions worked, they made

the finding aid jump to the appropriate spot. Some users appreciated the ability to search within a

finding aid and 63% used it as their first search tactic upon entering the page. However, many

were disconcerted by the fact that two windows popped open on the screen. (One remarked "pop-

ups are the spawn of Satan" when seeing the new window.) The search box failed for many users,

apparently because of a programming bug, leading to frustration. As one commented:

> First I tried using the search function which was in one of the two pop-up windows. It
> appeared that this would search within the finding aid, but apparently it was searching
> across all the finding aids. [This impression was mistaken.] Also, it loaded the results of
> my search in the same window (replacing the search engine screen with the finding aid).
> When I tried to go "back" and search again, the search function didn't work. So I went
> back to my old strategy of just searching using the browser "find" function, which worked
> fine.[39]

Another noted simply that this interface was "Just awful, I'm sorry. I would be a lot happier with

a straightforward web page using CTRL-F than with this setup."[40] Techincal problems explain its

high average search times.

Finally, archival terminology tended to confuse experienced and inexperienced

participants alike. When asked about the results screen provided during a search, one very

experienced archival user explained how even terms like "finding aid" can be ambiguous or

misleading:

> I mean I'm assuming when it says entire finding aid that every collection is on here. I
> mean that's what I'd assume by looking at that, that all the other finding aids are online. I
> mean it looks self-explanatory, there's 904 finding aids, and yet I'm thinking a finding aid
> is for a collection, it seems to me that there's more collections than that at the Bancroft. It
> seems like a very small number. What am I missing, that would be my question.[41]

Based on participant observations, the link labeled "Folder List" in another task caused

less confusion than "detailed description." While the latter carries only a faint tinge or jargon,

one participant noted "when I saw a link like that I assumed perhaps the full list is not on this page. I think I scrolled down for a second and there was really quite a bit of text there so . . . I clicked to folder list cause I thought it might take me to another html document."[42] Self-apparent layout and visual clues are as important as using simple terminology. As one user (an archivist) noted for the finding aid that included two separate windows, "I was very uncomfortable with this arrangement," and another though that "even though it worked for me, I don't find it particularly clear."[43]

Assessment

This study shines a dim crack of light in a very dark room. Archivists (at least in the published literature) have given relatively little systematic attention to gauging user interactions with and reactions to finding aids. And users' search behaviors are very complex.   Nevertheless, the results discussed above suggest some preliminary conclusions about how users navigate electronic finding aids and which features help them search most efficiently. This information, particularly when viewed alongside that of recommendations in the broader information seeking literature,[44] can help design better access systems.

Finding aids currently placed on line are most efficiently used by either archival and computer experts. Non-experts find them more difficult to work with. They need over 90 seconds on average to find a relevant item when participating in a canned search under observation. Results may vary in real life, and this is certainly much better than pre-web search times. But participants' interview comments suggest that few novice users will spend a minute or more trolling through a repository's website unless they are highly committed to finding relevant information. Novice users are hampered by their lack of archival expertise. But simplifying

display elements and limiting search choices will help all users, not just novices

While no one design will be perfect for both experts and non-experts, some features will help both groups accomplish their work more efficiently. In this respect, archivists and digital librarians should follow design practices which allow users to engage in structured browsing and keyword searches. While this advice may seem trite, other options are likely superfluous for most users. At the collection level, keyword search boxes using Google-like search rules and alphabetical lists will provide rapid access for users who have a good idea what they are looking for. This will facilitate efficient searches for those who use backwards chaining (footnote chasing) or name search techniques to gather basic information about a collection. Lists may also encourage serendipitous discoveries, but creator lists should not be the only option provided for collection-level searching on archival web sites.[45] As one user noted, it is easy to use lists "only if you know the title," so an institution should not rely on lists as the only access point.[46] Given this proviso, lists should be prominently featured on archival websites. They will help meet the information needs of users who rely on name collection techniques previously described by Charles Cole and those he cites by providing direct access to a list of names, without the filter of a search box.[47] They also match the expert user browsing tactics described in this study.

Users of electronic finding aids will find the information they need most quickly when presented with relatively few search paths and when those paths are clearly mapped out in the overall site design. As one participant noted for the Princeton finding aid (interface eight):

> [T]hey have a couple things on the left there that isn't the traditional frame style [i.e. the design used in the TARO stylesheet tested in task seven] but it sort of mimics it with the things that you click on the left side of the photo. They're just tags further down on the webpage but it's easy to see if you're looking for something in particular how to look [for it.][48]

The page's simple design provided subtle but powerful visual clues. The top right contained a photo of the collection's creator, drawing casual browsers into the body of the text. The top left included a bulleted list of six purple links, hinting to experienced web browsers that the document had no external links and could be searched using the browser's find-in-page function The links bore simple labels: "Introduction," "Biography," "Arrangement," "Description," "Folder List" and "Appendix: Timeline of Career." "Description" linked to short summaries of the collection's subseries. This allowed participants who utilized a browse-and-read strategy to quickly decide which parts of the collection might be most useful. Each subseries was linked directly to the corresponding folder list, further facilitating browsing. The finding aid for interface seven (the TARO/EAD Cookbook finding aid) exhibited many similar characteristics. Both illustrate the value of using simple display features to present archival description to both novice and experienced users.

Simplification is essential but must not come at the expense of completeness. Experienced archival users often browse through an entire finding aid, using inference to identify appropriate series or subseries in which information of interest to them might be located. They use headings within a tables of contents to jump quickly to appropriate sections of the finding aid. These users spend much time browsing the finding aid for items, and it is inadvisable to undermine or complicate their ability to browse. Institutions should not defeat the find-in-page (CTRL-F) function built into all major browsers[49]. At the same time, finding aids could benefit from an option for internal searches. This would allow experienced users to more readily target their searching. Many inexperienced archival users also want ready access to archival information and use search boxes to input search terms. Major software products with EAD

modules, such as DLXS, now allow searches limited to a single finding aid, and such capabilities should be a high priority for further development.

While this paper suggests some basic factors archivists might consider in designing finding aid systems, perhaps its most significant conclusion, in common with other user studies, is simply that a deeper understanding of users is needed. The findings given here were based on a controlled, canned test, given under lab conditions, to a relatively small number of people. This should not undermine their significance, but the findings must be qualified and refined by other user studies. There is ample room for research describing other kinds of archival searches (e.g. subject or provenance based searches) or for setting it more fully within the context of electronic information seeking behaviors undertaken by real users conducting real searches.

Even with the findings such studies might provide, it is unlikely that online finding aids will ever make the chaotic nature of archival systems wholly understandable to archival users. The well-known role of the archivist in mediating access to collections has been enhanced, not diminished, by the Internet.

Appendix: Survey Questions and Tabulated Responses

Responses	on-site(observed) = 35	off-site (unobserved) = 54	Total = 89

1.   I am a:

	Off-site	On-site	Total
University Faculty Member	11	3	14
University Academic	16	5	21
Professional/Staff Member			
Graduate Student	4	19	23
Undergraduate Student	3	8	11
Member of the Public	18	0	18
No answer	2	0	2

2.   My gender is:

	Off-site	On-site	Total
Male	15	13	28
Femal	37	22	59
No Answer	2	0	2

3.   My age is:

	Off-site	On-site	Total
under 26	7	9	16
26-35	9	13	32
36-45	12	8	20
46-55	13	2	15
55-60	3	2	5
No Answer	0	1	1

4.   If you are a university student, staff or faculty member, please provide your program

affiliation or department. Responses were coded by the project assistant as follows:

	Off-site	On-site	Total
Archivist	29	0	29
Librarian	14	3	17
Library School Student	4	6	10
Historian or History Student	3	12	15
Other Affiliation	4	14	18

5.   Do you use the Internet (World Wide Web) for personal use?

	Off-site	On-site	Total
Yes	53	35	88
No	1	0	1

6.   Do you use the Internet (World Wide Web) for research, course-related, or scholarly use?

	Off-site	On-site	Total
Yes	53	35	88
No	1	0	1

7.   From where do you most often access the Internet?

	Off-site	On-site	Total
Home	10	18	29
Work	1	11	12
School	43	6	49
I do not use the Internet	0	0	0

8.   In the past year, how many times have you used the University of Illinois Archives or

another archives or historical society?

	Off-site	**On-site**	**Total**
None	12	13	25
Once	1	6	7
2-5 times	12	7	19
6-20 times	7	7	14
21-100 times	11	1	12
More than 100 times	11	1	12

9. In the past year, how many times have you physically visited a public library or research

(university) library?

	Off-site	**On-site**	**Total**
None	1	0	1
Once	2	0	2
2-5 times	3	2	5
6-20 times	8	5	13
21-100 times	14	11	25
More than 100 times	25	17	42
No Answer	1	0	1

10. In the past year, how many times have you used electronic library resources such as

electronic journals, article abstract and indexing services, or library databases, including

library on-line catalogs?

	Off-site	**On-site**	**Total**
None	0	0	0
Once	0	1	1
2-5 times	2	3	5
6-20 times	6	4	10
21-100 times	10	9	19
More than 100 times	35	18	53
No Answer	1	0	1

11. Have you used archival finding aids in a paper format?

	Off-site	On-site	Total
Yes	46	15	61
No	7	11	18
Don't know	0	8	8
No Answer	1	1	2

12. Have you used archival finding aids in an electronic format?

	Off-site	On-site	Total
Yes	20	47	67
No	7	5	12
Don't know	8	1	9
No Answer	0	1	1

13. How would you characterize your skill in using computers (check one):

	Off-site	On-site	Total
I am a novice computer user; have difficulty using email, web browser, word processor.	0	0	0
I am an experienced computer user; have no difficulty using basic functions like email, web browser, word processor or spreadsheet	8	20	28
I am a highly experienced computer user, I can install software or hardware, develop web pages, use databases.	33	9	42
I am a "power" user; I can configure software, write macros or scripts, develop and manage databases	12	4	16
I can program a computer using a high-level language such as C or Java.	0	2	2

No answer	1	0	1

14.   When searching for information *on the entire Internet*, which of the following strategies

do you **MOST** prefer to use:

	Off-site	On-site	Total
Browsing and clicking links	3	3	6
Using search engines or forms	50	32	82
No answer	1	0	1

15.   When searching for information *on an individual website*, which of the following

strategies do you **MOST** prefer to use:

	Off-site	On-site	Total
Browsing and clicking links	42	26	68
Using search engines or forms	11	9	20
No answer	1	0	1

Endnotes

1.The author would like to acknowledge generous assistance from the University of Illinois Library Research and Publication Committee in funding the research upon which this article is based. Sarah Erekson, the project assistant, provided invaluable help in administering the survey, coding results, and compiling statistics. William Maher, Robert Burger, Elizabeth Yakel, Amy McCrory, John Rees, and Timothy Cole offered incisive comments upon earlier drafts of the article. Special thanks go to the 89 individuals who participated in the study and to Janet Olsen, Daniel Linke, Richard Szary, Kris Kiesling, Lynn Holdszkom, Janis Holder, and Julie Grob, who graciously granted me permission to critique their institutions' finding aids in this article.

2.William J.Maher, "Chaos and the Nature of Archival Systems," Presented at the Society of American Archivists, 56th Annual Meeting, Montreal, Canada, September 15, 1992. 1992. Available: http://web.library.uiuc.edu/ahx/workpap/chaosshort.pdf. (Accessed May 11, 2004)

3.Robert Spindler and Richard Pearce-Moses, "Does AMC Mean Archives Made Confusing? Patron Understanding of USMARC AMC Catalog Records," *American Archivist* 56 (1993): 331-32.

4.Kristina Southwell, "How Researchers Learn of Manuscript Resources at the Western History Collections," *Archival Issues* 26, no. 2 (2002): 91-109; Elizabeth Yakel, "Listening to Users." *Archival Issues* 26, no. 2 (2002): 111-127.

5.William Maher, "The Use of User Studies," *The Midwestern Archivist* 9, no. 1 (1986): 15-26; Paul Conway, "Facts and Frameworks: An Approach to Studying the Users of Archives," *The American Archivist* 49, no. 4 (1986): 393-407.

6.Paul Conway, *Partners In Research: Improving Access to the Nation's Archives*. Pittsburgh: Archives & Museum Informatics, 1994; Linda J. Long, "Question Negotiation in the Archival Setting: The User of Interpersonal Communciation Techniques in the Reference Interview," *The American Archivist* 53 (1989): 40-50; Helen R. Tibbo, "Interviewing Techniques for Remote Reference: Electronic versus Traditional Environments," *The American Archivist* 58 (1995): 294-310; Elizabeth Yakel and Laura Bost. "Understanding Administrative Use and Users in University Archives," *The American Archivist* 57 (1994): 596-615.

7.Frederick Miller, "Use, Appraisal, and Research: A Case Study of Social History," *The American Archivist* 49 (1986): 393-407; Jacqueline Goggin, "The Indirect Approach: A Study of Scholarly Users of Black and Women's Organizational Records in the Library of Congress Manuscripts Division," *Midwestern Archivist* 9, no. 1 (1986): 57-67; David Bearman, "User Presentation Language In Archives," *Archives & Museum Informatics* 3/4 (Winter 1989/90): 3-7; Wendy Duff and Catherine Johnson, "A Virtual Expression of Need: An Analysis of E-mail Reference Questions," *American Archivist* 64, no. 1 (2001): 43-60; Kristin Martin, "An Analysis of Remote Reference Correspondence at a Large Academic Manuscripts Collection," *American Archivist* 64 (2001): 17-42; Helen Tibbo, "Primarily History in America: How U.S. Historians

Search for Primary Materials, " *The American Archivist* 66, no. 1 (2003): 9-50; Elizabeth Yakel, and Deborah A. Torres, "AI: Archival Intelligence and User Expertise," *The American Archivist* 66, no. 1 (2003): 51-78.

8.Steven L. Hensen "Archival Cataloging and the Internet: The Implications and Impact of EAD," *Journal of Internet Cataloging* 4/3-4 (2001): 75-95; Steven. L. Hensen, "The First Shall be First: APPM and Its Impact on American Archival Description,"*Archivaria* 35 (Spring 1993): 64-70; Christopher J. Prom, "Does EAD Play Well with Other Metadata Standards? Searching and Retreiving EAD using the OAI Protocols." *Journal of Archival Organization* 1/3 (2002): 51-72; Christopher. J. Prom, "The EAD Cookbook: a Survey and Usability Study," *American Archivist* 65 (Fall/Winter 2002): 257-75; Clay Redding, "Reengineering Finding Aids Revisited: Current Archival Descriptive Practice and Its Effect on EAD Implementation," *Journal of Archival Organization* 1/3 (2002): 35-50; James M. Roth, "Serving up EAD: an Exploratory Study on the Deployment and Utilization of Encoded Archival Description Finding Aids." *American Archivist* 64 (Fall/Winter 2001): 214-237; Elizabeth J. Shaw, "Rethinking the EAD: Balancing Flexibility and Interoperability." *New Review of Information Networking*, 7 (2001): 117-131; and Jill Tatem, "EAD: Obstacles to Implementation, Opportunities for Understanding," *Archival Issues* 23, no. 2 (1998).

9.Rita L. H. Czeck, "Archival MARC Records and Finding Aids in the Context of End-User Subject Access to Archival Collections,"American Archivist 61/2 (Fall 1998): 426-440; Spindler and Pearce-Moses, "Archives Made Confusing?, 330-41.

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

11.Anne Gilliland-Swetland, "Popularizing the Finding Aid: Exploiting EAD to Enhance Online Discovery and Retrieval in Archival Information Systems by Diverse User Groups." (2001): 201. See also Matthew Young Eidson, "Describing Anything That Walks: The Problem Behind the Problem of EAD." *The Journal of Archival Organization* 1/4 (2002): 5-28.

12. She lays out some thought-provoking suggestions for systematically redesigning archival access (based on past users studies regarding user information-seeking, particularly the Getty project). See also Dennis Meissner, "First Things First: Reengineering Finding Aids for Implementation of EAD," *The American Archivist* 60/4 (1997): 372-87.

13.The Online Archive of California Evaluation Project made promising steps in this direction, laying out a detailed and systematic methodology for the evaluation of on-line archival resources, including an assessment of user needs and interface usability. However, only a planning report was issued. Anne J. Gilliland-Swetland, "Evaluation Design for Large-Scale Collaborative Online Archives: Interim Report of the Online Archive of California Evaluation Project," *Archives and Museum Informatics* 12 (1998): 177-203.

14.Domain experts are typically through of as those who understand the norms and content of a subject area. In other domains (for example, computer science, business, and law) domain

knowledge has been shown to play a significant role in facilitating faster, more effective electronic searches. (See Gary Marchinonini, Sandra Dwiggins, Andrew Katz, and Xia Lin. "Information Seeking in Full-Text End-User-Oriented Search Systems: the Roles of Domain and Search Expertise," *Library and Information Science Research* 15 (1993): 35-69.) Domain knowledge is often contrasted with systems knowledge. Systems knowledge entails an understanding of search techniques, data structures, and computer operating principles. Systems knowledge also significantly improves performance; and some literature suggests that in certain fields it may be even more important than domain knowledge. (See M. D. Kiestra, M. J. W. Stokmans, and J. Kamphuis. "End-Users Searching the Online Catalogue: The Influence of Domain and System Knowledge on Search Patterns." *The Electronic Library* 12, no. 6 (1994): 335-43.)

15. Historians or experienced archives users might be seen as domain experts, given their content knowledge in specific historical topics. But a domain expert in one field of history (say US Reconstruction Era) may be novice in another (say 20th Century US social policy), even if it seems to be related. Similarly, it is difficult to determine whether archivists should be considered domain experts, systems experts, or both. For example, archivists have always been the heaviest users of finding aids, but it is unclear whether familiarity with one or two finding aid systems breeds familiarity with other finding aid systems. Similar points could be made regarding archivists' historical knowledge.

16. Yakel and Torres have recently argued that "archival intelligence" plays a significant role in helping researchers work with archival materials. Archival intelligence includes an understanding of archival principles, theories, procedures (such as provenance), the application of specific strategies to reduce uncertainty (such as question asking), and the application of intellective skills (such as interpreting the relationship between finding aids and the primary sources) Yakel and Torres, "AI," 51-78.

17. As a starting point, the study assumes that users will likely come into contact with a repository's website at some point in their research, although even this most basic point has been untested with the exception of one published study which found that approximately two-thirds of US historians recently surveyed visit repository websites, with the highest rates of use (85%) among assistant professors. Tibbo, "Primarliy History," 23

18. Peiling Wang, "Methodolgies and Methods for User Behavioral Research," in *Annual Review of Information Science and Technology (ARIST)*, edited by Martha E. Williams (Medford, N. J.: Published for the American Society for Information Science by Information Today, Inc., 1999), 53-99.

19. Controls were placed on the search terms and interfaces in order to limit external factors which might skew the results, and users were given hints as to what prefatory label to look for (e.g. mss. #, record series #). The four search terms were controlled for spelling difficulty and the collection prominence, i.e. participants searched for the papers of individuals whose names would not be common knowledge to the average citizen. Each of the four finding aid retrieval systems complied with recommendations found in ISAD (G) (General International Standard

Archival Descriptions). In particular, they used the archival principle of multi-level description and included the basic elements needed to describe archival materials, such as creator, title, dates, extent, and scope/contents (i.e. a descriptive summary)

20.It was impossible to ask participants to search for materials on the same topic across each of the four collections, but the collections searched were approximately the same size, the finding aids were about the same length (30 printed pages), and participants were asked to search for a folder of material regarding a prominent organization. Folder titles corresponding to these organizations were located approximately midway through the finding aid, not at the beginning or end.

21.Although I briefly explored the possibility of using videoconferencing or mouse-tracking software for off-site participants, the technical challenges of doing so were too steep to overcome.

22.The decision to code an individual user as an expert in one or more categories was made based on her/his response to specific survey questions. While it might be argued that this method lacks the subtlety of correlating times to individual questions, segregating participants into three categories allowed for more understandable results.   In initial attempts to correlate data to specific questions, it proved difficult to determine which of the many variables tested proved most significant, although the data tended to indicate that archival and computer experience were important.   Participants qualifying as a self-reported "archival expert" met one or more of the following criteria: 1) Program affiliation as archivist; 2) Program affiliation as history AND status of faculty OR graduate student. 3) Used archives more than 5 times in the past year. Participants qualifying as a self-reported computer expert met one of more of the following criteria: 1) Characterized computer skills at "power user" level or above; 2) Used electronic library resources more than 100 times in the past year AND characterized self as a "highly experienced" computer user.

23.Marcia Bates, Deborah N. Wilde, and Susan Siegfried, "An Analysis of Search Terminology Used by Humanities Scholars: The Getty Online Searching Project Report Number 1," *The Library Quarterly* 63 (January 1993): 1-39. Marcia Bates, "The Design of Databses and other Information Resources for Humanties Scholars: The Getty End-User Online Searching Project Report No. 4," *Online and CD-ROM Review* 18: 6 (1994): 334; Stephen E. Wimberly and William G. Jones, "Patterns for Information Seeking in the Humanities," *College and Research Libraries* 50 (1999): 642.

24.The success rate for interface four (Princeton) would likely have been considerably higher, but the site tested was unavailable due to server problems during four of the 35 search sessions conducted on site; rather than discard the data it was thought better to cite it with qualification.

25. T-tests were performed to determine the probability that differences between the mean successful search times would be representative of the entire population of finding aid users. The t-test indicates the significance of a difference between the means of paired samples. One establishes a null hypothesis and if the difference is significant rejects it in favor of an alternative

hypothesis. In this case, the null hypothesis was posited that there is no significant difference in the search times for the interfaces between each of the four interfaces (H_0: $M_{t2}=M_{t3}=M_{t4}=M_{t5}$) and that the interfaces would therefore be searched with equal efficiency by the broader population of finding aid users.   The alternative hypothesis (H_A: $M_{t2}..M_{t3}..M_{t4}..M_{t5}$) stated that the interfaces would not be searched with equivalent time efficiency by the entire population of users.

26.After correlating the observed search strategies to the survey results, only six of the 26 who claimed to prefer browsing or clicking links actually used that as their primary technique.

27.Respondent number 15.

28.Respondent number 28.

29.Respondent number 43.

30.Wendy M. Duff and Catherine A. Johnson. "Where is the List with All the Names? Information-Seeking Behavior of Genealogists." *The American Archivist* 66, no. Spring/Summer 2003 (2003): 79-95.

31.Charles Cole, Name Collection by Ph.D. History Students: Inducing Expertise," *Journal of the American Society for Information Science* 51:5 (Mar 2000): 445.

32.Respondent number 61.

33.A smaller number tended to click hyperlinks to jump to positions of interest in the finding aid, a variation of the scrolling strategy which will be discussed in more detail below.

34. By contrast, only five (14%) of the on-site users used Adobe Acrobat's binoculars button, which provides a find-in-page function for PDF files as a backup to CTRL-F. Several off-site also participants commented on problems using the find-in-page with the PDF file.

35.As with collection-level searches, users' actual behaviors while performing folder-level tasks did not mirror their preferences stated on the survey. Only nine of the 35 participants claimed to prefer using search forms, but 22 used them when given the option alongside numerous browsing tools.

36.Respondent number 133; respondent number 135.

37.Respondent number 129.

38.Respondent number 110.

39.Respondent number 65.

40.Respondent number 115.

41.Respondent number 113.

42.Respondent number 110.

43.Respondent number 60; respondent number 129

44.See especially Gilliland-Swetland, "Popularizing the Finding Aid," 214-221 for some preliminary recommendations.

45.Title and subject lists may be similarly useful, although more research is needed to test their efficacy.

46.Respondent number 47.

47.Charles Cole, "Name Collection," 445.

48.Respondent number 129.

49.Depending on the exact combination of browser and Acrobat reader installed on a computer, PDF finding aids may not be searchable using the find-in-page (CTRL-F) function. Designers should also exercise caution in dividing a finding aid into more than one HTML page.