The Value of a Note: A Finding Aid Usability Study

Betts Coup
Houghton Library, Harvard University, elizabeth_coup@harvard.edu

The Value of a Note: A Finding Aid Usability Study

THE VALUE OF A NOTE: A FINDIaidNG AID USABILITY STUDY

Finding aids have long been an essential part of archivists' work. To create a finding aid is to create a surrogate of an archival collection. Multiple levels of description are used to distill information about the unique groupings and parts of a collection and to place its contents into context. Archivists make decisions about what to include in a finding aid based on their own judgment as trained professionals but also with the intent of creating a finding aid that will be genuinely helpful to researchers. Indeed, as the revised principles of Describing Archives: A Content Standard (DACS) state, "Users are the fundamental reason for archival description. . . . To make wise choices about descriptive practices, archivists must develop and maintain an awareness of user needs and behaviors."[1]

The practical challenge facing archivists is knowing what users will find helpful. Although archivists often seek to center users in their processes of description, these efforts are frequently based on guesses concerning what users will find useful. Very little research has been done to understand how users read and navigate archival description; there is a distinct disconnect between the intention of centering users and carrying out usability or user studies to understand user needs.[2] This study is intended to advance our comprehension on this matter.

The archival field has seen a shift in its discourse over the past fifteen years, emphasizing "More Product, Less Process" (MPLP).[3] Yet archivists rarely discuss the impact this shift has had on the users of our description. This study was developed with this fact in mind. It presents an analysis of how researchers use the content of our finding aids, deploying task analysis to shed light on what "usable" means for researchers in the post-MPLP descriptive landscape. This study is intended to help professional archivists better understand how users use finding aids and how reduced description impacts user experience.

Task analysis is "the process of learning about ordinary users by observing them in action to understand in detail how they perform their tasks and achieve their intended goals."[4] A design research method typically employed in usability studies for websites, task analysis identifies what users can do on a website, how they carry out tasks, and whether the navigability of the website needs improvement in order to be functional.[5] In this study, by contrast, we employed task analysis

---

[1] Technical Subcommittee on Describing Archives: A Content Standard, "DACS Statement of Principles," Describing Archives: A Content Standard. Version 2019.0.3, Society of American Archivists, https://saa-ts-dacs.github.io/-dacs/04_statement_of_principles.html#2-users-are-the-fundamental-reason-for-archival-description.

[2] There are likely many reasons for this. There is often a lack of resources and support for routinizing user experience work around finding aids, as well as a significant amount of education needed regarding the design and management of user experience studies.

[3] Mark A. Greene and Dennis Meissner's "More Product, Less Process" article catalyzed this shift. See Greene and Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68, no. 2 (January 2005): 208–63, https://doi.org/10.17723/aarc.68.2.c741823776k65863.

[4] "Task Analysis," Usability.Gov, https://www.usability.gov/how-to-and-tools/methods/task-analysis.html.

[5] User research is typically done to understand how design—particularly web design—impacts users and their interactions with (and use of) an interface, system, application, or website. Task analysis is just one method used in design research and usability studies to assess whether and how users complete specific tasks in a particular setting, and in this case was used to assess how users navigate and use finding aids in PDF form. "User Research Basics," Usability.Gov, https://www.usability.gov/what-and-why/user-research.html.

to assess the effectiveness not of the interface but of the content of finding aids—examining how users navigate the finding aid, how they make use of different levels of description, and how richer levels of description impact the usability of a finding aid.

Two questions are at the heart of this study. The first concerns what users read and use in a finding aid to determine if a collection is relevant to their needs. Do researchers use collection- or series-level notes valued by archivists, such as scope and content and arrangement notes, in order to understand what is in a collection? Or do users rely mostly on information found at lower levels, describing what might be in a single file or the content of a single item? The second question arises once we acknowledge that granular description of collection materials will not always be possible, especially in the post-MPLP world.[6] If granular description is not available in a finding aid, are users able to ascertain what is in a collection and whether it is relevant to their search?

Literature Review

Although finding aids have been available online for more than two decades and usability studies of web design have been around even longer, this study is distinct in that it applies user design research methods—specifically task analysis—not to the design of a web interface but to the multilevel content of archival description and to the impact on usability of the fullness of archival description at different levels.

To place this study into context, it is important to understand two intersecting but different research approaches: user studies and usability studies. As defined in Hea Lim Rhee's article "Reflections on Archival User Studies," a user study is "an archival investigative activity that collects, analyzes, and interprets data on users and use by empirical research methods," while a usability study "investigate[s] only the usability of archival access systems and websites, not users and use themselves."[7] This study applies usability study methods to finding aids and their content, rather than to websites or systems.

Much research has been done over the past twenty years regarding the usability and user experience of online finding aids, particularly those employing encoded archival description (EAD). Elizabeth Yakel's work has focused on user experience, and her articles have been based on both user studies and usability studies. In 2004, she completed a usability study of an EAD interface, and the analysis of her findings brought to light the ways archives-specific terms and the multilevel structure of description (and the challenges those levels raise for searching) were barriers for researchers when using online archival finding aids.[8] This was a continued examination of the challenges brought on by archival processing that she and Deborah A. Torres had outlined previously in the article "AI: Archival Intelligence and User Expertise," a user research study involving twenty-eight interviews with archival users.[9]

---

[6] See the Literature Review, below, for more detail about this argument.

[7] Hea Lim Rhee, "Reflections on Archival User Studies," *Reference and User Services Quarterly* 54, no. 4 (2015): 29–42, 30–31 (quotation).

[8] Elizabeth Yakel, "Encoded Archival Description: Are Finding Aids Boundary Spanners or Barriers for Users?" *Journal of Archival Organization* 2, nos. 1–2 (2004): 63–77, https://doi.org/10.1300/j201v02n01_06.

[9] Elizabeth Yakel and Deborah A. Torres, "AI: Archival Intelligence and User Expertise," *American Archivist* 66, no. 1 (2003): 51–78.

Yakel also published "Listening to Users" in 2003, which examines assumptions archivists make about users and their understanding of archives and archival access tools, such as finding aids. This study was based on a user study of undergraduate students and scholars.[10] In 2010, Morgan Daniels and Yakel completed a usability study of two online finding aid systems to determine what kinds of search strategies and approaches were more effective for users to identify the information they sought.[11] Yakel's work, both sole and coauthored, has been foundational to understanding how researchers use finding aids. However, this body of work focuses not on the content of the finding aids and how they—the various notes and fields—are being used but rather the challenges users experience with the structure and navigation of online finding aids. Yakel's work examines archives-specific language about arrangement and description, the multilevel organization of information, the structure EAD provides, how users search the systems through which online finding aids are accessed, and how these search strategies can be improved.

Other user studies have examined the audience for online finding aids and whether the costs of putting archival description online are worth it. In the 2004 article "Users of EAD Finding Aids: Who Are They and Are They Satisfied?" Lisa R. Coats reviews the contemporary literature on user studies of online finding aids. Coats discusses the user pool as it was understood at that time and how standardization due to encoding might improve user experience.[12] In the same year, Christina J. Hostetter called for the assessment of finding aid interfaces through usability testing, questioning whether the costs, labor, and related practicalities of encoding and maintaining finding aids in an online environment was beneficial.[13]

Additional studies have investigated online finding aid interfaces. A usability study by Jihyun Kim published in 2004 examined seventeen EAD finding aid interfaces. Kim considered data elements, labeling terminology, navigation, browsing, and searching, and found that despite the standards being used for encoding, finding aid functionality—including navigability, searching, and terminology—differed across repositories and that the long textual notes used within finding aids generated navigational difficulties.[14] Wendy Scheir's 2006 study of novice users working with online finding aids continued this line of inquiry. Scheir had a set of users complete tasks with finding aids online and found that, with experience, users developed a self-education process which led to better search and use. Scheir's study suggests that users become more proficient with finding aids over time, once familiar with the language and structure.[15] Xiaomu Zhou assessed fifty-eight EAD interfaces in a 2007 study and found that only a few were aided by searching functions and were taking advantage of searching technologies already readily available.[16]

---

[10] Elizabeth Yakel, "Listening to Users," *Archival Issues* 26, no. 2 (2003): 111–17.

[11] Morgan Daniels and Elizabeth Yakel, "Seek and You May Find: Successful Search in Online Finding Aid Systems," *American Archivist* 73, no. 2 (2010): 535–68, https://doi.org/10.17723/aarc.73.2.p578900680650357.

[12] Lisa R. Coats, "Users of EAD Finding Aids: Who Are They and Are They Satisfied?" *Journal of Archival Organization* 2, no. 3 (June 2004): 25–39, https://doi.org/10.1300/j201v02n03_03.

[13] Christina J. Hostetter, "Online Finding Aids: Are They Practical?" *Journal of Archival Organization* 2, nos. 1–2 (2004): 117–45, https://doi.org/10.1300/j201v02n01_09.

[14] Jihyun Kim, "EAD Encoding and Display: A Content Analysis," *Journal of Archival Organization* 2, no. 3 (June 2004): 41–55, https://doi.org/10.1300/j201v02n03_04.

[15] 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 (2006): 49–85, https://doi.org/10.1300/j201v03n04_04.

[16] Xiaomu Zhou, "Examining Search Functions of EAD Finding Aids Web Sites," *Journal of Archival Organization* 4, nos. 3–4 (February 2007): 99–118, https://doi.org/10.1300/j201v04n03_06.

By the late 2000s, user and usability studies of online finding aids had become increasingly critical of earlier research. Cory Nimer and J. Gordon Daines's "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective" reviewed both the literature and their personal and professional experiences as they prepared to redesign the finding aid interface at Brigham Young University. They listed inconsistency, archives-specific terminology, searching, and confusing hierarchical organization as significant issues for users.[17] Richard J. Cox's 2008 "Revisiting the Archival Finding Aid" explores the utility of archival finding aids and the ways researchers from design, museum, and other fields use them, intending to lead to a shift in the way archivists consider and construct archival description and how descriptive standards define our work.[18] In 2015, Luanne Freund and Elaine G. Toms studied how genealogists and historians completed tasks with finding aids and explored the challenges presented by the structure of the online finding aid, examining the processes and interactions used to complete tasks with the available online tools and information.[19]

Usability work in the archives field has turned to methods of searching and navigability of online finding aids. Junte Zhang's 2011 dissertation, rewritten into a published article entitled "System Evaluation of Archival Description and Access," explores how users search EAD finding aids and how additional searching and organizational systems for finding aids might improve user access.[20] Rachel D. Walton's master's thesis, entitled "Looking for Answers: A Usability Study of Online Finding Aid Navigation," examines Princeton University's online finding aids and notes the same challenges that Yakel had identified nearly fifteen years earlier: archives-specific language and multilevel description navigation in an online environment are significant challenges for users. Walton recommends giving users "a way to visually explore and browse through collection contents" and to "provide easy and quick access to individual items within a collection." Walton also suggests keyword searching across levels and simplifying the visual experience of a finding aid, recommendations that appear remarkably similar to those from one of the earliest usability studies of finding aids by Burt Altman and John R. Nemmers in 2001, assessing the Pepper OnLine Archival Retrieval and Information System (POLARIS) at Florida State University.[21]

Other similar studies and examinations demonstrate noticeable trends relating to terminology, standardization, navigability and users' understanding of location within the hierarchy of description (often referred to as "where they are"), and searchability. Walton notes in her literature review that since EAD's emergence in the mid-1990s, more than thirty articles about online finding aids generally and EAD more specifically have been published in *American Archivist* alone, not to mention others across major journals in the field. The number of usability studies focused on the

---

[17] Cory Nimer and J. Gordon Daines, "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective," *Journal of Archival Organization* 6, no. 4 (2008): 216–32, https://doi.org/10.1080/-15332740802533214.

[18] Richard J. Cox, "Revisiting the Archival Finding Aid," *Journal of Archival Organization* 5, no. 4 (2008): 5–32, https://doi.org/10.1080/15332 740802153245.

[19] Luanne Freund and Elaine G. Toms, "Interacting with Archival Finding Aids," *Journal of the Association for Information Science and Technology* 67, no. 4 (2015): 994–1008, https://doi.org/10.1002/asi.23436.

[20] Junte Zhang, "System Evaluation of Archival Description and Access," *ACM SIGIR Forum* 45, no. 2 (September 2012): 109–10, https://doi.org/10.1145/2093346.2093367.

[21] Rachel D. Walton, "Looking for Answers: A Usability Study of Online Finding Aid Navigation" (M.A. thesis, University of North Carolina, 2015), 43–44; Burt Altman and John R. Nemmers, "The Usability of On-Line Archival Resources: The Polaris Finding Aid Project," *American Archivist* 64, no. 1 (2001): 121–31.

interfaces that display finding aid content to users ranges well into the teens, as can be seen in Walton's and Zhang's work on searchability and navigability.[22]

However, none of the user or usability studies related to online finding aids have examined how users interact and make use of the *content* of finding aids. The existing studies are more typical applications of user research methods, rather than applying usability study methods to content and levels of description or to how researchers use the various DACS-compliant notes.

Usability research has not been done to examine which portions of a finding aid users typically seek out and whether they rely on top-level aggregate descriptive notes or more granular description for locating relevant collections and materials within collections. Archivists spend a great deal of time creating description for entire collections, as well as further information about different groupings and sometimes about individual items within collections. Yet little research has been done to understand how often the granular parts of a description are used or whether they are considered necessary by researchers. One goal of this study is to use design research methodology to understand not how the *format* of these documents (structure, coding, search functions, and navigability) impact researchers but instead how users read and use the *content* of finding aids.

This study should also be placed within the larger context of usability research in the library and information science field. Library website usability studies, often of online public access catalogs, have been carried out across a variety of library types and sizes over the past several decades. Usability within this context has been defined as "a system [that] has visible working functionality familiar to its users, maximum reliability, and useful content that is supported by its environment and aligned with context of use."[23] An early usability study of the University of Buffalo library system's website, using task analysis methods, discussed the importance of understanding if and how users can and do use online sites to access library materials, from books to journal articles and more.[24] Similar studies have been carried out over the following two decades, and many large academic libraries include a User Experience Center or have user researchers on staff. Although the work is prevalent across university libraries, such user experience efforts can be challenging and challenged: a 2013 examination of usability testing in libraries by Jennifer Emanuel looked at studies already carried out across the field and assessed methods for making the work more rigorous, along the lines of social science research.[25]

Usability research should also be understood in its widest application, across web design. Although the same methods are being applied in this study not to a web interface but to content, it remains important to understand the context from which this approach developed. Jakob Nielsen, a web usability consultant whose work has been foundational to the field, describes usability as "a quality

---

[22] Ibid., 2, 3.

[23] Yu-Hui Chen, Carol Anne Germain, and Abebe Rorissa, "Defining Usability: How Library Practice Differs from Published Research," *Portal: Libraries and the Academy* 11, no. 2 (2011): 599–628, https://doi.org/10.1353/-pla.2011.0020.

[24] Brenda Battleson, Austin Booth, and Jane Weintrop, "Usability Testing of an Academic Library Web Site: A Case Study," *Journal of Academic Librarianship* 27, no. 3 (2001): 188–98, https://doi.org/10.1016/s0099-1333(01)00180-x.

[25] Jennifer Emanuel, "Usability Testing in Libraries: Methods, Limitations, and Implications," *OCLC Systems and Services: International Digital Library Perspectives* 29, no. 4 (2013): 204–17, https://doi.org/10.1108/oclc-02-2013-0009.

attribute that assesses how easy user interfaces are to use." Nielsen goes on to note that "on the Web, usability is a necessary condition for survival. If a website is difficult to use, people leave. . . . If users get lost on a website, they leave. If a website's information is hard to read or doesn't answer users' key questions, they leave. . . . There are plenty of other websites available; leaving is the first line of defense when users encounter a difficulty."[26] The difference, of course, between website content and finding aid content is that users don't have another finding aid to go to. And yet, if frustrated, they too may give up and go elsewhere with their research, or miss important information about their area of study. Hence, the importance of ensuring that finding aid content, and not only structure, is usable.

Just as usability and design research have been advancing in the archives and library field, there has been an evolution in archival description as well. Over the past fifteen years, there has been a trend toward providing more aggregate archival description and reducing time spent on detailed information at more granular levels (files and items), attempting to reduce ever-growing backlogs and providing access (even if minimal) to more archival materials. This trend began with Green's and Meissner's widely cited article, which introduced the acronym "MPLP" into our shared archival vocabulary. Their argument for MPLP was based on the simultaneous need to provide archival access to users in an efficient and adequate manner while reducing sizable backlogs of inaccessible material.[27] Considerable research has been done into the success of MPLP in terms of its impact on processors, processing practice, and backlog and researcher access both in the United States and abroad.[28] Much has also been written on the impact on preservation of materials and applying required restrictions to access when carrying out processing that adheres to an MPLP framework.[29]

---

[26] Jakob Nielsen, "Usability 101," UseIt.com, 2012, http://www.useit.com/alertbox/20030825.html. Amy Deschenes, the head of UX and digital assessment at Harvard Library, pointed me to the work of Jakob Nielsen as a foundational leader in user experience studies.

[27] Greene and Meissner, "More Product, Less Process."

[28] For examples of available case studies on this topic, see Donna E. McCrea, "Getting More for Less: Testing a New Processing Model at the University of Montana," *American Archivist* 69, no. 2 (2006): 284–90; Rachel Anchor, "'More Product, Less Process': Method, Madness or Practice?" *Archives and Records* 34, no. 2 (2013): 156–74, https://doi.org/10.1080/23257962.2013.818937; Adrienne Harling, "MPLP as Intentional, Not Necessarily Minimal, Processing: The Rudolf W. Becking Collection at Humboldt State University," *American Archivist* 77 (2014): 489–98; Stephanie H. Crowe and Karen Spilman, "MPLP @ 5: More Access, Less Backlog?" *Journal of Archival Organization* 8, no. 2 (2010): 110–33, https://doi,org/10.1080/15332748.2010.518079; Matt Gorzalski and Marcella Wiget, "'More Access, Less Backlog': How the Kansas Historical Society Got Its Groove Back," *Archival Issues* 33, no. 1 (2011): 7–24; and Janet Huack, Rose Sliger Krause, and Kyna Herzinger, "MPLP Ten Years Later: The Adventure of Being among the First," *Provenance: Journal of the Society of Georgia Archivists* 35, no. 2 (2019), https://digital-commons.kennesaw.edu/provenance/vol35/iss2/5.

[29] For examples of available case studies on this topic, see Gerald Chaudron, "To MPLP or Not to MPLP: That Is the Question with Photographs," *Journal for the Society of North Carolina Archivists* 10, no. 1 (2012): 2–19; Anne L. Foster, "Minimum Standards Processing and Photograph Collections," *Archival Issues* 30, no. 2 (2006): 107–18; and Laura McCann, "Preservation as Obstacle or Opportunity? Rethinking the Preservation-Access Model in the Age of MPLP," *Journal of Archival Organization* 11 (2013): 1–2, 23–48, https://doi.org/10.1080/15332748.2013.871972; among others. For an example of an available case study on this topic, see Steven M. Gentry, "A Gentle Approach to 'Gentle Ren': Processing the Papers of Former College President Renwick Jackson," *Provenance: Journal of the Society of Georgia Archivists* 32, no. 1 (2014), https://digitalcommons.kennesaw.edu/provenance/vol32/iss1/7.

Considerably less research has been done into how MPLP-framed processing and description have impacted researchers' use of description, and if users find aggregate description adequate. A second goal of this study is to assess how users read and use aggregate description, how they react when various levels of a collection are described minimally, and whether this limits users' ability to understand the contents and contexts of collections.

Methodology

User research and task analysis.

This study makes use of task analysis. In contrast to other social scientific research methods, task analysis is not intended to generate causal inferences or to establish correlations with statistical significance. Instead, the usability research presented in this article identifies trends and tendencies for further investigation and empirical study. Taking inspiration from how web designers have deployed task analysis to understand the ways users navigate websites, this study uses task analysis to investigate how researchers read and use the different elements of a finding aid.

Despites its rich potential, task analysis has so far not been used to study how researchers read and use the contents of a finding aid. Task analysis is particularly well suited for this purpose because trends and tendencies can be identified with a workable number of participants, and the human subject interview portions of the research can be completed in a reasonable amount of time. Since this project was designed to be completed within a single calendar year, task analysis proved to be a practical approach.

Finding aid selection.

To archivists, the finding aid represents carefully constructed data that might have been encoded in EAD or input into a database system such as ArchivesSpace, with content placed in specific fields and notes as required by content and data structure standards such as DACS and EAD. What users typically see on the front end, however, is a webpage or perhaps an exported PDF.[30] In both formats, the user interacts with various pieces of information spread across the finding aid—from administrative information and collection-level scope and content data at the top to more detailed descriptions of smaller groupings (series, subseries, files) or items below. The user must be able to navigate through this material to identify whatever information speaks to their needs or research questions.

The three finding aids selected for this study were gathered from a survey of repositories across the United States, with advice drawn from multiple colleagues, including Dorothy Berry, Kate Donovan, Adrien Hilton, Jessica Sedgwick, and Melanie Wisner. The finding aids were purposefully selected from beyond the institution where I work, Harvard Library, in order to expand the scope of the study from being a comparative study of internal practices at one or many local repositories to a broader view of various models of description used at Harvard libraries and beyond.

---

[30] Because this study specifically was *not* intended to review the interfaces of the institutions included, PDFs were downloaded and used for the testing.

I sought good examples of description on three different models—one traditional or granular, and two that represent varieties of "minimal" description in the post-MPLP world. All three models can be found in use at Harvard's Houghton Library, where I am a processing archivist, but also at innumerable other institutions across the country:

Model 1: A finding aid with aggregate contextual information at the collection level but fairly minimal file- or item-level description. A finding aid on this model typically uses information from a survey of the collection and from the creator/donor/dealer in order to create the description. At Houghton, finding aids of this type are commonly used for accession-level processing. Model 1 perhaps best reflects the transition toward aggregate description that has taken place in the archival field in the fifteen years following the publication of Greene and Meissner's article.

Model 2: A traditionally granular, in-depth description of a collection, including thorough aggregate description at the collection and series levels, as well as detailed lower-level description at the file or item levels. The description is typically based on an in-depth analysis of the collection's contents and documentation as provided by the creator/donor/dealer, as well as external research into the creator and their work and life.

Model 3: A finding aid with more minimal aggregate contextual information containing some collection-level description but with more detailed description of the collection at lower levels. At all levels, this description may be based on a collection survey and/or information provided by the creator/donor/dealer. Model 3 is used at Houghton Library for accession-level processing when a dealer or donor provides a contents list or other more detailed information that can be easily transformed into archival description.

Each model of a finding aid has its strengths and weaknesses. However, the goal of the present study was to investigate whether all three can serve as useful surrogates for archival collections and be used by researchers to understand their contents and place them into context. The secondary goal was to see whether there is a difference in usability among the models.

In selecting finding aids for testing, I sought descriptions of collections that encompassed multiple types of creators, including both personal papers and organizational records. The initial piloting in November and December 2019 involved the finding aids for four collections—two collections of personal papers and two of organizational records—on a mix of the models described above. However, after discussing the first draft of the study script with Amy Deschenes in the User Research Center at Harvard Library, she suggested that the study include three finding aids to keep testing under an hour. Thus, the final selection included two finding aids for personal papers collections and one for organizational records. Two of the finding aids were from New York City–based repositories, while the third was from a Midwestern repository.[31]

The three finding aids used are:

Model 1: Peter Weiss Papers, Tamiment Library, New York University (Finding Aid 1)

---

[31] For full disclosure, I have lived in both New York City and the Minneapolis area, which may have impacted my selections subconsciously, but the basis for those finding aids' selections was the fact they were strong examples of the three models of description sought.

Model 2: Tom Wolfe Papers, New York Public Library (Finding Aid 2)

Model 3: Ampersand Club Records, Minnesota Historical Society (Finding Aid 3)

The finding aids were exported from their holding repositories' websites in October 2019 as PDFs and were used throughout testing in this format. They were not tested live on those sites in order to remove the question of usability of each display system. This information was explained to each participant at the beginning of the testing process. In deference to Internal Review Board requirements, the PDFs were not updated at any point during the study, although it is possible that changes were made at the repositories during the eight months from the time of download to the completion of the testing.

User pool and online testing atmosphere.

The goal for the user pool was to have a variety of users, with differing levels of experience in archival research and different fields of interest, tested during the single year allocated for this project. The original intention was to have fifteen to twenty users complete the test, as this is typically an acceptable number for usability studies.[32] The first ten users were tested from December 2019 to April 2020, recruited in-person prior to the COVID-19 pandemic's impact on daily work in the United States. In April 2020, I posted a call for participants to H-Net, a listserv for history and the social sciences, and received a sizable influx of interested participants, thus the user pool was expanded to approximately fifty users. A total of fifty-five user studies were scheduled; three were experienced archivists and were not included in final assessments. Unfortunately, one user's test recording file was corrupted and unrecoverable. Thus, a pool of fifty-one users makes up the data set for analysis.

The form that potential applicants completed asked only for contact information and whether the user had archival experience, as I sought a group of users who understood the purpose of finding aids while allowing for any level of research experience.
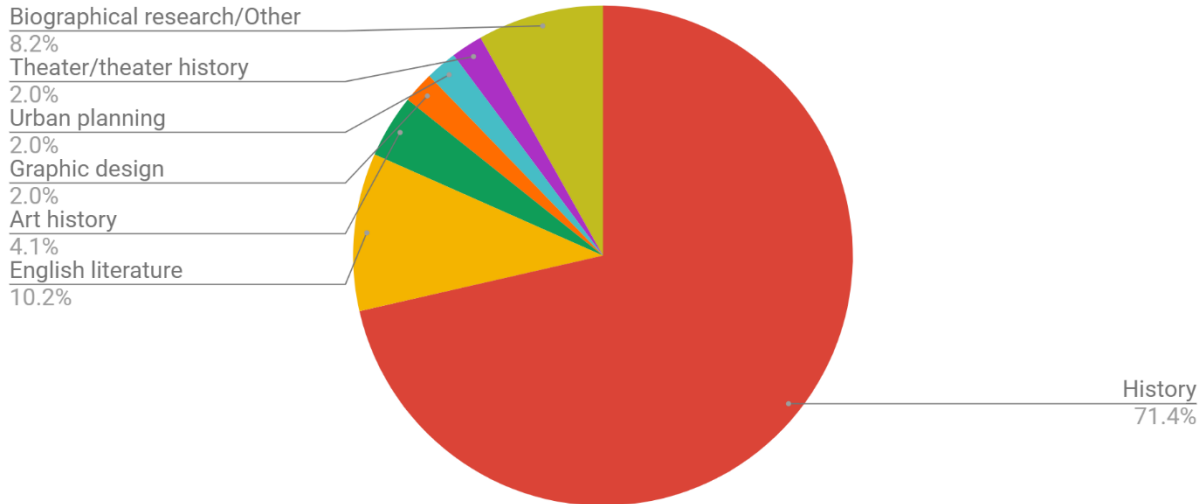
---

[32] Janet M Six and Ritch Macefield, "How to Determine the Right Number of Participants for Usability Studies," UXmatters, January 4, 2016, https://www.uxmatters.com/mt/archives/2016/01/how-to-determine-the-right-number-of-participants-for-usability-studies.php.

Figure 1: User Pool Experience Levels
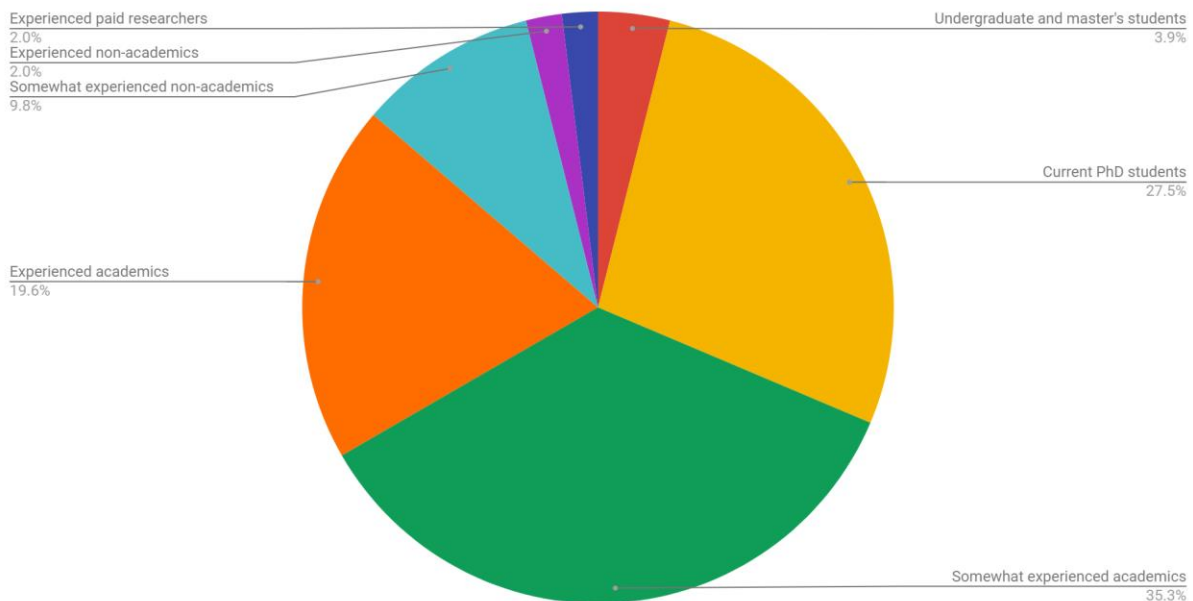
The user pool was heavily weighted toward those in the academic profession, whether as professors at colleges or universities or graduate students. As can be seen in figure 1, the largest share of participants was "somewhat experienced academics," meaning those who teach at colleges and universities and who have been doing archival research for one to ten years. This subgroup comprised nineteen users, or 35.3 percent of the total user pool. Many of these users' archival research experience came from completing their dissertations and/or reworking dissertations into books. The second and third largest subgroups of researchers were "experienced academics," meaning professors or instructors who have been using archives for more than ten years, and current doctoral students working to complete their dissertations. Experienced academics made up roughly 20 percent of the total user pool with a total of ten, while there were fourteen current doctoral student participants, or 27.5 percent. All in all, approximately 82 percent of the user population were graduate students or professional academics. A minority of users were from off the beaten academic research track, with a total of nine nonacademic users (five "somewhat experienced," one "experienced"), a single paid independent researcher, as well as one undergraduate and one master's student.

Figure 2: User Pool Academic Fields

Biographical research/Other
8.2%
Theater/theater history
2.0%
Urban planning
2.0%
Graphic design
2.0%
Art history
4.1%
English literature
10.2%

History
71.4%

The user pool was heavily concentrated in certain academic fields, particularly history and English literature (fig. 2). A few other fields were included, such as art history, graphic design, urban planning, theater history, and biographical research. Three users had experience working as employees in archival repositories, two were current employees also conducting their own historical research, while one was retired from archival work and part of a research team. The archivists without separate research projects who volunteered to participate were not included in the final user group.

The fact that the user group was concentrated so heavily in history and English literature likely had to do with the study notice being posted to H-Net, a location suggested by multiple reference librarians after in-person library use ceased in March 2020 due to the COVID-19 pandemic. Prior to that date, table tents and cards referencing the study had been placed in the reading rooms of Houghton Library, Schlesinger Library, Harvard University Archives, Loeb Music Library, the Center for the History of Medicine (all at Harvard University), as well as Northeastern University Archives and Special Collections and the University of Massachusetts Boston University Archives and Special Collections. If respondents had been pooled from these in-person outlets instead, the cohort might have been more student-centered and likely not as weighted toward just two academic fields but due to the COVID-19 pandemic, this was unavoidable. I was surprised by the sheer number of respondents to a single H-Net post, having sought only five to ten additional participants. Capping the study at fifty participants was determined to be a reasonable size for an independent project set to terminate after one year.

The user pool was also mainly based in the United States, though there were two academics working in Ireland as well as users from Portugal, the United Kingdom, Canada, and Australia. Most users had archival experience primarily in the United States, but ten had done research in the United Kingdom, Ireland, Sweden, Jamaica, the British Virgin Islands, Portugal, Spain, India, Australia, Japan, and the Philippines. The users were not asked about racial or gender identities, though this could be an interesting avenue for assessment in future research.

The testing environment undeniably altered the way users experience each finding aid and how they were able to complete the tasks, as well as how difficult they perceived each task to be. First,

the testing was scheduled for about one hour. Some users took little more than half of the allotted time, whereas others took slightly more than the planned hour. It is possible that some users felt they could complete these tasks quickly; it is also conceivable that completing tasks in an online, recorded environment made users feel nervous and rushed. To mitigate the latter concern, I encouraged users to take all the time they needed, but of course some users may have continued to feel pressed for time or nervous.

Second, the online, recorded environment was far from ideal for performing these tests. The testing was done through the sharing of my screen on Zoom, on which three tabs for the finding aids appeared. The users were asked to scroll through the first finding aid selected and share their first impressions, which also gave them the opportunity to become familiar with the creator, collection content and context, and format of the finding aid. In the next stage, users were asked to complete each of the assigned tasks, to assess the difficulty of each task, and to move on to the next finding aid. The finding aids' order in the testing process was randomized for each user.

Two problems arose while completing the test. First, in the Zoom shared-screen environment, scrolling can sometimes be delayed or jumpy, depending on Wi-Fi/internet access, making it difficult at times for users to thoroughly read notes or gain a sense of the overall document. Again, I made efforts to accommodate users who experienced this problem. In most cases, the problem was no more than a mild annoyance, but it turned out to be more frustrating for a small number of users.

Second, in the Zoom shared-screen environment, users were not able to search the PDF as they normally would using the Ctrl-F function. This limitation was explained to users from the outset. To address this limitation, I asked researchers to explain whether searching would be part of their approach to a given task in a more ideal environment. Most users were happy to explain when they would have deployed the search function. These users completed the assigned tasks by scrolling up and down the document. A few users found the inability to search more frustrating. These users were less likely to scroll back and forth throughout the testing process.

Although these online interface issues were somewhat concerning, in light of the COVID-19 pandemic they were also unavoidable. Luckily, most of the participants were experienced Zoom users and likely much more comfortable in the Zoom environment than they would have been before the pandemic started. I worked to mitigate the interface issues and asked for patience on the part of the participants. The vast majority of users were happy to work within the confines of the Zoom shared-screen environment in order to complete the study tasks.

Research scenarios.

Users were asked to imagine the following research scenarios, specifically created for the finding aids and intended to frame the study tasks.

Finding Aid 1—Research Scenario: Peter Weiss is a New York attorney with expertise in international law. He has a history working with human rights organizations and involvement in the field of human rights law. Imagine that you are interested in looking into the work these organizations did and the legal work Weiss did for them.

Finding Aid 2—Research Scenario: Imagine that you are researching the life and career of the American writer Tom Wolfe, preparing to write a biographical sketch and ascertaining important influences on his work. You head to his archival papers to see whether you can find information that you have struggled to find elsewhere.

Finding Aid 3—Research Scenario: Imagine that you are a member of the Ampersand Club in Minneapolis, interested in writing a history of the club for the website, and that you have come to the archives for additional resources.

Tasks and data levels.

Each user completed three tasks with each of the three finding aids. The tasks typically assessed what portions and levels of a finding aid the researcher turned to in order to understand the content of the collection.

An example of a task was discovering whether a particular area of the creator's research is included in the collection and where users might find it, which would require a user to read both a collection-level scope and content note as well as to look within lower levels of description to series-level notes and file- or item-level listings. Another task involved discovering when women were invited to become members of an organization and where evidence of this might be, which required information from the historical note to guide the user to files by date. Other tasks related to understanding whether the materials relating to a topic were restricted, in order to see how users interacted with multiple levels of description and understood access restrictions notes.[33]

The degree of completion of each task and its level of difficulty were assessed in the following way: First, each user supplied their own assessment of the difficulty of completing each task with each finding aid. These user-provided assessments were then used to furnish responses to the two study research questions: how do users experience and use various levels of description, and, in cases where full description is lacking at one or more levels, does the absence of full description impact how well users are able to use finding aids?

Degrees of task completion:

- 1 suggests a task that was fully and easily accomplished;
- 2 suggests that the user was mostly successful in completing the task, with some complications or questions, and might have missed a portion of the relevant information in the finding aid;
- 3 suggests the user struggled to complete the task but had ideas for how to handle it and made strides toward completing it;
- 4 suggests the user was not able to complete the task but had some thoughts for how to approach it;
- 5 suggests the user was unable to finish the task, was not sure how to begin, and would resort to asking an archivist or calling all boxes. In these cases, the use of the finding aid has not been successful.

---

[33] For more information about the specific tasks, see the appendix.

Levels of difficulty:

- 1 is a task that was very easy to complete;
- 2 is a task that was somewhat easy to complete;
- 3 is a task that was neither easy nor difficult;
- 4 is a task that was somewhat difficult to complete;
- 5 is a task that was very difficult to complete.

Key Findings

This study attempted to answer two related research questions:

1. What are the elements researchers use in order to know if they would want to come to the reading room and go through several boxes? Are the notes archivists use to be compliant with archival field standards, such as scope and content and arrangement notes, in fact notes that people expect to find and read to discover what is included in a collection? Are item-level or other granular levels of description expected?
2. In cases where there is a lack of granular description, are users able to ascertain what materials are in a collection? In these cases, what notes are being used to glean this information?

The second question was the easier one to answer. Users were able to complete all of the tasks at least partially with all three finding aids—demonstrating that even the finding aids offering less than full description were usable. The finding aid with the highest rate of average usability was Finding Aid 2, the finding aid with detailed description at all levels. Finding Aid 1, the finding aid with aggregate description and a thorough scope and content and arrangement note, was nearly as usable on average; users were able to decide whether the collection included the material related to each specific research question. Finding Aid 3, which had detailed lower-level components accompanied by more minimal collection-level description, was still moderately usable but the least usable of the three.

The study also examined how users perceived the difficulty of working with each finding aid. The result, illustrated in figure 3, was similar: all three finding aids were moderately easy to use as a basis for completing tasks and responding to specific research questions. Once again, users believed the thorough-at-all-levels finding aid (Finding Aid 2) was easiest to work with. The user pool also identified the aggregate description-heavy finding aid as being the next easiest to use (Finding Aid 1), and the third model as being just slightly more challenging (Finding Aid 3).

Figure 3: User Completion Average and Perceived Difficulty Average

Averages Compared Across Finding Aids
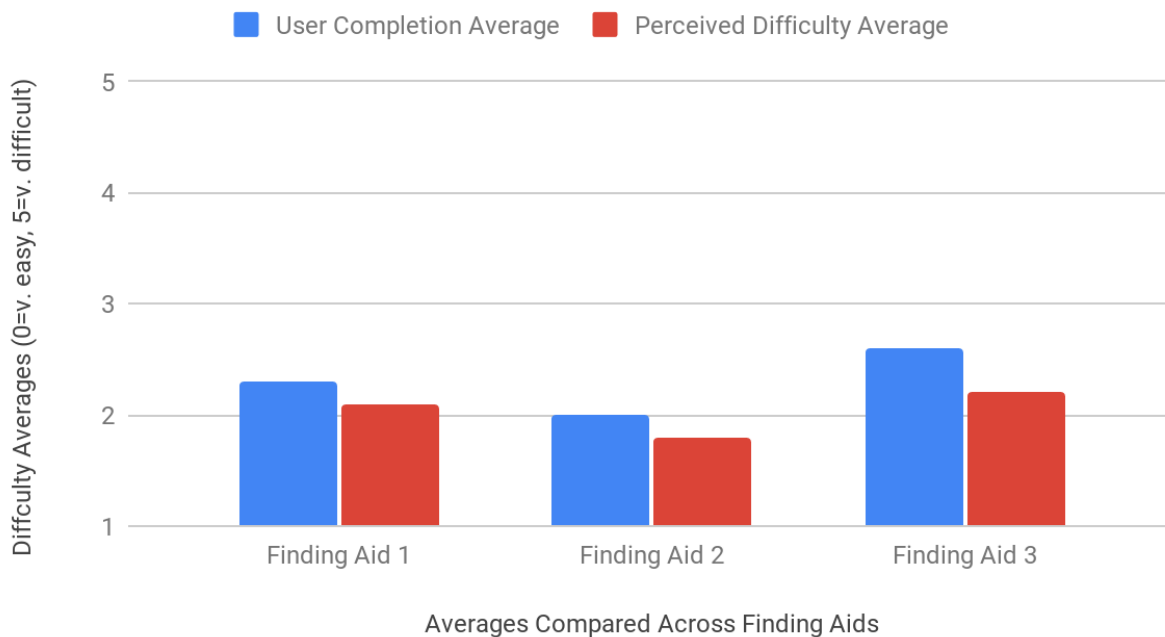
Inferring an answer to the study's first research question—what parts of a finding aid do researchers use and do they read notes—was somewhat more complicated. Multiple tasks were formulated to address this question. Throughout the study, it became clear that the notes used most frequently were the scope and content and the biographical/historical notes at the collection level, with well more than half of all users mentioning these notes when asked about their impressions of each finding aid. However, when users were asked to complete tasks requiring information included in these notes and nowhere else, only half of the users turned to them to identify relevant information.

Moreover, what was made available to users and how often they turned to other levels of description seemed to be interdependent. In cases where lower-level description was minimal, users were likely to turn to the notes. By contrast, when the lower-level description was thorough and collection-level information was less so, information included only in the notes was more likely to be missed entirely. Finally, in the finding aid with rich description at every level, users had a tendency to look to series-level notes to identify what is in each series.

With Finding Aid 1, it became clear that slightly more than half of users tended to go to the notes to respond to questions; the remainder attempted to use minimal lower-level description. Full completion relied on users identifying information in both the scope and content and biographical notes. Only 25 percent of users were able to do this, but slightly more than half were able to use context from the notes coupled with lower-level description to complete the task. In another task with the same finding aid, 57 percent were able to identify important information in the scope and content note that defines how to use lower-level description, and 51 percent used both the scope and content note and lower-level description to fully complete the final task.

With Finding Aid 2, 63 percent of the users read the scope and content and biographical notes to respond to one task, though only a few identified particular details. In a case with thorough description, users had the tendency to turn to these notes but were not always terribly precise in looking through the entirety of either collection- or series-level notes. Still, more users relied on the notes in this longer finding aid, particularly since they could not turn to the Ctrl-F function to search for terms in the testing environment. For two separate tasks, 90 percent or more identified information in series-level scope and content notes, demonstrating that users do turn to these shorter aggregate notes when they are made available, rather than even more detailed lower-level description.

For Finding Aid 3, in testing whether users would still turn to the notes when using this model of description, about one-quarter of the users relied on the historical note as well as the lower-level description to complete tasks (25% to complete one task; 27% to complete another). Users were more likely to skip these notes when rich lower-level content was available, relying on the lower levels to fulfill tasks.[34]

As noted above, users were asked to complete tasks related to the conditions governing access notes for all three finding aids. About 30–40 percent of users were able to find these notes and understand their meaning across the three finding aids. Thus, the majority of users (60–70%) were not able to locate or interpret access notes and how they might impact a user's experience of a collection, a troubling result related to finding aid usability.

Free Response Assessment

When the test participants initially encountered each finding aid, they were asked to share their first impressions of the overall document and to point out the portions they typically read and those they usually skipped. These first impressions were mixed, with some users liking and others disliking the same aspects of a finding aid. At the same time, they suggested ways of adjusting archival description to reduce confusion.

The free responses pointed to elements of the finding aids that users typically read and use, particularly the collection-level notes such as scope and content and biographical/historical notes, as well as basic descriptive content (creator, title, dates, extent). The responses showed that some users have a tendency to examine much of the administrative and explanatory content beyond the notes, such as provenance, processing, and access information, particularly when this information is different from what they are accustomed to seeing.

The responses also provided further support for one of the main conclusions drawn above: the degree to which researchers use notes depends on the available description. When aggregate description is rich and lower-level content minimal, as in the first finding aid, users tend to read and consider aggregate notes. When content is thorough at all levels, users are split in what they choose to mention in their first impression, with some referencing aggregate-level information and others mentioning information from lower levels. Finally, when lower-level content is richer and collection-level description more minimal, users consistently focus their attention on the lower levels.

---

[34] A lengthy internal report is also available with further details regarding tasks and responses for all three finding aids. For in-depth reports on these findings, please contact the author.

Figure 4: Users Mentioning Finding Aid 1 Sections—First Impressions
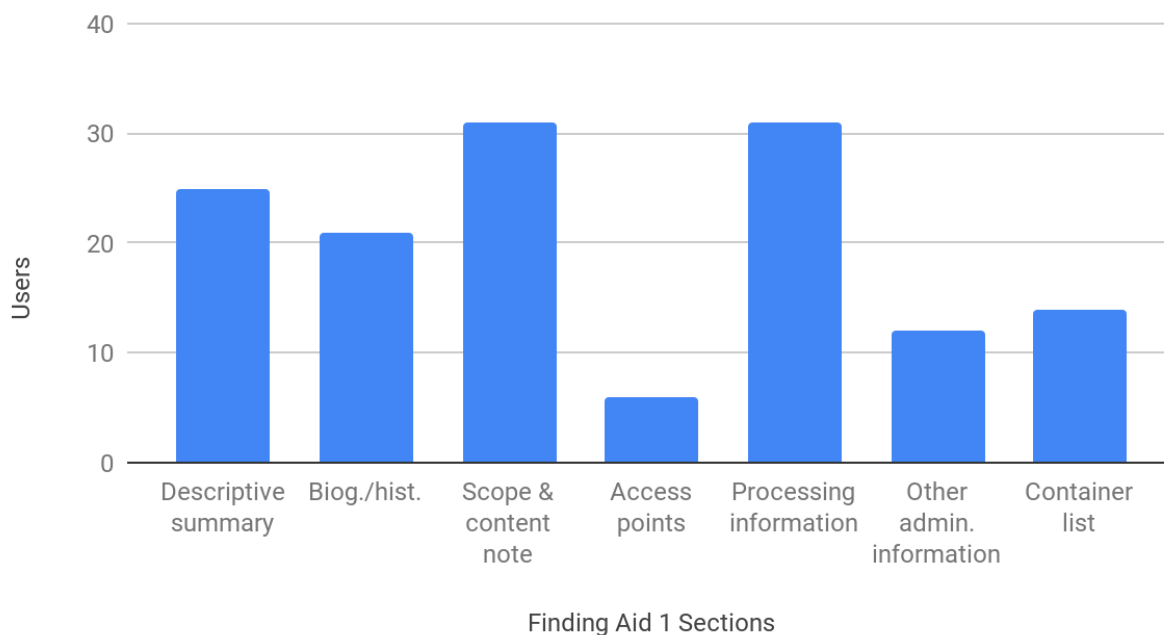
As can be seen in figure 4, the most frequently discussed portions of Finding Aid 1 were the descriptive summary section (which includes creator-donor information, title, inclusive dates, the abstract, quantity, location, and languages), the scope and content note, the processing information note, and the biographical note. More than thirty users (61%) chose to mention the scope and content note, the most in-depth portion of this finding aid. Forty-one percent mentioned the biographical note. Users mentioned that they would often skip or skim these to start but return to them for content and context. The other collection-level descriptive note that was discussed by 61 percent of the users was the processing information note. This note explained that the materials were rehoused but that the creator's original arrangement and titles were maintained. Some were surprised or confused by this information, while others were interested in both the impact and meaning of this organizational structure. No matter the reaction, the information processing note was clearly read and interpreted by many researchers as they did their first reading.

Figure 5: Users Mentioning Finding Aid 2 Sections—First Impressions
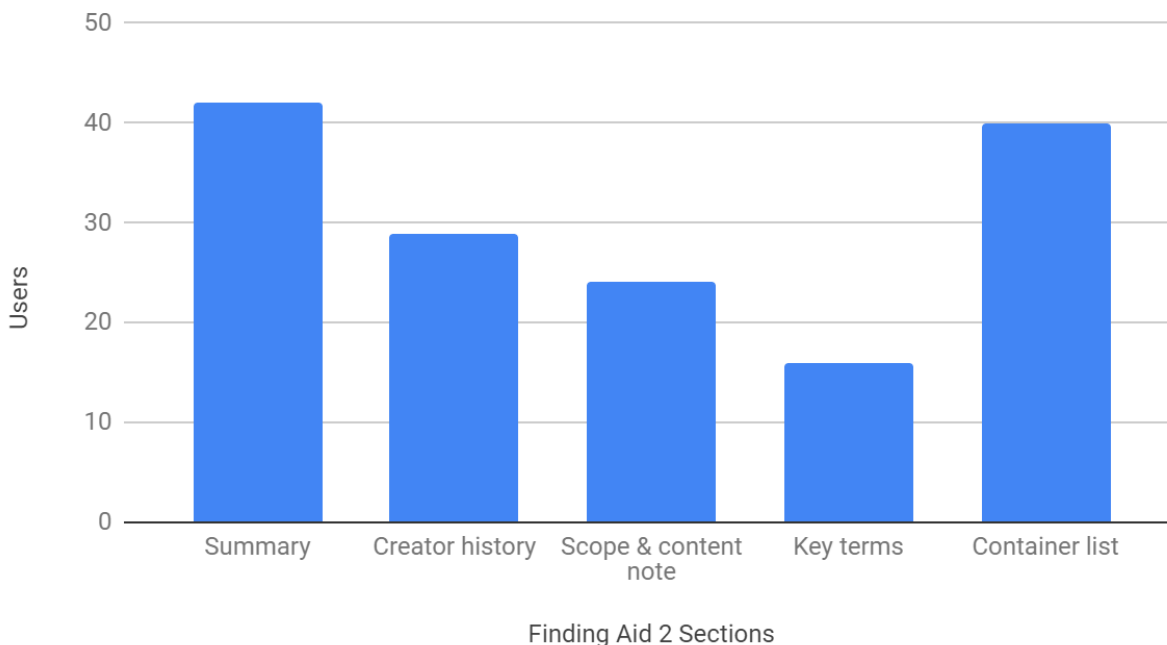
For Finding Aid 2, the summary—which includes creator name, title, dates, size, source, abstract, access and conditions governing access, preferred citation, and processing note—received the most attention from users, who sought brief descriptive information to understand the scope and breadth of the collection (fig. 5). Eighty-two percent discussed their use of the summary section. Similarly, users also mentioned the biographical note (57%) and the collection-level scope and content note (47%). Some said both were too dense to read carefully, but many also appreciated the depth, detail, and lists of names included in each note. Seventy-eight percent of the users brought up the container list, which in this finding aid included both series-level scope notes as well as file-level descriptions. Thus, although slightly more than half of the users were interested in the collection-level notes, when lower-level descriptions were supplied, users were heavily drawn to those. Interestingly, the most frequently discussed portions of the finding aid were the (briefest) description section—the summary section—and the (long) container list, including eighteen pages of detailed description.

In the case of Finding Aid 3, though some users looked at higher-level components, the sizable lower-level description weighed most heavily on their minds (fig. 6). This was the richest level of description and points to users' tendency to focus on the information as it is provided. About twenty users (39%) commented on the overview (including the creator, title, dates, abstract, quantity, and location information) and the biographical note, which were the most detailed portions of the collection-level description. About one-third of users discussed the scope and content note, often commenting on its brevity.

At the same time, 98 percent of users discussed the detailed description found mostly in the lower-level content. In this finding aid, this detailed portion of the description received considerably more attention from users. While the lack of organizational structure in this section bothered some

users, the perceived benefit of the detail provided outweighed these concerns. Many users commented that the lack of structure would not be a problem if they were able to do searches via the Ctrl-F function, as they normally would.[35]

Figure 6: Users Mentioning Finding Aid 3 Sections—First Impressions


Finding Aid 3 Sections
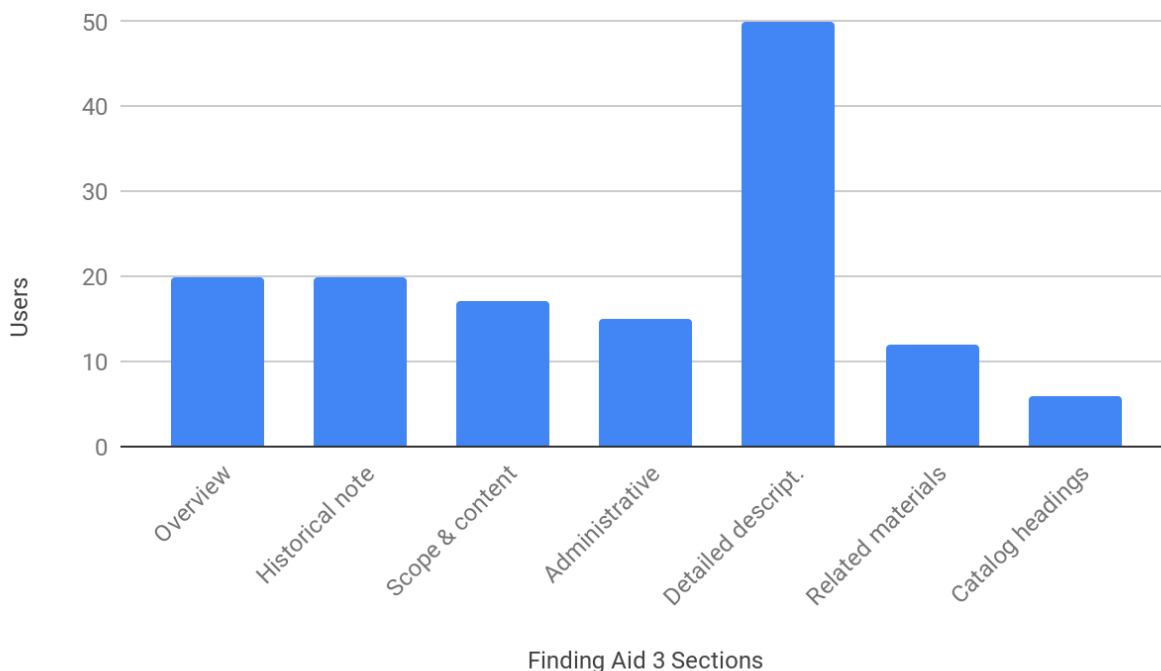
Following the completion of the task-based testing, users were asked what struck them as particularly important content in a finding aid. Again, there was not a great deal of consistency across the responses, with a mix of users feeling that detailed lower-level description or collection-level notes were most important (fig. 7).

The largest group of users stated that detail at the lower levels was the most important aspect of a finding aid, with twenty-eight users saying so (55%). Types of detailed information users mentioned included dates, names, titles of documents, and the chronological relationship between materials. Twenty-two users (43%) mentioned the importance of strong collection-level notes, with eighteen (35%) mentioning specifically the importance of the scope and content note. An additional 18 percent felt a balanced finding aid with organization and some detail is critical. Smaller groups of users identified other issues and elements as particularly necessary for their use, from subjects/keywords to processing information, related materials, access restrictions, citation guidance, and more.

---

[35] A lengthier report with details about users' comments and overall impressions of all three finding aids is available. For in-depth reports on these findings, please contact the author.

Figure 7: What Users Find Important
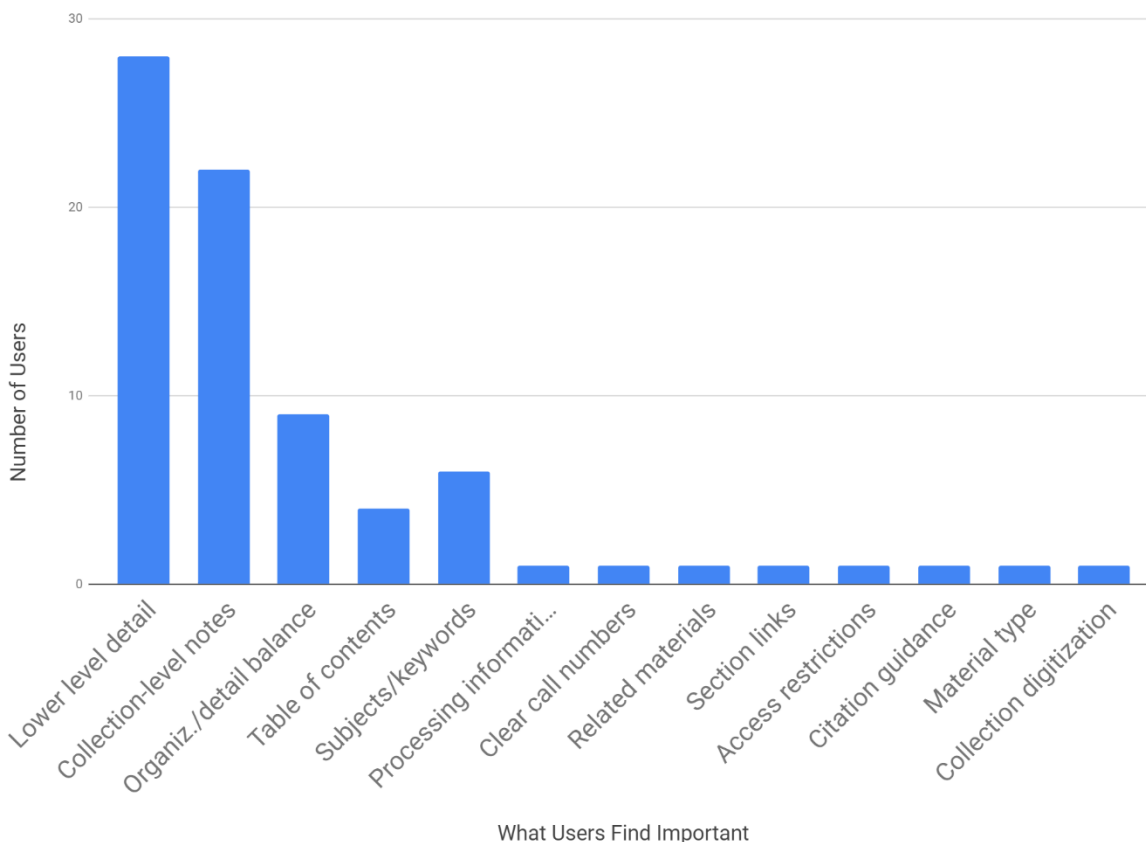
The final question of each testing experience was what additional information users would like to see included in finding aids. As with the first impressions, users were invited to respond to this final question in a free and unstructured manner. Accordingly, their answers could mention information that the user saw in the test finding aids but had not encountered elsewhere, information that the user would have liked to see in the finding aid given their prior research experience, or information that the user had never before seen in a finding aid but felt would be valuable to include. From the responses to this question, it became apparent that many users would love to have more and easier access to information about related materials. A large number of participants would like to see additional descriptive information about a collection, more thorough and expansive subject analysis and keyword usage, as well as details about processing and provenance.

It was interesting that only one user mentioned wanting to see collections digitized. Digitization is a common request from users and, in review of requests at Houghton Library, has been found to be the most common user request. The lack of user focus on digitization in the present study may simply have been the result of asking users to focus specifically on content and use of description in the course of testing. In other contexts, it is likely that a significantly larger proportion of users would be interested in digitized materials being made available in finding aids.

Figure 8: What Users Want to See
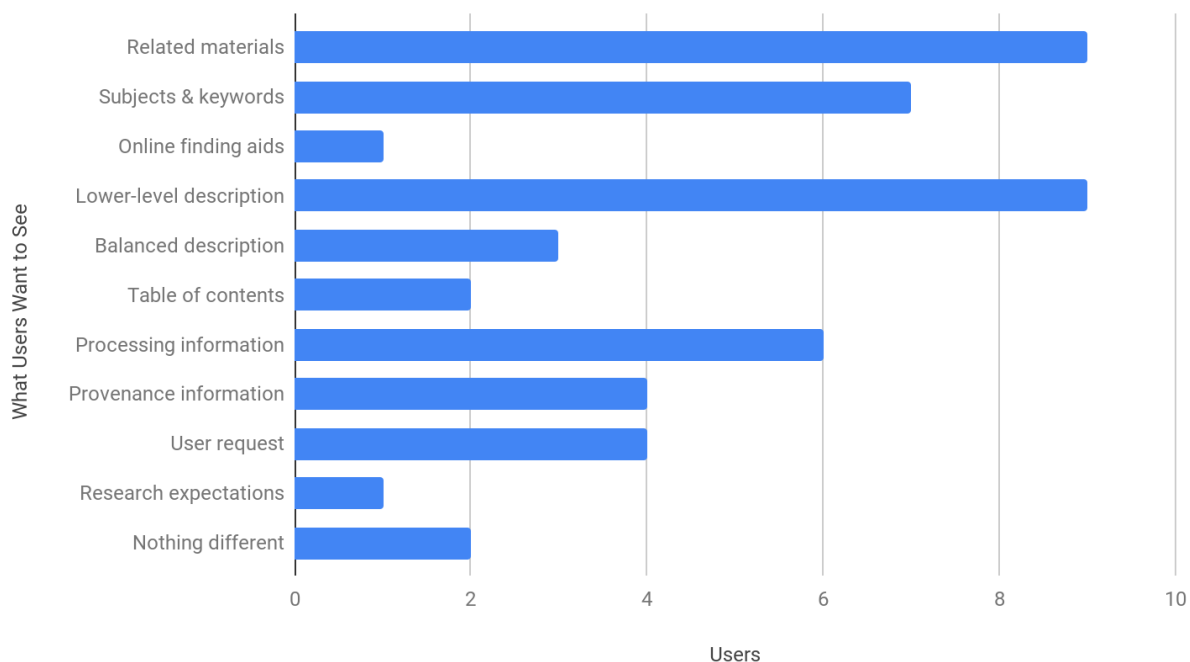
Although what users would like to be able to see in a finding aid is spread across many elements, there is clear interest in reading description at multiple levels, expanded subject analysis and methods for using linked keywords to provide access to related collections and repositories, and more general information about related materials. Some of these user interests relate to work that is currently being considered, or already being done, in newer conceptual models such as Records in Context (RiC), Social Networks in Archival Contexts (SNAC), and Archives Portal Europe. User interests in these respects underscore the value of using consistent terminology, so that already existing description may be easily transformable for use in future models of archival description or in a future data environment.

Conclusions

Based on the data gathered from the user pool, this study suggests that all three models of archival description tested are similarly usable for researchers. Irrespective of the descriptive richness of the finding aid, tasks assigned to users were completed to an average completion score of 2 to 2.5 on a scale of 1 to 5, with a score of 1 denoting full completion of the task and a score of 5 denoting that the user was unable to start. This result suggests that all three models are usable, albeit not perfectly so. In interpreting this result, it must be kept in mind that users varied in their level of thoroughness, and that some of this variation likely explains why some tasks were completed to a score of 2 rather than a perfect score of 1. The average completion score of 2 may also reflect the challenge of navigating multiple fields and levels of description, a suggestion that is consistent with other findings in the user and usability studies literature.

Turning now to users' assessments of the difficulty of working with each finding aid, users assigned difficulty scores ranging from 1.8 to 2.2 on a scale of 1 to 5. Thus, users on average found

the finding aids slightly easier to use than their task completion rates would suggest. Importantly, the variance of perceived difficulty was minimal and suggests that users found all three finding aids easy to use, or even very easy to use.

User experience level was examined to check whether experience level significantly impacted the user's success rate in completing each task. Interestingly, there were no discernible trends for any experience level or group of users. Relative newcomers to archival research were as successful as those with multiple decades of experience—or in some cases, faced the same challenges. In one respect, this is an encouraging finding, as it shows that years of experience are not required for success with archival description.

It is hard to say how usable a finding aid is, as researchers will approach each collection and its description with different interests, questions, and ideas for how to use it. However, based on the tasks accomplished in this usability study and users' perception of the difficulty in accomplishing them, it appears that all three models are reasonably easy to use. All three models are functioning as intended, providing access to and context for archival collections. The second model of description is only slightly easier to use, judging from the task completion score and the users' perceptions of difficulty. Each of the two finding aids containing some level of minimal description appear to be nearly as successful in providing information about what is in a collection, and whether it is relevant to a set of research questions, as the model that includes thorough description at all levels. This result is a truly exciting finding—suggesting not merely that finding aids created within an MPLP framework are *usable* but that such nontraditional finding aids can be *nearly as usable* as finding aids created with thorough but more resource-intensive description at every level.

In connection with this result, it is worth noting that the usability of the model finding aids in this study depended on researchers' willingness to be flexible about the portions of the finding aid they used. To be successful with the various tasks assigned, users had to be willing to examine any section of the finding aid for potentially useful information. In the cases of Finding Aids 1 and 3 (in which thorough description was not available at all levels), this meant focusing attention on several different portions of the finding aid to complete a single task. When using Finding Aid 1, users needed to be more attentive to collection-level description in order to respond to an inquiry, whereas Finding Aid 3 required users to focus on the detailed description as well as the briefer collection-level notes to complete the tasks. Users who were able to adjust their expectations concerning what they would find in each finding aid were more successful at completing tasks. Conversely, users who expected full description at all levels struggled to work with the two less traditional finding aids.

These observations are not meant to suggest that the onus for understanding the content of archival collections should be on the researcher. Instead, these observations should be kept in mind when writing description and planning where and how to place detailed information about a collection. As archivists and librarians, we may want to make clearer to users that finding aids often look quite different and are constructed on a variety of models. Approaching the diversity of finding aids with this knowledge in mind should make their use somewhat easier. Still, it is exciting to note that when users are flexible in their expectations and approaches to finding aids, even finding aids with minimal description can be nearly as usable for—and judged nearly as easy to use by—researchers.

Future Work

Additional work is required to understand how various limitations of this study might have impacted it: how users beyond the humanities and social sciences use finding aid content, and how content and interface work together to facilitate—or impede—user experience. It might be interesting to gather data often used in task analysis, such as tracking navigation and clicks, to see precisely how users navigate a finding aid and move through levels of description. This seemed unnecessary in the PDF format, but if carried out on live websites, the data could show interesting trends in how users navigate through content. Timing the tasks could produce interesting findings about levels of difficulty. It would also be illustrative to see if users' age, gender, or racial identities impact their experiences interacting with finding aid content and the various models of description.

In addition, we need methods for understanding when users need additional description in order to use a finding aid. Although this study found that all three models are usable, this result should not be taken to imply that all finding aids are usable for all researchers. We need to develop methods for receiving feedback when there is not enough description, when there are mistakes or clarifications are needed, and to note historical or recent uses of harmful, euphemistic, or offensive language. All of these issues directly impact the usability of archival description. It is for this reason that a second year-long study is being carried out, this time looking at existing reference data to assess what researchers are typically not able to find, what questions they ask about description, and whether and how they request additional information about a collection. This second study is a collaborative effort with Zoe Hill, a member of the Public Services staff at Houghton Library, and focuses specifically on reference data from this repository. This additional work will also include the development and implementation of a feedback form, designed to solicit user requests for editing or additional description, in order to directly respond to users' needs regarding the content of finding aids.

This usability study was developed in order to understand more about how researchers interact with finding aids, with the aim of making our description easier to understand and informing our decisions about the depth of description necessary for access. While a variety of approaches can be used to shed light on these questions, this study used task analysis to examine the usability of three models of description, and found that finding aids with MPLP-influenced description can be nearly as usable as traditional finding aids containing full description. We are now beginning the process of considering how this result should impact the practicalities of creating new finding aids and improving existing ones. Further usability and user studies—even routinizing the practice of performing these studies on archival description—are necessary for archivists to do exactly as we achieve our aim: center users in our descriptive choices and provide access to all who are interested in the materials we steward.

Appendix

For access to the testing script used in this study, please visit https://docs.google.com/document/d/-16m6h1jlWCHczVKFqUU-fPGvziCsGjTP2evq6nnJzBnw/edit?usp=sharing. For in-depth reports with further details, please contact the author.