Speaking on the Record: Combining Interviews with Search Log Analysis in User Research

Lynn Silipigni Connaway
OCLC Research

Brittany Brannon
OCLC Research

Christopher Cyr
Morning Consult

Peggy Gallagher
OCLC

This article represents work done on behalf of OCLC Research. When this article was written, Christopher Cyr was an Associate Research Scientist at OCLC. At the time of publication Christopher had left OCLC to become the Director of Data Science for Morning Consult.

Preprint citation

Journal citation:

Speaking on the Record: Combining Interviews with Search Log Analysis in User Research

Lynn Silipigni Connaway, Brittany Brannon, Christopher Cyr, Peggy Gallagher

OCLC Research

Author Note

Lynn Silipigni Connaway, OCLC Research, connawal@oclc.org

Brittany Brannon, OCLC Research, brannonb@oclc.org

Christopher Cyr, OCLC Research, ccyr85@gmail.com

Peggy Gallagher, OCLC Research, gallaghp@oclc.org

Correspondence concerning this article should be addressed to Lynn Silipigni Connaway, 6565 Kilgour Pl, Dublin, OH 43017. Email: connawal@oclc.org

Abstract

This paper reports on a novel sequential mixed methods approach to study user search behavior within a library discovery system. Search logs were used to reconstruct each participant's search session and create customized semi-structured interview protocols using the critical incident technique. Findings from the interviews informed statistical log analysis, which identified features of search sessions that made accessing resources more likely. Using this sequential mixed methods approach provides a more holistic way to study library discovery systems. Individual interviews provide more precise data when protocols are created using the participants' logs. Similarly, statistical log analyses are enhanced with users' descriptions of their behaviors in library discovery systems. While prior studies have employed both interviews and log analysis, using the methods to inform one another reduces the limitations and enhances the benefits of each.

Speaking on the Record: Combining Interviews with Search Log Analysis in User Research

1. Introduction

A novel sequential mixed methods approach was used to identify how academic library users navigate an online discovery layer. Semi-structured individual interviews were personalized based on a re-creation of individuals' search logs. This helped to mitigate the biases common in interview methods. Interviews provided rich data about how and why users engaged in different search behavior and shed light on their perceptions of successful search experiences. The findings from the interviews were then used to inform the development of search success categories for quantitative analysis of search logs. Log analysis provided data collection at scale to model probabilities of different factors that influence search success and failure within the system.

2. Problem statement

Identifying how academic library users navigate online discovery systems is crucial for developing and implementing library systems. Discovery systems are an important step toward aligning users' experiences searching library systems with the expectations they have developed while using search engines and other commercial applications (Fons, 2016). Bossaller and Moulaison Sandy (2017, p. 602) explain that:

> "...discovery systems are highly interconnected systems that sit on top of and seamlessly bring together the results from disparate library databases such as the library catalog, electronic article databases, and e-book packages. Discovery systems were developed in 2009, essentially uniting the user-friendliness of

discovery interfaces with the functionality of multidatabase search offered

through federated search, simplifying search of library materials through the

addition of a central index."

Discovery systems remove obstacles for academic libraries users, providing them access to curated and vetted resources in more familiar and convenient ways.

Researching library discovery systems provides special challenges because the systems integrate a number of other information systems, including commercially and locally developed catalogs, databases, and document delivery services. Library discovery systems are unique because library resources come in a variety of physical and digital formats, and can be accessed through a variety of methods, channels, and systems. Discovery systems are most common at academic libraries, and academic library discovery systems are also the most studied (Bossaller & Moulaison Sandy, 2017). This project focused on academic library users of a cloud-based library discovery system that allows users to find resources available at their library and in other libraries worldwide, most commonly books and articles.

3. Previous Methods for Studying Discovery System Use

The present paper combines search logs to directly measure user behavior with semi-structured individual interviews to gather self-reported data. Search logs provide a system-based perspective on what users do within a system, while interviews can provide a user-based perspective on the decisions, contexts, motivations, and success of a search that are not limited to interaction within a single system. Each method has benefits and limitations that the use of them combined can help to mitigate. Both of these methods are relatively uncommon in LIS.

In a meta-review of 58 LIS methods review articles, Ullah & Ameen (2018) found that interviews account for approximately 4% of the methods used and log analysis accounts for approximately 1.5% of the methods used (Ullah & Ameen, 2018). In a review of research articles published from 2001-2010 in three LIS journals, Chu and Ke (2017) found that interviews accounted for approximately 9% of the methods used, while log analysis accounted for approximately 4% of the methods used (Chu & Ke, 2017). While these numbers roughly capture the use of interviews and log analysis in published LIS research, both methods are beneficial for a variety of uses in assessment and planning and are likely used within institutions in ways that are not captured in the literature.

In a systematic review of 80 peer-reviewed articles about library discovery layers published between 2009 and 2013, Bossaller and Moulaison Sandy (2017) found that case studies were the predominant method used, accounting for 65% of all methods. Usability testing and surveys followed, accounting for 26% and 21% of the methods respectively. Log analysis accounted for only 9% of the methods used (Bossaller & Moulaison Sandy, 2017). The majority of these studies' findings were focused on the use and usability of library discovery layers (51%) and discovery layer selection and implementation (46%).

Searching behavior is a difficult topic to study because there are many factors that impact how users find and select resources, factors which must often be observed and measured in quite different ways. For example, a searcher's mood, previous experience, and goal(s) when searching (Ross, 1999), or any recommendations that they've received (Chevalier & Mayzlin, 2006; Liu, 2006), are all factors that the searcher

must report. Similarly, search success and the enjoyment of the search experience must be reported by the searcher (Tang & Jhang, 2020) because there is not necessarily a direct correspondence between a searcher accessing an information resource and that resource meeting their information need. Other factors, such as cultural background and influence (Khosrowjerdi et al., 2020), are difficult to measure empirically.

Some factors, such as the usage of system features or interaction with information resources, can be directly observed, either through log data or observational methods. However, library discovery systems present especially large challenges for researchers. In enclosed discovery systems, user access can be measured directly. Library discovery systems unite a variety of databases, bibliographic, and other information systems, connecting users to both physical and digital resources. Access can happen inside the library discovery system, in a connected system, or through outside methods such as external webpages or finding the physical item on a library shelf. As a result, it is not possible to directly measure all types of access in a library discovery system.

3.1 Log Analysis

Agosti et al. (2012) review the use of log analysis from 1999-2009 and find that it separates into two branches: web search engine log analysis and digital library system log analysis. This paper belongs to the second branch, which is a research tradition that predates the World Wide Web, beginning with transaction log analysis of online public access catalogs (OPACs) (Agosti et al., 2012; Jones et al., 2000). Agosti et al. (2012) identify an early example of transaction log analysis of OPACs carried out from 1981-

1983 at OCLC, where researchers analyzed transaction logs recorded on tapes to study system feature usage. With the emerging popularity of digital libraries, the techniques of log analysis were applied to digital library systems beginning in the late 1990s (Agosti et al., 2012; Jones et al., 2000).

As library discovery layers have become more ubiquitous, particularly in academic libraries, they are now being studied using log analysis, although other methods remain far more common (Bossaller & Moulaison Sandy, 2017). Library discovery systems record every user request made to the servers. This provides an unobtrusive way to collect large amounts of user data (Agosti et al., 2012; Dumais et al., 2014; Jansen, 2006). Log analysis involves looking at these recorded transactions to understand system performance and make inferences about user behavior. For example, other LIS studies have used logs to improve system performance and configuration, to study users' search strategies, system feature usage, and search failure and success, and to categorize different types of information behaviors within these systems (Agosti et al., 2012; Fischer et al., 2020; Hunter, 1991; Jones et al., 2000; Lown et al., 2013; Nouvellet et al., 2019).

Analysis of search logs offers many benefits. In their review of log analysis studies on the use of electronic journals, Jamali et al. (2005) highlight the potential for log analysis to help understand information behavior. Logs capture a variety of types of usage that other measures do not capture, and show what people actually do rather than what they think they do (Dumais et al., 2014; Jamali et al., 2005). Connaway and Radford concur, stating that the log "clearly demonstrates how users employ their search strategies rather than how users describe their search strategies" (Connaway &

Radford, 2021, p. 213). Since log events are captured directly from user behavior in a natural environment, there is no interference from the researcher (Dumais et al., 2014; Jansen, 2006).

In addition, log analysis allows the collection of a large sample size relatively inexpensively. Sample sizes for log analysis typically are in the hundreds of thousands to millions (Dumais et al., 2014), which would be prohibitively expensive or simply impossible for many other data collection methods. This means that it is less resource dependent to gather a representative sample of logs than it would be for most other methodologies. It also provides an opportunity to capture unusual or less common behaviors that may not appear in smaller sample sizes (Dumais et al., 2014).

Log analysis also has several limitations. Log analysis requires that researchers make use of the data that they have available, instead of deciding what data to collect. Samples are by default non-random and might not be representative of the population under consideration, since they only capture people who use the system under study, even though the large sample size allows researchers to generate a random sample within that population. Additionally, all of the observations collected are limited to existing interactions with the existing system, or "what people do with the tools they have" (Dumais et al., 2014, p. 356). This means that logs cannot reveal what users wished to do or what they would have preferred. Log data generally do not contain any user demographic information (Connaway & Snyder, 2005; Dumais et al., 2014), and it can be difficult or impossible to tell which log events belong to a single user. This means that while logs are good for generalizing about the behavior of an entire population (Fischer et al., 2020), they may not say much about individual behavior.

Log events are often ambiguous and can be prone to data inconsistencies. From a data perspective, server logs do not capture events that are handled browser-side, such as using the browser's back button to revisit cached pages. Depending on the size and scale of the system, the logs may also be prone to dropped data and log events whose names and actions are unintelligible (Dumais et al., 2014). Understanding timing and sequencing within log events also comes with inherent challenges, such as network lag and non-chronological log event capture (e.g., multi-tabbing) (Dumais et al., 2014).

From a behavioral perspective, it is often difficult to fully understand what log events mean in terms of user behaviors and perceptions. In part, this is because there is not a direct one-to-one correspondence between log events and user behaviors, either low-level or complex. For example, some log events are generated by user action while others are system-generated and occur regardless of user behavior. Similarly, the same observed log events may have very different real-world meanings, such as quickly abandoned searches that can mean either that the user found exactly what they wanted immediately or that they did not get anything relevant. Dumais et al. (2014) argue that due to these ambiguities, deep analysis of complex user behavior should be done first by a hand examination or interpretation before trying to perform large-scale analysis.

Importantly, logs also cannot capture a user's context, perspectives, or states of mind during interactions with the system. This means that log analysis cannot directly address how or why people engaged in the recorded behaviors or what their motivations, attitudes, and perceptions of success were (Dumais et al., 2014). In the context of discovery systems, log analysis does not explain why they were searching or if they found and were able to access what they were looking for (Connaway & Snyder,

2005). Dumais et al. (2014) suggest that researchers should employ complementary methods, such as usability testing, eye tracking, field studies, and surveys, to ensure a fuller understanding of user behavior. Although they do not mention interviews specifically, interviews allow for the collection of rich, contextualized data that addresses many of the limitations of log analysis.

3.2 Individual Semi-Structured Interviews

User interviews can help capture search and discovery behavior as the user understands it, rather than as a computer system understands it. Interviews can gather rich data about how and why users engage in particular behaviors, not just what behaviors they have engaged in. Interviews have been used in several studies of search behavior (e.g., Burt & Liew, 2012; Fu, 2019; Hu, 2019; Kipp & Campbell, 2010). However, Bossaller and Moulaison Sandy (2017) do not specifically mention interviews at all in their review of library discovery system research methods.

Semi-structured interviews are one method of user interview, where the interviewer asks a set list of questions with follow-up and probing questions. The flow of the interview is the primary determinant of the follow-up questions, which allows researchers to probe into areas of interest and confusion (Connaway & Radford, 2021). This offers flexibility in the interview, depending on the information that the interviewee provides (Bailey, 2017). The critical incident technique (CIT) developed by Flanagan is a common method for creating semi-structured interview protocols (Connaway & Radford, 2021; Flanagan, 1954). This technique asks participants to reflect on a specific event in order to elicit insight into their behavior in different situations (Connaway &

Radford, 2021). CIT is particularly helpful for identifying behaviors that lead to success or failure and placing those behaviors in the appropriate context (Redmann et al., 2000).

Individual semi-structured interviews do have drawbacks. They tend to involve a relatively small number of observations. There are scheduling and logistical hurdles that make semi-structured interviews challenging. They are time-consuming to conduct and analyze, so most organizations do not have the resources to conduct a large number of semi-structured interviews. They also can be quite expensive, since many semi-structured interviews involve participant incentives in addition to researcher time to schedule, collect, and analyze the data. For these reasons, it can be difficult to obtain a representative sample.

Semi-structured interviews can be somewhat unreliable. Since these interviews involve having interviewees report their own behavior, they can suffer several biases. Social desirability bias occurs when users tell researchers what they believe the researcher wants to hear (Babbie, 2021). This can be an especially great concern when the researcher is affiliated with the organization that created the discovery system being researched. Users can forget critical pieces of information. This is particularly important for search behavior, as interviews are generally conducted after a search session ends and users might not remember every step that they took. Finally, researchers inadvertently can bias interviews themselves with the way that they ask questions and probes (Connaway & Radford, 2021).

3.3 Combining Log Analysis and Interviews

Because of the challenges discussed above, a novel mixed methods approach was used to capture the complexities of user experience. This methodology, to the

authors' knowledge, has not been used previously and is an update to a methodology developed by Connaway, Budd, and Kochtanek (1995). It builds on recent work on log analysis (Garcia-Gathright et al., 2018; Lamkhede & Das, 2019). A recent study of a music streaming service used interviews to inform statistical modeling of logs, but used the same interview protocol for each participant (Garcia-Gathright et al., 2018). Because the current methodology directly asks users about trace behavior found in search logs, rather than their perceived search behavior, it allows for more precise and thorough understanding of the meaning of search log events.

4. Methodology

This study employed sequential mixed methods with each method's analysis informing the next method. The research design was novel because it used semi-structured individual interviews that were personalized to each participant based on the re-creation of their search session from log events. These personalized interviews allowed the users to explain their actions and perceptions during their search session, including how and why they conducted their search in the way that they did and whether they thought their session was successful. These user explanations informed translation between user behaviors and transaction log events prior to conducting a log analysis. The log analysis was used to determine the probability of different factors affecting search success.

The first stage of data collection was a screening survey used to recruit semi-structured interview participants. Survey respondents identified the purpose, goals, and success of their search and provided an ID that identified their search session within otherwise anonymous logs. The survey also included demographic data, such as

14

education level and discipline, that was used to select the sample. Survey responses were collected from April 5-26, 2018.

The second stage of data collection was personalized semi-structured individual interviews. Once participants were selected and scheduled for an interview, the log events associated with the ID they provided in the survey were reviewed to reconstruct their search session. The reconstructed search session and details from the survey data were used to create personalized interview protocols that asked each individual about events that happened during the search and why they considered the experience successful or unsuccessful. Interviews were conducted from April 27-May 29, 2018. Interview data were coded using a codebook that included themes that emerged from the data.

The third stage of data collection was the log analysis. The individual interviews provided insight into how the log events correlated to participants' real-life behavior. This insight was used to create categories of log events that represented meaningful user options and actions. In September and October of 2018, these categories were used to analyze a sample of the discovery system logs for patterns that predict the success and failure of search sessions. To coincide with the timing of the previous interview data collection, logs of searches that took place in April 2018 were analyzed. The log analysis provided a broader picture of the nature of discovery and access, including the probability of accessing materials based on the search and the items retrieved.

4.1 Recruitment

A list of potential interviewees was identified through a screening survey that used a purposive convenience sample. Contacts at 15 U.S. academic libraries and one library consortium that used the discovery layer under study were invited to participate. Five university libraries in different US regions agreed to participate, including four at small private academic institutions and one at a large public institution. All participating universities had their institutional review board (IRB) approve the study methodology. Librarians at each university sent a recruitment email to their regular patrons that explained the project and incentive for participation and gave a link for interested participants to access the screening survey after they had completed a search. Respondents were offered a $20 Amazon gift card if they both completed the survey and participated in an interview.

One challenge of recruiting in this way was that participants were not prompted to complete the survey at the time they finished their search. Instead, they received the email request and then had to remember to take the survey later, after they had completed a search. This likely is the explanation for the low response rate. The original intent was to integrate a pop-up into the discovery layer that would remind users to take the survey during their search. Given the way that the discovery layer was created and configured, the labor and time needed for this approach was not feasible. While a pop-up in the discovery layer during or immediately following the search would have been optimal, even a pop-up on the library webpage would have likely produced a better response rate by prompting users to take the survey close to when they were conducting searches.

The screening survey asked respondents to provide key information about their search session, including what they were looking for and if they felt that they had accomplished their purpose. They were then asked to paste the Request ID from their search, which was used to match their survey with the log of their search session. The survey provided demographic information about the respondents, which included gender, age, academic level, discipline, parent or guardians' education level, and the educational institution where they performed the search.

The intention was to use these demographics to select individuals to interview from different educational levels, disciplines, parents' or guardians' educational levels, gender, and age. However, with the limited number of respondents and attrition between the survey and interviews, it was not possible to perfectly meet those selection criteria. As a result, representation of participants from a variety of disciplines across undergraduate, graduate, and faculty levels was prioritized for the individual interviews. During the course of the interviews, saturation was reached, which is when ideas or concepts are repeated by different participants (Connaway & Radford, 2021). Fourteen interviews were completed in total.

4.2 Reconstructing Searches and Personalizing Semi-Structured Interviews

Once a participant was selected and scheduled for an interview, a customized interview protocol was developed for that participant using details of the search session found in the logs. To the authors' knowledge, this methodology of creating interview protocols based on session logs has not been done before. The Request IDs that participants provided in their screening surveys enabled the research team to identify the users' search session, and all related log events were then extracted and cleaned of

traces unassociated with user behavior. This resulted in a log of every command and click event associated with the search session.

Log events were not always recorded in sequence. As a result, it was necessary to closely look at the time stamps and logical flow of the events in the logs to reconstruct the search. Two members of the research team reconstructed each user's behavior based on details found in the log. This information was used to create a summary of the major actions taken by the participant, including search terms and limiters used, clicks into individual items, and clicks on elements such as descriptions and full-text access.

Interview protocols were developed using CIT, where users were reminded of the specific steps that they took throughout their search session and asked to explain them. Interviews were divided into four sections: purpose of the search, review of the search log, success of the search, and concluding questions. The full interview protocol can be found in Appendix 1.

In purpose of the search, participants were asked to elaborate on what they were searching for and why they had conducted the online search. The protocols served as a reminder of what they searched for, as interviewees were asked about their specific search terms. In the review of the search log, participants were asked about specific steps that the search log indicated that they had taken, why they had taken them, their perception of the search results, and whether they had used external tools not reflected in the search log. Users were, for example, asked about limiters such as title and author, and in some cases, they did not remember using these limiters. These reminders helped to generate detailed responses about search behavior that otherwise

would not have been mentioned in user interviews. In success of the search, participants were asked to explain whether they found what they were looking for, how they felt about the experience, and if there were points of delight or frustration. The specificity of the questions about their search helped users identify the specific points in the search that impacted their overall perceptions of the search experience and its success.

Some user behavior that was articulated during the interviews cannot be captured by logs. Users might use a discovery system to find a book but use a different search engine to look up reviews of the book. Alternatively, users might leave a browser window open on search results for several hours while they have walked away from their computer. The log would indicate that the search session lasted for several hours, when in reality it only lasted a few minutes. Just as online user behavior outside of the system will not be captured by the logs, physical search behaviors will not be captured by the log. Users who consult a human source, such as a librarian or professor, or go to the shelves to look at a book will not have that behavior captured by the log but it can emerge during interviews.

A codebook was developed based on the common themes emerging from the interviews. These themes were organized to capture user search strategies, decision-making factors, preferred resource formats, feelings of frustration and delight, and other relevant search behaviors and preferences. The researchers coded several interviews together to refine the codebook and resolve problems of ambiguity in code definitions. Each interview was coded by two team members, with an average of 84% agreement between paired coders.

4.3 Developing Categories for Log Analysis

The user interviews helped identify the actions represented by the log events. While the intention for the interviews was to better understand the contexts, motivations, and perceived success of searches, the process of creating the personalized protocols and conducting the interviews served as the kind of hand interpretation that Dumais et al. (2014) recommend before trying to analyze logs. For example, during the interviews, there were users who did not recollect requesting holdings of an item, despite log records indicating that they had. Upon investigation, it was discovered that the holdings might have been automatically generated on the search results page.

Based on the user interviews, all log requests and click events were classified into six categories: 1) search results, 2) online access attempt, 3) physical access attempt, 4) attempt to save, 5) physical access option, and 6) other (a small number of events that did not fit cleanly into the other categories). The *search results* category indicated when the user searched for something. Any *attempt* category indicated that the user had definitively clicked on a link or button to try to get or save a resource. Both *access attempt* categories required users to click to get access to either an online or physical resource. This included clicking on a "full-text access" link for digital resources and clicking on a "place hold" link for access to a physical resource, among other options. The *attempt to save* category indicated that the user saved the item or citation.

The *physical access option* category included log events where the user had the option of accessing a physical item but did not unambiguously indicate that the user attempted to access the item. This categorization came to be as a direct result of the user interviews that were conducted. For example, one user's session log suggested

that they abandoned the search on the results page. The user, however, said that they found the call number for a book on the results page, and then walked to the library stack to pick up the item. As described above, system requests for resource holdings might indicate that users discovered resources and began the process of accessing them—or could have been automatically generated by the system for each search result. These holdings provide sufficient information for users to retrieve a physical item from the shelf. Because it is impossible to definitively say whether some of these events indicated user action or not, it was important to classify them as physical access options.

Approximately 325,000 search sessions occurred in the library discovery system during April of 2018. Of these, approximately 282,000 were from academic libraries. Because the individual interviews focused on users of academic libraries, the log analysis excluded non-academic libraries. This is consistent with the focus on academic libraries in the literature on discovery systems (Bossaller & Moulaison Sandy, 2017). The academic libraries in the log sample included academic research libraries (ARLs), 4-year colleges, and community colleges. Search logs included all requests and click events that were sent to the discovery system. For each search session, an analysis of the session length, number of searches, average words per search, number of results per search, and the search refinements and limiters used during the session was conducted.

Additionally, the six categories were used to conduct two different types of analysis on the logs. First, aggregate descriptive statistics of the variables in the data were generated. Second, a logistic regression analysis was used to determine the

impact that several factors had on the likelihood that a search session contained an access attempt or option. This identified which features of the system and types of search behaviors were most important in determining access. The large sample size of the logs allows for probability predictions, which cannot be generated from the smaller sample sizes of qualitative data.

5. Reflections on the Methods

Studying users' information-seeking behavior within a discovery layer through the combination of log analysis with semi-structured individual interviews provided detailed and targeted information specific to the search session. Individual search session logs helped to create customized interview protocols, which elicited more detailed insights into the search session than traditional interview protocols that ask general questions about a search session. In addition, user interviews enhanced log analysis by clarifying the meaning of log events that were otherwise ambiguous. In many cases, these clarifications would not have been otherwise obvious, and likely would not have been correctly classified if not for the customized user interviews.

Future researchers should keep this methodology in mind when conducting user experience studies on systems that leave digital traces. Preparing interview protocols based on log records can enhance the specificity and reliability of user interviews, which in turn can increase the accuracy of conclusions that come from log analysis. The insights that these combined methods can yield provide a more holistic understanding of the interaction between user search behaviors and features of the discovery system. This can help library staff to further develop their discovery systems in ways that help to improve access. It also enables library staff to help users improve their search

behaviors to more consistently discover and access resources that meet their needs.
Search success is nuanced and somewhat elusive. While it may never be possible to
determine search success from logs, this combined methodology provides one way to
identify log events that represent access from a user perspective. Interviewing and other
self-reporting methods can allow researchers to better elicit users' perceptions of search
success.

5.1 Disadvantages

While the methodology of this study yielded many benefits, it did have
disadvantages. In particular, the plurality of methodologies required specialized skills
that might not be available to every research team. This project was conducted with
team members who had the programming skills required to query search logs, both to
identify specific search sessions and to combine search statistics. It also required
having team members trained in semi-structured interview methodology and codebook
development. Finally, it required having team members trained in statistical analysis.
The methodology also was very time consuming and required many hours of time and
commitment by the project team.

Two interviewees expressed privacy concerns when they were presented with
logs of their search sessions and asked specific questions about their search behavior.
Users were reassured that no personal information was recorded in their search logs,
and the only way search logs were matched with them was the Request ID that they
provided in the survey. This touches on a bigger issue when researching search
behavior, as there is a fundamental tension between search systems respecting user
privacy, and systems helping users access information more efficiently (Bawden &

Robinson, 2020; Dumais et al., 2014). This methodology offers one possibility for addressing user privacy while collecting individual search information. The discovery system search sessions are anonymous, and only recorded through Request IDs, which contain no personal information and were voluntarily provided by interviewees. This enables a deeper understanding of user behavior without revealing any information that users are uncomfortable revealing and have not consented to reveal and share. However, anonymity in logs does not guarantee user privacy (Dumais et al., 2014), and log data should be handled and shared carefully to ensure user privacy is respected.

5.2 Advantages

Using logs to personalize interview protocols reduces the biases inherent in qualitative semi-structured interviews. Users are unlikely to remember every step that they took during a search session, especially when the session might have occurred a month earlier. Mentioning specific log events helped remind users of parts of the search that they had forgotten. In this way, creating the protocols from the log events helped to mitigate recall bias and focused on more specific aspects of the user experience. Additionally, because the searches were conducted independently of the research and recreated from the log events, that behavior was not influenced by the data collection. While there still is an element of social desirability bias possible as users explain why they took particular actions, the risk of social desirability bias is reduced because they are not reporting the sequence of events of their search. The users only are explaining the how and why of the actions they are asked about.

The semi-structured interviews captured data about two things that logs alone cannot capture: users' perceptions of the search process and any search behaviors

happening outside the discovery system. The interviews provided rich data about how and why users search, and how they perceive search experiences. This includes the motivations for their search, their perceptions of their information needs, and specific points of frustration and delight throughout the experience. The interviews also revealed search behaviors that are not captured in the library discovery system logs. Many participants started with or used web search engines, and others talked about the positive effect that previous library instruction had on their search experiences.

By using the interviews to identify the user behaviors that different log events could represent, the categories of log events more accurately reflected users' behaviors and perceptions of their search. Log analysis is blind to any behavior happening outside of the system. In a library discovery system, where resources are heterogenous and can be accessed in a variety of ways, this poses particular challenges for understanding search success. For example, some users reported going to the library shelves to pick up physical items, where reviewing the logs would have suggested an abandoned search.

Using this sequential mixed methods approach provides a more holistic way to study library discovery systems. Individual interviews can provide more precise data when protocols are created using the participants' logs. Similarly, the accuracy of statistical modeling in log analysis was improved by creating categories based on users' descriptions of their behaviors when a certain log event was present. While prior studies have employed both interviews and log analysis, using the methods to inform one another reduces the limitations and enhances the benefits of each.

References

25

Agosti, M., Crivellari, F., & Di Nunzio, G. M. (2012). Web log analysis: A review of a decade of studies about information acquisition, inspection and interpretation of user interaction. *Data Mining and Knowledge Discovery*, *24*(3), 663–696. https://doi.org/10.1007/s10618-011-0228-8

Babbie, E. R. (2021). *The practice of social research* (Fifteenth edition). Cengage.

Bailey, C. R. (2017). *A Guide to Qualitative Field Research* (Third edition). SAGE Publications, Inc.

Bawden, D., & Robinson, L. (2020). "The dearest of our possessions": Applying Floridi's information privacy concept in models of information behavior and information literacy. *Journal of the Association for Information Science and Technology*, *71*(9), 1030–1043.

Bossaller, J. S., & Moulaison Sandy, H. (2017). *Documenting the Conversation: A Systematic Review of Library Discovery Layers*. https://doi.org/10.5860/crl.78.5.602

Burt, M., & Liew, C. L. (2012). Searching with clustering: An investigation into the effects on users' search experience and satisfaction. *Online Information Review*.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354.

Connaway, L. S., Budd, J. M., & Kochtanek, T. R. (1995). An investigation of the use of an online catalog: User characteristics and transaction log analysis. *Library Resources & Technical Services*, *39*(2), 142–152.

Connaway, L. S., & Radford, M. L. (2021). *Research methods in library and information science* (7th ed.). Libraries Unlimited, an imprint of ABC-CLIO, LLC.

Connaway, L. S., & Snyder, C. (2005). Transaction log analyses of electronic book (ebook) usage. *Against the Grain*, *17*(1), 85–93.

Dumais, S., Jeffries, R., Russell, D. M., Tang, D., & Teevan, J. (2014). Understanding User Behavior Through Log Data and Analysis. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of Knowing in HCI* (pp. 349–372). Springer New York. http://link.springer.com/10.1007/978-1-4939-0378-8_14

Fischer, R. K., Iglesias, A., Daugherty, A. L., & Jiang, Z. (2020). A transaction log analysis of EBSCO Discovery Service using Google Analytics: The methodology. *Library Hi Tech*, *39*(1), 249–262. https://doi.org/10.1108/LHT-09-2019-0199

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*(4), 327.

Fons, T. (2016). *Improving web visibility: Into the hands of readers*. ALA TechSource alatechsource. org-American Library Association.

Fu, H. (2019). Mixed language queries in online searches. *Aslib Journal of Information Management*.

Garcia-Gathright, J., St Thomas, B., Hosey, C., Nazari, Z., & Diaz, F. (2018). Understanding and evaluating user satisfaction with music discovery. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 55–64.

Hu, X. (2019). Evaluating mobile music services in China: An exploration in user experience. *Journal of Information Science*, *45*(1), 16–28. https://doi.org/10.1177/0165551518762070

Hunter, R. N. (1991). Successes and failures of patrons searching the online catalog at a large academic library: A transaction log analysis. *RQ*, 395–402.

Jamali, H. R., Nicholas, D., & Huntington, P. (2005). The use and users of scholarly e-journals: A review of log analysis studies. *Aslib Proceedings*, *57*, 554–571.

Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. *Library & Information Science Research*, *28*(3), 407–432.

Jones, S., Cunningham, S. J., McNab, R., & Boddie, S. (2000). A transaction log analysis of a digital library. *International Journal on Digital Libraries*, *3*(2), 152. https://doi.org/10.1007/s007999900022

Khosrowjerdi, M., Sundqvist, A., & Byström, K. (2020). Cultural Patterns of Information Source Use: A Global Study of 47 Countries. *Journal of the Association for Information Science and Technology*, *71*(6), 711–724. https://doi.org/10.1002/asi.24292

Kipp, M. E., & Campbell, D. G. (2010). Searching with tags: Do tags help users find things? *KO KNOWLEDGE ORGANIZATION*, *37*(4), 239–255.

Lamkhede, S., & Das, S. (2019). Challenges in Search on Streaming Services: Netflix Case Study. *ArXiv Preprint ArXiv:1903.04638*.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89.

Lown, C., Sierra, T., & Boyer, J. (2013). How Users Search the Library from a Single Search Box. *College & Research Libraries*, *74*(3), 227–241.

Nouvellet, A., D'Alché-Buc, F., Baudouin, V., Prieur, C., & Roueff, F. (2019). A

    Quantitative analysis of digital library user behaviour based on access logs.

    Qualitative and Quantitative Methods in Libraries, *7*(1), 1–13.

Redmann, D. H., Stitt-Gohdes, W. L., & Lambrecht, J. J. (2000). The Critical Incident

    Technique: A Tool for Qualitative Research. *Delta Pi Epsilon Journal*, *42*(3), 132–

    153.

Ross, C. S. (1999). Finding without seeking: The information encounter in the context of

    reading for pleasure. *Information Processing & Management*, *35*(6), 783–799.

Tang, M.-C., & Jhang, P.-S. (2020). Music discovery and revisiting behaviors of

    individuals with different preference characteristics: An experience sampling

    approach. *Journal of the Association for Information Science and Technology*,

    71(5), 540–552. https://doi.org/10.1002/asi.24259

Appendix 1: Interview Protocol

Purpose for Doing Online Search

Instructions for Interviewer: Refer to the interviewee's survey responses

Question set-up:

Let's first quickly review the responses you gave on the brief survey you took immediately after conducting your search.

You indicated that you [read answer from interviewee's survey response]:

    1. Knew the exact item or material you wanted, and you were using this site to locate or access it.

    2. Knew the exact item or material you wanted, and you were using this site to get details about it (e.g., to check the publication date or author of a book).

    3. Didn't know the exact item or material you wanted but you knew a topic you wanted to explore.

4. Other:

1. Please tell us what you were looking for and why you decided to do an online search.

    Probe: Was this for a class assignment or research paper or something other purpose?

Review of Search Log

Instructions for Interviewer: Refer to the interviewee's search log

Question set-up:

Now, let's look at the actual search you conducted and talk about your search strategy and the steps you took. It looks like you began your search with [topic]:

2. Tell me a little bit about whether your search worked the way you thought it would.

    What was the next step you took (and why did you take that)?

        [protocol log summary here]

    Did you try narrow down your search in any way?

        Why?

    How did you narrow your search?

        Which of the boxes on the side did you check to narrow your search (e.g., by date or author or format, etc.?)

Success of Search

Instructions for Interviewer: Refer to the interviewee's survey responses

Question set-up:

Going back to the survey you completed, we already know that you [read answer from interviewee's survey response]:

1. Knew the exact item or material you wanted, and you were using this site to locate or access it.

2. Knew the exact item or material you wanted, and you were using this site to get details about it (e.g., to check the publication date or author of a book).

3. Didn't know the exact item or material you wanted but you knew a topic you wanted to explore.

4. Other:

You also indicated that you were hoping to [read answer(s) from interviewee's survey response]:

- Borrow an item from the library

- Read/download/listen to what I was searching for immediately online.

- Purchase an item.

- Finish a citation or find details about the item (e.g., publication date or author).

- Not sure.

- Other:

Let's talk about that for a few minutes.

If searcher was looking for a "known item" (option "a" or "b" above)

3a. Did the item you were searching for come up in your search results? In other words, did you find it?

[If yes, item was in search results]: Were you able to get the item?

If yes, able to get item	If no, not able to get item
How did you get it? Tell me about that experience. Probe, if needed: How long did it take? Would you say it was easy?	What did you do next any why? Probe, if needed: Did you just give up on finding the item? Did you search for a different item? Did you search for the same item someplace else (e.g., Amazon)?

[If no, item was not in search results]: What did you do next and why? Probe, if needed: Did you just give up on finding the item? Did you search for a different item? Did you search for the same item someplace else (e.g., Amazon)?

If searcher was looking for a "topic" or other "unknown item" (option "c" from above)

3b. Did your search turn up something that you thought would meet your needs?

[If yes, found something that met needs]: Were you able to get what you found?

If yes, able to get item	If no, not able to get item
How did you get it? Tell me about that experience. Probe, if needed: How long did it take? Would you say it was easy?	What did you do next and why? Probe, if needed: Did you just give up on getting what you found? Did you search for a different item? Did you search for the same item someplace else (e.g., Amazon)?

[If no, did not find anything that met needs]: What did you do next and why? Probe, if needed: Did you just give up on finding what you were looking for in the library's search? Did you search for something different? Did you search for the same thing someplace else (e.g., Amazon)?

Instructions for Interviewer: Refer to the interviewee's survey responses
Question set-up:
Going back to the survey you completed one more time, you indicated that you: [read answer from interviewee's survey response]:

- Accomplished what you were hoping to do
- Did not accomplish what you were hoping to do

[If searcher accomplished]:

4a. I'd like to understand how you felt about your search experience overall. Would you say you were delighted with your search experience?

      [If yes] What made it delightful?

      [If no] What do you think would have made it delightful?

[If searcher did not accomplish]:

4b. I'd like to understand how you felt about your search experience overall. Was there anything about the search experience that you would say delighted you?

      [If yes] What was it that delighted you and why?

      [If no] What do you think would have made it delightful?

5. Were you frustrated at any time with your search experience?

      [If yes] Where in your search were you frustrated, and what made it frustrating?

          What did you do next? (Probe, if needed: Did you find a work-around?

          Abandon the search? Start over?)

Conclusion of Interview

6. What else, if anything, would you like to share about your search experience?

7. What questions do you have for me?