

University of Denver

Digital Commons @ DU

Electronic Theses and Dissertations

Graduate Studies

2022

Digital Searching: A Grounded Theory Study on the Modern Search Experience

Nicolas Armando Parés

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Computer Sciences Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), and the [Information Literacy Commons](#)

Digital Searching: A Grounded Theory Study on the Modern Search Experience

A Thesis

Presented to

the Faculty of the Morgridge College of Education

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Nicolas Armando Parés

November 2022

Advisor: Peter Organisciak

©Copyright by Nicolas Armando Parés 2022

All Rights Reserved

Author: Nicolas Armando Parés
Title: Digital Searching: A Grounded Theory Study on the Modern Search Experience
Advisor: Peter Organisciak
Degree Date: December 2022

Abstract

This Grounded theory study explores US adults' modern information search process as they pursue information through digital search user interfaces and tools. To study the current search process, a systematic grounded theory methodology and two data collection methods, a think-aloud protocol and semi-structured interviews, are used to develop the theory. The emerging theory addressed two tightly connected research questions that asked, “What is the process by which humans search and discover information?” and “What is the process by which search and discovery interfaces and tools support the modern search process?”

The study collects participant data from US adults who have recently completed internet-based searches or were employed as reference librarians or information specialists that regularly support internet-based searches. The 40 data collection events, 20 interviews, and 20 think-aloud search activities are constant comparatively analyzed. This analysis shows the importance of validating information during the search process as 1) establishing accurate information needs and 2) verifying information results. This theory and study findings can be applied to modern search interface designs or used to update modern information literacy curriculums.

Table of Contents

Chapter One Introduction.....	1
Background	1
Modern Need for Information	2
The Gap in the Literature	3
The Process of Search and Discovery	4
Studying the Process with Grounded Theory.....	6
Studying the Process with Authentic Exposure.....	6
A Modern Understanding.....	7
Chapter Two Literature Review	8
The State of Internet-based Searching.....	8
The Digital Search Divide as Access	8
Digital Search Literacy.....	11
Information Behavior and Information Retrieval.....	13
Information Retrieval	14
Discoverability & Findability.....	15
Human Search and Sense-making.....	16
The Digital Search Experience.....	17
Design Principles.....	18
Information Seeking as a Process.....	21
Theories of Information Behavior	22
Measuring Process with Usability Testing.....	24
A Qualitative Approach to Studying Human Process	25
Problem Statement	26
Chapter Three Methods	28
Restatement of purpose	28
Research Approach	28
Research Question.....	29
Sampling.....	30
Theoretical & Determinant Sampling	31
Data collection.....	31
Procedures	32
Zoom Web Conferencing for Data Collection	33
Pre-interview Protocol.....	33
Think-aloud Protocol.....	34
Interview Questions.....	36
Probing Questions	37
Data Analysis	37
Combining Data Collection Methods.....	38
Constant Comparative Analysis	39
Coding	39

Open Coding.....	40
Axial Coding	42
Selective Coding.....	43
Generating Theory.....	43
Methods for Verification and Saturation.....	44
Chapter Four Theory and Findings.....	46
Developing a Grounded Theory as a Process.....	46
The Theory	46
What is the process by which human search and discover information?.....	48
Concept of Starting a Search	49
Concept of Refocusing Searches	50
Concept of a Finished or Complete Search	51
Concept of Importance	52
Category of Establishing Accurate Information Need	53
What is the process by which search and discovery interfaces and tools support the modern search process?.....	55
Concept of Ease of Use	56
Concept of Familiarity.....	56
Concept of Analyzing Results	57
Category of Verifying Search Results.....	58
Chapter Five Conclusion.....	60
The Emerging Quality of Validating Information in a Search.....	60
Return to the Research Questions.....	61
Establishing an Accurate Information Need.....	61
Verifying Information Results.....	62
Discussion	62
Limitations	64
Future Directions.....	65
Final Thoughts.....	66
Reference	68
Appendices.....	75
Appendix A	75
Appendix B	79
Appendix C	81
Appendix D	84

List of Tables

Chapter One.....	1
Table 1.....	10
Table 2.....	11
Chapter Three.....	28
Table 3.....	38
Table 4.....	41
Chapter Four.....	46
Table 5.....	53
Table 6.....	54
Table 7.....	57
Table 8.....	58

Chapter One Introduction

Background

Searching for information, sense-making, and retrieving information are unique human processes. Our searches may come from simple curiosities about the world around us or grow to be much more complex, but they are bound by the searcher's search process and the tools they select to support their search. As the field of information retrieval grows to be multi-disciplinary, many scholars have theorized and developed frameworks that cover different stages, tasks, or steps in the search process. These frameworks and theories cover the information-seeking process from web-based perspectives, design principles that are applied to search user interfaces, and frameworks to make information and documents more findable or discoverable by users. While these frameworks, theories, and design principles use the search process as effectiveness assessments, they only apply to specific steps or interactions in a user's more extensive search process. Few of these theories and frameworks start from and center the search process, and few have studied the modern search process (Dervin, 2004).

Information seeking and search have been theorized about in information science for many years. Over time these models have been tested using humans and protocol

analyses to assess their accuracy given new technology or changes to information literacy (Hyldegard, 2006). With constant change and technological advances, other procedural or staged models have been theorized and again tested by users (Bokhari, Adhami, & Ahmad, 2021; Duarte, Oliveira, Cogo, & Pereira, 2015; Huurdeman & Kamps, 2016). Usability assessments often test models and theorized search procedures for precision, accuracy, and satisfaction. These models start from a search tool or theorized model, not a study of the modern search process and user's information needs. This grounded theory builds a deeper understanding of the search process from the perspective of the searcher and then considers how search tools might support the searcher's experiences and needs.

Modern Need for Information

The vast amount of information available over the internet makes access and a user's ability to seek answers to questions or discover new information incredibly easy. Network growth through recent years reveals our continuous search for information (Dos Santos Pergentino, 2020). With the urge to consume information, search engine popularity rises, becoming our browser's primary function when typing in the navigation bar. Over the past 20 years, information specialists have tested different approaches to making digital objects and information more findable and discoverable with technology-centered factors like advanced search features and facets. Still, these solutions are insufficient without understanding the modern search process from a searcher's perspective (Cleverley, Burnett, 2015). Advanced search filters may go unused as they

do not support the searcher's tasks of asking questions, crafting key terms, analyzing search results, or refocusing searches when results do not match their information needs.

Wikipedia, one of the most popular websites for finding information, developed an initial interface with interactive information retrieval features covering the aspects of the search box, categories, navigation, layout, and search result views but ultimately modified that design based on a set of human tests. Randomly selected subjects tested the interface in the scenario of simulated search tasks and provided their feedback via a post-task questionnaire. This human-based insight led to the redesign of that webpage to meet users' search tasks and preferences (Sandhu & Liu, 2014). This design example demonstrates the need to understand the modern search process before designing the technology to support that process.

The Gap in the Literature

Theories grounded in the modern search experience and search process should be the starting place for any information behavior or search interface design project. The information needs of today's adults can lead them to search using many different search engines and interfaces. Users may interact with an artificial intelligence tool like voice assistants or pose questions to a chatbot on a manufacturer's website. While this technology is prevalent today, these technologies were only just being developed when Barbara Dervin was writing about sense-making, and Russell-Rose & Tate were writing about the search experience (2004, 2012). More recently, Information scholars have theorized and tested new models for search processes like the DICO or BRIDGES

models (Cleverley & Burnett, 2015; Duarte et al., 2015). These models begin to theorize about how users may pursue information but are limited to specific audiences and professional communities with particular user characteristics. Theory grounded in general search users and individuals who support modern searches may bring more applicable theory when designing search tools and search literacy curriculum.

Information retrieval is a multidisciplinary field. This field draws from many industries and disciplines today. Human-centered design, people-centered design, and socially-centered design are approaches to product, UI/UX, and services where individuals or societies are centered when making design decisions. These lead to focusing decisions on human needs, societal needs, and the individual's journey through the process, product, UI, or service. These design considerations apply to interfaces, interactions, or experiences that represent only a step or stage of a human-centered search process and are often tested by humans in the context of a full search. While these principles may guide individual interface and interaction designs, they may not represent the entire human search process.

The Process of Search and Discovery

The field of information retrieval started in the library and information science discipline but has grown to become a multidisciplinary domain covering information science, computer science, social sciences, and medicine. Many of these approaches are grounded in the industry's technology, like health informatics or data science, or grounded in making digital objects or information more findable by expanding metadata.

In contrast, other approaches focus on user interface design or advanced search features to guide the search query. When these search solutions are developed, they often start with a needs analysis or a known design framework. They are then tested with human subjects to validate their search interface or interaction design. While these search solutions are designed outside the process and refined after human testing, other information retrieval experts centralize the search process.

In Hurdeman and Kamps' (2016) "Supporting the Process: Adapting Search Systems to Search Stages," the search process was theorized, and many recommendations were made based on the stage of the process(pp. 394-403). The more complex tasks performed using search engines, involving learning and construction, may have multiple stages, potentially affecting searchers' feelings, thoughts, and actions. While the Hurdeman and Kamp's model posited that the search process might affect human emotions and thoughts, other information needs models found that human characteristics affected the process (Johnson & Finn, 2017). These studies and theorized search process models present interesting findings of how influential the human search process might be on the search tools and interfaces used to complete a search. These studies come closest to identifying theories of the searcher's experience for searching and how the process, user, and tool may impact the search but are not grounded in the modern search process using modern search tools.

Studying the Process with Grounded Theory

This study shares insight into the modern search process as humans experience the phenomena of completing a search. Grounded theory is appropriate when studying social interactions or experiences and aims to explain a process, not to test or verify an existing theory (Lingard et al., 2008). In today's modern world, searches begin with curiosity, question, or need and develop further using technology, discovery tools, search engines, and possibly additional fact-checking or other deeper searches. This emerging theory brings integral insight into the modern search experience and search process with select concepts of tool familiarity in design and results, the user's ability to verify information, the need for accurate questions or key terms, and refocusing searches when results do not match the need. This emerging theory posits that, in an information-rich environment like the modern internet, the modern search process emphasizes verifying the information used in the search by validating results and establishing an accurate information need.

Studying the Process with Authentic Exposure

There are many ways to study human processes and develop theories of how they might work. In user interface design and user experience design, protocol analyses are often used as human testing. These protocols expose researchers to the user's processes which provides the opportunity to make observations. One common protocol analysis is the think-aloud protocol, where users verbally share their thoughts during tasks. The researchers collect observations and take notes on the user's actions and thinking as they

complete a process. Think-aloud protocols and usability tests allow researchers to gather observational data on human processes. This data collection method and recording of the participant through zoom provided authentic and reliable observational data on the search process.

A Modern Understanding

The observational data from the think-aloud search activity and interview data provided a thorough amount of information that was systematically analyzed, coded using Nvivo, and then developed into the grounded theory. The emerging theory developed through this study is the importance of validating information during the search process as 1) establishing accurate information needs and 2) verifying information results. This study's combination of data provided a needed picture of the modern search process by including diverse participants with varying levels of search skills while providing participant preference for search tools. The comparative analysis, resulting theory, and findings could inform search interface and search tool design or develop a modern information literacy curriculum to reflect the importance of validating information during the search process.

Chapter Two Literature Review

The State of Internet-based Searching

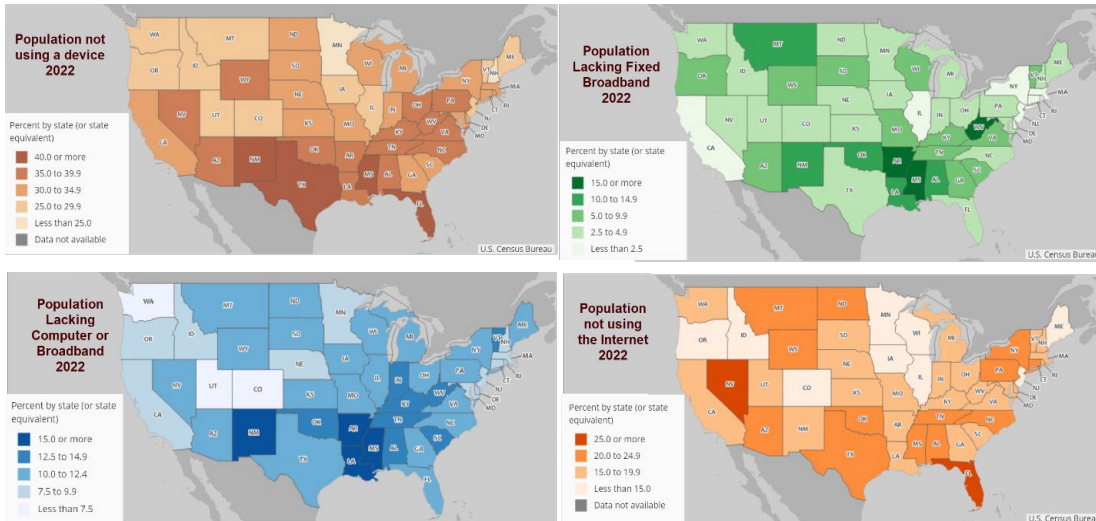
The COVID pandemic of 2020 brought added attention to a national conversation on the digital divide between the US population with access to reliable internet and a reliable device and those without access to the internet, devices, or needed network infrastructure. This conversation is particularly relevant to the modern search process when we look at studies evaluating the tasks we use the internet for, like searching. Several studies by the Pew Internet and American Life Project indicate that search engine use has become one of the most common online activities (Fallows, 2005, 2008). Access to technology and internet infrastructure, along with domain-specific or digital literacy, can change a person's search process or affect the tools they use to complete their search tasks.

The Digital Search Divide as Access

Data collected by the National Telecommunications and Information Administration of the United States Department of Commerce presents a picture of the change and inequalities across the US (2022). Figure 1 below presents four geographic maps covering the percentage of State populations that do not have devices, access to fixed broadband, lack computer access at home, and are not internet users.

Figure 1

2022 NTIA Data on Computer, Device, Broadband, and Internet Access in 48 States



Note. 2022 NTIA Data on Computer, Device, Broadband, and Internet Access in 48 States.

The differing capabilities to access internet infrastructure and tools are an important factor as we continue to develop an understanding of the modern US adult population’s search process. We can see that 16 states have 35% or more of the population without a device, 12 states have populations of 12.5% or more that lack a computer, and only nine states have less than 15% of their population that do not use the internet at all (National Telecommunications and Information Administration, 2022). The following table provides some details and highlights the change from 2019 to 2022 during the Pandemic (see table 1).

Table 1

2022 Main Reason for Household Not Online at Home

Main Reason for Household Not Online at Home	2019(%)	2022(%)	Change (%)
Can Use Elsewhere	0.029362	0.003824	-0.02554
Don't Need It or Not Interested	0.599748	0.104393	-0.49536
No/Inadequate Computer	0.029129	0.004909	-0.02422
Not Available in Area	0.032441	0.007029	-0.02541
Privacy or Security Concerns	0.020019	0.004423	-0.0156
Too Expensive	0.18756	0.032942	-0.15462

Note. Reasons for no internet at home. NTIA 2022 US Internet Use Survey Data

The final column shows the change in these proportions through the pandemic. The negative change shows a reduction in reasons for no internet access at home. These proportions of the US population dropped significantly from 2019 to 2022 as access increased and services and infrastructure reached more of the US. While these proportions show increased access to the internet and devices, it also presented a need to understand alternative methods of accessing the internet for a search, like mobile use and alternate locations like libraries, internet cafes, and more. The following NTIA 2022 survey data presents more nuance to the US population accessing the internet or internet-based search (see table 2).

Table 2

2022 Location of Access to the Internet

US Population - Location of Access to the Internet (%)	2022 (%)
Uses Mobile data plan from Any Location	0.741438
Uses the Internet While Traveling Between Places	0.446523
Uses the Internet at Home	0.756437
Uses the Internet at School	0.154658
Uses the Internet at Someone Else's Home	0.230385
Uses the Internet at Work	0.294678
Uses the Internet at a Coffee Shop or Other Business	0.119889

Note. NTIA 2022 US Internet Use Survey Data. Location of Access to the Internet.

Seventy-four percent of the US population has access to a mobile device and data plan, while 75% have access at home. These data points begin to paint a unique and diverse internet access picture across the US and beg the following questions. What does the US adult population’s search process look like when not all searches are done on a PC? How can search tools support these varying search contexts?

Digital Search Literacy

In addition to access, we should consider humans and their ability to use the internet and information. Many studies have examined information retrieval and search processes of specific industries with populations with similar characteristics. Cleverley and Burnett’s (2015) proposed BRIDGE model for information retrieval was initially tested on and designed using industry membership content and then tested with 54 practicing petroleum engineers from 32 organizations (p.97). While this model could be

used and tested on other populations, considerations for a more diverse population with digital skills and access are needed.

Health informatics has long studied information retrieval as health care providers have needed precise returns when searching on the job. The newest models propose a machine-learning search process that could support the industry's improved search process and precision (Demo and Sedig, 2021). When considering this model for larger, more public search process considerations, some variability on what technology or tools are available and what level of ability the user has become a potential challenge to the model.

Outside of domain- or industry-specific populations, search has been studied as a task and part of several different literacies; information, media, or digital. A considerable increase in users' abilities to search for, produce and disseminate information has increased the focus and emergence of these types of literacies (Wuyckens, Landry & Fastrez, 2022). These literacies or skills have been studied and defined in many ways. A recent meta-analysis provides a clearer definition for each of these literacies that have implications for search. Wuyckens et al. (2022) define information literacy as,

“Thus, information literacy generally pertains to the acquisition of certain abilities associated with using information search tools (technological or otherwise), knowledge of the search process, and the ability to create, evaluate and share information. In contrast, the abilities considered to constitute media literacy primarily deal with the concept of media, which is associated with issues of access, comprehension, analysis, and creation. As a concept, media literacy emerges from a different tradition than the one that gave rise to the concept of digital literacy, originally anchored in computer science. The latter concept first focused on basic technical

competencies in digital technologies, then gradually expanded to include a much more extensive set of abilities deemed essential to societal integration” (p.173)

These definitions blend when individuals use different technologies or devices to complete search tasks where the item may be information, visuals, or data. These skills may also differ between industries, communities, and levels of education.

The search process of the US adult population may look very different depending on access to the internet and infrastructure, availability of technology, or a population’s literacy skills. These variabilities should be considered and studied because they may affect the human-salient design of a search interface or the discoverability of information, objects, or documents.

Information Behavior and Information Retrieval

Today there are many information behaviors used to search out information. A major step in addressing an information need is deciding which tools to use and which collections to search over, a process that is sometimes referred to as source selection. Many source selection choices include phoning, emailing, texting a friend, reaching for a physical book, going to a physical library, sitting down at a networked computer, and starting a Web browser (Hearst, 2009, p. 174). In 2022, there are even more ways to search for information, including intelligent voice assistants (IVA) and mobile devices as powerful as PCs. Newer technologies like augmented reality (AR) and virtual reality (VR) are beginning to push information search behavior into three-dimensional space, with humans needing to understand and discover meaningful gestures, an evolving field

called guessability. Guessability is essential for symbolic input, in which users enter gestures or keywords to indicate characters or commands or rely on labels or icons to access features (Wobbrock, Aung, Rothrock, & Myers, 2005). With all these new modes and mediums for carrying out search process tasks, a review of information search behaviors, information retrieval, complex search processes, and information search processes should explore this new and more connected information world.

Information Retrieval

Information Retrieval is the tools and frameworks that include search, indexing, analysis, and evaluation for applications such as the web, social and streaming media, recommender systems, and text archives. This includes research on human factors in search, bridging artificial intelligence and information retrieval, and domain-specific search applications (Information Retrieval Journal, n.d). The standard information retrieval (IR) model selects and ranks the relevant information with respect to a user's query. The classic example of these systems is the discovery tools used on library websites to query based on key terms and then return ranked lists of possible matches, often measured by precision and recall. Today, these systems are tasked by the human search process to include new and different forms of information like datasets and news sources.

The field of data science has led to the development of dataset search and retrieval that broadly encompasses frameworks, methods, and tools that help match a user's data need against a collection of datasets. Services like Github and open data policies have led

to more publicly available datasets for search and use (Chapman, Simperl, Koesten, Konstantinidis, Ibáñez, Kacprzak & Groth, 2020). News sources and information provided by the web are not new but have become more commonplace. Information retrieval systems provide a relevant query to a news search engine; the user returns a single news result page consisting of various news articles aggregated from thousands of online news sources available on the World Wide Web. The availability and use of major news search engines like Microsoft News, Google News, and Newslookup have created new modes for information retrieval and have found their way into the human search process (Bokhari, Adhami, & Ahmad, 2021).

Discoverability & Findability

Making information and digital objects findable and discoverable is a longstanding practice in the information science field. Discoverability is the ability of users to find and execute features through a user interface. Discoverability in academic library websites often provides advanced features with Boolean values (or not) to refine a search query. Findability includes the practice of refining and providing metadata for information, documents, and digital objects that can also interact with advanced discovery features. These extra data points are then used in search tool algorithms as key terms for finding those items. Discoverability is a recurrent problem with Voice User Interface (VUI) design. VUI makes it difficult for users to understand what commands are supported by a newly encountered system (Kirschthaler, Porcheron & Fischer, 2020). Users end up "guessing" commands that a system might support, often leading to

interpretation errors and frustration. One solution to this problem is to display contextually relevant command examples as users interact with a system (Srinivasan et al., 2019). Recently, discoverability has reached into today's multimodal and context-rich search processes by providing search solutions in verbal follow-up or gesture support in the modern multi-modal search context.

Human Search and Sense-making

Sensemaking or sense-making is the information behavior theory by which people give meaning to their collective experiences (Dervin, 2004). Sensemaking refers to an iterative process of formulating a conceptual representation of a large volume of information. Search plays only one part in this process; some sensemaking activities involve search throughout, while others consist of doing a batch of searches followed by a batch of analysis and synthesis (Dervin, 2004). Sensemaking is often applied to information-intensive tasks such as intelligence analysis, scientific research, and legal discovery (Hearst, 2009). In this theorized framework for sense-making, search tasks and search interfaces are designed to aide this cognitive process by providing precision of search results and maximum result recall. The human then makes sense of these results in their process. Sense-making inherently represents the human in search, emphasizing the interpretive and context-driven aspects of meaning-making when people interact with systems. It is not just that we seek information but also make sense of it.

The Digital Search Experience

Marti Hearst (2009) explains, “Search is a means towards some other end, rather than a goal in itself. When a person is looking for information, they are usually engaged in some larger task and do not want their flow of thought interrupted by an intrusive interface.”(p.18). This fundamental human-centered principle for designing a search user interface explains the role of a search user interface as one specific interaction in a larger search process.

Search user interfaces are designed with usability principles and some search-specific interface guidelines. Ben Shneiderman first listed these principles:

- Offer informative feedback.
- Support user control.
- Reduce short-term memory load.
- Provide shortcuts for skilled users.
- Reduce errors; offer simple error handling.
- Strive for consistency.
- Permit easy reversal of actions.
- Design for closure. (1997)

These principles include additional user-centered and process-oriented design principles that acknowledge the multiple tasks or steps humans may take in their search process.

While these principles improve guidance for search interface design, the interface is only one stage of the search process.

The Search Experience by Russel-Rose and Tate (2012) broadens the search interface set of considerations to include a much more human-centered framework.

“The most fundamental step is to recognize that the opinions are based on a set of assumptions—in particular, assumptions about who is doing the searching, what they are trying to achieve and under what circumstances, and how they are going about it. Each assumption corresponds to a separate dimension by which we can define the search experience”(p.2).

These three dimensions, user, context, and modes of search, include the process and the more modern and multi-modal reality of 2022. When humans have differing devices, web access, and literacies, their search process may look significantly different than others. Designing for these multimodal methods by including strategies for facets, mobile search, and social searching is a necessity today.

Design Principles

Within human search experiences, sense-making, and the search process, we interact with various platforms, browsers, digital interactions, and user interfaces. Hearst (2009) writes,

"There is no exact formula for producing a good user interface, but interface design indisputably requires the involvement of representative users. Before any design starts, prospective users should be interviewed or observed in field studies doing the tasks which the interface must support." (p.51)

Human-centered design, human factor design principles, socially-centered design, and user experience design principles cover ways that search and discovery specialists can build search interfaces and item discovery data with the user-centered. These frameworks and approaches apply human or user-centered principles to search interfaces and discovery tools to facilitate a more successful search interaction.

Human factor design principles refer to how users interact with systems, machines, platforms, or tasks. These design factors are internationally recognized through the ISO and Human-Computer Interactions (HCI) interdisciplinary field. These principles include physical ergonomics, which refers to the human body's responses to physical work demands; consistency in a system which is how it looks and works the same throughout; control of information by the user; and more, but do these principles make the experience more human?

Alongside human-centered or human-factor design principles is an emerging approach to design called socially-centered design. Socially-centered design or society-centered design holds principles aimed at framing the user as a beneficiary and not just a user (Gold, 2020). The guiding principles at play in socially-centered design, according to the Society-centered Manifesto, are

- Place care first
- Earn trust
- Empower collective agency
- Reimagine public value

- Design for people's rights
- Ensure fair and just oversight
- Redistribute the power of technology
- Create compassion at scale
- Design for regenerative action
- Confront uncertainty (Society Centered, 2020)

When designing search and discovery tools and interfaces with these principles, do users experience a more human search process? The potential for these principles is already being used to clean longstanding metadata structures that carry inappropriate references like 'alien' for immigrants, but have they been applied thoroughly enough to the search process overall?

These design principles are pervasive in product and UI design. Product creators look for ways to reduce the cognitive load on users by using empathy and human input in the design process. It is understood in the field that the implications for human factor-salient design can make or break a search and discovery process (Babich, 2020). This leads to focusing decisions on human needs, societal needs, and the individual's journey through the process, product, UI, or service.

Today, humans interact with technology and search tools in almost every context and with completely new types of interaction. Users may begin a search process or reach the stage where they interact with a search interface in their automobile. This incredibly unique use case highlights the need for human-centered considerations. Searching for an

item on an interface is a critical concern because a poorly designed interface that draws attention to unwanted information can extend drivers' search for items of interest and pull attention away from roadway events (Lee et al., 2016). Another very human search interface need is location-specific discovery while balancing personal privacy. For some users sharing their precise location supersedes their need for a location-based search. To serve this specific human need, models have been theorized to provide location-based services and privacy while balancing access under these distracted conditions (Dewri & Thurimella, 2016). These unique models apply human needs and theorized information retrieval processes to build search interfaces.

Another emerging search and discovery field is data science and data sets. Google recently beta-released a search service for datasets, which allows users to discover data stored in various online repositories via keyword queries. These developments foreshadow an emerging research field around dataset search or retrieval that broadly encompasses frameworks, methods, and tools that help match a user data need against a collection of datasets (Chapman et al., 2020). These searches for raw data are another unique example of making objects more discoverable by expanding discoverable metadata. This example of making datasets more discoverable may improve the initial search step, but how might this solution improve the search process at other stages?

Information Seeking as a Process

Marti Hearst (2009) describes the importance of ISPs and explains, "In order to design successful search user interfaces, it is necessary to understand the human

information-seeking process, including the strategies people employ when engaged in search” (p. 81). Several theories for information behavior have been proposed and theorized to address the process of searching.

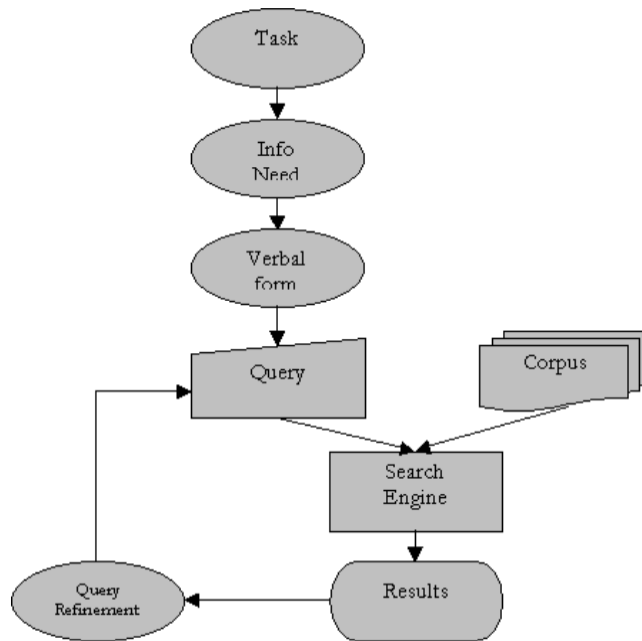
Theories of Information Behavior

The most common information-seeking process or ISP written and theorized about is a process where identifying an information need, followed by the activities of query specification, examination of retrieval results, and, if needed, reformulation of the query, repeating the cycle until a satisfactory result set is found (Hearst, 2009). This fairly simple process model only includes the cognitive steps involved in a search process and leaves out contextual information that might affect the search process. This theory was expanded upon in the late 1990s and early 2000s to include the role of the searcher’s knowledge, the system, the information collection, and searching in general (Hearst, 2009).

This holistic view was expanded upon further by Marchionini and White’s 2008 model. Marchionini and White (2008) theorized that the information-seeking process covered recognizing a need for information, accepting the challenge to take action to fulfill the need, formulating the problem, Expressing the information need in a search system, examination of the results, reformulation of the problem and its expression, and use of the results (p. 208). This model, which centers more around search as a set of tasks, has been theorized and applied further when studying information retrieval from a non-human perspective. Figure 2 shows this web-based approach.

Figure 2

Broder's Information Seeking Process



Note. Broder's Standard Information-Seeking Process

Broder (2002) described this web-oriented view of the process: "In the web context, the "need behind the query" is often not informational in nature. We classify web queries according to their intent into three classes:

1. Navigational. The immediate intent is to reach a particular site.
2. Informational. The intent is to acquire some information assumed to be present on one or more web pages.

3. Transactional. The intent is to perform some web-mediated activity.” (p.5).

While this web-oriented process provides some clear guidance when designing search query tools, these three intents do not represent the entire user’s information needs or search goals. These intents operate without knowledge of the context in which the user is undertaking the search process. Additional methods of evaluating the process are needed to fully understand the human need, information needs, and the search process.

Measuring Process with Usability Testing

The measures of a good usability product, service, or interface are learnability, the efficiency of use, memorability, subjective satisfaction, and non-catastrophic errors (Nielsen, 1993). Protocol analyses are used to measure these through a process that might include multiple browser windows, devices, or steps. Protocol analysis is a psychological research method that elicits verbal reports from research participants. It has found application in the design of surveys and interviews (Sudman, Bradburn & Schwarz, 1996), usability testing (Henderson, Smith, Podd & Varela-Alvarez, 1995), and design research (Gero & McNeill, 1998). Protocol analysis has even been used to describe and theorize about the design process itself (Gero & McNeill, 1998).

One of the most widely used methods for evaluating the usability of a product, software, or system is a think-aloud protocol (Doi, 2021). This method helps to discover usability problems and to examine improvements because it provides direct information on a user’s thinking and cognitive processes, which are challenging to examine through

mere observation or questionnaire surveys. This method was first proposed in 1980 by Ericsson and Simon but was not introduced into the usability field until 1982 (Lewis, 1982).

While this method does lend itself to collecting data on processes that go beyond usability, there are some limitations to consider (Fan, Shi & Truong, 2020). Data analysis of think-aloud protocols can be challenging. One proposed solution to better support the use of verbal data, such as utterances obtained via think-aloud and notetaking, was to interpret it as qualitative data. Qualitative data analysis methods used in qualitative research in the social sciences, like grounded theory, can provide analysis methods that lead to more efficient and meaningful think-aloud protocol results (Doi, 2021).

A Qualitative Approach to Studying Human Process

There is reason to believe that the human search process has changed today. With more emphasis on testing search user interfaces and user interactions with humans than starting from the human process, search may be designed more for the technology than the humans using them. Studying human processes is not new in the qualitative research field.

Grounded theory research design explores processes, interactions, or experiences and forms a theory for those experiences. The researcher focuses on a process or action with distinct steps or phases over time. Thus, a grounded theory study has “movement” or some action the researcher attempts to explain. A process might be “developing a general

education program” or the process of “supporting faculty to become good researchers” (Creswell, 2019). Researchers draw from notes and memos to build categories that connect. With the ultimate aim of developing a theory of this process or action. Grounded theory methodology is often used when little is known about a subject or problem area, allowing a theory to emerge.

Grounded theory has been used in the user experience field to analyze experiences and processes through different websites and digital services. Broberg studied the impact of design and usability guidelines on several tribal websites (2011), and Young et al. used grounded theory to evaluate the usability of restaurant websites (2007). Grounded theory was also used to study the perception of usability characteristics in smartphone apps and software (Lee et al., 2012). Grounded theory centers on the process to best understand a more diverse user base. Although a grounded theory researcher develops a theory by examining many individuals who share the same process, action, or interaction, the study participants are not likely to be located in the same place or interact on so frequent a basis that they develop shared patterns of behavior, beliefs, and language (Creswell, 2019). This qualitative research method approach lends itself to developing a theory for the process of searching for information that applies beyond select samples or specific tools.

Problem Statement

The state of internet-based search has changed with the variability of needed technology, access, and literacies. New modes of information seeking and queries like intelligent virtual assistants and new information formats like datasets have also changed

the information need for which humans search. While designers are working to implement human-centered, user-centered, and socially-centered design principles into the stages, interactions, and interfaces of our search process, the search process itself needs to be re-evaluated and explored now that the contexts and goals of our searches have changed. These significant changes to the human search context and the lack of current literature leave a rich space of inquiry to explore. What is the nature of the modern search process? What can our current research methods and usability assessments tell us about the search and discovery process? How can a modern theory of the search process improve the design of search interfaces, experiences, and interactions?

Chapter Three Methods

Restatement of purpose

This study aims to explore the search process for adult populations in the United States that have experienced using a search interface to find specific information related to their life. This study used the systematic procedures approach to grounded theory by collecting data from interviews and observed search activities using think-aloud protocols to generate a theory that explores the modern search process.

Research Approach

The grounded theory study approach used in this study is the systematic procedures approach of Strauss and Corbin (2008). With grounded theory, the selection of subjects, data collection, and analysis are linked from the beginning of the research, proceeding in parallel and interacting continuously. The research design uses data collection which is immediately followed by analysis. Practical implications must be considered, given the close relationship between data collection and analysis. It is important that researchers not be over-enthusiastic about data collection and conduct a series of interviews without concurrent analysis, possibly resulting in a missed opportunity to sample based on emerging concepts (Duffy, Ferguson, Watson, 2004)

This research design approach was selected because of the complexity of the research questions and the unknown nature of how impactful search literacy, modern technology advancement, and search tools are on search and discovery. Grounded theory, by nature, provides a more pragmatic or practical discussion once the culminating substantive-level theory has been identified. Qualitative interviewing is not only focused on capturing experiences and perspectives from participants, but it also allows for the interpretation and, in some cases, explanation of those experiences—in relation to each other and the contexts in which they arise (Edwards & Holland, 2013). This interviewing practice and observational data generated off the think-aloud protocol improved the concurrent analysis by providing more data to code and leading to quicker or more reliable coding and the emergence of categories. By sequencing the think-aloud search activity before the interview protocol, the researcher was able to ask follow-up questions, and additional interview questions created the opportunity for a connection between the observational data and specific interview questions. The semi-structured interview approach permitted space for the researcher to ask clarifying questions to participants based on observations.

Research Question

1. What is the process by which humans search and discover information?
2. What is the process by which search and discovery interfaces and tools support the modern search process?

Sampling

This study recruited 20 participants to complete an interview and a think-aloud search activity. The inclusion criteria required that participants:

- Recently completed a search experience or were employed as search specialists in roles like Reference Librarians
- Be over the age of 18
- Currently live in the US

The final sample for the study included eight participants with roles or recent experience in reference librarianship and twelve participants categorized as general searchers. This broad group of individuals and perspectives provided a rich theoretical sampling based on various search experiences and expertise (Creswell, 2019). These participants reached out to the researcher and were subsequently recruited through communication and messaging posted to professional library listservs, professional-oriented social media, and individuals known to the researcher that fit the inclusion criteria.

The sample also included a diverse use of search tools, including the Google search engine, Wikipedia, government websites, data set repositories like Kaggle.com, and applications on mobile devices like voice search. Technological devices used during the think-aloud protocol were predominantly laptops and PCs, with sixteen participants using them to complete the think-aloud activity. The other four participants used mobile devices or tablets to complete their search activities.

Theoretical & Determinant Sampling

Theoretical sampling is commonly understood as a means for sampling participants with a set of theoretical considerations in mind. However, theoretical sampling also occurs in the actual data-generating process, for example, through interviewing (Foley et al., 2021). Theoretical sampling across the two participant criteria complemented the research design and acted to reinforce and check emerging categories, memos, and axial coding. A discriminant-like sampling strategy was used to overcome the common grounded theory challenge of theoretical ideas or notions leading the systematic analysis and limiting the analytic, substantive theories and open categories to emerge (Creswell, 2019). A determinant sampling strategy is where researchers gather information from individuals different from the initial interviewees to determine if the theory holds.

For this study, an alternation between expert and general search user participants with differing search expertise provided the opportunity to verify the theory and core concepts. The order of interviews alternated by interviewing 1-2 general search users and then interviewing a search expert. This approach built confidence in the resulting grounded theory and analysis, which led to a faster saturation.

Data collection

This study used two data collection methods, a semi-structured interview protocol and a think-aloud protocol analysis. Both methods occurred in one scheduled meeting with participants. The researcher collected twenty think-aloud search activities from the

twenty participants and completed twenty interviews for forty data collection events. The observations from think-aloud activities, interview transcripts, and memos were analyzed and coded using Nvivo 20.

Procedures

The Zoom web conferencing tool was used to facilitate the interviews and think-aloud search activities, and the events were recorded. Recording the data collection provided an opportunity to review coding and findings, which improved reliability. Also, participants could complete the interview and data collection from their own spaces using their own devices via web conferencing. This was important for two reasons, first, to meet the participant in their own space and second, to let them share their authentic story via the tools and context they would normally use to complete searches. Prior to the data collection, two pilot interviews and think-aloud activities were completed. This pilot provided insight into interview protocol issues. The pilot also provided opportunity to practice interviewing and observing activities through the web conferencing tool.

The informed consent form collected consent for recording, and participant comprehension of the recording request was confirmed during the pre-interview stage. The data collection event began with a pre-interview protocol followed by the think-aloud search activity and the semi-structured interview. Interviewing is the most frequently deployed data collection method in grounded theory (Charmaz & Belgrave, 2012), either alone or, in this study, with other data collection methods like observations

(Bryant & Charmaz, 2019; Charmaz, 2014; Corbin & Strauss, 2008). For this study, a think-aloud protocol was used to observe a search activity.

Zoom Web Conferencing for Data Collection

The pre-interview protocol, think-aloud search activity, and interviews were carried out over Zoom, a web conferencing tool. Participants agreed to the web conferencing medium when signing the consent form and agreed to the recording of the interview portion and think-aloud search activity. Using a web conferencing tool benefitted both the researcher and the participants. Participants were interviewed where they felt most comfortable and using the same technology that they authentically completed searches. Additionally, the researcher was able to record the activities which improved reliability of the data analysis. The researcher was able to review and check the recordings at later stages for focal categories and the emergent theory.

Pre-interview Protocol

Pre-interview information and procedures were provided before the interview began and then reiterated at the very beginning of each interview. This step included confirmation of consent forms and an explanation of data treatment and confidentiality. The consent form and pre-interview procedure were where the researcher gathered needed information to secure informed consent, provide assurances of confidentiality, and confirm participant understanding of the study. This section also included three brief screening questions aimed at ruling out vulnerable populations and ensuring that

participants fit the inclusion criteria. The applicable Institutional Review Board approved this protocol design.

The pre-interview protocol was also used to address ethical considerations like autonomy and language justice. According to the *Sage Handbook of Qualitative Research Ethics*, the following areas of potential ethical issues are addressed in this research design (Olsen, 2008). Autonomy refers to the research participant's right to choose whether to participate in research based on the information provided. Autonomy was communicated in the study at several stages, including the pre-interview procedures and the IRB-approved communication messages that include details of participants' rights to refuse participation. Additionally, the nature of think-aloud protocols required that the informed consent include language on the potential stress of completing a process under observation. Finally, the researcher applied and supported ethics of justice and language justice, including definitions of technical terms and language in the pre-interview procedure and providing space for questions about the interview and study in that step (see appendix A).

Think-aloud Protocol

The study of thinking-aloud protocols has a long tradition in cognitive psychology, the field of education, and the industrial-organizational context (Guss, 2018). The think-aloud protocol comes from the protocol analysis family of research methods. Protocol analysis is a psychological research method that elicits verbal reports from research participants. It has found application in the design of surveys and interviews

(Sudman, Bradburn & Schwarz, 1996), usability testing (Henderson, Smith, Podd, & Varela-Alvarez, 1995), and design research (Gero & McNeill, 1998).

Think-aloud protocols have been used in qualitative research but come with some considerations. Charters(2010) explained that requesting verbal statements from a participant during a normally non-verbal task may change the participants thinking, but conclude that even if their view of thought processes is necessarily incomplete, verbal reports such as those from think-aloud data are a “thoroughly reliable” sources of information about thought processes. Before designing a think-aloud method, the researcher needs to decide on the type and level of difficulty of the research task, the degree of prompting which is appropriate, the use of other data to support inferences from think-aloud protocols, and the method of analysis (p.71).

For this specific study, a concurrent think-aloud protocol was selected. The participants shared their thinking, and their search process was observed as they moved through the search and discovery task (see Appendix B). The researcher collected observational notes on participant search behavior and noted their verbal utterances. The use of Zoom web conferencing and recording the activities supported accurate observational notetaking while recording verbal utterances through transcription. These observations were included in the coding and data analysis stage.

The think-aloud task was designed with four information retrieval topics to choose from, ensuring that the participant selected a topic to search for that is unfamiliar

(see appendix B). Participants were able to choose a device and search tool of their choosing to complete their search activity.

Pallab and Kuusela (2000) compared the use of concurrent and retrospective note-taking during a protocol analysis and found that,

“In general, the concurrent protocol analysis method outperformed the retrospective method not only was the number of concurrent protocol segments listed higher than that of retrospective protocol segments, but concurrent data provided more insights into the decision-making steps occurring between stimulus introduction and the final choice outcome”(p.397).

A concurrent note-taking protocol analysis was undertaken to maximize the amount of data to code openly through the think-aloud protocol, which led to memos and observational notes.

Interview Questions

The interview questions were designed to address research questions (Table 3) and provide the opportunity for conversation. This was done by employing unstructured or lightly structured interview guides consisting of open-ended questions. Remaining flexible throughout the data collection in order to be able to capture data in a maximally open way can generate more data (Timonen et al., 2018). The interview questions reflect some concepts of the information-seeking processes but are in no way limiting. Additionally, these interview questions are open-ended enough to interpret differently by different participant search experiences. For example, search specialists like reference

librarians could elaborate on supporting patrons just the same as a general search user was able to share stories of searching for everyday information.

Probing Questions

Probing questions were used to have participants clarify the process. The interviewer is to varying degrees a “co-constructor” of the data because the researcher has the capacity to deploy knowledge and insight to sensitize concepts in data and to steer the course of the inquiry (Timonen et al., 2018). Probing questions used to have participants expand on codes and concepts that were developing later in the study or to elaborate further on observations from the participant's think-aloud were employed. At other times, probing questions were used to validate a code or concept in participants' interview responses.

Data Analysis

This study used data from transcripts of the semi-structured interviews and observational notes from the think-aloud protocol in the constant comparative analysis approach to grounded theory. The combination of data collection tools, observation, and transcripts are not new to grounded theory studies. Grounded theory is best utilized when a large amount of qualitative data has formed in unpredictable ways. The researcher creates meaning of the data through systematic analysis. (Lawrence & Tar, 2013). By using observational data and interview data, the potential for drawing connections and capturing emerging categories is increased, and the verification of saturated categories

can occur faster when collecting data from a more sense-rich data collection event. Additionally, the interviews were recorded through the Zoom Web conferencing tool. These recordings and transcripts improved the reliability of the coding and analyses (see Appendix D).

Combining Data Collection Methods

The following table aligns the research questions with the data collection methods. This intentional design supports the validity and reliability of the study’s exploration of the research questions. Interview questions(I#) and observational data(O) helped to draw connections and comparisons during the analysis stage. According to Anfara et al., (2002) “keeping in mind that research questions provide the scaffolding for the investigation and the cornerstone of analysis” (p.30).

Table 3

Research Question associated with Data Collection Method

#	RESEARCH QUESTION	INTERVIEW QUESTION (I#) OR THINK-ALoud OBSERVATION (O)
1	What is the process by which humans search and discover information?	I1, I2, I3, I6, I7, O
2	What is the process by which search and discovery interfaces and tools support the modern search process?	I4, I5, I8, I9, O

TABLE 3. Documentation of Research Question to Data Collection Method

Constant Comparative Analysis

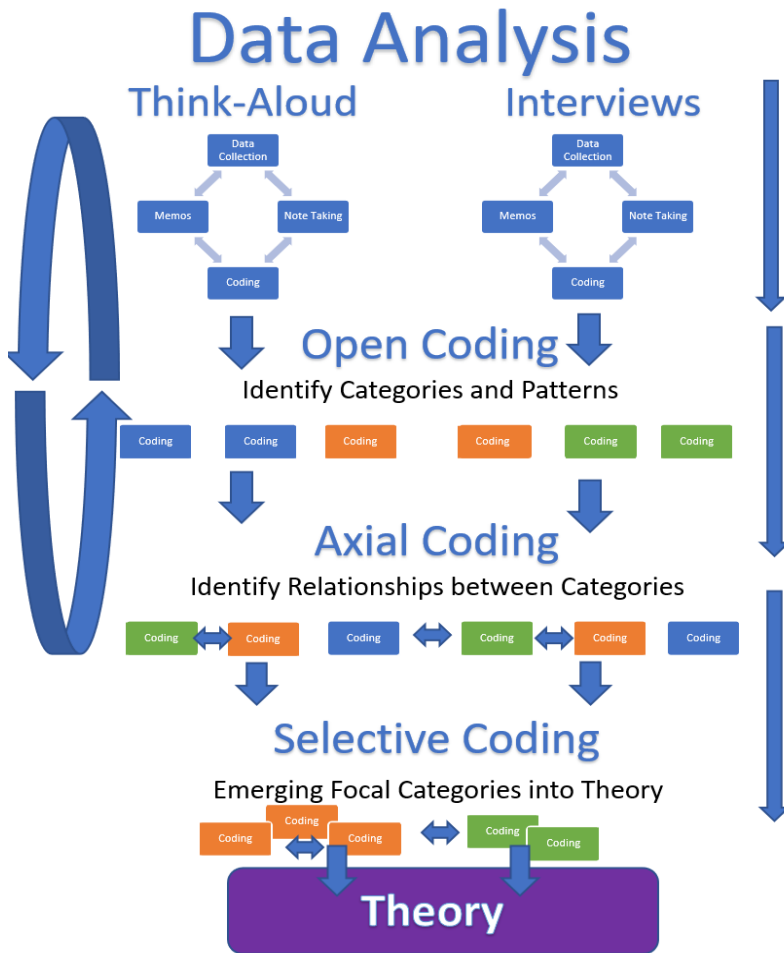
This study used a constant comparative analysis approach. In grounded-theory research, the systematic methodology for analyzing data is called constant comparative analysis. Researchers are constantly and from the beginning of data collection, analyzing, notetaking, and coding. Hallberg (2006) describes the comparative as “that every part of data, i.e., emerging codes, categories, properties, and dimensions as well as different parts of the data, are constantly compared with all other parts of the data to explore variations, similarities and differences in data”(p. 143). In between each interview, the researcher would write down memos, open code, and axial code the transcripts. This core concept of grounded theory methodologically guided this thesis to its theory.

Coding

Coding as a three-phased approach is recommended by Corbin and Strauss (Strauss, 1987; Corbin & Strauss, 1998). The three coding stages are open coding, axial coding, and selective coding. Figure 3 presents the coding process and constant comparative analysis approach.

Figure 3

Systematic, Constant Comparative Data Analysis



Note. Open coding, axial coding, and selective coding leading to a theory

Open Coding

Open coding entails a close examination of the data, breaking it down into parts, making comparisons, and questioning. The indicators are identified bits of data collected and data resulting from breaking down the data. Open coding uses “indicators”—words,

phrases, sentences, paragraphs, statements from the data, or entire observations—to develop “concepts.” Indicators are constantly compared as the researcher works to identify new insights until theoretical saturation is reached (Corbin, 2008). The researcher is looking for terms study subjects use—terms that can be thoroughly and “minutely” analyzed (Strauss, 1987, p. 31).

In this grounded theory study, transcripts and observational notes from the think-aloud search activities were analyzed and openly coded in Nvivo 20. The researcher used or identified twenty codes to draw comparisons and see connections between code concepts. The following table presents the set of codes that acted as the foundation for the grounded theory.

Table 4

The Frequency of Open Codes in Data

Code	Frequency of Reference	Code	Frequency of Reference
Complete Search	69	Ease of use	152
Difficulty	31	Familiarity	130
Importance	53	Filters and sorts	72
Search Expert	8	Google	97
Search Generalist	12	Ease of use	152
Search Literacy	66	Unwanted Information	22
Analyze Results	34		
Ask Questions	107		
Key Terms	76		
Refocus	21		
Specificity	13		
Verify	50		
Data	132		

Note. Frequency of Code Reference in Data

In addition to open coding, memo writing was used (see Appendix C). Charmaz (2006) describes this as

“Memo writing is the methodological link, the distillation process, through which the researcher transforms data into theory. In the memo-writing process, the researcher analytically interprets data. The Grounded Theorist discovers emergent social patterns by sorting, analyzing, and coding the ‘raw’ data in memos. It is the methodological practice of memo writing that roots the researcher in the analyses of the data while simultaneously increasing the level of abstraction of his/her analytical ideas” (p.163).

Thorough and detailed memos supported the development of categories and potential connections between categories. By writing memos continuously throughout the research process, the researcher explored, explicated, and theorized the emergent patterns, categories, and, ultimately, the theory (Charmaz, 2006). As readers will see in the results section of this study, the memos supported connections between selected concepts like key terms and asking questions that shared search process properties.

Axial Coding

Axial coding is further coding within a category, involving analysis of paradigm conditions and other subcategories relating to the category: who, what, when, where, and why (Strauss, 1987). Strauss(1987) explains that “Axial coding begins to link categories and will lead to finding core categories. Strauss recommended that beginning analysts use a coding paradigm, which is a reminder to code according to what he calls “paradigm items” such as conditions, consequences, relations among actors, and strategies” (p. 27). For this study, categories like establishing information needs or verifying search results had consistent occurrence with the core concepts of completeness of a search and the

concept of importance in a search. Identifying these shared conditional or consequential relations with other core concepts led to the substantive theory.

Selective Coding

The final form of coding was selective coding which, according to Strauss (1987), “pertains to coding systematically and concertedly for the core category” (p. 33). As the data collection events continued, the core categories defined by denser saturation became the primary coding process as the study moved to generate a theory. The interviews and observational data were then systematically reviewed and coded for the categories and core concepts that make up the theory. This study used codes supporting and sharing properties of the search process or the search context to develop the theory. Although some codes like ‘google’ or ‘data’ had very high frequencies, those codes did not hold a consistent property or definition and were not included in core categories.

Generating Theory

At this stage, axial coding, selective coding, and memos developed into the emergence of core concepts, categories, and theories (Corbin & Strauss, 2008). Corbin and Strauss add that researchers should write detailed and lengthy memos to help generate a more abstract theory. The most important part of this stage is to permit the emergence of a theory that comes from data, not preconceived notions or prior logic. Glaser and Strauss (10967) also noted the importance of emergence, explaining, "In

short, our focus on the emergence of categories solves the problems of fit, relevance, forcing, and richness.” The researcher seeks a “theory that ‘fits or works’”.

Methods for Verification and Saturation

Several aspects of this grounded theory study support verification and act as measurements for reaching saturation, including theoretical sampling with the intent to sample a diverse set of participants, multiple types of data collected, and concurrent analysis.

According to Glaser and Strauss (1967), “An important point about the generated concepts used for categorizing is that they must have two essential features: that they are “analytic” and “sensitizing” (p. 38). When the researcher has finished coding and identifying categories, categories need to be specific properties or characteristics and must be bound to the participants' personal experience. To ensure that categories are analytical and sensitizing, the interview protocol and observations from the think-aloud protocol overlapped to act as analytical verification. Additionally, probing questions helped sensitize responses by clarifying statements or asking for more response details.

Another strategy for verification is that the sample must be appropriate, consisting of participants who best represent or have knowledge of the research topic. The theoretical sampling and balance of search users and search specialists ensured an appropriate sample. Concurrent analysis creates a balance of ‘what is needed’ and “what is known.” This pacing and the iterative interaction between data and analysis is the

practice grounding this approach in reliability and validity (Morse et al., 2002). The use of a concurrent analysis also supported the verification of the theory and study findings.

Saturation is the point at which no additional data emerge to enable the researcher to develop further properties (Glaser and Strauss, 1967). The theoretical sampling used in this study helped reach saturation by leveraging a diverse sample. The multiple data collection methods provided no new data points that would add to the axial coding and properties or dimensions of the core categories and emerging theory.

Chapter Four Theory and Findings

Developing a Grounded Theory as a Process

This chapter contains the substantive theory and the study's findings, organized by the research questions and the systematic analysis process completed to explore those research questions. The findings begin with the selective codes, presented as concepts. Then the core categories that emerged from the selective codes acted as the core categories' properties and dimensions; finally, the core categories are described, which make up the major dimensions of the emerging theory.

To best ground the findings in participants' experiences as they complete internet-based searches, observations and transcript quotes were captured in memos and shared below to exemplify the concepts, categories, and theory. Additionally, a code's frequency or sum count across the 40 observational and interview data events is shared to highlight internal validity (see table 4). The transcripts were pulled from the Zoom recordings, which maintained the coding reliability as the researcher replayed recordings and read the transcripts when completing data analysis.

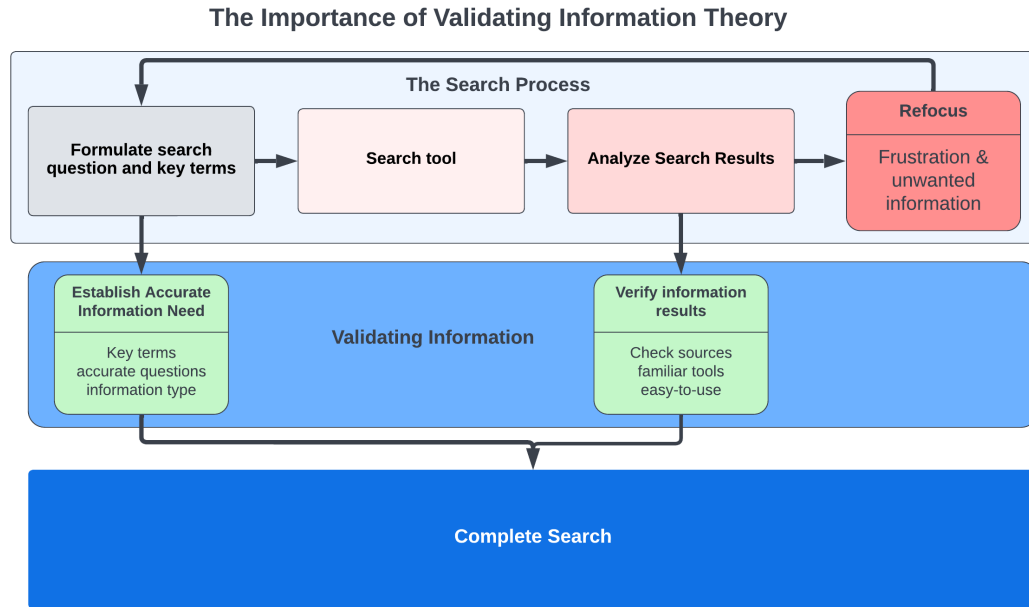
The Theory

A constant comparative analysis of the data led to the emergence of the theory of the importance of validating information during the search process as 1) establishing

accurate information needs and 2) verifying information results. This emergent theory states that modern search users experience a need to validate their search information during the search process, which presents itself in two notable ways. The first way to validate search information in a search process is to establish accurate information needs. Users require accurate key terms or word choices in their initial search inquiries and need an understanding of the information type or purpose to reach their search goals. The second way of validating search information in a search process is to verify information results. The information retrieved by a search tool can support the verification of information by presenting information in a familiar fashion or design, providing filters and facets to sort information by measures deemed relevant, and providing accurate information without unwanted information. Figure 4 shows the connection between the importance of validation and methods for validating search information in relation to a search process.

Figure 4

Validating Search Information Theory Diagram



Note. The diagram shows the relationship between the search process and the core categories to make up the complete theory.

To further unpack and exemplify this substantive theory, the research questions will be addressed by two core categories and selective concepts that act as properties and characteristics of the ultimate theory.

What is the process by which human search and discover information?

This important research question (R1) highlighted the complexity of modern search behaviors and presented some core concepts and categories that led to the importance of establishing accurate information needs as validating search information. Here is an example of the modern search process described by a participant: “typing in

my search term, whatever I am looking for, and then filtering through results”(Participant 20). This short but common description of steps in a search is more involved than we would initially think. The search terms must be somewhat established, and the filtering of results must be informed by some understanding of the information type or goal. It was participant responses like this that began to develop this study’s findings on emerging qualities of the modern search process.

Concept of Starting a Search

The process of starting a search occurred for all participants regardless of the search tool or device. Two selective concepts occurred as starting points for participants. The first was asking questions which was observed in all twenty of the interviews (see appendices C, Memo 1). “step one is what do you know already? Um, even if it that's just like the idea of it and then from there you just ask and ask and ask in different ways.” (participant 1). The second concept that started a search was formulating accurate key terms. Key term creation was observed in eighteen of the twenty think-aloud search activities and twelve inclusions in response to the question, “What is your process of searching or discovering information? Moreover, it was coded with frequency in response to the question, “What was most important to your search activity?” When considering the participant’s process and the context, these codes and concepts began to share the dimension of starting a search.

Concept of Refocusing Searches

The concept of refocusing a search using new key terms, previewing other search results, or moving a search to a different tool occurred in twelve think-aloud search activities, most commonly after an initial set of search results. Refocusing searches was described as restarting a search due to unwanted result information or inaccurate search results. “if I really wanted to pursue this on my own, I would have had to have reformulated. perhaps, as I said, look at one event and then added more detail” (participant 2). This concept exemplified that when there is a lack of established information need, the search would need to be reset or refocused (see Appendix C Memo 4). When a search lacked established information needs, then refocusing was common.

Concept of Difficulty

Searches were most commonly described as difficult or complex when the information type or specificity of information made the search feel difficult. This concept began to emerge as a property of a search process when establishing an information need was complicated by information type or specificity. Difficulty was coded in twelve interviews and eight think-aloud activities where the search activity requested specific information or information types like data.

“So, for example, I work at an engineering school and a lot of my students. They want to find information on these very complex processes and things like that. I don't necessarily understand um. And when I try to, you know, follow my usual search process to get more

information. It's extremely difficult to find it because it's just not something that the general public is really interested in, or maybe it's too new, and they just don't It's not more widespread.”

Another primary code that indicated complex searches was information type when searching for certain types of information like data. Participants described data as “A lot of times with data it's hard to find” or “I was first excited that it was data since I’m a data librarian, So it's like, Yes, I’m winning. Um. But Then I went to the census page, and I was like, oh, I'll look through all those tables and reports for the answer (specific data).” These two examples of selective codes occurred at the first stage of the search process.

Concept of a Finished or Complete Search

Question three of the interview protocol posed the question, “when is a search complete?” The resulting completeness code seemed at first to not be valuable. However, as participants began to elaborate on the search activity and clarify when a search was finished, this concept began to relate to establishing an accurate information need. For example, a participant states,

“find the information that I’m looking for it answers my question so on the hike that wouldn't have been complete yet until I found where it was on a map and found out how far it was and where we park, and then it would be done other times.” (Participant 10)

This example shows that establishing an accurate information need is integral to completing a search. Alternatively, when a lack of established information needs completeness, or searches are not perceived as complete, the search would have continued through refocusing the search.

Concept of Importance

Similar to the concept of a complete search, the interview question "what was most important to their search experience?" did not initially feel meaningful. However, as more interviews were collected, it became clear that importance was tied to accurate initial searches based on asking good questions or accurate key terms. Additionally, the property of importance and its coding occurred frequently with verifying results. This property of importance began to support the emergence of the core categories. For example, "Um, if I know that i'm looking for, say, a specific book, or I need good bibliographic data." This example demonstrates the connection between importance and an accurate understanding of the information need.

Table 5

Examples of Importance Coding

#	Transcript or Observation of Importance
1	“The most important thing is obtaining the information I was looking for.”
2	“I think it would be two things, getting the accuracy of the information and also timely.”
3	“That it will be useful and point me to a few trails to choose among that are where I want them to be. How long do I want them to be. How difficult I want them to be, et cetera,”
4	“What was it? Mostly for the things that come up first on Google, or were the most important for me, since I didn't know exactly what they knew I was looking for.”

Note. Excerpts from interview scripts that exemplify the importance concept.

Table 5 shows several examples of this concept's importance as it supports identifying the accuracy of information as important and core categories like verifying information results.

Category of Establishing Accurate Information Need

As the analysis continued, there became a clear connection between these core concepts of starting a search, asking accurate questions with key terms, and the concepts of importance and completion of a search. These concepts began to emerge as properties or characteristics of the first core category, establishing accurate information needs. Accurate key terms and questions based on properties of the information need, like information type and accurate terms, were associated with complete searches and perceptions of importance. Table 6 provides examples of this core category.

Table 6

Establishing Accurate Information Need Examples

#	Transcript or Observation of Establishing Accurate Information Need
1	“I'm trying to think of the keywords of what I want to find rather than the specific of what I'm looking for. It's like in that last one, like places, Denver, events, food trucks, September two thousand and twenty-two.”
2	“Okay, and also, because I'm not a great speller. I tend to do that to make sure they have the correct title and that I'm spelling it correctly.”
3	“They just care about the words that you put in it. So that's like when I need to use descriptive words and add everything I know I need.”
4	“You have to craft the words right enough to get what you want right, and there are times when finding out those words are pretty difficult.”

Note. Interview excerpts of the Establishing information need category saturation.

Participants described not having established information needs as difficult. For example, “You have to craft the words right enough to get what you want and there are times when finding out those words are pretty difficult.” (Participant 8). Without establishing a well-defined information need, searches became complex. For example,

“I wanted to find out the purpose of a device that I found. Google instantly took me to like a shopping site. You know a series of shopping websites like, Oh, here's how you can buy it. It was written as if I was a plumber, and I already knew what I was looking for, so I had to back out. and then I went back, and I read a couple of other sites to try to really understand not only what this valve was, but why would you use this valve instead of just having a pipe? But with this situation, you know, I had to back out and try again and again”.

These examples explain the importance of establishing accurate information needs by understanding the information type needed and using accurate questions or key terms to reach the correct information (see Appendix C memo 7). This category and its core concepts saturated the data each time the researcher analyzed new and old data bringing a degree of validity and reliability to the category as a key component of the substantive theory (Morse et al., 2002).

What is the process by which search and discovery interfaces and tools support the modern search process?

The second research question and the associated interview questions and observations developed the other integral category to the emerging theory, the theory of the importance of validating information during the search process as 1) establishing accurate information needs and 2) verifying information results. The second way of validating search information in a search process is to verify information results. The information retrieved by a search tool can support the verification of information by holding these properties, presenting information in a familiar way, providing filters and facets to sort information by measures deemed relevant, and providing accurate information without unwanted information. The following concepts began as code and, after comparative analysis and rounds of data collection, were compiled as properties and core concepts of the core category, which explains how search tools can support validating search information.

Concept of Ease of Use

This property of search tool use occurred in twelve interviews and occurred as features, filters, sorting, web designs, and devices. Two participants used a virtual assistant to complete their think-aloud search activities, and they described use as “ we use our virtual assistant and ask her a question about something that we're thinking about when at home.” (Participant 6) Additionally, search tools were not used when features and filters did not function as expected, for example, “One of those frustrating sites. where I've done all the filters and everything, and then it keeps losing my info, and I hate that.” (Participant 14). This example demonstrates how the opposite concept of ease of use can bring a search to an early end.

Concept of Familiarity

Participants most often responded to the interview prompt of “What are your reasons for selecting search interfaces or tools over others?” with the property of familiarity. Familiarity occurred in 19 of the 20 interviews. Participants were driven to use tools and interfaces that they were most familiar with, “I like Google Chrome and Google is a search engine. I'm: not a fan of Microsoft. You know their search stuff. I use it if I have to. Um, but it's really personal preference”(Participant 12). When probing deeper, participants described familiarity with tools as, “I know the information that's going to be brought up.” This familiarity with the information retrieved and how to use the display of the search tool were core reasons for selecting search tools over others (see

Appendix C Memo 5). Table 7 provides additional examples of the concept of familiarity and its relationship to verifying information results.

Table 7

Familiarity and Verifying Information Results

#	Transcript or Observation of Familiarity and Verifying Information Results
1	“I’m real familiar with it. I kind of know what I get. You know. I know what I will get when I click on a product that brings up the out page. Um. So yeah, I just know what I’m getting.”
2	“but also aesthetically, I am the most comfortable with them because I've used them.”
3	“I know enough about Google that I don't like how they make people and companies and businesses pay to acquire me as a customer. So if I know and am familiar with the place where I would like to end up, I try to cut Google out of it.”
4	“I don't know. I've never heard of Do303, so I’m going to Eventbrite. I'm more familiar with it.”

Note. Examples of data that connected familiarity and verifying results.

Concept of Analyzing Results

The concept of analyzing results occurred most commonly as the second step in a search process. Users need to check or verify the information by looking at timestamps or reviewing the information format or type. “So I always check just to make sure that in Google and then enter it in the search bar in our library system, which we use Polaris.” (Participant 2). The concept of verifying results occurred in eighteen interviews and included codes marking for trust, checking, or validity. One participant explained it as

“I’ll go. Well, yeah, I remember using that site before or with a reputation. And I go. Oh, yeah, I’ve heard of this site or somebody with just a really good presentation, and then I will go into that.” (Participant 17)

Category of Verifying Search Results

These core concepts began to connect when analyzing data and thinking about the process and context. The core concepts emerged as properties of the core category of verifying search results. Search tools that support the search process were easy to use, familiar, but more importantly they presented ways to check the results. These tools supported the search process best when verifying search results. A common application of verifying results is to use different search tools to verify and validate results (see Appendix C Memo 2). Table 8 presents several examples of the category.

Table 8

Verifying Search Results Coded Examples

#	Transcript or Observation of Verifying Information Search Results
1	“But sometimes you're looking to see who wrote it, and there would be certain news outlets that look like news.”
2	“I try to look for something that has a decent rep. If I’m looking for medical, I’ll look at the Cleveland Clinic, or I will look at and NIH or the CDC. I don't just go hunting around and go. Oh, here's some guy with an opinion, and that's the same thing for my veterinary care.”
3	“If I’m looking for something specific, I want to see it in multiple sources, and I am careful to look at the exact source.”
4	"Eric, and from my own experience, I know that Eric is a reliable source to use for this"

Note. Excerpts of the Verification of Search information Results category saturation

This category was present in all of the think-aloud search activities, including voice assistant and phone-based searches and browser or PC-based searches.

The inverse of this category was explored in the data to validate this category—this analysis brought forward coding for unwanted information. A participant explained: “but search can also be frustrating, and that's often because of the ads and pop-ups that interfere with one search. And you know, I think we've all learned to notice what part of the search results page, to avoid and start looking at more meaningful results. But it's still a frustration.” (Participant 9). This analysis helped to validate the importance of verifying search results as a category and integral component of the substantive theory.

Chapter Five Conclusion

The Emerging Quality of Validating Information in a Search

The substantive theory of the importance of validating information during a search by 1) establishing accurate information needs and 2) verifying the search results begins to reveal modern changes and qualities of today's search process. This emerging quality is the result of the immense amount of information on the internet and the increased access to the internet since 2020. The modern search experience requires the validation of search information used in the initial inquiry or used as an evaluative end step once the query has returned results.

As technology continues to change and offer more searchable information and new search tools like voice assistants, our need to complete searches or find valid information will continue to grow. While finding trustworthy sources or crosschecking results is not new to search process models, it is important to recognize that it occurs at more than one stage. Validating search information occurs when identifying key terms, formulating accurate questions, or analyzing results. Users experience difficult or complex searches when validating does not occur, leading to unwanted information and frustration

Return to the Research Questions

The research questions explored by this grounded theory were:

1. What is the process by which humans search and discover information?
2. What is the process by which search and discovery interfaces and tools support the modern search process?

Validating search information emerged as an integral quality of the modern search experience. This emerging concept has implications for a user's search experience and was categorized as important and connected to complete searches. This emerging theory occurred when considering what search tool support was and was not and occurred when discussing importance and completeness of a search. Through axial coding and considering the studied context and search process, it became clear that validating search information influenced and impacted both search tools and search processes.

Establishing an Accurate Information Need

Validating search information occurs earliest in the search process to establish accurate information needs. Establishing an accurate information need includes formulating key terms, identifying an information type, or asking questions with a degree of accuracy and knowledge for the end result. This core category occurs early in the search process before using a search tool. However, it can be affected by search tools and devices like the limiting input of voice-based search tools or the beneficial use of filters and facets to refine searches by the information need (see Appendices C, Memo 7).

Verifying Information Results

Validating search information occurs later in the search process once the initial query has been executed in a search tool. Verifying information results includes dimensions of familiarity with the tool and how its information is retrieved and visualized, the ease of use on validating and interpreting the results, and finally, check the results for the accuracy of the search purpose. Search tools affect this category in positive and negative ways. Tools that are visually familiar with result page design and features like filters can support the verification of information results. In contrast, other search tools may produce unwanted information and create difficulty in the search process.

Discussion

To make more sense of the findings of this study, it is important to look at the emerging theory in relation to the information-seeking processes. For the purpose of this study, the modern search process was described, and stages were ordered sequentially based on an analysis of the interview protocol's responses to “what is your process of searching or discovering information?”. Participant 19 described it most representatively as,

“So typically, the first thing I want to do, I'll just you start off with just seeing if what is in my head or how to describe something is how it's actually typically described. Um, if not, I'll usually use that first source to try and determine what is the actual, you know. Ah! Verbiage is used to

describe what I'm looking for, and then I need to do a second search. Kind of refine it. Looking for, like, you know again, based on what I'm seeing, the results are. Is this getting closer to what I need? Um! It has going through results on that on that basis and kind of making sure that I can kind of narrow it down as much as possible to what I'm actually looking for, and see if that's um partially my description is not getting me there, or is the information just doesn't exist, or what."

This typical process shared many similarities with older information-seeking processes. Comparisons between Hearst's (2009) model of "The process where identifying an information need, followed by the activities of query specification, examination of retrieval results, and if needed, reformulation of the query, repeating the cycle until a satisfactory result set is found." and the process that emerged here was very similar (p.2). Additionally, participants' steps to completing the think-aloud search activity followed Broder's model (2002), "The process of the task, info need, verbal formulation, query, search engine, results" (p. 4).

One last meaningful comparison is that of Sense-making and the importance of the searcher. Dervin explains, "The human then makes sense of these results in their process. Sense-making inherently represents the human in search, emphasizing the interpretive and context-driven aspects of meaning-making when people interact with systems." (Dervin, 2004, p. 22). This substantive theory is not contrary to Dervin's definition but quite possibly clarifies the human act of interpreting or making sense

through validation. To validate searches is, in some ways, making sense and answering the question that participants posed in their process, “Is this getting closer to what I need?”(Participant 18)

Limitations

The first limitation is that of the sample. The inclusion criteria collected data from individuals living in the US who are above the age of 18. Additionally, during recruitment of participants, the researcher did not collect demographic information. While participants were evaluated for how much or how little experience searching the internet for information that they had, no other participant descriptors or demographics were collected.

The second possible limitation of this study is the use of a concurrent think-aloud protocol. While this study used the protocol practice to generate observations on general search behavior and search process, the protocol has been challenged when applied for usability testing. When using think-aloud to test a search tool or web page design’s usability, the concurrent tasks of speaking about what the subject is doing and completing the requested task could interfere with each other (Van den Haak, De Jong, M., & Jan Schellens, P., 2003). In the case of this study, the search activity and participants' ability to complete a more authentic search process might have been limited. Although this may be a common limitation to think-aloud protocols applied to usability testing or studying processes alone, this study included interview questions as a second collection of related information on the process.

The third possible limitation is that of the sole researcher. The researcher must give the needed space in the research study to permit the theory to emerge. Existing biases of the researcher's own experience with searching for information and the development of a literature review prior to the commencement of the grounded theory study might have biased the study (Glaser, 1992). However, later developments in the Straussian grounded theory methodology called for literature reviews explaining that Corbin and Strauss (1990) acknowledged that an initial literature review may be needed and can be used without precluding the researcher's open approach to data collection and theory-generating. This possible limitation was expanded upon in this study's thesis. An argument was made for how the importance of validating search information expands and develops known information-seeking processes further.

Future Directions

Most participants learned to search through trial and error, but some of the US population still needs opportunities to learn and get help. The concept of search literacy was present throughout this study with a very high frequency and often highlighted gaps in participants' abilities to complete searches.

“So I think they are really leaving behind like the fifty and up population. Where they might have kind of dabbled in technology. But it just kind of skyrocketed. And now, this generation is just more in tune with how to put things together and look for things. That's because they've grown up with it. Trial and error, because we have a lot of young, older people, Ah!

Patrons who come in struggle with tools when just searching for things. And I think it's also a dialect where their language is different than what is being coded now.” (Participant 4).

This memo highlights a specific case where a participant who works as a reference specialist described the common occurrence of assisting others with establishing an accurate information need. This theory could guide the evaluation or creation of modern information literacy curriculums where the emphasis is placed on developing the searcher’s initial search process step to establish accurate information needs and to develop strategies for verifying search results.

Another future direction is to evaluate search tools and features for how they might support this emerging quality of the modern search process. Does sorting features, facets, and filters improve confidence in search results? How can information retrieval be presented by search tools that provide verification or develop trust? Can tools with inputs like voice search support the establishment of accurate information need? How the technology can support validating searches could improve the use and lead to more efficient searches.

Final Thoughts

This study employed a unique but promising research design. The combination of grounded theory research design, think-aloud protocol, and a semi-structured interview provided a large amount of data for quick and reliable analysis when studying a human

but digital process. The data collection also allowed participants to elaborate on their think-aloud search activity and complete the activity using various tools and approaches. This acted as a type of member checking but was completed within one synchronous data collection event. Combining the qualitative research method, data collection approach, and theoretical sampling practices improved reliability and strengthened the theory's validity.

The emerging theory of validating search information should be applied in this information-rich digital age. The need to validate search information throughout the search process and to have search tools that support the validation of search information is quickly becoming a major need for information consumers today. To navigate the modern internet with its abundance of information in various forms, search tools should support searchers with validating information, and our search process should reflect the importance of information validity at each step.

Reference

- Anfara, Brown, K. M., & Mangione, T. L. (2002). Qualitative Analysis on Stage: Making the Research Process More Public. *Educational Researcher*, 31(7), 28–38.
<https://doi.org/10.3102/0013189X031007028>
- Babich, N. (2020, October 20). Human factor principles in UX design: Adobe XD ideas. *Human Factor Principles in UX Design*. Retrieved June 10, 2022, from <https://xd.adobe.com/ideas/principles/human-computer-interaction/human-factors-ux-design/>
- Belgrave, & Charmaz, K. (2012). Qualitative Interviewing and Grounded Theory Analysis. In *The SAGE Handbook of Interview Research : The Complexity of the Craft* (2nd ed., pp. 347–365). SAGE Publications, Inc.
<https://doi.org/10.4135/9781452218403.n25>
- Bokhari, Adhami, M. K., & Ahmad, A. (2021). Evaluation of News Search Engines Based On Information Retrieval Models. *Operations Research Forum*, 2(3).
<https://doi.org/10.1007/s43069-021-00081-0>
- Broberg, L. L. (2011). A grounded theory approach to examining design and usability guidelines for four-year tribal college web sites (Order No. 3460923). Available from ProQuest Dissertations & Theses Global. (879053198).
<https://du.idm.oclc.org/login?url=https://www-proquest-com.du.idm.oclc.org/dissertations-theses/grounded-theory-approach-examining-design/docview/879053198/se-2?accountid=14608>
- Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, 36(2), 3–10.
<https://doi.org/10.1145/792550.792552>
- Bryant, A. & Charmaz, K. (2019). The SAGE Handbook of Current Developments in Grounded Theory. In *The SAGE Handbook of Current Developments in Grounded Theory*. SAGE Publications Ltd.
<https://doi.org/10.4135/9781526436061>
- Charmaz, K. (2014). Grounded Theory in Global Perspective: Reviews by International Researchers. *Qualitative Inquiry*, 20(9), 1074–1084.
<https://doi.org/10.1177/1077800414545235>
- Chapman, Simperl, E., Koesten, L., Konstantinidis, G., Ibáñez, L.-D., Kacprzak, E., & Groth, P. (2020). Dataset search: a survey. *The VLDB Journal*, 29(1), 251–272.
<https://doi.org/10.1007/s00778-019-00564-x>
- Charters, E. (2010). The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brock Education*, 12(2).

- Cleverley, P., Burnett, S. (2015) Retrieving haystacks: a data driven information needs model for faceted search. *Journal of Information Science* 41:1, 97-113. Online publication date: 5-Nov-2014.
- Corbin, Strauss, Strauss, & Anselm L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* / Juliet Corbin and Anselm Strauss. (3rd ed.). Los Angeles, Calif.: Sage Publications.
- Creswell, J. W. & Cheryl Poth (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4rd Ed.). Thousand Oaks, CA: Sage. Third edition (2012) is also acceptable.
- Dervin, B (2004) *Sense-Making Methodology Reader: Selected Writings of Brenda Dervin*. (2004). *Reference & Research Book News*, 19(1), 210-. Ringgold, Inc.
- Doi, T. (2021). Usability Textual Data Analysis: A Formulaic Coding Think-aloud Protocol Method for Usability Evaluation. *Applied Sciences*, 11(15), 7047-. <https://doi.org/10.3390/app11157047>
- dos Santos Pergentino, Canedo, E. D., Lima, F., & de Mendonça, F. L. L. (2020). Usability Heuristics Evaluation in Search Engine. In *Design, User Experience, and Usability. Interaction Design* (pp. 351–369). Springer International Publishing. https://doi.org/10.1007/978-3-030-49713-2_25
- Demelo, & Sedig, K. (2021). Design of Generalized Search Interfaces for Health Informatics. *Information (Basel)*, 12(8), 317-. <https://doi.org/10.3390/info12080317>
- Dewri, Thurimella, R. (2016). Mobile local search with noisy locations. *Pervasive and Mobile Computing*, 32, 78–92. <https://doi.org/10.1016/j.pmcj.2016.04.014>
- Duarte, Oliveira, E., Cogo, F. R., & Pereira, R. (2015). Dico: A Conceptual Model to Support the Design and Evaluation of Advanced Search Features for Exploratory Search. *Human-Computer Interaction – INTERACT 2015*, 87–104. https://doi.org/10.1007/978-3-319-22723-8_8
- Duffy, K., Ferguson, C., & Watson, H. (2004). Data collecting in grounded theory--some practical issues. *Nurse Researcher*, 11(4), 67-78.
- Ericsson, & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Fathi Najafi, T., Latifnejad Roudsari, R., Ebrahimipour, H., & Bahri, N. (2016). Observation in Grounded Theory and Ethnography: What are the Differences?. *Iranian Red Crescent medical journal*, 18(11), e40786. <https://doi.org/10.5812/ircmj.40786>

- Fallows, D. (2005). How Women and Men Use the Internet. Pew Internet and American Life Project 28.
- Fan, M., Shi, S., & Truong, K. N. (2020). Practices and Challenges of Using Think-aloud Protocols in Industry: An International Survey. *Journal of Usability Studies*, 15(2).
- Foley, Timonen, V., Conlon, C., & O'Dare, C. E. (2021). Interviewing as a Vehicle for Theoretical Sampling in Grounded Theory. *International Journal of Qualitative Methods*, 20. <https://doi.org/10.1177/1609406920980957>
- Gero, & Mc Neill, T. (1998). An approach to the analysis of design protocols. *Design Studies*, 19(1), 21–61. [https://doi.org/10.1016/S0142-694X\(97\)00015-X](https://doi.org/10.1016/S0142-694X(97)00015-X)
- Gold, S. (2020, July 14). Society-Centered Design. Retrieved June 10, 2022, from <https://www.oreilly.com/radar/society-centered-design>
- Guss, C. (2018). What Is Going Through Your Mind? Thinking Aloud as a Method in Cross-Cultural Psychology. *Frontiers in Psychology*, 9, 1292–1292. <https://doi.org/10.3389/fpsyg.2018.01292>
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies or qualitative research*. New Brunswick: Aldine Transaction.
- Haider, & Sundin, O. (2019). *Invisible Search and Online Search Engines : The Ubiquity of Search in Everyday Life (Edition 1) (1st ed.)*. Routledge. <https://doi.org/10.4324/9780429448546>
- Hallberg. (2006). The “core category” of grounded theory: Making constant comparisons. *International Journal of Qualitative Studies on Health and Well-Being*, 1(3), 141–148. <https://doi.org/10.1080/17482620600858399>
- Hearst, M. (2009). *Search user interfaces*. Cambridge University Press.
- Henderson, R. D., Smith, M. C., Podd, J., & Varela-Alvarez, H. (1995). A comparison of the four prominent user-based methods for evaluating the usability of computer software. *Ergonomics*, 39(10), 2030–2044. <https://doi.org/10.1080/00140139508925248>
- Husband, G. (2020). Ethical Data Collection and Recognizing the Impact of Semi-Structured Interviews on Research Respondents. *Education Sciences*, 10(8), 206–. <https://doi.org/10.3390/educsci10080206>
- Huurdean, & Kamps, J. (2016). Supporting the Process: Adapting Search Systems to Search Stages. In *Information Literacy: Moving Toward Sustainability* (pp. 394–

404). Springer International Publishing. https://doi.org/10.1007/978-3-319-28197-1_40

- Hyldegard, J. (2006). Collaborative information behaviour—exploring Kuhlthau’s Information Search Process model in a group-based educational setting. *Information Processing & Management*, 42(1), 276–298. <https://doi.org/10.1016/j.ipm.2004.06.013>
- Information retrieval journal. (n.d.). Springer. Retrieved June 22, 2022, from <https://www.springer.com/journal/10791/aims-and-scope>
- Iordache, E., Mariën, I., & Baelden, D. (2017). Developing Digital Skills and Competences: A Quick-Scan Analysis of 13 Digital Literacy Models. *Italian Journal of Sociology of Education*, 9(1), 6-30. doi: 10.14658/pupj-ijse-2017-1-2
- Johnson, & Finn, K. (2017). *Designing User Interfaces for an Aging Population: Towards Universal Design*. Elsevier Science & Technology.
- Kirschthaler, Porcheron, M., & Fischer, J. (2020). What Can I Say?: Effects of Discoverability in VUIs on Task Performance and User Experience. *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–9. <https://doi.org/10.1145/3405755.3406119>
- Kuhlthau. (1989). *Facilitating Information Seeking through Cognitive Modeling of the Search Process*. A Library Studies Research Project. Distributed by ERIC Clearinghouse.
- Kuusela, & Pallab, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology*, 113(3), 387–404. <https://doi.org/10.2307/1423365>
- Lawrence, J. and Tar, U. (2013) *The Use of Grounded Theory Techniques as a Practical Tool for Qualitative Data Collection and Analysis*. *The Electronic Journal of Business Research Methods*, 11, 29-40.
- Lee, J.-W., Im, H.-H., Kim, J.-H., Kang, S.-J., & Kim, M.-S. (2012). Exploring User Perceived Usability Characteristics of Applications on Smart Phones: A Grounded Theory Analysis of User Reviews. *Journal of the Korea Academia-Industrial cooperation Society*. The Korea Academia-Industrial Cooperation Society. <https://doi.org/10.5762/kais.2012.13.2.615>
- Lee, Lee, J., & Lee, J. D. (2016). A Visual Search Model for In-Vehicle Interface Design. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1874–1878. <https://doi.org/10.1177/1541931213601427>

- Lewis, C. (1982) Using the “Thinking-Aloud” Method in Cognitive Interface Design; IBM TJ Watson Research Center.
- Lingard, Albert, M., & Levinson, W. (2008). Qualitative Research: Grounded Theory, Mixed Methods, and Action Research. *BMJ (Online)*, 337(7667), 459–461. <https://doi.org/10.1136/bmj.39602.690162.47>
- Liu, & Özsu, M. T. (2009). *Encyclopedia of Database Systems*. Springer.
- Liu, C., Liu, Y.-H., Liu, J., & Bierig, R. (2021). Search Interface Design and Evaluation. *Foundations and Trends® in Information Retrieval*, 15(3–4), 243–416. <https://doi.org/10.1561/15000000073>
- Marchionini, G., & White, R. (2007). Find what you need, understand what you find. *International Journal of Human-Computer Interaction*, 23(3), 205–237. <https://doi.org/10.1080/10447310701702352>
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods* 1(2), Article 2.
- National Telecommunications and Information Administration. (2022) NTIA Data Explorer. <https://www.ntia.gov/data/explorer#sel=internetUser&disp=map>
- Nielsen, J.(1993) *Usability Engineering*; Academic Press, Inc.
- Olson, K. (2018). Ethical issues in grounded theory. In *The sage handbook of qualitative research ethics* (pp. 486-494). SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781526435446>
- Rieh, Collins-Thompson, K., Hansen, P., & Lee, H.-J. (2016). Towards searching as a learning process: A review of current perspectives and future directions. *Journal of Information Science*, 42(1), 19–34. <https://doi.org/10.1177/0165551515615841>
- Russell-Rose, & Tate, T. (2012). *Designing the Search Experience* (1st edition). Morgan Kaufmann.
- Sandhu, A.K., & Liu, T. (2014). Wikipedia search engine: Interactive information retrieval interface design. 2014 3rd International Conference on User Science and Engineering (i-USER), 18-23.
- Strauss, Corbin, & Corbin, Juliet M. (1998). *Basics of qualitative research : Techniques and procedures for developing grounded theory / Anselm Strauss, Juliet Corbin.* (2nd ed.). Thousand Oaks: Sage Publications.
- Strauss, A. (1987). *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge University Press.

- Srinivasan, Dontcheva, M., Adar, E., & Walker, S. (2019). Discovering natural language commands in multimodal interfaces. *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 661–672. <https://doi.org/10.1145/3301275.3302292>
- Sutcliffe, & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10(3), 321–351. [https://doi.org/10.1016/S0953-5438\(98\)00013-7](https://doi.org/10.1016/S0953-5438(98)00013-7)
- Sudman, Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers : the application of cognitive processes to survey methodology*. Jossey-Bass Publishers.
- Syed, & Collins-Thompson, K. (2017). Optimizing search results for human learning goals. *Information Retrieval (Boston)*, 20(5), 506–523. <https://doi.org/10.1007/s10791-017-9303-0>
- Society centered design. (2020). Retrieved June 22, 2022, from <https://societycentered.design/>
- Timonen, V., Foley, G., & Conlon, C. (2018). Challenges When Using Grounded Theory: A Pragmatic Introduction to Doing GT Research. *International Journal of Qualitative Methods*. <https://doi.org/10.1177/1609406918758086>
- Van den Haak, De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology*, 22(5), 339–351. <https://doi.org/10.1080/0044929031000>
- Winckler, Cava, R., Barboni, E., Palanque, P., & Freitas, C. (2015). Usability Aspects of the Inside-in Approach for Ancillary Search Tasks on the Web. In *Human-Computer Interaction – INTERACT 2015* (pp. 211–230). Springer International Publishing. https://doi.org/10.1007/978-3-319-22668-2_18
- Weick, K., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16(4): 409–421.
- Wobbrock, Aung, H., Rothrock, B., & Myers, B. (2005). Maximizing the guessability of symbolic input. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, 1869–1872. <https://doi.org/10.1145/1056808.1057043>
- Wuyckens, G., Landry, N., & Fastrez, P. (2022). Untangling media literacy, information literacy, and digital literacy: A systematic meta-review of core concepts in media education. *Journal of Media Literacy Education*, 14(1), 168-182. <https://doi.org/10.23860/JMLE-2022-14-1-12>

Young Namkung MS, RD, PhD, Seo-Young Shin PhD, RD & Il-Sun Yang PhD,
RD (2007) A Grounded Theory Approach to Understanding the Website
Experiences of Restaurant Customers, Journal of Foodservice Business
Research, 10:1, 77-99, DOI: [10.1300/J369v10n01_05](https://doi.org/10.1300/J369v10n01_05)

Appendices

Appendix A

Data Collection Protocol

Pre-interview Procedures

- Introductions: introduce myself, explain the process, how long the interview will take, and the general format of questions.
- Study purpose: Explain the purpose statement and why this is important
 - Clarify ‘search experience and search process.’
 - Clarify ‘search tools and discovery interfaces’
- Screening questions:
 - Are you over 18?
 - Are you currently living in the US?
 - Have you used search interfaces on library websites or google before?
 - Do you have the ability to fully consent to participate in this study?
- The consent form link will be provided again. Consent form approval check and verification of consent comprehension.
 - Did you complete the consent form?
 - If no. Then please follow the link that I am about share and read through and sign the consent form now.
 - Do you understand that we will record this session?
 - Do you understand that this interview contains a search activity?
- Treatment of data explanation:

- Participant data will be confidential. The link between your identifiers and the research data will be destroyed after the records retention period required by state and/or federal law. With your permission, I would like to record this interview so that I can make an accurate transcript. Once I have made the transcript, I will delete the recordings. Your name will not be in the transcript or my notes.
- Questions, Comments, or concerns?
- I will begin the Zoom recording at this point.

Think-aloud Exercise

Pre-task verbal commands:

Before we begin with the task, let's practice thinking out loud. Talk me through your three steps of preparing your device and how you will search for any topic. Say whatever you are thinking and do the best that you can.

Task Instructions:

Let's select one of the following topics to complete a search using your device. When a topic, please begin and remember to think aloud through your search process. When you have found the ideal information, the search activity will end.

Topic 1: using a search tool, social media, or a local public search tool, find a local event with food trucks in your area in the next week or two.

Topic 2: Using a website or tool of your choice like google, Kaggle, or Data.gov, search for US socio-economic data that is interesting to you. (socio-economic data) (income)

Topic 3: Using a search tool of your choice, find the best Colorado hike that you, as a hiker, would consider doing.

Topic 4: Using the internet, please find venue information on a local concert in the Toronto area for November.

When you have found the ideal information, stop. Remember to think aloud through the process of searching. (r1,r2)

Introductory Questions

Rapport building opening questions

1. What is your process of searching for or discovering information? (r1)
2. How do you feel during the typical internet search process?
 - Are searches easy or hard? (r1)
3. How do you decide that a search is complete? (r1)

Key Interview Questions

Now, let's think back on the think-aloud search and discuss it.

4. What was most important to that search experience? (r1, r2)
 - What is your level of familiarity with the tools that you used?(r1, r2)

5. What are your reasons for selecting search interfaces or tools over others? (r2)
6. How did you feel about the modern search process? (r1)
7. How did you use the search tool/interface in this search? (r1)
8. How did the tools support your search process? How did they not? (r2)

Probing Questions:

- Could you clarify?
- Could you speak more about _____ interface/interaction/services?
- Tell me more about your experience.

Conclusion Questions

As we bring this interview to an end, I have one closing question.

- To obtain your final thoughts, is there anything else you would like to share about searching for information?

Thank you for your time today! Please feel free to reach out to me with any questions.

Appendix B

Think-aloud Protocol

The search task design was based on the following criteria:

The topics were selected from different areas of general knowledge with enough variety to avoid the possibility that the participant could know about all the topics. Additionally, the topics are findable through common search and discovery tools like Google or Compass discovery tools on public library websites. Each topic and task requested the user to find a specific detail or instance of the topic in a digital page or artifact like a journal, video, or website.

Pre-task verbal commands:

Before we begin with the task, let's practice thinking out loud. Talk me through your steps of preparing your device and how you will search for this topic. Say whatever you are thinking and do the best that you can.

Topic 1: Please search for information on recycling. When you have found the ideal information, stop. Remember to think aloud through the process of searching. (r1,r2)

Topic 2: Please search for information on a local market. When you have found the ideal information, stop. Remember to think aloud through the process of searching. (r1,r2)

Topic 3: Please search for information on a famous abstract painter. When you have found the ideal information, stop. Remember to think aloud through the process of searching. (r1,r2)

Topic 4: Please search for information on a jazz musician. When you have found the ideal information, stop. Remember to think aloud through the process of searching.

(r1,r2)

Appendix C

Memo Samples

Memo 1

Process: Asking questions

Participant Quote/Observation:

Researcher: great will be perfect. Okay. So the prompt is using a website or tool of your choosing like cable data. Go, Google, search for us socioeconomic data that is of interest to you.

Participant: Okay Ah, you see this? Would, uh Hispanic serving institutions be a valid? Ah, a looking for a percentage uh Hispanic serving institutions.

Quick Analysis: This participant is a reference librarian and added extra and specific details to the search. Expanding it with HIS and information types containing percentages. This expands the initial query to cover mores specific search terms.

Memo 2

Process: Verifying to figure it out

Participant Quote/Observation:

“Yesterday we were trying to figure out why the slogan for Dairy Queen was different in Texas than the rest by it. Yeah, brand new news to me as well. So I did a couple of Google searches on that, and I felt complete. Once I had like a little bit of data like numbers from empirical data. I had some qualitative data. Yes, that's the other one. And um! Then I could just verify a couple of things on Wiki as well just like, make sure. And then I felt like it was done, and I shared that with the group when I was a guy figured this out,”

Quick Analysis: The searcher uses other results to verify their findings prior to reaching a state completeness.

Memo 3

Condition: Accuracy of key terms and questions as important

Participant Quote/Observation:

“um and tap in the search bar and try not as detailed, I guess, but as simply, I can phrase it, into that search bar, and then from there, if I get too broad um to broad of searches or results. Then I try and go as detailed as possible, either using quotation marks or um stuff to kind of exclude anything that it's giving me that isn't associated with what I want to look for.”

Quick Analysis: Accuracy of terms used in search is important to searcher's with search expertise like this reference librarian. The interview included a story about helping individuals with searches and gathering as much information at the first stage as possible.

Memo 4

Process: Refocusing for lack of established need

Participant Quote/Observation:

“and if given the time, or if I really wanted to pursue this on my own, I would have had to have reformulated. So it really focused on word, choice, and not necessarily helping you to formulate the question or accuracy of of the info retrieved. Yeah, and it seems like it wouldn't have been that difficult.”

Quick Analysis: The search needed to be refocused and new key terms or questions had to be formulated to deliver more closely related results to the information need. Difficulty was also associated when searches were refocused.

Memo 5

Condition: Familiarity as credibility

Participant Quote/Observation:

“The article was actually that kind of credibility of that article. So I wanted to know if the person was, if it was a post that somebody of just posting, copying what someone else wrote, or if they had actually been there by, there weren't a lot of trashy choices to choose from. But sometimes a search will have like you're looking You're looking to see who wrote it, and there would be certain like news outlets that look like news. But i'm not familiar with that. I wouldn't use”.

Quick Analysis: Searchers reach for familiar sources and follow credible sources. This familiarity is slightly different than search tool familiarity but is of the same core familiar concept.

Memo 6

Condition: validity as trusting the search tool

Participant Quote:

“Say I've been looking at concerts a lot, you know, getting tickets, and you have to scroll. Say, probably like three to six results down to not click on scale like scam sites and scalpers, and that it's frustrating. I don't know if Google should be responsible for you know, putting more trustworthy results first. But it's frustrating that they aren't farther up. I think it's very misleading.”

Analysis: Validating information presented at the results stage should fall on the part of the search tool and not just the searcher. Searchers become frustrated with unwanted results. This statement and corresponding search activity observation reaches saturation for the core category of verifying search information results.

Memo 7

Condition: Establishing accurate information needs

Participant Quote/Observation:

“I think it's it would be two things, getting accuracy of information, and also timely, because I know, with that patron with the clay shooting. Sometimes she can get the dates wrong, and so getting that accurate information into on her. Oh, no, it's actually on Saturday, you know, at four o'clock, and so sometimes you have to kind of fudge even, and question what they're giving you at the same time. Because if you're not getting those search results fast, or you can't find them, it might be the misinformation that they gave to you. So I think it's also doing it”

Quick Analysis: As a response to the importance question, establishing the right or correct information is a common occurrence for this public reference librarian. The story highlights the need for establishing accurate information need in the form of details.

Appendix D

Sample Search Task & Interview

Researcher: So the first thing we'll do is we'll begin with a search task. I'll give you a prompt and something to search for. You can complete your search on a cell phone. You can use social media, Google, whatever search tool or search engine you might use. Naturally, Research, If you are using your browser and you're comfortable sharing your screen, please feel free to do that. And we'll do a little bit of for practice if you's thinking out loud as kind of an unnatural things.

Participant: So yeah,

Researcher: talk me through your three the first three steps of preparing your device for a search,

Participant: I guess, opening my browser if I haven't already.

Participant: Ah, Yes,

Participant: typing in my search term, whatever I'm looking for, and then filtering through results

Researcher: awesome, perfect. So that's great. So say whatever you're thinking and doing the best you can. You could speak of feelings, actions, things that you're seeing on the screen. And so we can end the search when you find the most ideal information fitting that prompt. So finding the

Participant: that it would be relevant. So let me give you your search task all right,

Researcher: using a search tool, social media or local public search for like library search, find a local event with food trucks in your area in the next week or two

Participant: i'll share the screen first. Oh, you can select a single application here. Yeah, I haven't really used Zoom at all,all right. So my browser is already open, so I don't need this tab. I'm just gonna

Participant: i'm just going to type in my search terms.

Participant: Um do Denver food trucks. Um!

Participant: We like food, truck events. And then

Participant: I think my first thought is to just put in like the next week's date

Participant: something like that.

Participant: That's what? right away.

Participant: Okay. So i'm looking at some of these results here, and it looks like there's a food truck festival

Participant: I didn't know about.

Participant: No,

Participant: it shows dates here already.

Participant: If I were to go with this, I would look at I would go through the web page and actually verify those I Don't really trust Google to always have the right date.

Researcher: Yeah. And so we can end the search when you find the most ideal information fitting that prompt. So finding the truck.

Participant: all right. Well

Participant: don't know who I've never heard of. Do three hundred and three, so i'm just gonna go to event Bright. I'm more familiar with it,

Participant: Nann. Near the right.

Participant: Well, let's see.

Participant: I'm going to use there.

Participant: Ah

Participant: filters Here

Participant: it is a food truck hub

Participant: the nowadays

Participant: Yeah,

Participant: all right. So I know that's not far from me, and

Participant: it looks like a good mix of things.

Researcher: Great. That's the ideal information.

Participant: And

Participant: I mean if i'm looking for events in my area with food trucks the next week, then this sounds like a good option,

Researcher: great. So let me. You can go ahead and stop sharing,

Researcher: and we'll switch into the second Interview part here.

Researcher: So the first questions are going to be pretty broad about any search you've done searching in general online. Um, for the first three questions, and i'll let you know when we shift into the the final four or five questions that have to do specifically with that search, or any other search that would be helpful to share, and as like an example when you answer the questions.

Researcher: So the first question is, What is your process of searching or discovering information?

Participant: Okay, that is very broad.

Participant: Ah,

Participant: I guess, using like a few key terms.

Participant: And then

Participant: I just i'm just seeing what comes up uh

Participant: I already mentioned it a little bit, but I tend to go towards sources that I've already

Participant: you've heard of, or have used before. First,

Participant: then, you know, if I can't find what i'm looking for on those, Then i'll use newer ones that i'm not familiar with. But

Researcher: how do you feel during the typical Internet search process?

Participant: Yes, they're pretty

Participant: easy for the most part,

Participant: you know, especially things more like that that are, you know,

Participant: not super technical or specific.

Researcher: How do you decide that a search is complete?

Participant: It's. Oh, yes, if I'm happy with the results I've found, or occasionally I'll check a few other sources, too, and try to verify whatever I've read.

Researcher: All right. So now we're going to switch into some questions related to that search experience. So, think about that one

Researcher: or others more recently that you've done that might help answer the question better. What was the most and what was most important to that search experience.

Participant: It's

Participant: no, I guess, correct information, because you know it's like the when I clicked on that page. Originally I had to then use their search filters to find what I was actually looking for.

Participant: Ah,

Participant: the information to be what I was looking for right away.

Researcher: What are your reasons for selecting search interfaces or tools over others.

Participant: I mean, you

Participant: like Google over Bing or whatever. Yeah,

Participant: I guess, for most things I just use Google because it's the most

Participant: don't know. I guess it's just the most well used. Feel like That's the easiest to to

Participant: just jump into right away.

Participant: Most familiar with it.

Researcher: It's a familiarity.

Researcher: Okay,

Researcher: great. How did you feel about this specific search process?

Participant: And it's a little cluttered.

Participant: You know Google puts up those little windows of information that they extract from websites, and

Participant: I I even mentioned it when I was searching it. I don't necessarily trust those.

Participant: I find that often those don't pull the right info.

Researcher: How do you use a search tool? How did you use the search tool or interface in this search,

Participant: I guess I used Google to find, like a broad selection of

Participant: events to look for, and then I use the internal search functions on the website. I chose to

Participant: narrow it down.

Researcher: It's a

Researcher: How did the tool support your search process, or how did it not?

Participant: I mean, Google did give me

Participant: lot of information that was around what I had put in. It did have dates around the date that I selected

Participant: ah

Participant: shown in the search results. So I think it was trying to be helpful.

Researcher: All right.

Researcher: So uh,

Researcher: as we bring this interview to an end, I have one closing question to obtain your final thoughts. Is there anything else you'd like to share about searching for information today?

Participant: Say I've I've been looking at concerts a lot, you know, getting tickets, and you have to scroll. Say, probably like

Participant: three to

Participant: six results down to not click on scale like scam sites and scalpers,

Participant: and that

Participant: it's frustrating. I don't know if Google should be responsible for you know, putting more trustworthy results first. But it's frustrating that they aren't

Participant: farther up. I think it's very misleading.

Researcher: great.

Researcher: Thank you for your time today. Please feel free to reach out to me with any questions I will end the recording.