The query is just the beginning: Exploring search-related decision-making of young adults

Veronica Maidel
Syracuse University, USA

Dmitry Epstein
Cornell University, USA

ABSTRACT

Web search has become an integral part of everyday online activity. Existing research on search behavior offers an extensive and detailed account of what searchers do on the search results pages. Yet, there is limited inquiry into what drives the particular search decisions are being made and what contextual factors drive this behavior. This study provides a user-centric inquiry focused on in-depth detailed investigation of search-related decision-making processes. It builds on data collected through analysis of structured observations of young adults performing search on their personal laptops. It focuses explicitly on the decisions the users make after completing a query and facing a list of search results. The study reveals a pattern of sophisticated use of a variety of explicit cues, tacit and contextual knowledge, as well as employment of an incremental search strategy.

INTRODUCTION

The continuously growing volumes of information pose significant challenges to evaluating credibility, quality, and relevance of online content. In recent years search has emerged as a major tool through which people reach information and experience the online world. Search engines have been referred to as the gatekeepers of online information, as they apply algorithms that make decisions about which content to present to the user, out of the millions of available options. They have been criticized for enabling the reproduction of the traditional media landscape where a handful of large, wealthy websites are accounting for the most of the web content and traffic (Granka, 2010). Yet search engines are not a purely technical phenomenon. Hargittai (2007) described them and their uses as "embedded in a myriad of social processes that are important for social scientists to consider in their research in order to understand the social implications of these important tools of our time" (p.775). Moreover, she emphasizes that "[g]iven their popularity, search engines are important brokers of information, and knowing more about how they represent content and how they are used is vital to understanding patterns of information access in a digital age" (p. 775-776).

There is a variety of studies focused on search engines and information seeking behavior. Depending on the home discipline, they range from system-focused studies (development and improvement of search algorithms and indexing techniques) to the human-focused studies (dealing with users' information needs and their information seeking behavior) (Kelly, 2009). Studies that acknowledge the human agency typically focus on user behavior, with an emphasis on identifying patterns of how searchers interact with

search engines. Yet, there is limited inquiry into what facilitates the searcher's decision-making as to which search results to follow.

There is also a variety of studies dealing with the questions of credibility and quality of online information. Some of the early scholarship in this area focused on evaluation of the elements of web pages and their content (Flanagin & Metzger, 2000; Fogg et al., 2001, 2003). Another line of scholarship focused on the cognitive processes involved in evaluation of information (Hilligoss & Rieh, 2008; Metzger, 2007). Recently, there is a growing focus on the social aspects of credibility judgments (Metzger, Flanagin, & Medders, 2010). Although on its face useful, there is limited application of this body of knowledge for furthering understanding of the searcher's decision-making process when encountered a page with a set of results.

Bridging the two areas of inquiry, this study aims to shed light on what the users do know about search engines and the search process, and how they integrate this knowledge with their perceptions of credibility in their use of the search systems. It builds on data collected through analysis of structured observations of young adults in the US, performing search on their personal laptops. It focuses explicitly on the decisions users make after completing a query and facing a list of search results. In other words, this paper asks not only what elements of the search results people pay attention to or what search results they actually follow, but also (1) how they interpret the various elements of the search results and (2) what aspects of context of their search activity influence their decision about which result to click on. Our goal is to advance a more holistic view of the Web as a longitudinal social experience.

BACKGROUND

Understanding how people search, particularly how they navigate through their search results is important for those designing the search engines, teaching digital literacy skills, as well as for those concerned with the social implications of search. Yet there is limited research into what facilitates people's decision to choose certain search results over others. There is extensive literature on what decisions people make when faced with a list of search results or on the information needs that underlie their search behavior, and how to predict those (Agichtein, Brill, & Dumais, 2006; Agichtein, Brill, Dumais, & Ragno, 2006; Downey, Dumais, Liebling, & Horvitz, 2008; Jansen, 2006; Jansen, Booth, & Spink, 2008; Jansen & Spink, 2006; Rieh & Xie, 2006; Silverstein, Marais, Henzinger, & Moricz, 1999). There are numerous studies about the path the users take through the search results until the moment they decide to click on a particular link (Cutrell & Guan, 2007; Dumais, Buscher, & Cutrell, 2010; Granka, Joachims, & Gay, 2004; Guan & Cutrell, 2007; Lorigo et al., 2008); in other words how they physically reach a link on a search engine results page (SERP). But there is limited inquiry into what facilitates particular decisions and what contextual factors drive this behavior (Hargittai, 2002; Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010; Teevan, Alvarado, Ackerman, & Karger, 2004). We identify the main barriers to fathering this line of research to lie in the area of methodology and limited conceptual repertoire.

The methodological challenge

Some approaches to studying search behavior, such as transaction log analysis, are predominantly search-data-driven. This type of inquiry is well suited for identifying patterns of search behavior. It provides information about the length of the query, frequency of its reformulation, number of results consulted on the SERP (Silverstein et al., 1999), length of the search session, percentage of single-term queries, and the use of query operators (Jansen & Spink, 2006); it also allows identifying categories of query reformulations (Jansen, Booth, & Spink, 2009; Rieh & Xie, 2006) and stability of search behavior over time and across search platforms (Jansen & Spink, 2006). The search-data-driven approach, however, is inherently limited in explaining the decision making process behind that behavior—a transaction log

does not record the users' motivation and does not track their cognitive processes, such as reasons for the search or the decision-making process when it comes to selecting a search result (Jansen, 2006).

Other approaches to studying search behavior are more problem-driven, which in turn requires more controlled research environments. Particularly interesting in the context of decision-making process about search results is the more recent and more sophisticated method of eye tracking. Eye tracking is used to obtain a deeper understanding of where people invest attention on the SERP, for how long, and in what order, before they click on a search result (Hornof & Halverson, 2003). The controlled nature of eye tracking analysis allows drawing a rather detailed map of the users' search behavior, including the order in which the search results are viewed (Lorigo et al., 2006), styles of search results evaluation as a function of the amount of information consulted on a SERP (Aula, Majaranta, & Räihä, 2005), differences in search results' scanning behavior across different platforms (Lorigo et al., 2008), the influence of the snippet length on search performance (Cutrell & Guan, 2007), as well as identifying which search results get the most attention (Dumais et al., 2010; Granka et al., 2004; Guan & Cutrell, 2007; Pan et al., 2007).

The experimental nature of the eye tracking studies provides more data on information behavior of the users compared to transaction log analysis, but those studies still use behavior elicited through eye tracking analysis as a proxy for the user's cognitive processes. Gaining a deeper understanding of the decision-making processes using eye tracking is problematic, because the difficulty in analyzing and interpreting eye tracking data and the difficulty in integrating eye tracking methods with other usability testing techniques, such as think-aloud (Jacob & Karn, 2003; Poole & Ball, 2005).

The conceptual challenge

One of the main research areas in the field of Information Retrieval (IR) pertaining to decision-making activity of a searcher on the SERP is the study of relevance. There is an ongoing effort to understand relevance and its manifestations. Saracevic (2007) classified a series of criteria used by the searchers when making judgments about the degree of relevancy. Those included: content (topic, quality, depth, scope, currency, treatment, and clarity), object (characteristics of the document, such as representation, availability, costs, etc.), validity (accuracy, authority, trustworthiness of sources, verifiability), situational match (appropriateness to situation, usability, urgency), cognitive match (novelty, mental effort), affective match (emotional responses, frustration), and belief match (personal credence given to information, confidence). He stressed that "user perception of topicality seems still to be the major criterion, but clearly not the only one in relevance inferences" (p.2130). The growing recognition of the social factors influencing search behavior places more weight on criteria other than content. Focusing on criteria such as validity and belief match, offers an avenue connecting research on search decision-making to the research on credibility of online content.

Research on credibility judgments of online content focuses primarily on the user's interaction with the content of target web pages. Early research in this area suggested that people judge online content mostly based on appearance. Fogg et al. (2003), for example found that "design look" was by far the most frequently attributed feature in credibility judgment across different types of website and Flanagin and Metzger (2000) found that web users do not invest much in verifying online information. There are numerous models explaining the process of credibility evaluation. Hilligos and Rieh (2008), for example, offered a credibility judging model consisting of three layers (construct, heuristics, and interaction) where credibility judgment is built through continuous and repeated engagement with online content. Metzger (2007) proposed a dual processing model of credibility assessment under which more motivated and able users will engage in a more thorough evaluation of credibility compared to the less motivated and less capable users. Similarly Pirolli's (2005) information foraging theory suggest that people ''tend to optimize the utility of information gained as a function of interaction cost'' (p.351). Taken together,

these models suggest that web users strive to optimize they cognitive effort and time to perceived quality of the outcome ratio.

Most recently, there is a growing focus on the contextual and social aspects of credibility judgments. Flanagin and Metzger (2000) as well as Hilligos and Rieh (2008), for example, found that the type of task influences how much people are willing to invest in verifying the credibility of information. Metzger et al. investigated how social affordances of the Web are integrated in credibility judgments; they identified four main strategies: social information pooling, social confirmation of personal opinion, enthusiast endorsements, and resource sharing via interpersonal exchange. Although the way scholars think about and study credibility has significantly changed over time and grew into a more sophisticated and nuanced approach, most of this research still focuses on evaluation of online content on a particular web page.

We argue that when people rely on search engines as their primary gate into the Web, credibility judgments often take place before users reach a particular web page—they happen on the SERP. Unlike target pages, however, SERPs contain limited cues that searchers can utilize in their assessment. Limited knowledge of the users regarding how search engines work (i.e. how the ranking system works including prioritizing of sponsored and non-sponsored links) further complicates credibility judgments at this preliminary stage (e.g. Hargittai, 2007; Jansen & Spink, 2007). Arguably, this limitation prevents optimal use of the search systems, but ubiquity of search in everyday online activities, suggests that the users still derive utility from their imperfect use of the systems. In other words, the users make sense of the search systems based on what they do know, their past experience with search engines, as well as numerous contextual factors.

Overcoming the challenges

This study belongs to a line of research asking to address the limitations of the primarily data-driven and highly controlled approaches. This line of research builds on the studies of relevance in Information Retrieval (IR) (Cooper, 1971; Park, 1993; T. Saracevic, 1975) and the attempts to establish links between the real information needs of the users and their subsequent search behavior (Kelly, 2009). This line of research ventures beyond the realm of search platform and involves contextual factors and tacit knowledge. Teevan et al. (2004) is a good example of a study focused on people's search behavior in their natural settings. Following the logic of diary studies, the authors conducted semi-structured interviews in which participants reported their most recent search activity. Teevan et al. observed that instead of jumping directly to their information target using keywords, the participants navigated to their target with small, incremental steps, using their contextual knowledge as a guide, even when they knew exactly what they were looking for in advance.

Hargittai (2002) used a similarly contextually-rich, yet more controlled, approach where she conducted structured observations of how people find information online. Her study suggested that people's ability to find information on the Web is a function of complex set of contextual factors including their technical and informational environments and the level of their relevant skills (e.g. ability to use browser's navigational features, ability to enter valid search terms, etc.). In a more recent study, Hargittai et al. (2010) combined structured observations with elements of think-aloud technique to study the entire process of information seeking of young adults: from search engine selection, through the evaluation of search results, all the way to the final destination. The researchers found that "the process of information-seeking is often as important as verifying the results when it comes to assessing the credibility of online content" (p.479), specifically that the participants tend to rely extensively on the search engine rankings. Moreover, the researchers suggest that the lower levels of information literacy are associated with higher trust in the search engines. Another finding of the study suggests that the participants have information seeking routines and that those routines are built primarily around brands, such as specific search engines or information repositories (e.g. Wikipedia).

While studies of information-seeking behavior in electronic environments and in context shed some light on what facilitates particular search choices, there are still more questions than there are answers in this domain. One factor that limits the ability of current research unpack the search-related decision-making process is the fact that these studies do not actually aspire to do that. For example, Teevan et al. (2004) aimed to explore the range of orienteering behaviors, Hargittai (2002) explored the question of internet skills, and Hargitai et al. (2010) focused on online content evaluation. Another limiting aspect of the existing inquiry lies in methodology. While studies like Teevan et al. pay a lot of attention to the actual information environment of their participants and the real-life information seeking tasks, they rely primarily on user-reported behavior. At the same time, studies like Hargittai et al. do collect both user-reported and actual behavior of their participants, but this comes at the cost of context – first, the participants are forced to operate in an alien information environment and second, as much as the tasks are close to the real-life situations, they are still artificial for the individual participants. In this study we aspire to overcome some of these limitations.

METHOD AND DATA

Conceptually, we focus specifically on instances of decision-making where the users decide which search results to click on. Methodologically, our study is based on recall of actual activities the participants have gone through and their enactment of those activities in their own technical environment. While this approach has its own limitations and drawbacks, which we will discuss later, it allows investigating real-life situations in information environment our participants are familiar with and are trustful of.

The data discussed in this paper were collected in May-July 2010 as part of a larger international study in an urban private university in North-East of the US. The participants were recruited among undergraduate students, who met a series of criteria, namely: 25 years old or younger, native speakers of English, and owned a laptop, which they were willing to use in the study. Earlier research found different variations of these factors to be related to people's abilities to navigate online information and search (Eamon, 2004; Hargittai et al., 2010; Kralisch & Berendt, 2004; Zhang & Chignell, 2001).

The participants were recruited through campus advertisements and classroom announcements. They were offered a $15 compensation for their participation in the study. While we are aware of the limitations of our sampling, particularly when it comes to generalizability, we believe it captures a rather diverse group. A total of 32 participants took part in the study, split evenly between male and female participants. The range of ages span from 18 to 25 years old with an average age of 20. Almost half of the participants (47%) were Caucasian, 20% African-American, 17% Asian-American, and 13% Hispanic. 38% of our participants came from households where at least one of the parents has completed at least some college education, 28% came from households where at least one of the parents held a graduate degree, and 34% from households where none of the parents had any college education. 66% of the participants lived in dorms at the moment of the study, 22% rented, and the rest had a different living arrangement. As mentioned above, all of the participants owned a laptop. The oldest laptop was purchased in 2005, the newest in 2010, and on average the participants used machines that were about 1 year old. Among the participants, 56% were Mac users and 44% were Windows users. None of the participants used Linux or other operating systems.

An average data collection session lasted between 45 and 60 minutes and included seven parts. Part one consisted of a structured questionnaire about the technical environment of the participant including the kind of equipment used in the session, operating system, browser, email client, etc. The questionnaire was administered by the researcher. Part two consisted of a structured questionnaire about internet use focused on time spent online and frequency of visiting various types of websites and engaging in different online activities. This part was also administered by the researcher and focused on activities related to school work, religion, hobbies, politics, health, and more. Part three of the session included a

structured, self administered questionnaire aimed to survey the participants' online skills. The questionnaire included knowledge questions as well as questions found to be good proxies of digital literacy (Hargittai, 2005, 2009).

Parts four and five of each session included a semi-structured observation of the participant's behavior (Hargittai, 2002) based on their recall of past activities (Teevan et al., 2004). Our method draws on what Kelly (2009) calls "spontaneous and prompted self-report" (p.89), which is a technique of collecting data from subjects while they engage in search and the observer tries to elicit feedback about their search behavior. Subjects are not required to continuously verbalize their thoughts (as with think-aloud), but are instead asked to provide feedback at fixed intervals or when they think it is appropriate. The purpose of this technique is to get more refined feedback about the search that can be associated with particular events, rather than summative feedback at the end of the search.

Part four of the data collection session focused on the participants' recall of their online routines. The participants were asked to walk the researcher through their daily online routines, enacting their browsing behaviors and explaining them out loud. Part five utilized the participants' responses to the questionnaire in part three to prompt them about various instances of visiting websites and engaging in different online activities. As before, the participants were asked to think-aloud as they enact their online behavior. In regard to the users' search activities, the researcher asked the users about their decisions to follow links from the SERPs they reached when executing a query as part of their routine or as part of recalling their past online activities. Following the observation, the participants were asked to answer a few more questions related to their digital literacy. These questions were organized in a structured questionnaire administered by the researcher. Finally, the participants filled out a self-administered questionnaire about their demographics.

Parts four and five of the session were captured on both video and audio, including the participants' laptop screen. These steps later allowed extracting both quantitative and qualitative data about the participants' behavior and their explanation of thereof. This paper builds on the qualitative analysis of over 30 hours of recorded material. On average, each participant had 7 instances of search behavior during an observation session, which provided us with a rich account, not only of their behavior, but of their rational of search-related decision-making. While we have not requested use of a particular search engine, all our participants used Google as their search engine of choice.

USER DECISION-MAKING ON SERP

As mentioned above, this study focuses explicitly on the decisions the users make after completing a query and facing a list of search results. While our goal was to capture a wide range of search-related decision-making patters, there are some behaviors that stood out in light of existing literature. For example, we could repeatedly observe our participants focusing on the top results and rarely venturing beyond the first SERP. This observation is consistent with numerous eye-tracking studies, such as Lorigo et al. (2008), and in-person observations, such as Hargittai et al. (2010). One participant openly stated, "I would go to the first hit, rarely I would look at the title or the URL, I just use Google's decision making." But other participants, based on their commentary and responses to the question of what helped them make the decision to follow through with a particular search result, seemed to utilize ranking order as only one piece of information to be considered, as opposed to trusting it blindly. We found that our participants take into account a variety of elements on the SERP, before making a decision, especially when they deal with the more complex informational queries. Moreover, we observed an extensive use of contextual and tacit knowledge when it comes to evaluation of search results.

Interpretation of SERP Elements

Among our participants, the visual focus on the top search results was consistent across the board, but the decision to follow through with a result included a number of additional elements. When analyzing a SERP, our participants would often factor in ranking, but would also pay attention to discrepancies between their information need and the title of the result or the short description accompanying it (the snippet). Specifically, they would expect to find their exact search string emphasized in the title of the result or in the snippet. For example, a male participant searching for flu symptoms stated, "some of the titles seem retarded… they don't have the things I'm looking for ... I'm judging it by what the title says…I'm not interested in swine flu, only in cold, so I wouldn't go to that one." Similarly, when searching for information about privacy, a female participant pointed at the title of a result and stated, "I'm going to choose the first one because they have my exact [query] that I typed in there." In other words, search string formulation and assessment of the results are tightly intertwined and are not seen as two separate steps.

While title and snippet of the results are usually examined together, titles seem to attract the initial attention. A male participant, looking for names of two scientists, explained his decision-making processes in the following way, "You look at the title and the few sentences underneath to try and decide." When asked explicitly why he did not follow the top search result, the participant explained, "Because the title on the second one contained my research topic and the first one had 'Newman' in it – I don't know who 'Newman' is and 1999 is a bit old." Similarly, another male participant suggested that he will read the snippet only "if the title isn't bluntly obvious." A female participant, who was looking for differences between mosquito and bed bugs bites, explained her decision to follow the fifth result on the SERP by pointing at the title and saying, "It's exactly what I Googled – 'mosquito bites vs. bed bug bites'." As before, the participants were looking for an exact match between their search string and terms in the search results.

When the participants explored the snippets, they also looked for a close match to their query. For example, a male participant conducting research about presidential campaigns explained his decision-making process:

> I look for my search, bolded. I look at the bolded [words] for my search title, for my search criteria...I look at the links what's bolded and what's not... I'll see a [term] over here a [term] over there... they are not actually what I am looking for [hand-gesturing that the words are not adjacent]... So I'll click on links that have the actual phrase that I was searching for.

Similarly, absence of relevant terms in the snippet leads to a decision to reformulate a query. The following example of a female participant not to click on any of the search results for information about an artist, illustrates this point well:

> I don't think any of this stuff does [have any relevant information about the artist]. This is why I am not clicking on it. What's on the top line… [pointing at the titles] None of this has [the name of the artist] in it.

This explanation was followed by a query reformulation.

While an expectation to find their exact search string in the title of the result or in the snippet was a relatively common occurrence, the decision-making process on SERP was not always as straight forward. In many cases, our participants would pay attention to the position of their search terms in the snippet and the immediately adjacent words. For example when a male participant searched for an option to watch free movies online, he explained:

> I would go through the results and see which one says ‘free’ because not all of them say ‘free’ in the snippet... I won’t click on this one [pointing at a result] because it says ‘free trailer.’ So I will go to the one that says ‘watch movies online for free’.

In other words, the term ‘trailer’ adjacent to the term ‘free’ in this case served as a cue for lack of relevance of a particular search result.

In other instances, when examining the words adjacent to the query terms, the participant brought in additional, relevant to the search, knowledge in order to interpret the results. They would pay attention to terms, which were not originally part of their query, but when appeared in the snippet they were interpreted as an indication of relevance or lack of thereof. For a example, a male participant searching for summary of a book, explained why he felt confident in following through with a particular search result:

> I knew things about the book, but I didn’t know anything in the book. But I knew it was a novel. I knew the plot. And the work was set in 1959 [pointing at the year on the snippet]. And the setting of the area is Chesapeake Bay [pointing at the name of the area on the snippet].

Another male participant, who looked for information about Buddhism, offered the following explanation to his decision to click on a lower ranked result: “it says ‘tolerance’… so… it means how to deal with them, and that’s not what I was looking for… I was looking for their belief system and what they stand for.” A female participant who searched for an answer to question about relationships clarified her decision-making process by explaining, “It just sounds more informative. I want something that’s more social, like kids around my age… it’s like ‘my’ and it’s not ‘your’ it’s not, so formal. I like informal information.” In this case, she interpreted the words ‘my’ and ‘your’ as a qualifying characteristic of expected content. Since she wanted a peer advice she chose to opt for website that told individual stories as opposed to providing advice from an adult or a professional. We will come back to discussing the function of contextual knowledge in the decision-making process later.

Another element of the search results presented on SEPR and utilized by our participants was the URL. One participant in fact referred to the URL as the first attribute of the search results she tends to examine. Other participants referred to a number of ways in which they utilized the URLs in their search results evaluation process. On the one hand, they looked for simplicity in the URL. For example, a female participant, who searched for information about psychology, explained her behavior in these exact words: “I am mostly looking at things that have a simple URL.” On the other hand, when looking for a particular type of information, the participants interpreted the actual or the top level domain as a proxy for the quality of the content. For example, a female student looking for statistics on immigration visas, explained her intent to use a government website, “I usually want a government statistic… The URL is a part of my decision… I don’t want just .com [or] .net.” In another case, a female participant looked for a story about a horse that broke its leg. She preferred to click on a particular result because she was convinced it is a newspaper, “I will look both at the URL and what they say here [pointing at the snippet]… so this is a newspaper.” We can see how the participants use additional cues, not just ranking, to make sure the results make sense in the particular context of their search.

In our observations, what may appear as “blind” trust in the search engine rankings, happened in cases where the participants lacked or did not understand cues described above. For example, a male participant who searched for information about a medical condition explained:

> I read the short summary [snippet] down here and it looked like it was going to explain what [name of disease] is. [after clicking on it] Turns out it does nothing except it gives me a different name, so I went back and still flipped through a couple [of search results].

Following a clarifying question whether he checked the results in the order of their appearance, he explained, “As long as they looked relevant… [based on] the short description.” In other words, to state

that that the participants were mindlessly trusting the ranking system would not always be correct. Instead, systematic examination of the websites in the order of their appearance, while utilizing other available cues, becomes part of the evaluation process of the search results. For example a female participant looking for cheap clothes online, who came to a list of unfamiliar websites explained, "I just go through them all because I don't know any of these names" in order to familiarize herself with the available options. In other case, a male participant, who looked for material for a philosophy paper, explained how he deals with lack of familiar cues on SERP

> I would go to the first one and read through it and if I still don't understand what he's saying then I would go to different website [in the order of their appearance] until I'm satisfied and until I have a clear idea what he's saying and I will write it in a paper.

These findings are in fact consistent with earlier eye tracking studies, which suggested that when lacking clarity about relevance of the results, the users employ more scrutiny examining the SERP (Pan et al., 2007). They also allude to earlier findings in the field of IR, such as those described in Bates' (1989) "Berrypicking" model, which we come back to later in our discussion. One may also choose to view these findings through the lens of credibility research. In this sense, they are consistent with the expectancy violation and the consistency heuristics (Metzger et al., 2010), whereby the match between the searcher's expectations and the search results trumps the ranking suggestion by the search algorithm and triggers a more thorough review of the results or a query reformulation.

Although usually our participants made very quick decisions regarding which search results to follow on a SERP, when prompted, they exposed a complex system of decision making. A typical decision to follow through with a search result would include a combination of cues and strategies, which incorporates both explicit cues available on the SERP and what we can view as tacit and contextual knowledge, which we discuss in greater detail in the next subsection.

Tacit and Contextual Knowledge

As the last observations suggests, in making their decisions regarding a set of search results, our participants relied not only on the explicit cues, but also on implicit factors, which we describe as tacit and contextual knowledge. By tacit knowledge we mean the participants' prior experience with various websites and their expectations regarding the type and quality of the content they can expect based on that experience or other explicit cues. By contextual knowledge we mean the participants' awareness of the purpose of a particular search and its expected outcome. This classification, in fact, resonates with Park's (1993) categories of variables affecting relevance assessments. Our definition of tacit knowledge draws on Park's idea of 'internal context', which "indicates various sources deeply rooted in an individual's previous experience with content in the field and perceptions or beliefs about the problem area" (p.333). The definition of contextual knowledge, on the other hand, draws from Park's notion of 'external context', which "indicates factors that stem from an individual's search and current research," which in turn "tend to originate from the individual's view about the search goal, search process, research stage, or research product" (p.336).

Wikipedia emerged as one of the more prominent examples of our participants' use of both tacit and contextual knowledge. Even though there are numerous studies suggesting that the quality and reliability of Wikipedia information is reasonably good (Chesney, 2006), there is still prevailing opposition to the use of this resource in education (Denning, Horning, Parnas, & Weinstein, 2005). College students are known to use Wikipedia for academic purposes, but when asked they report skepticism about the quality of its content (Lim, 2009). As one of our participants explained: "The only .org I don't trust is Wikipedia. That's because they let people edit the stuff."

Consistent with previous research, most of the participants in our sample had experience with Wikipedia, were familiar with the type of content available on the site, and could recognize it easily when it

appeared on a SERP. At the same time, their view of how useful this website can be, varied by the context of their search. Thus for example a male participant, who looked for a technical solution to throttle bandwidth, explained:

> If it's Wikipedia I might go into it if I want to learn more about it [process of throttling bandwidth], but if I'm looking for a product or tool – I won't… If it's someone asking a question on a forum, this is what I usually like to go to.

In this case, the participant had expectations from content he could find on Wikipedia and employed this knowledge in the context of a particular task. A number of participants explained under what conditions they would choose Wikipedia among other search results. As one female participant explained it, "I trust Wikipedia, not to cite, but for a general idea of what I am looking for." Another female participant, who searched for information on bioluminescence for a research paper, said, "Since this is a research report I probably won't go to Wikipedia." Alternatively, she explained:

> If I want to quote something for my paper… this one looks like a good one: JSTORE. I know JSTORE … This to me looks a lot more official… If it's more of a common knowledge type thing that I wouldn't have to quote from, then Wikipedia is a good quick thing.

Similarly, another female participant, who wrote a paper on a local theatre, explained: "Wikipedia just to get general overview what it is... and [website of the theater] is their website because I figured I would get accurate information from the source." A male participant, also searching for information for a course paper, presented his justification, "Sometimes I go to Wikipedia and sometimes I don't... Cause Wikipedia – teachers don't like Wikipedia… I use it when I need biography, but don't use it for school." At the same time, another male participant, who was generally interested in Buddhism, explained, "Buddhism – I would go to Wikipedia because Wikipedia outlines everything for you even before you read the article." From the point of view of credibility analysis, our participants appear to rely heavily on predictive judgment (Rieh, 2002) and reputation heuristic (Metzger et al., 2010)

Our observations about Wikipedia also resonate with Hargittai et al. (2010) finding about young adults placing trust in brands when they search for and evaluate information online. At the same time, our observations suggest that while the participants recognize brands, they make instrumental use of them based on their prior experience and in the context of a particular search task. For example, a female participant looking for an answer to a relationship-related question, explained: "I like stuff from answerbag, yahoo answers... because other questions that I've asked brought me to this website." Similarly, a male participant, who searched how to deal with a particular medical condition, stated, "Obviously I would go to ehow.com before I will go to dizziness-and-balance.com because I know ehow.com and I know it's organized well and I'm probably gonna find the information quickly." Similar brand recognition of websites such as AllRecipes, Amazon, eBay, NY Times, About.com, etc. was continuously utilized in the decision-making process regarding the results presented on a SERP. Here again, our participants enacted predictive judgment and utilized reputation heuristics (Metzger et al., 2010; Rieh, 2002).

It is important to emphasize that brand recognition was used not only in cases where participants chose to go to a particular website, but also in cases where they chose to avoid a particular link. For example, a female participant who was looking for company information for her internship, explained why she avoided one of the top ranked results by saying, "...and then the second one is LinkedIn and I know what LinkedIn is." Another female participant, who looked for a recipe, explained:

> There are certain cooking websites that I trust more than others, for example I would be much more apt to go to the JoyOfBaking website because I know it's a cookbook, a well known cookbook. Whereas AllRecipies.com… I know everyone can submit their recipes and I don't know how good they are going to be.

In some cases, the participants seemed to be aware of the low credibility of the source they were referring to, but they choose to proceed, because of the particular information need they were asking to fulfill. For example, a female participant explained her decision to use answers.com, "It is not a credible source, but I just like to get opinionated answers." In a way, such behavior is consistent with what Metzger et al. (2010) refer to as social information pooling. The same participant later also explained why she decided not to follow the first result on the SERP, "Because I want something that's more social. This probably won't be considered credible until actual research go over there, but sometimes I like to get firsthand accounts."

As we can see from the quotes above, the participants make instrumental decisions depending on their contextual knowledge (the purpose of their search) and their tacit knowledge (prior experience with the particular site in question or other websites). This is consistent with previous research such as Park (1993), who found that an individual researcher's perceptions and knowledge about journals, as well as about authors and their previous works and affiliated academic programs or institutions, can influence a decision on which bibliographic citation is more relevant. This is also suggests that the iterative view of credibility judgments (Hilligoss & Rieh, 2008; Rieh, 2002) can be applied not only to the evaluation of the content of the web pages themselves, but also the evaluation of SERPs at the point of decision making about which result to follow.

Jumping-off Points

Another search behavior that came out of our observations is similar to the "Berrypicking" model proposed by Bates (1989) more than 20 years ago. This model suggests that "the query is satisfied not by a single final retrieved set, but by a series of selections of individual references and bits of information at each stage of the ever-modifying search" (p.409). We touched on this behavior when we discussed cases where our participants did not have enough cues to commit to a single search result, and instead resolved to systematic examination of a number of them in sequence. In addition to that observation, we noticed that for certain searches, when the participants had a gap in their knowledge, they would make use of what a number of them referred to as a "jumping off point." In these cases, the participants would visit a website to gather initial terms to be used in subsequent queries. This would usually be at an initial stage of their research, where they still did not have enough substantive knowledge on the topic in order to perform an effective search that fully satisfies their information need. For example, one male participant noted:

> It's just the beginning of my research… I'm studying presidential campaigns [typing in query 'presidential campaigns'] so I'll just go here and this is just to get background information on what's going on and see where I can go from here… it's kind of like a funnel."

Contextual and tacit knowledge come into play in this type of searches as well, when our participants knew what type of websites would be able to assist them in establishing a "jumping off point". Once again, Wikipedia was mentioned as a website that can serve this purpose. A female participant searching for biology related information said:

> "…and even though you know not to always trust what you read on Wikipedia it's a good jumping off point… I would come here and find out… two different types of beetle families… and so now I had this term, I would probably copy and paste that and then do a Google search on that to get better and more specific info".

Another participant, who was writing a psychology paper, noted, "Wikipedia I can't use for a paper, but I'll read about it and then if it tells me something that I could search, then I'll search that…" The same participant, while performing a different search on foreign policies, went to Wikipedia first and explained: "… just see if they have general information because I don't know enough and I'm just looking for keywords to take me to the next step." The participants seemed to be cognizant of the popular

critique of Wikipedia, yet given the ease of use and the great utility offered by this resource, they found a way to integrate it in their search practices (see Lim, 2009 for an extended discussion).

Being able to verify search results through consistency checks also came up as an important aspect, especially in searches related to health. Metzger et al. (2010) referred to this strategy as the consistency heuristics for judging credibility. When our participants were not certain about the relevance or credibility of answers they were getting from the search, they would look for consistency across several websites. For example, a male participant, who was looking to resolve a problem of excessive cough, said:

> "I start with the first [result] and try to look for consistency… if I see the same thing over on each page then it must be… I look up 15 websites and then 10 say the same thing and 5 are all different from themselves, I think that other 10 are legitimate, who have a doctor writing these things."

A female participant, who wanted to look for side effects of a certain medication, searched 'headaches' and explained:

> "I've had headaches, so I'll search 'headaches' or I'll search side effects… I'll go to Wikipedia, I'll go to a couple of blog sites, drugs.com, and then the things I see all across the board… If I only see it on one of the websites I won't take into account, but just the consistency through the websites… just because… if Wikipedia has the same as everything else it's just gonna help me keep searching for different things."

As we can see from the examples in this and the previous sections, the process of checking for consistency would usually be manifested in serial clicking on links, one after another in the order of their appearance on the SERP, especially if the users are unfamiliar with the websites that came up in the search. This finding is in line with previous research by O'Day and Jeffries (1993), who characterized the information seeking process by presenting "triggers" and "stop conditions" that guide people's search behaviors and who found that people often perform comparisons between the results in order to find consistency.

CONCLUSIONS AND FUTURE RESEARCH

In this chapter we reported on a study of how college students interact with SERPs and what drives their decision making process while this interaction occurs. We found that there are various elements in the snippet and the title that the participants are taking into account in order to decide which search results to follow. They had certain expectations for the results, such as finding their exact search string in the title of the result or in the snippet. Looking at the position of their search terms in the snippet and the immediately adjacent words was also a rather common practice. Their decisions were also influenced by terms which did not appear in the query but did appear in the snippets or the titles. In addition to these elements, there was an aspect of tacit and contextual knowledge that guided the decision making process. We found that the participants' familiarity with the nature of the websites that appeared in the SERP affected their decisions based on the expectations that they had from the content on these websites. Some of these websites were considered by the participants to be good "jumping off points" for further exploration and reformulation of their queries.

As noted in the methods section, the data for this study was collected as part of a larger study on the young adults' online routines. As such it encompasses a number of limitations that offer concrete avenues for future research and additional analysis of existing data, which did not fit the scope of the current paper. Since this study was based on recall of actual past activities, we had limited control over the type of queries that were executed by the participants. Therefore, many of the queries were simple

navigational queries, as opposed to more complex informational ones. On the one hand, this behavior reflects the participants' real life search behavior, thus adding to the external validity of our findings. On the other hand, it could be interesting to see what cues the participants would incorporate into their decision making process for more complex queries that require substantive effort and multiple reformulations. Moreover, it could be beneficial to investigate how these cues assist in guiding the reformulation process. Also, the fact that the data collection was based on recall, forced the participants into trying to remember the exact query they used, which may have introduced bias and may not adequately reflect their first-time interactions with the query and the search results. Incorporating these lessons in future studies will help producing a more robust inquiry. In addition, future research should look into the difference in the elements that users pay attention to as a function of their expertise in search engines and their online skills.

This project offers a detailed insight into the decision-making processes of young English-speaking searchers. It highlights that credibility judgments do not happen solely on the content web pages, but they are employed already at the stage of sifting through information sources via a search engine. The study demonstrated that many of the credibility judgment models and heuristics identified as suitable for evaluation of online content, can be also applied to the evaluation of the search results on a SERP. Given the limited information about the target websites provided on the SERP, the reputation heuristic seems to be the dominant strategy, but searchers also incorporate other approaches in their search decision-making. It is important to view this decision-making process as a part of a broader online experience where searchers form impressions about the reputation of websites (or online brands) and develop knowledge about the variety of needs that can be met by different websites. The decision-making process on the SERP is a manifestation of that accumulated experience and the ability to assess credibility based on a limited number of cues provided on the SERP. As such, we are asking to advance an experiential view of the Web, as an environment one needs to engage with in order to make smart choices in navigating it.

We hope our study will serve as a stepping stone for better understanding of search behavior and the related decision-making and credibility judgment processes, thus reinforcing human-focused research in IR. At the same time, we hope our findings can be also used to inform future studies of better SERP design and improvement of the search experience. Particularly, we want to highlight the need for elements that offer signals for the identity and credibility of the source on the other end of the link, as people engage Iin credibility judgments already on the SERP. For educators, this study asks to challenge the commonly-sound claim that young adults have an almost blind belief in the ranking algorithms of search engines, particularly that of Google. Instead, it draws a picture of rapid, yet complex decision-making, which utilizes numerous cues and invokes tacit and contextual knowledge. Thinking about educating for critical and thoughtful use of the search engines, the emphasis should be on creating opportunities for the users to accumulate search experience in environments where they can receive feedback on cues interpretation. Those seeking to promote thoughtful engagement with the web should avoid categorically dismissing websites or categories of website, such as Wikipedia or user-generated content, but instead allude to the various uses different sites can serve in the process of information discovery and query reformulation.

ACKNOWLEDGMENTS

The authors are grateful to the former Information+Innovation Policy Research Institute at the Lee Kuan Yew School of Public Policy, National University of Singapore, and especially to Professor Viktor Mayer-Schönberger, for their support and funding of this research.

REFERENCES

Agichtein, E., Brill, E., & Dumais, S. (2006). Improving web search ranking by incorporating user behavior information. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 19–26). Seattle, Washington, USA: ACM. doi:10.1145/1148170.1148177

Agichtein, E., Brill, E., Dumais, S., & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 3–10).

Aula, A., Majaranta, P., & Räihä, K. J. (2005). Eye-tracking reveals the personal styles for search result evaluation. *Human-Computer Interaction-INTERACT*, 1058–1061.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, *13*(5), 407–424.

Chesney, T. (2006). An empirical examination of Wikipedia's credibility. *First Monday*, *11*(11). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1413/1331

Cooper, W. S. (1971). A definition of relevance for information retrieval. *Information storage and retrieval*, *7*(1), 19–37.

Cutrell, E., & Guan, Z. (2007). What are you looking for?: an eye-tracking study of information usage in web search. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 407 – 416).

Denning, P., Horning, J., Parnas, D., & Weinstein, L. (2005). Wikipedia risks. *Communications of the ACM*, *48*(12), 152. doi:10.1145/1101779.1101804

Downey, D., Dumais, S., Liebling, D., & Horvitz, E. (2008). Understanding the relationship between searchers' queries and information goals. *Proceeding of the 17th ACM conference on Information and knowledge management* (pp. 449–458).

Dumais, S., Buscher, G., & Cutrell, E. (2010). Individual Differences in Gaze Patterns for Web Search. Presented at the IIiX, New Brunswick, New Jersey.

Eamon, M. K. (2004). Digital divide in computer access and use between poor and non-poor youth. *Journal of Sociology & Social Welfare*, *31*(2), 91–113.

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism and Mass Communication Quarterly*, *77*, 515–540.

Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., et al. (2001). What makes web sites credible? A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 61–68). Seattle, Washington, United States: ACM. doi:10.1145/365024.365037

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of web sites? A study with over 2,500 participants. *Proceedings of the 2003 Conference on Designing for User Experiences* (pp. 1–15). San Francisco, California: ACM. doi:10.1145/997078.997097

Granka, L. (2010). The politics of search: A decade retrospective. *The Information Society*, *26*(5), 364–374. doi:10.1080/01972243.2010.511560

Granka, L., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 25–29).

Guan, Z., & Cutrell, E. (2007). An eye tracking study of the effect of target rank on web search. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 417–420).

Hargittai, E. (2002). Beyond logs and surveys: In-depth measures of people's web use skills. *Journal of the American Society for Information Science and Technology*, *53*(14), 1239–1244.

Hargittai, E. (2005). Survey measures of web-oriented digital literacy. *Social Science Computer Review*, *23*(3), 371–379.

Hargittai, E. (2007). The social, political, economic, and cultural dimensions of search engines: An introduction. *Journal of Computer-Mediated Communication*, *12*(3), 769–777.

Hargittai, E. (2009). An update on survey measures of web-oriented digital literacy. *Social Science Computer Review*, *27*(1), 130–137.

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication*, *4*, 468–494.

Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, *44*(4), 1467–1484. doi:10.1016/j.ipm.2007.10.001

Hornof, A. J., & Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays (pp. 249–256). Presented at the ACM CHI'03.

Jacob, R. J. ., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573–605). Elsevier Science.

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

Jansen, B. J., Booth, D., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, *44*(3), 1251–1266. doi:10.1016/j.ipm.2007.07.015

Jansen, B. J., Booth, D., & Spink, A. (2009). Patterns of query reformulation during web searching. *Journal of the American Society for Information Science and Technology*, *60*(7), 1358–1371.

Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, *42*(1), 248–263.

Jansen, B. J., & Spink, A. (2007). The effect on click-through of combining sponsored and non-sponsored search engine results in a single listing. *Proceedings of the 2007 Workshop on Sponsored Search Auctions*. Presented at the WWW Conference.

Kelly, D. (2009). Methods for Evaluating Interactive Information Retrieval Systems with Users. *Foundations and Trends in Information Retrieval*, *3*(1—2), 1–224.

Kralisch, A., & Berendt, B. (2004). Linguistic determinants of search behaviour on websites. *Proceedings of the Fourth International Conference on Cultural Attitudes towards Technology and Communication, Karlstad, Sweden* (pp. 599–613).

Lim, S. (2009). How and why do college students use Wikipedia? *Journal of the American Society for Information Science and Technology*, *60*(11), 2189–2202. doi:10.1002/asi.21142

Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., et al. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, *59*(7), 1041–1052.

Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management*, *42*(4), 1123–1131.

Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, *58*(13), 2078–2091.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439. doi:10.1111/j.1460-2466.2010.01488.x

O'Day, V., & Jeffries, R. (1993). Orienteering in an information landscape: how information seekers get from here to there. *CHI '93: Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems* (pp. 438–445). ACM.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823.

Park, T. K. (1993). The Nature of Relevance in Information Retrieval: An Empirical Study. *The Library Quarterly*, *63*(3), 318–351.

Pirolli, P. (2005). Rational analyses of information foraging on the Web. *Cognitive Science*, *29*(3), 343–373. doi:10.1207/s15516709cog0000_20

Poole, A., & Ball, L. J. (2005). Eye tracking in human-computer interaction and usability research: current status and future prospects. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). Pennsylvania, PA: Idea Group Inc (IGI).

Rieh, S. Y. (2002). Judgement of information quality and cognitive authority in the Web. *J. Am. Soc. Inf. Sci. Technol.*, *53*(2), 145–161. doi:10.1002/asi.10017.abs

Rieh, S. Y., & Xie, I. (2006). Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Information Processing & Management*, *42*(3), 751–768. doi:10.1016/j.ipm.2005.05.005

Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, *26*(6), 321–343.

Saracevic, Tefko. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, *58*(13), 2126–2144. doi:10.1002/asi.20681

Silverstein, C., Marais, H., Henzinger, M., & Moricz, M. (1999). Analysis of a very large web search engine query log. *ACM SIGIR Forum* (Vol. 33, pp. 6–12).

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: a study of orienteering behavior in directed search. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 415–422). Vienna, Austria: ACM. doi:10.1145/985692.985745

Zhang, X., & Chignell, M. (2001). Assessment of the effects of user characteristics on mental models of information retrieval systems. *Journal of the American society for information science and technology*, *52*(6), 445–459.

ADDITIONAL READING SECTION

KEWORDS

search, search engines, user behavior, decision making, structured observations, contextual knowledge, tacit knowledge