Characterizing and Enhancing Dual-Process Information Behaviors for Exploratory Search

Michael Zarro
College of Information Science
and Technology
Drexel University
Philadelphia PA

mzarro@drexel.edu

1. INTRODUCTION

In this work I investigate human search behaviors in the consumer health domain, where searchers are likely to need assistance from intermediaries to find useful information. Dual-process theory from social psychology is introduced to help understand exploratory search behaviors and relate them to user processing of information resources. This research is intended develop a dual-process model of exploratory search and to improve interactive search and browse tools similar to those found in digital libraries like MedlinePlus or PubMed. While the health domain is used for this work, it is hoped the findings will be transferable to other search environments.

Health information seeking is often an *exploratory search* task [18] where the information sought is ambiguous and unknown to the searcher. In contrast to single-item, or known-item, lookup tasks exploratory search features "information-seeking processes that are opportunistic, iterative, and multi-tactical" [31]. Health searchers visit many resources across the unstructured Web [30, 14], following a *berrypicking* path [1] using digital libraries hosted by commercial, government, and medical organizations while also accessing personal blogs, forums, and social media sites. Exploratory searchers learn as they go, using acquired knowledge to guide future search strategies and tactics. Improving intermediary capabilities of digital libraries with interactive design and health vocabularies should help searchers find the resources they need and understand what they find. However, the methods health consumers use to process information while performing exploratory search tasks are not yet fully understood.

Searchers are likely to engage with information resources through two processing modes explained by *dual-process theory* from the social psychology domain; the *Elaboration Likelihood Model (ELM)* [22] and the *Heuristic Systematic Model (HSM)* [6]. Dual-process theory proposes that humans cognitively analyze information in parallel via two concurrent processes. The first process, *central route* (ELM) or *systematic processing* (HSM) is cognitively intensive and requires logical analysis of content. The second process, *peripheral route* (ELM) or *heuristic processing* (HSM) is less cognitively demanding, relying on rules stored in memory, pre-existing ideas, and peripheral cues to influence analytical activity.

Systematic processing requires more effort than heuristic processing, thus the processor must possess both sufficient *motivation* and *ability*, which can be influenced by variables such as intelligence, previous knowledge, time, and environmental distractions [23]. A less motivated or less able searcher will rely more on the comparatively easy heuristic processing than the higher effort systematic processing. Health consumers are likely to have high motivation but low ability when using health information resources, affecting the balance between these two processing modes. HSM and ELM have been used to guide research for public health campaigns and to build models of trust and credibility for websites. A natural progression is using them to investigate information seeking behaviors on the Web and in digital libraries.

The research proposed herein aims to investigate the relationship of dual-process theory to exploratory search behaviors. First it will ***characterize*** information seeking behaviors drawing from information sciences, health informatics, and social psychology. Second, it will investigate opportunities to ***enhance*** the performance of information seeking environments through information system design, including developing interactive user-interfaces that utilize health vocabularies. Such systems have been shown to help searchers gain new knowledge by improving search strategy and result retrieval [21]. When design is combined with expanded insight into human information processing, it should prove possible to create digital library technologies that augment behavior by improving query formulation, search and browsing performance, and ultimately acquisition of new knowledge.

2. CONCEPTUAL FRAMEWORK

2.1 Health Information Search

A focus on web-based consumer health support began in the late 1990's around the time PubMed and MedlinePlus were launched by the National Library of Medicine (NLM). The number of health resources online has since exploded with the adoption of interactive Web2.0 technologies like blogs and wikis, presenting challenges for consumers looking for trusted sources [20]. Health information retrieved from the Web can have an impact on health outcomes [10]. However, the health information search process can be one of trial and error [28]. Searchers may use "suboptimal" search strategies [12], and encounter incorrect information, or misunderstand the information they find [17]. Recent efforts like the Consumer Health Vocabulary [35] (added to the Unified Medical Language System in 2011) and interactive system studies [21] aim to help searchers find, understand, and use information resources addressing their needs.

2.2 Exploratory Search

Exploratory search "blends querying and browsing strategies" that help a user "lookup, learn, and investigate" using parallel processes or tasks [18]. Exploratory searchers are generally: "(1) unfamiliar with the domain of their goal (i.e., need to learn about the topic in order to understand how to achieve their goal); (2) unsure about the ways to achieve their goals (either the technology or the process); and/or even (3) unsure about their goals" [31]. Health consumers starting a search are often unfamiliar with the medical topics they are researching, unaware of the resources and terminology needed to fully investigate their need, and not sure what information will satisfy their goal. They

move through iterative collecting, gathering, and learning activities in a "berrypicking" fashion [1], learning new concepts and developing expertise over time. As they pick up new pieces of information, the original query is modified or replaced with a new information need based on what was learned earlier.

Fundamental to many models of information seeking is the concept of a "gap" in knowledge as a motivation for using an information system. The information seeking process may take place over a single session, or many weeks or months, as the user moves from an anomalous [2] or confused state of knowledge to a more coherent understanding of their information need. Information system use helps close the gap through learning and exploring. Success or failure of the information seeking process occurs not through system measures like relevance or precision, but rather within the mind of the user.

Information seeking is often represented as an iterative process: 1) recognizing the need for information, 2) formulating and running a query, 3) examining the results of this query, 4) then refining the query and repeating or exiting the process. The query is a searcher's imperfect expression of the information need, and in the exploratory search context the system user learns as they iterate through search stages. As the searcher discovers new knowledge, they gain expertise in the form of learned heuristics and increased ability to systematically process information that was previously inaccessible to them. Over time increased learning and ability can lead to changes in search behavior. The information seeking process in electronic environments is "both systematic and opportunistic. The degree to which a search exhibits algorithms, heuristics, and serendipity depends on the strategic decisions that the information seeker makes and how the information-seeking factors interact" [19]. The present work aims to explore the interaction of cognitive components that influence parallel search and learning activities by examining them through the lens of dual-process theory.

2.3 Dual-Process Theory

Dual-process theory emerged in the mid 1980's to characterize human judgment and information behavior in two parallel tracks, the heuristic or peripheral route, and the systematic processing or central route. Heuristic and systematic processing modes are used in unison while investigating information, thus the recipient of a message is said to engage in dual-processing. Dual-process theory can help explain human evaluation and utilization of resources in an information system.

In this work, I use the Elaboration Likelihood Model of Persuasion [22, 23] and similar Heuristic-Systematic Processing Model [6] to examine information seeking behaviors. Although born of social psychology, dual-process theory has since become viewed as "applicable beyond this one persuasion context" [5], and [9] suggested it could help explain health consumers' search behavior. These theories have also been used to guide public health research [8, 25] and investigate trust issues for health websites [24].

Dual-process theory asserts that humans process information through two concurrent channels. First, "heuristic processing" that requires comparatively little cognitive effort. In this mode, the information recipient uses judgmental rules saved in memory, pre-existing ideas, and superficial qualities of a resource to make decisions, such as whether or not to accept the message in an informative source. Examples of heuristic processing rules include *experts are usually correct*, *well-designed websites are credible,* and *the first result on a search results page is the best result*. Second, there is "systematic processing," where the health consumer takes a more cognitively intensive approach. This requires that the message be evaluated logically, necessitating more time and effort on the part of the human receiving the message. Searchers use both processes in parallel while using an information system.

A key component of dual-process theory is the "elaboration continuum" of the ELM that describes the impact of a person's motivation and ability on their elaboration, or cognitive effort. The ELM and HSM both hold that the human processor is constrained by motivation and ability to analyze information. As motivation and ability change, the searcher will move up or down the elaboration continuum with heuristic or systematic processing taking a predominate role.

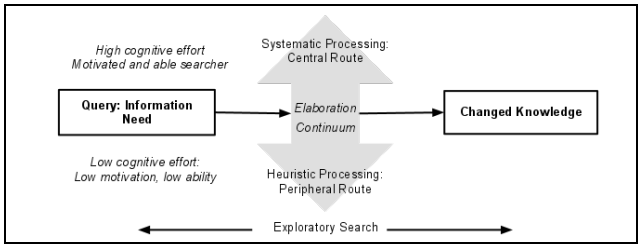

Figure 1. Dual-process theory in exploratory search.

2.3.1 Heuristic Processing: The Peripheral Route

Heuristic processing is "a more limited processing mode that demands much less cognitive effort and capacity than systematic processing" [5], requiring a minimum of effort and can be active when motivation or ability is low. Because the motivation is likely high for many health information searchers, and the ability low, it is expected that heuristics are a part of their information seeking behaviors.

Users store heuristics as knowledge structures, or decision rules, in their memory that help reduce cognitive load as they traverse an information domain. Successful use of heuristics requires their 1) availability – the user has stored them in memory, 2) accessibility – the user can retrieve the heuristic from memory, and 3) applicability – the relevancy of the heuristic to the present task [5]. Previous research has shown that heuristics are used by web searchers in many domains, including public health [13, 27]. Organizations like the NLM have published lists of rules for evaluating information that can help health consumers form heuristics for use when evaluating health resources, and as serve as guidelines for system designers.

2.3.2 Systematic Processing: The Central Route

Systematic processing is a "comprehensive, analytical orientation in which perceivers access and scrutinize all informational input for its relevance and importance to their judgment task, and integrate all useful information in forming their judgments" that requires "more than marginal levels of effort and cognitive capacity" [5]. When engaged in systematic processing, users exert substantial effort at cognitively analyzing information, and attempt to understand and evaluate the message in order to respond in a manner that advances their task goals. Attitudes developed through systematic processing are more stable than those developed through heuristic processing.

Health consumers possessing previously held information about a topic use this knowledge when comparing new arguments in the domain. For example, [32] suggests that health consumers develop expertise when managing a long-term illness. As the searcher gains experience with a chronic condition they increase

their systematic processing abilities. Additional time and cognitive effort is likely spent relating new information to previously held notions. Thus, information seekers may spend more time on later sources of information than the first Web pages they view, as found in [16]. Observations of user actions and analysis of dual-processing activity may reveal patterns that can be leveraged to improve interactive search tools and enhance searcher capabilities in future digital libraries.

2.4 Interface Design and Vocabulary

A key challenge for designers is helping non-experts use non-medical terminology to find useful and credible medical information. Information system researchers have created many experimental search and browse user interfaces that allow users to sort, filter, expand, and otherwise modify search results. The primary goal of these systems is helping guide searchers to resources while supporting exploration and learning in a concept space. Interactive search is an active and growing component for medical digital libraries. For example, the NLM sites MedlinePlus and PubMed both include features like faceted search and search suggestions while the National Institutes of Health Library incorporates an interactive semantic interface in their online search page.

Health consumers and medical experts often represent concepts differently on the lexical layer, despite an overlap on the conceptual layer [29]. Vocabularies like Medical Subject Headings (MeSH) have been found to support health consumers as part of interactive interfaces in exploratory search situations [26]. Interactive systems utilizing faceted search interfaces and health vocabularies help bridge the "vocabulary gap" between searcher and information provider that inhibits finding useful health information [35]. The "vocabulary problem" is well known to information professionals [15], and continues to be a challenge to system designers.

3. RESEARCH QUESTIONS

The conceptual framework discussed above led to the following research questions that will guide future work:

1a) How are observed information seeking behaviors characterized as simultaneous heuristic and systematic processing?

1b) What is the relationship of search stage (early to later iterations) to prevalence of heuristic and systematic processing?

2) What features of an interactive search system, such as information visualization and utilizing health vocabularies, help fulfill the intermediary function and support higher levels of exploratory search and knowledge acquisition by consumers engaged in dual-process information behaviors?

4. RESEARCH PLAN

The proposed research design is expected to be a two-part naturalistic and directed study. Participants are expected to have real-life health information needs. In part one, participants will be observed searching for real-life health needs in natural setting. Data will be recorded using unobtrusive screen-capture tools and transaction logs. Longitudinal data of search behaviors may also be collected via diaries and installation of a browser plugin to capture longitudinal data of web use over several days or weeks [4]. The second part will test interactive elements of a search system using an experimental prototype built to support behaviors observed in phase one. Experimental tasks will be structured to elicit exploratory search needs. Use of this system will be compared to a baseline system.

Study design will be influenced by dual-process theory research [5, 6, 22, 23], work in public health [8, 25], and Web research [24, 33]. Key independent or quasi-independent variables include the participant, task, system, and setting. Questionnaires and subjective user reports like self-ratings, think-aloud, and interviews can help measure cognitive effort and learning. Eye-tracking data can measure the participant's fixation on (and processing of) various experimental features [7]. Interaction data will also be captured, including; queries entered, hyperlinks clicked, and pages viewed. Experimental features may be rated qualitatively by collecting user feedback in a think-aloud protocol, stimulated recall, questionnaire, or interview. Data from different sources will be triangulated to verify observations and strengthen conclusions. Relationships among variables can help build models of dual-process behaviors for searching and browsing. It may be found that some factors may have higher influence than others during different search stages, and certain factors may mediate learning and processing while others are moderate or limit learning and search performance.

I conducted a prior research study [34] to investigate the use interactive filters in an experimental health search and browse interface. Medically trained and lay participants both reported finding terms that helped expand their searches. Users were also observed reading the list of search filters, then using terms they acquired when scanning through the search results (i.e. "*It's nice, it gives you an idea of what it is that you need to look out for"*). Lay-user and medically trained participants reported the search filters helpful and transaction logs revealed 75% of filters clicked added new information to the original search. This exploratory research suggests interactive features support health information search sessions.

5. CONTRIBUTION AND CONCLUSION

According to Marchionini, "the aim of information seeking is to get relevant information into one's head and use it in conjunction with known information to take some action or integrate it into the knowledge base. This is accomplished by coordinating information-seeking factors in systematic and heuristic ways" [19]. Human information behaviors are the mechanism by which we may observe systematic and heuristic information processing in action. This work introduces dual-process theory as a way to gain insight into information seeking practices. Improved understanding of searchers' heuristic and systematic processing can lead to responsive, interactive digital library systems that anticipate and enhance user activity.

Self-directed health information search often proves difficult due to the complexity of medical concepts. Intermediary functions are now performed in large part by search and browse features in digital libraries. Eysenbach's *apomediation* model [11] describes how health information consumers find information online using resources like search tools and social media sites in the absence of traditional intermediaries like doctors and nurses. The searcher in this model bears the final responsibility for selecting and processing information. Lost in the move to self-directed information seeking are the systematic and heuristic processing skills of human intermediaries (librarians, nurses, and doctors) with expertise in the health domain or use of information systems. As predicted by Belkin, Brooks, and Daniels, computer-based systems today "perform at least some the functions that human intermediaries perform" [3]. Digital libraries now serve as

intermediaries, supporting the user as they move through multiple search iterations. Advances in digital library technology should improve exploratory search and learning outcomes.

Currently, the interaction of heuristic and systematic processing with ability and motivation is not well understood in the digital libraries or Web search domains. A contribution of this work will be to improve the understanding of the search process and the relationship of heuristic and systematic processing to information seeking strategies and tactics. This research is intended to develop theory and models useful for contributing to the knowledge and understanding of human information processing in exploratory search.

There are practical and theoretical implications to the present work. Digital libraries incorporating technologies that enhance search behavior should help users learn and modify search tactics and strategies, leading to improved results. Researchers may build on the findings to develop future models and frameworks. Designers can make use of the findings to implement new system features supporting information retrieval. The design of these systems should be guided by state of art knowledge in design and human behavior. The research proposed herein is intended to develop theory and models to broaden knowledge and understanding of information seeking behaviors.

6. REFERENCES

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

[2] Belkin, N.J. et al. 1982. ASK for information retrieval: Part I. Background and theory. *Journal of documentation*. 38, 2 (1982), 61–71.

[3] Belkin, N.J. et al. 1987. Knowledge elicitation using discourse analysis. *International Journal of Man-Machine Studies*. 27, 2 (1987), 127–144.

[4] Capra, R. 2011. HCI Browser: A Tool for Administration and Data Collection for Studies of Web Search Behaviors. *Design, User Experience, and Usability. Theory, Methods, Tools and Practice*. A. Marcus, ed. Springer Berlin / Heidelberg. 259–268.

[5] Chaiken, S. et al. 1989. Heuristic and Systematic Information Processing within and beyond the Persuasion Context. *Unintended Thought*. J. Uleman and J. Bargh, eds. The Guilford Press.

[6] Chaiken, S. 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*. 39, 5 (1980), 752.

[7] Cole, M.J. et al. 2011. Dynamic assessment of information acquisition effort during interactive search. *Proceedings of the American Society for Information Science and Technology*. 48, 1 (2011), 1–10.

[8] Dinoff, B.L. et al. 1999. Reducing AIDS risk behavior: the combined efficacy of protection motivation theory and the elaboration likelihood model. *Journal of Social and Clinical Psychology*. 18, 2 (1999), 223–239.

[9] Dutta, M.J. and Bodie, G.D. 2008. Web Searching for Health: Theoretical Foundations and Connections to Health Related Outcomes. *Web Search*. A. Spink and M. Zimmer, eds. Springer Berlin Heidelberg. 235–254.

[10] Dutta-Bergman, M. 2006. Media use theory and internet use for health care. *The Internet and health care: Theory, research and practice*. (2006), 83–103.

[11] Eysenbach, G. 2007. From intermediation to disintermediation and apomediation: new models for consumers to access and assess the credibility of health information in the age of Web2.0. *Studies in Health Technology and Informatics*. 129, Pt 1 (2007), 162–166.

[12] Eysenbach, G. and Köhler, C. 2002. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*. 324, 7337 (Mar. 2002), 573 –577.

[13] Fogg, B.J. et al. 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* (San Francisco, California, 2003), 1–15.

[14] Fox, S. 2006. *Online Health Search 2006*. Pew Research Center's Internet & American Life Project.

[15] Furnas, G. et al. 1987. The vocabulary problem in human-system communication. *Communications of the ACM*. 30, (Nov. 1987), 964–971.

[16] Kules, B. et al. 2009. What do exploratory searchers look at in a faceted search interface? *Proceedings of the 2009 joint international conference on Digital libraries - JCDL '09* (Austin, TX, USA, 2009), 313.

[17] Leroy, G. 2008. Improving Consumer Health Literacy with Information Technology. *Encyclopedia of Healthcare Information Systems*. IGI Global.

[18] Marchionini, G. 2006. Exploratory Search: From finding to understanding. *Communications of the ACM*. 49, 4 (Apr. 2006), 41–46.

[19] Marchionini, G. 1995. *Information seeking in electronic environments*. Cambridge Univ Pr.

[20] 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 (2007), 2078–2091.

[21] Mu, X. et al. 2010. Search strategies on a new health information retrieval system. *Online Information Review*. 34, 3 (2010), 440–456.

[22] Petty, R.E. and Cacioppo, J.T. 1986. The elaboration likelihood model of persuasion. *Advances in experimental social psychology*. 19, (1986), 123–205.

[23] Petty, R.E. and Wegener, D.T. 1999. The Elaboration Likelihood Model: Current Status and Controversies. *Dual Process Theories in Social Psychology*. S. Chaiken and Y. Trope, eds. The Guilford Press. 41–72.

[24] Sillence, E. et al. 2004. Trust and mistrust of online health sites. *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04* (Vienna, Austria, 2004), 663–670.

[25] Steginga, S.K. and Occhipinti, S. 2004. The Application of the Heuristic-Systematic Processing Model to Treatment Decision Making about Prostate Cancer. *Medical Decision Making*. 24, 6 (2004), 573 –583.

[26] Tang, M.-C. 2007. Browsing and searching in a faceted information space: A naturalistic study of PubMed users' interaction with a display tool. *Journal of the American Society for Information Science and Technology*. 58, 13 (Nov. 2007), 1998–2006.

[27] Tombros, A. et al. 2005. How users assess web pages for information seeking. *Journal of the American society for Information Science and Technology*. 56, 4 (2005), 327–344.

[28] Toms, E.G. and Latter, C. 2007. How consumers search for health information. *Health Informatics Journal*. 13, 3 (Sep. 2007), 223–235.

[29] Tse, T. and Soergel, D. 2003. Exploring medical expressions used by consumers and the media: an emerging view of consumer health vocabularies. *AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium*. (2003), 674–678.

[30] White, R.W. and Horvitz, E. 2009. Cyberchondria: Studies of the escalation of medical concerns in Web search. *ACM Transactions on Information Systems*. 27, 4 (Nov. 2009), 1–37.

[31] White, R.W. and Roth, R.A. 2009. Exploratory Search: Beyond the Query-Response Paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*. 1, (Jan. 2009), 1–98.

[32] Wilson, J. 1999. Acknowledging the expertise of patients and their organisations. *BMJ : British Medical Journal*. 319, 7212 (Sep. 1999), 771–774.

[33] Wirth, W. et al. 2007. Heuristic and systematic use of search engines. *Journal of Computer-Mediated Communication*. 12, 3 (2007), 778–800.

[34] Zarro, M. and Lin, X. 2011. Using Social Tags and Controlled Vocabularies As Filters for Searching and Browsing: A Health Science Experiment. (Mountain View, CA, 2011).

[35] Zeng, Q. and Tse, T. 2006. Exploring and developing consumer health vocabularies. *Journal of the American Medical Informatics Association: JAMIA*. 13, 1 (Feb. 2006), 24–29.