Searching for Specific Health-related Information in MedlinePlus: Behavioral Patterns and User Experience

Yan Zhang
School of Information, University of Texas at Austin, 1616 Guadalupe, Austin, TX 78701. yanz@ischool.utexas.edu.

Searches for specific factual health information constitute a significant part of consumer health information requests, but little is known about how users search for such information. This study attempts to fill this gap by observing users' behavior while using MedlinePlus to search for specific health information. Nineteen students participated in the study, and each performed 12 specific tasks. During the search process, they submitted short queries or complete questions, and they examined less than 1 result per search. Participants rarely reformulated queries; when they did, they tended to make a query more specific or more general, or iterate in different ways. Participants also browsed, primarily relying on the alphabetical list and the anatomical classification, to navigate to specific health topics. Participants overall had a positive experience with MedlinePlus, and the experience was significantly correlated with task difficulty and participants' spatial abilities. The results suggest that, to better support specific item search in the health domain, systems could provide a more "natural" interface to encourage users to ask questions; effective conceptual hierarchies could be implemented to help users reformulate queries; and the search results page should be reconceptualized as a place for accessing answers rather than documents. Moreover, multiple schemas should be provided to help users navigate to a health topic. The results also suggest that users' experience with information systems in general and health-related systems in particular should be evaluated in relation to contextual factors, such as task features and individual differences.

Introduction

According to the Pew Internet & American Life Project, searching for health information is the third most popular online activity across all age groups, after e-mail and search engine use (Zickuhr, 2010). As of 2010, about 80% of Internet users in the U.S. have looked online for a wide range of health topics (Fox, 2011). Given the highly specialized nature of medicine, many questions that drive consumers to search for information are vague and exploratory, requiring the information seeker to filter, interpret, evaluate, and make inferences about the information found. In this process a great deal of uncertainty is often involved; for example, it is not uncommon for consumers to review information on serious and rare conditions in web searches even if their symptoms are in reality innocuous (White & Horvitz, 2009).

However, there is also no lack of health queries that are specific and with more clearly defined search goals. For example, what are the "optimum doses of cytomil and levoxil for a 45-year-old female" (Zeng et al., 2004, p. 49). "How long do you fast before a cholesterol test" (Zhang, 2010, p. 213). What vaccines need to be taken before a trip to Africa? What is the content of the dietary guideline recently put forth by the USDA? In information science (IS) these questions are often termed *simple tasks*, *specific tasks*, *factual tasks*, *known-item search tasks*, or *simple lookup tasks* (Marchionini, 2006; Wildemuth & Freund, 2009). In different instances, there may be variations in the completeness or accuracy of users' knowledge about answers to these questions (Cole, 2012), but the answers tend to be "discrete and well-structured objects such as numbers, names, short statements, or specific files of text or other media" (Marchionini, 2006, p. 42). Looking for specific information presents a major category of user information needs (Taylor, 1991) and is instrumental for people to gain basic health-related knowledge.

Most existing studies on consumer health information searching examined how users perform exploratory searches to learn about or investigate a health concern (Cartright, White, & Horvitz, 2011; Keselman, Browne, & Kaufman, 2008; Sillence, Briggs, Fishwick, & Harris, 2004; Zhang, Wang, Heaton, & Winkler, 2012). Few studies have paid attention to those simple and specific searches (termed *specific tasks* hereafter) that consumers carry out for concrete,

Received November 2, 2012; revised February 4, 2013; accepted February 5, 2013

factual information. Guided by the framework of task-based information searching, which postulates that variations in needs lead to variations in search behaviors (Vakkari, 2003), this study intends to fill this gap by studying users' behavior of and experience with searching for specific health information in MedlinePlus, a web-based consumer health information system developed and maintained by the National Library of Medicine (NLM). Two research questions will be addressed: (a) How do users use MedlinePlus to find answers for specific health-related tasks? (b) How do users feel about their experience with MedlinePlus for this purpose?

MedlinePlus was chosen as the platform is because it is an authoritative and credible source of health information for consumers. Great efforts have been consistently made to improve the site's information architecture to better support search and navigation (Marill, Miller, & Kitendaugh, 2006). The results based on the study of such a system could more effectively inform the design of consumer health information systems.

Related Literature

Search Tasks

A task is not only a reflection of users' information needs (Belkin, Brooks, & Oddy, 1982), but also an activity to be performed to accomplish the need (Hackos & Redish, 1998). Thus, *task* is an important contextual element that affects how users search for information and, subsequently, how systems should be designed to support the task (Cool & Spink, 2002; Vakkari, 2003). As a result, users' behavior while interacting with a system and their experience with that system need to be studied in relation to particular task attributes (Byström, 2002). Tasks have been characterized based on different attributes (Kim & Soergel, 2005; Li & Belkin, 2009), among which specificity was a popular one. Specificity is often defined in relation to task goals. Specific tasks are those with clearly defined goals (Wildemuth & Freund, 2009). In existing studies, specific tasks often equate with close-ended tasks (Marchionini, 1989b), goal-directed searches (Sutcliffe & Ennis, 1998), factual tasks (Freund, 2008), known-item searches (Kim & Allen, 2002), specific-item searches (Qiu, 1993), simple lookup tasks (Marchionini, 2006), or simple tasks (Borgman, 1986).

Although there may be variations in the length and complexity of answers to these tasks, as well as in users' knowledge concerning the accuracy of the answers (perfect match search vs. best match search) (Cole, 2012), the answers to these questions are often discrete objects (e.g., numbers, names, or short statements) that can be readily extracted from the retrieved information. In most cases, no high-level information processing, such as interpreting and synthesizing, needs to be performed (Marchionini, 2006). Thus, the specific tasks used in this study were operationalized based on three factors: certainty, distribution of answers, and cognitive effort. *Certainty* indicates that the tasks have clearly defined goals; when the answers appear, they can be easily recognized. *Distribution of answers* indicates that the answer to one particular task is located in one place; users do not need to integrate information from different pages. *Cognitive effort* refers to the fact that the tasks do not require users to interpret or synthesize the information found.

Empirical studies have suggested that task specificity has a significant impact on the search process. For example, Rouet et al. (2001) found that, in answering specific questions, students used a locate-and-memorize strategy; that is, they skimmed a number of paragraphs and made a lengthy pause at the paragraph containing the answer. In answering general questions, on the other hand, students demonstrated a revise-and-integrate pattern; that is, they examined several portions of the text and went back and forth between them to establish connections. In a later study, the same researcher found that participants had faster and more precise searches when seeking answers to specific questions as opposed to general questions, and they looked back at the questions less frequently (Rouet, 2003). Additionally, Kim and Allen (2002) found that users viewed fewer web pages, accessed fewer embedded links, and performed fewer keyword searches when completing specific known-item searches versus general subject searches.

Health Information Searching Behavior

People look for a wide variety of health-related information on the web, including information about specific diseases or conditions (e.g., stroke, depression, and pregnancy), symptoms, treatments, and lifestyles (e.g., diets and fitness), as well as information about doctors, hospitals, and health insurance (Fox, 2011; Spink et al., 2004; Zhang, Wolfram, Wang, Hong, & Gillis, 2008). One of the most extensively researched aspects of health information searching may well be query behavior. Numerous studies have consistently revealed that consumers' health queries are simple and short, with the majority containing fewer than three terms and the average query length being two terms (Hong, Cruz, Marnas, Early, & Gillis, 2002; Spink et al., 2004; Zeng, Kogan, Ash, Greenes, & Boxwala, 2002). Formulating queries is a difficult cognitive task for many consumers. Three levels of difficulty have been identified. At the term level, they misspelled medical terms; at the conceptual level, they were not able to formulate specific keywords to describe their intentions; and at the mental model level, they tended to misunderstand their current conditions (Boden, 2009; Zeng et al., 2004; Zhang, 2010). These difficulties could lead to failed searches and subsequent user frustration (Zeng et al., 2004).

Some researchers have examined the process or pattern of health information searching. Search patterns are a set of temporally and semantically related moves (Wildemuth, 2004) that reflect processing differences (Borgman, 1986). One common means to characterize search patterns is to examine users' transitions between different states of a system or between different actions, such as

DOI: 10.1002/asi

search, display, new term, and reorder (Borgman, 1986; Marchionini, 1989a). To capture various actions in temporal order, participant observation has often been used. For example, Toms and Latter (2007), by observing participants searching for health information using web search engines, characterized the search as a trial-and-error process. Sillence et al. (2004), after observing menopause patients searching for disease-specific information on the web, reported that participants often began with simple searches using single keywords, then modified their searches by making use of Boolean operators or by altering search terms. Similarly, by observing users accessing MedlinePlus to search for exploratory tasks, Zhang et al. (2012) analyzed participants' search patterns by examining transitions between the two basic methods of interaction: browsing and searching (Belkin, Marchetti, & Cool, 1993). They found that when the complexity of a task increased, the number of transitions between the two techniques also increased.

The other means to explore search patterns is to analyze the path of query reformulations (Wildemuth, 2004; Zhang, 2008). Because this type of study focuses on changes in queries, the query log analysis method is a natural choice. For example, Cartright, White, and Horvitz (2011) analyzed a large number of logged search queries targeted at diagnosing illness based on symptoms. They found that users oscillated between different foci during the search, and demonstrated two search patterns: evidence-based and hypothesis-directed. The former centered on looking for details and relevance of symptoms, whereas the latter focused on pursuing content on facets of one or more illness (i.e., risk factors, treatments, and therapies) and on the differentiation between different diseases. The two patterns were not independent of each other, and transitions between them could occur.

Accessing results is another aspect of health information search behavior that has been examined. There have been conflicting results concerning this behavior. Eysenbach and Kohler (2002) observed that their participants tended to choose one of the first results, with only a few checking out results beyond the first page. But Toms and Latter (2007) found that their participants accessed, on average, 5.4 pages of results, with only 16% examining just the first page. Users employ various strategies in determining which result to click on. Toms and Latter (2007) also reported that their participants relied on scanning the titles, summaries, and URLs, while paying little attention to information such as dates, size of the site, and types of files. Sillence et al. (2004) also reported that menopause patients examined source identifiers that indicate quality, such as URLs. Zhang (2012) reported that college-age consumers with specific health concerns looked for medical terms closely related to their intentions, such as symptoms, treatments, side effects, or terms indicating source credibility, like *MD* and *doctors*. Numerous studies also revealed that when examining search results, users simultaneously evaluated both the design features of health websites (e.g., layout and visual appearance of a site) and the clarity and simplicity of the language as indicators of the credibility and usability of the information (Eysenbach & Kohler, 2002; Sillence et al., 2004; Toms & Latter, 2007).

Knowledge about users' behavior while using existing consumer health websites is crucial for designing high-quality health information systems (Eysenbach & Kohler, 2002). Because search tasks are part of the context shaping and assigning meanings to information searching behaviors, the behaviors should be studied in relation to task characteristics. This study is an attempt to examine users' behavior while using MedlinePlus to solve search tasks with specific and clearly defined goals.

User Experience (UX)

The information retrieval (IR) research community has a long tradition of evaluating IR systems. At the early stage, the most common evaluation measure was system performance, particularly the system's ability to retrieve documents matching the subject matter of a query, in terms of precision, recall, efficiency (e.g., search time), and utility (worth and value of search results) (Harter & Hert, 1997; Su, 1992). As IR systems became increasingly interactive and the general public became the main users, the goal of IR system design began focusing more on developing systems that could actively engage users in the search process (Marchionini, 2006). At the same time, a user-centered and holistic evaluation approach began to gain momentum. User-centered measures, such as contextual attributes (e.g., characteristics of users and information tasks) and users' behavior while interacting with systems, were adopted in IR evaluation (Kelly, 2009; Su, 2003).

In recent years, UX has become a central target for interactive system design. UX is an overarching concept that subsumes an individual's entire interaction with a system, including the thoughts, feelings, and perceptions resulting from the interaction itself (Tullis & Albert, 2008). The rise of this concept has prompted both researchers and practitioners to evaluate interactive systems by going beyond traditional pragmatic functionality and usability to include hedonic aspects of the systems—for example, whether the system is stimulating, beautiful, engaging, and pleasurable (Karapanos, Hassenzahl, & Martens, 2008). In IR studies, common UX measures include usability, cognitive load, and users' feelings, such as enjoyment and engagement (Capra, Marchionini, Oh, Stutzman, & Zhang, 2007; Hu, Ma, & Chau, 1999; Kelly, 2009; O'Brien & Toms, 2008). Most of these evaluations were elicited from users' self-reports, commonly in the form of anchoring on a Likert scale (Kelly, 2009). In this study, to achieve a more holistic view of users' experience with MedlinePlus, in addition to the two pragmatic aspects of UX—usability and usefulness—three additional aspects were measured: users' perceptions of the interface, understanding of the system's working mechanisms, and emotional reactions to the system (engagement and enjoyment).

DOI: 10.1002/asi

Research Methods

The participant observation method was adopted because it allows researchers to examine in detail users' behavior while using MedlinePlus. This section describes the setup of the study.

Participants

On IRB approval, a recruitment e-mail inviting the participation of undergraduate students without MedlinePlus experience was sent to a campus-wide listserv. As a result, 19 students, 10 males and 9 females, were recruited. They were from 14 different major fields of study, such as art history, economics, English, education, geography, history, math, and psychology. Each participant was compensated with $20.

System: MedlinePlus

MedlinePlus (www.medlineplus.gov) is created and maintained by the NLM to provide consumer health information to the public. The information on the site includes publications of medical research sponsored by the National Institutes of Health (NIH) and publications of professional medical societies. The site has also licensed sources, including medical dictionaries, encyclopedias, information about drugs and supplements, directories of healthcare providers, and health-related news, from third parties. Its homepage at the time of the study is shown in Figure 1.

These major information sources were listed on the left column of the homepage and information in most of these sources was organized in an alphabetical order or by subject. A simple site-wide search was placed beneath the banner.

Search Tasks

Participants were asked to complete 12 specific tasks in this study. To derive tasks that reflect real information needs, 30 questions/tasks were selected from the health section of

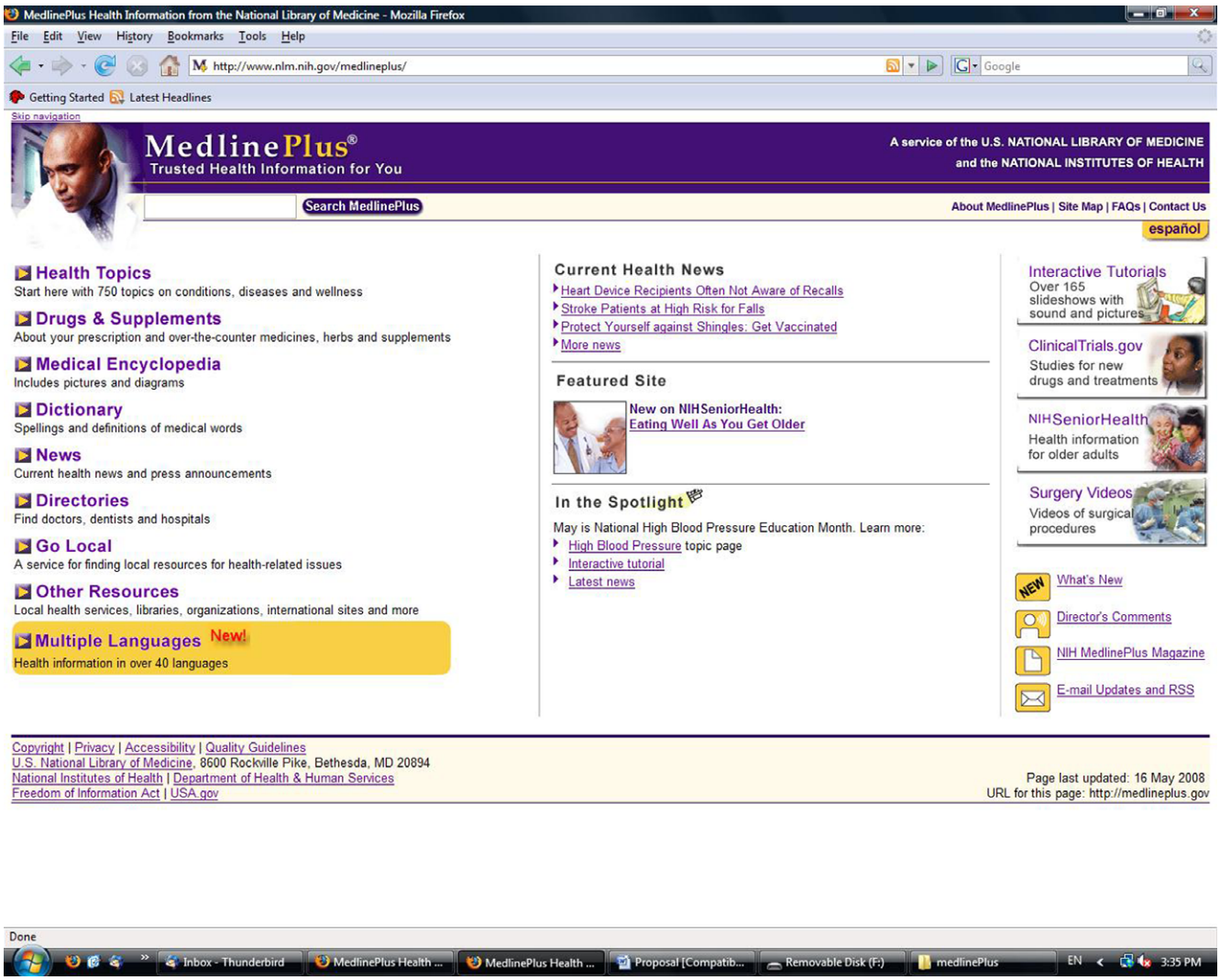

FIG. 1. Homepage of MedlinePlus. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

DOI: 10.1002/asi

Yahoo! Answers, a social Q&A site where general consumers ask questions of other consumers. For each task, a scenario was provided to make it more realistic for the participants (Borlund, 2000). To ensure that the tasks were correctly categorized as simple tasks that required specific information, two medical information professionals were asked to perform the tasks using MedlinePlus and rate the specificity of each based on three criteria: (a) the clarity of the information required to answer the question; (b) the distribution of the answer on the site; and (c) the extent to which high-level cognitive activity is required to complete the task. The tasks included in this study were those rated by both information professionals as specific tasks. That is, the information required to answer these questions was clear and easy to recognize; the answers were located on one page; and little high-level information processing was required to complete the tasks. Here are three example tasks:

> One of your friends told you that she been exposed to hepatitis B because her husband is positive with hepatitis B. She was immunized some time ago with a hepatitis B vaccine. But now she is wondering how long a hepatitis B vaccine is usually good for.
>
> Protein is a "building block" nutrient. Your body uses protein to build tissue, such as white and red blood cells, other cells in the immune system, skin, hair, and muscle. Given the importance of protein in your body, you want to find out how much protein an average person needs each day.
>
> Years ago, if a woman was HIV+, family planning was the last thing on her mind; HIV and having a child just didn't mix. The fear of transmitting HIV to her unborn baby was too great for most women. But with the advent of HIV medications, you wonder whether it might be possible for an HIV positive woman to carry a baby. You decide to find out what are the possible ways to prevent an HIV positive mom from passing HIV to her unborn child.

User Experience

Guided by the human-computer interaction (HCI) perspective that UX is a concept subsuming every aspect of user-system interaction, the UX with MedlinePlus was measured using a questionnaire that covered six distinct aspects: usability, usefulness, interface design, cognitive understanding of the system, enjoyment, and engagement. These aspects can be categorized into three dimensions:

- Dimension 1: Perceptual experience
  - Usability: (1) Easy to learn; (2) Easy to use
  - Usefulness of the content
  - Interface design: Whether users feel that the interface of MedlinePlus has a clear design
- Dimension 2: Cognitive experience
  - Cognitive understanding of how the website works: Whether users feel that they can develop a coherent mental model of MedlinePlus
- Dimension 3: Emotional experience
  - Enjoyment of the website
  - Engagement of the website

The questionnaire consisted of a series of 25 statements adapted from Davis (1989) and Capra et al. (2007). Participants were asked to rate their degree of agreement with each statement on a 5-point Likert scale (1- *strong disagreement*; 3- neutral; 5- *strong agreement*).

Procedure

All participants were tested individually and the data collection sessions took place at a private lab. A desktop computer with mouse and keyboard as input devices was used for data collection. The computer was equipped with the Windows Vista operating system and the default web browser was Internet Explorer (IE). The starting page in IE was set to the homepage of MedlinePlus.

On arrival, the participant was given a brief introduction to the study. After reviewing and signing a consent form, the participant finished a VZ-2 paper-folding test (Eckstrom, French, Harman, & Derman, 1976), measuring his/her spatial ability, because previous research suggested that spatial ability might affect people's perceptions of hypertext environments (Chen, 2000; Dillon, 2000). The participant then completed a demographic questionnaire, reporting their major field of study, computer experience, and experience with health information searching. Then he or she was asked to spend 5 minutes exploring the MedlinePlus system.

After the free-form exploration, the participant was presented with 12 search tasks. The order in which the tasks were presented was randomized. No time limit was imposed. On completion of each task, the participant rated the difficulty of the task, the mental effort required to accomplish the task, and satisfaction with his/her performance, on a five-point Likert scale. The search session was recorded using Camtasia. After completing all the tasks, the participant finished the UX questionnaire and was interviewed about his or her overall perceptions of the tasks, the MedlinePlus system, and the search process.

Data Analysis

Four types of data were collected: demographics and spatial ability, video-recorded search sessions, UX ratings, and interviews. The demographics, spatial ability, and UX ratings were imported into SPSS for statistical analysis. The video recordings were transcribed. Each webpage was an analytic unit: when a participant moved to a new page, a record was created. Actions the participant performed on this page, such as typing in a query, modifying a query, or clicking on a hyperlink, were coded into the record. The content of the queries was also transcribed.

Participants' interaction behaviors with MedlinePlus were analyzed at two levels: (a) specific interaction method—specifically, searching and browsing (Belkin, Marchetti, & Cool, 1993); and (b) task—specifically, the sequence of the deployment of the two interaction methods

DOI: 10.1002/asi

in completing a task. For the searching behavior, two major aspects—query reformulation and accessing results—were examined. Query reformulation involves both operational actions (physical) and conceptual changes (conceptual) (Bates, 1990; Fidel, 1990). Three types of operational actions were identified: actions that change the concepts involved in a query, actions that change the form of a query, and actions that change the relationships between terms. Table 1a shows the coding schema.

The query reformulation schema developed by Rieh and Xie (2006, p. 756) was adapted as the initial framework for coding the conceptual changes in query reformulations, because the schema focuses on "users' intentions toward a query's meaning." This initial coding schema consisted of four types of conceptual changes: specification, generalization, replacement with synonyms, and parallel movement. At the beginning of the coding, the author coded all the query reformulation instances based on this schema. At the same time, an open coding process was adopted to allow new types of conceptual changes to emerge from the data. As a result, a new type of conceptual change—switch to a new concept—was identified and added to the schema. Replacement with synonyms, a type of conceptual change set forth in the original schema, was removed because no such instances were identified in the queries generated in this study. Table 1b shows the resulting coding schema. The coding unit for both the physical actions and the conceptual changes was a query reformulation instance (i.e., $Q_i \rightarrow Q_{i+1}$).

To examine the validity, as well as the reliability, of the coding, a second coder independently coded all the query reformulation instances using the schema (Table 1b). The percentage agreement was 95.2%. The discrepancies were resolved by discussion. The other aspect of participants' searching behavior, that of accessing results, was analyzed in terms of the number and position of results clicked and the paths through which the results were accessed.

In this study, *browsing* refers to participants' behavior of following hyperlinks to access information. MedlinePlus is a metadata-rich system. Its information architecture not only allows users to navigate easily from one source (e.g., health topics, news, and dictionaries) to the other, but also allows them to access a particular health topic through various organization schemas, such as alphabetical lists, anatomic classifications, and subject hierarchies. The analysis focused

TABLE 1a. Coding schema for operational actions involved in query reformulations.

Actions	Type	Definition	Example
Related to Concept	Add a concept	Users add a new concept to the previous query.	hepatitis b → hepatitis b vaccine
	Delete a concept	Users delete a concept from the previous query.	kidney failure heritability → kidney failure
	Replace a concept (D..A..)	Users replace a concept in the previous query with a new concept and the new query shares at least one concept with the previous query.	low blood pressure diet → low blood pressure nutrition
	Change to new concept(s)	Users discard all the concepts in the previous query and begin with new concept(s).	supplements → recommended amount of vitamin A
	Repeat	Users repeat the previous query.	hiv → hiv
Related to form	Reformat a query into a question	Users reformat the query into a question, beginning with interrogating terms, such as what, how, and when, while the meaning of the query remains the same.	hepatitis B vaccine length → how long do hepatitis B vaccines lasts?
	Change term sequence	Users change the sequence of the terms.	low blood pressure diet → diet low blood pressure
	Error correction	Users change query terms in order to correct a typing or spelling error that have made.	low blood presure dietary needs → low blood pressure dietary needs
Related to relationship	Add operators	Users add operators, such as Boolean operators, plus sign, minus sign, or quotation marks.	low blood pressure diet → low blood pressure and diet
	Remove operators	Users remove operators, such as Boolean operators, plus signs, minus signs, or quotation marks.	"low blood pressure" → low blood pressure

TABLE 1b. Coding schema for conceptual changes involved in query reformulations.

Conceptual changes	Definition	Example
Specification	Users specify the meaning of the query by adding more terms or replacing terms with those that have more specific meaning.	amoxicillin → amoxicillin side effects
Generalization	Users generalize the meaning of the query by deleting terms or replacing terms with those that have more general meaning.	kidney failure heritability → kidney failure
Parallel movement	Users do not narrow or broaden the previous query. The previous query and the follow-up query have partial overlap in meaning, or the two queries are dealing with somewhat different aspects of one concept.	heart attach symptoms → heart attack care
Switch to new concept(s)	Users discard all the concepts in the previous query and switch to new concept(s); the new query does not have overlap with the previous query.	CPR → heart attack

DOI: 10.1002/asi

on participants' behaviors of using different types of sources and accessing particular health topics.

Participants' interaction behavior at the task level examines the sequential transitions between the two interaction methods, searching and browsing, as well as the transitions between sources (e.g., encyclopedias and health topics) in the process of completing a task.

The interviews were transcribed and analyzed. Because this paper focuses on behavioral patterns and the UX, the transcripts were used mainly to help interpret certain behaviors and will not be systematically reported here.

Results

Characteristics of the Participants

The age of the participants ranged from 19 to 22 years, mean (M) = 20.95; standard deviation (SD) = 0.85. Their spatial ability score ranged from 8 to 18.8, $M = 12.99$; $SD = 3.56$, and their Internet experience ranged from 6 to 12 years, $M = 9.89$; $SD = 1.76$. Among them, nine (47.4%) searched for health information on a monthly basis, eight (42.1%) on a yearly basis, and two (10.5%) had never searched online for health information. The major sources that they referred to for health information, in the order of the frequency of mentioning, were Google, family and friends, doctors, WebMD, Wikipedia, and the university's online health sources.

Session Length and Assessment of Search Tasks

The participants spent, on average, 20.11 minutes to complete the 12 assigned search tasks, $SD = 4.68$, with about 1.68 minutes for each task. Overall, they felt that the tasks were easy, $M = 2.07$ (1—*very easy*, 5—*very difficult*), and required a small amount of mental effort, $M = 2.11$ (1—*very small amount*, 5—*very large amount*). Participants were satisfied with their performance, $M = 4.04$ (1—*very disappointed*, 5—*very satisfied*). Spearman's rho tests suggest that participants' perception of task difficulty was positively correlated with their perception of the mental effort required to complete the task, $r(17) = .57$, $p < .05$, and negatively correlated with satisfaction with their performance, $r(17) = -.593$, $p < .01$.

Searching Behavior

Searching specifically refers to interactions with MedlinePlus using queries. This behavior was examined from four aspects: basic query features, query terms, query reformulation, and accessing results.

Basic features of queries: Number of queries and query length. A total of 237 queries were submitted. Table 2 shows the breakdown of the number of queries and query terms by tasks.

Table 2 shows that the number of participants who used the searching method varied from 11 to 16 across tasks. For each task, participants issued, on average, 1 to 2.63 queries, and the average query length ranged from 1.79 to 4 terms.

Query terms. Terms are the building blocks of a query (Fidel, 1985). Table 3 shows search terms as they were submitted, as well as the frequency of appearance by task. Three types of terms were identified: keywords that convey semantic meanings; stop words that appear with high frequency and thus do not help characterize a document; and operators that connect multiple search terms.

For each task, only a few key terms were used by the majority of participants as search terms, whereas the rest constituted a long tail and were used by only one or two participants. An examination of the search terms in relation to the task descriptions revealed that close to 70% of the terms were drawn from the task descriptions. The participant-generated terms were mostly synonyms of the terms or rewordings of the interrogating phrases in the descriptions. For example, for Task 8, participants used *nutrition*, *foods*, and *dietary* as search terms, while *diet* was the term that appeared in the description. For Task 5,

TABLE 2. Features of queries: Number of participants who submitted queries, number of queries, and query length.

	No. of Participants	No. of Queries (per participant)	No. of Terms (per query)
Task 1	13 (68.4%)	29 (2.23)	100 (3.45)
Task 2	13 (68.4%)	13 (1.00)	24 (1.85)
Task 3	13 (68.4%)	20 (1.54)	80 (4.00)
Task 4	12 (63.2%)	14 (1.17)	25 (1.79)
Task 5	11 (57.9%)	17 (1.55)	40 (2.35)
Task 6	12 (63.2%)	20 (1.67)	52 (2.60)
Task 7	14 (73.7%)	15 (1.07)	28 (1.87)
Task 8	16 (84.2%)	42 (2.63)	161 (3.83)
Task 9	12 (63.2%)	15 (1.25)	30 (2.00)
Task 10	11 (57.9%)	16 (1.45)	47 (2.94)
Task 11	14 (73.7%)	17 (1.21)	40 (2.35)
Task 12	12 (63.2%)	19 (1.58)	38 (2.00)
Total		*237*	*665 (2.81)*

DOI: 10.1002/asi

TABLE 3. Search terms.

	Keywords			Stop words		Operators
Task 1	hepatitis (21)/hepatitus (8)	length (2)	frequency (1)	is (2)	the (1)	? (2)
	vaccine (21)/vaccines (2)	duration (2)	years (1)	do (1)		
	B (14)/b (12)	effective (2)	time (1)			
	how long (3)	last (1)				
Task 2	creatine (13)	side (5)	effects (5)	of (1)		
Task 3	heart (17)	association (2)	someone (1)	a (5)	what (2)	+ (3)
	attack (14)/aattack (1)	care (1)	instructions (1)	do (3)	when (1)	" " (1)
	national (3)	cpr (1)/CPR (1)	aid (1)	is (1)	of (3)	
	institutes (3)	medical (1)	first (1)	has (1)	to (2)	
	American (2)	symptoms (1)	treat (1)	how (1)		
	health (3)	during (1)	treatments (1)			
Task 4	ALS (8)/als (2)	lateral (3)	gehrig's (1)			
	symptoms (2)	sclerosis (3)	disease (1)			
	amyotrophic (3)	lou (1)	test (1)			
Task 5	protein (16)/protien (1)	recommended (1)	per (2)			of (1)
	daily (7)	amount (1)	day (2)			for (1)
	needed (3)	average (1)	intake (1)			
	dosage (2)		person (1)			
Task 6	kidney (20)	genetic (6)	heritability (1)			
	failure (13)	link (2)	risk (1)			
	disease (7)	heredity (1)	genetics (1)			
Task 7	HIV (9)/hiv (5)	transmission (1)	mother (1)			and (2)
	pregnancy (8)/pregancy (1)		positive (1)			
Task 8	blood (37)/blodd (1)	needs (5)	foods (1)	to (4)	for (2)	" " (4)
	pressure (37)/presure (2)	eat (3)	raise (1)	what (3)		and (1)
	low (36)	nutrition (2)	recommended (1)			
	diet (12)	hypotension (2)	treatment (1)			
	dietary (5)	bold (1)				
Task 9	amoxicillin (13)	effect (1)	Amoxicillin (1)	of (1)		
	side (7)	effects (6)	amixicillin (1)			
Task 10	vitamin (14)/Vitamin (1)	amount (2)	supplements (1)	for (2)	of (2)	
	A (11)/a (4)	recommendations (2)	risk (1)			
	children (4)	recommended (2)	RDA (1)			
Task 11	aerobic (16)/aerobics (1)/anaerobic (1)	excercise (3)	health (2)			and (3)
	exercise (10)/exercises (1)	excersize (1)	benefits (1)			
		excerscise (1)				
Task 12	HIV (12)/hiv (6)	AIDS (1)	overview (1)	how (1)	does (1)	and (1)
	immune (7)	affect (1)	suppress (1)			
	system (4)/systems (2)					

when participants were prompted to find out "how much protein an average person needs each day," some participants used *amount* and *dosage* to express the quantity factor. For Task 1, participants were asked "how long a hepatitis B vaccine is usually good for," and they used *duration*, *length*, *lasts*, *years*, and *time* to express the time factor.

It is worthwhile to note that some participants also referred to their personal medical knowledge base for search terms related to the main concepts in the tasks. For example, for Task 3, a question about heart attacks, one participant used *CPR* (cardiopulmonary resuscitation) as a search term. For Task 8, two participants used *hypotension* as a substitute for *low blood pressure* in their search queries. For Task 10, a question about the intake of vitamin A, one participant used *RDA* (Recommended Dietary Allowance) as a search term. Several participants also expressed the types of information they expected to find in queries, including *instruction*, *first aid*, and *overview*.

Participants used natural language in searches, as illustrated by the use of many stop words. Six participants also directly imposed questions as queries (eight queries in total). Boolean operators were rarely used. The term *and* appeared about a half dozen times, but as it was always lowercase, it is plausible that most of them were intended to be a natural language expression rather than a Boolean operator. The search symbol " " was used by several participants in two tasks and the + sign was used by one participant in one task. Although written task descriptions were provided, the participants still misspelled terms (e.g., hepatitus, aattack, protien, excersize, and blodd).

Query reformulation. The first query that a user submits to an IR system represents his/her initial understanding of the problem at hand. As the search proceeds, the understanding may change and queries may be modified to reflect such changes. In this study, query reformulation was not pervasive.

DOI: 10.1002/asi

For 10 tasks, fewer than five participants reformulated queries. For the remaining two tasks, Task 3 and Task 8, six and ten participants, respectively, reformulated queries. Query reformulation was analyzed from two aspects: operational actions and conceptual changes. Each pair of queries, Q_i and Q_{i+1}, was compared to identify the changes. There were 83 query reformulation instances, in total.

Operational actions. As mentioned in the Data Analysis section, three types of operational (or physical) actions were identified in the query reformulation instances: actions that change concept(s), actions that change the form of a query, and actions that change the relationships between terms. In some cases, one query reformulation instance encompasses two actions. For example, in the case of "amixicillin" → "amoxicillin side effects," two actions were involved: add a concept (side effects) and correct a misspelling. Table 4a shows the actions related to the change of concepts across the tasks, and Table 4b displays the actions related to the change of the query form and of relationships between terms.

About 90.4% of the query reformulation instances involved actions that changed concepts in the queries. The most performed action was adding concept(s) to the previous query (A), followed by deleting concept(s) (D) and replacing a concept in the previous query with another concept when the rest of the query remained unchanged (D. . ., A. . .). There were only five instances in which participants discarded concepts in the previous query altogether and started with new concept(s), and two instances in which they repeated the previous query verbatim.

About 15.7% of the query reformulations involved the change of the form of a query or the relationship between terms. Among these changes, the most popular one was correcting misspellings in the previous query, followed by adding operators to specify relationships between terms.

Conceptual changes. Four types of conceptual changes were identified: specification (the query becomes more specific); generalization (the query becomes more general); parallel movement (the reformulated query has a partial overlap with the previous query); and switching to new concept(s) (the reformulated query has no overlap with the previous query). Table 5 shows these changes across tasks.

About 83.1% of all the query reformulation instances involved conceptual changes. Among these changes, the majority moved along a conceptual hierarchy to make a query more specific (58.0%) or more general (33.3%). Less than 10%, in total, moved toward a parallel but partially overlapping conceptual space or moved toward a new conceptual space.

Query iterations. The remaining 16.9% of query reformulation instances did not involve conceptual changes and were termed *query iterations*. One type of iteration was replacing

TABLE 4a. Actions that changed concepts in query reformulations.

	Add (A)	Delete (D)	Replace (D . . . , A . . .)	Change to new concept(s)	Repeat
Task 1	9	4	3		
Task 3	3	1	2	1	
Task 4	2				
Task 5	3	2			
Task 6	2	2	3		
Task 7		1			
Task 8	8	6	3	2	1
Task 9	2	1			
Task 10	2	2		1	
Task 11	1		1		
Task 12	3	2		1	1
Total	*35 (46.7%)*	*21 (28.0%)*	*12 (16.0%)*	*5 (6.7%)*	*2 (2.6%)*

TABLE 4b. Actions that changed the form of a query and the relationship between terms.

	Form			Relationship between terms	
	Reform a query into a question	Change term sequence	Correct misspellings	Add operators	Remove operators
Task 1	2		2		
Task 5			1		
Task 7			1		
Task 8		1	1	3 (quotation marks) 1 (Boolean and)	1 (quotation marks)
Task 9			1		
Task11			1		
Total	*2 (13.3%)*	*1 (6.7%)*	*7 (46.6%)*	*4 (26.7%)*	*1 (6.7%)*

DOI: 10.1002/asi

TABLE 5. Conceptual changes in query reformulations.

	Specification	Generalization	Parallel	Switch to new concept(s)
Task 1	11	4	1	
Task 3	3	1	1	1
Task 4	2			
Task 5	3	2		
Task 6	3	4		
Task 7		1		
Task 8	8	6	2	
Task 9	2	1		
Task 10	3	2		
Task 11	2			
Task 12	3	2		1
Total	*40 (58.0%)*	*23 (33.3%)*	*4 (5.8%)*	*2 (2.9%)*

a concept in a query with its synonym, as synonyms are different expressions of the same concept (e.g., "low blood pressure" → hypotension) (four instances). The other type was repeating the previous query verbatim (two instances) or changing only the form of a query, such as correcting misspellings and adding quotation marks (eight instances).

Accessing results. **Number and position of the results being examined.** Participants examined 209 results from a total of 237 searches. The analysis shows that, in 71.7% of the searches, only one result was explored and, in 21%, zero were explored. Results that participants examined tended to be those ranked high on the result list. The analysis shows that 56.0% of the accessed results were the first on the result list, 18.7% were the second, 9.6% were the third, and 8.6% were the fourth. The remaining (7.1%) ranged from the fifth to the tenth. None of the participants accessed results beyond the first page.

The result list was not the only feature that participants examined on a search result page. In MedlinePlus, when searching for a specific health topic (e.g., diabetes), a brief summary of the topic (diabetes) will appear in a blurb above the result list. Both the title of the blurb (in this case, Diabetes) and a link termed "Read more" at the end of the text in the blurb link to the health topic page where all the information about the topic (diabetes) is listed. Across search tasks, the participants clicked on the title of the blurb three times and the "Read more" link eight times to access the topic page of particular conditions. It is worth noting that three participants found the answer for a couple of tasks by reading the summary.

Paths of accessing results. On a search result page, users can access results through three different paths (Figure 2):

- Accessing results directly from the search result list.
- Refining a result list by selecting a particular resource, such as Health Topics and News, from the Collection panel, then accessing results from the refined result list.
- Refining a result list by selecting a topic from a system-generated topic list in the Cluster panel, then accessing results from the refined result list.

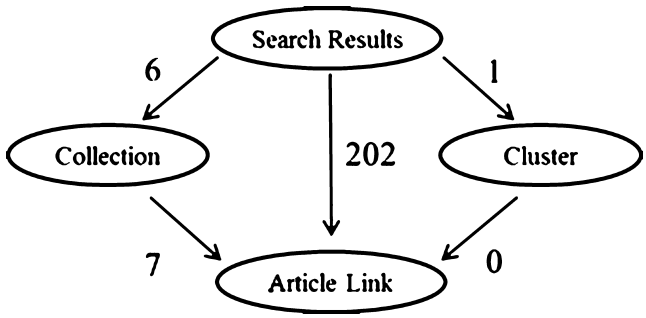

FIG. 2. Paths of accessing results.

Figure 2 shows that almost all the results were accessed directly from the result list. Types of resources listed in the Collection panel were used only six times to refine search results (Health Topics five times and Drugs & Supplements one time). Only one participant selected a system-generated topic to refine a search result, but s/he did not explore any results from the refined result list.

Browsing Behavior

Browsing refers to participants' behavior of following the information structure laid out by various information organization schemas, such as alphabetical lists and subject hierarchies. Two major browsing instances were of interest in this study: browsing to access a particular resource (e.g., the Medical Encyclopedia, Health Topics, and News), and browsing to access topics related to a health subject of interest (Related Topics).

Accessing different resources. MedlinePlus offered six major types of resources: Health Topics, Encyclopedia, Drugs & Supplements, News, Directories, and Dictionary. Of these, only the first three were accessed by the participants. Entries in both the Drugs & Supplements and Encyclopedia were organized in alphabetical order. In completing the tasks, ten participants accessed Drugs & Supplements and three accessed Encyclopedia by following the alphabetical organization schema.

DOI: 10.1002/asi

TABLE 6. Browsing to access a particular health topic (No. of participants/No. of visits).

	Alphabetical	Body location/system	Disorders and conditions	Health and wellness	Frequently requested topics
Task1	4/4	1/1	1/1		
Task 2	2/2			1/1	
Task 3	2/2	9/9			1/1
Task 4	4/4	2/3			
Task 5	2/2			4/4	
Task 6	1/1	7/7			
Task 7	4/4	1/1	1/1		
Task 8	3/3	3/3			
Task 9	2/2				
Task 10	1/2			1/1	
Task 11	3/4			4/4	
Task 12	6/6	3/3			
Total	*7/36*	*12/27*	*2/2*	*4/10*	*1/1*

TABLE 7. Browsing to access related topics (No. of participants/No. of visits).

	In-text links in summary	"Related Topics" links
Task 7	3/3	
Task 10	1/1	
Task 12		1/1
Total	*4/4*	*1/1*

Entries in Health Topics could be accessed through six schemas: (a) Alphabetical, (b) Body Location/Systems, (c) Disorders and Conditions, (d) Diagnosis and Therapy, (e) Demographic Groups, and (f) Health and Wellness. In addition, a set of links (e.g., Frequently Requested Topics, Interactive Health Tutorials, and Easy to Read) was provided to enable quick access to certain information. Among these schemas and links, five were utilized by the participants; the usage is detailed in Table 6.

The table demonstrates that the alphabetical list was used most frequently by participants, and was used in every task. The second most frequently used schema was the Body Location/Systems. However, the number of the participants who used this schema was almost double the number who used the alphabetical list. The third most used schema was Health and Wellness: four participants used it ten times while completing four tasks. The schema Disorders and Conditions, and the link Frequently Requested Topics, were used by only a couple of participants.

Accessing related topics. In MedlinePlus, each health topic page contains a summary of the topic, in which related organs, conditions, or diseases are mentioned and hyperlinked. Thus, users could access related topics through the in-text links in the summary. The second way of navigating to a related topic was to follow links in a system-provided Related Topics list. Table 7 shows the participants' usage of these two means to access related topics.

Participants did not trace related topics very often. Four participants followed in-text links when performing Task 7 and Task 10, and one used links in the Related Topics list when performing Task 12.

Interaction Path for Completing the Tasks

As has been mentioned, searching (S) the website and browsing (B) different resources were the two main strategies to accessing information in MedlinePlus. The sequence in which the two strategies were deployed in completing a task is defined as an *interaction path*. Table 8 shows the interaction paths that participants followed to complete each task. Numbers in parentheses indicate the number of participants who took a particular path.

The table indicates that the majority of the participants used either searching or browsing as the sole interaction strategy in completing the tasks, with more employing the former than the latter. Browsing was mainly used to access one of three resources: Health Topics (HT), Drugs & Supplements (DS), and the medical Encyclopedia (E). In most cases, participants were able to complete a task by browsing one resource, mainly Health Topics; in only five cases, participants had to browse two resources. For example, in completing Task 12, one participant browsed the Encyclopedia, did not find the answer, and moved on to browse Health Topics (B[E] → B[HT]). Among these resources, Health Topics was used the most. Drugs & Supplements was used only when the task explicitly involved a drug or supplement, such as Creatine (Task 2), Amoxicillin (Task 9), and Vitamin A (Task 10).

For each task (except Tasks 2 and 5), only a few participants transitioned between searching and browsing. The majority of the resultant interaction paths were short, with half beginning with browsing a source and ending with searching (e.g., B[E] → S), and the other half beginning with searching (one searched the Dictionary) and ending with browsing the Health Topics or Related Topics list. Only two interaction paths involved two transitions (S → B[HT] → S) and one involved three (B[HT] → S → B[E] → S).

DOI: 10.1002/asi

TABLE 8. Interaction paths for completing the tasks.

	N	Searching	Browsing	Transitions between searching and browsing	
Task1	19	S (12)	B[HT] (6)	B[E] → S (1)	
Task 2	19	S (10)	B[HT] (3) B[DS] (6)		
Task 3	19	S (8)	B[HT] (6)	B[HT] → S → B[E] → S(1) B[HT] → S(2)	S → B[HT] (1) S[D] → B[HT] (1)
Task 4	19	S (12)	B[HT] (5) B[E] (1)		S → B[HT] (1)
Task 5	19	S (11)	B[HT] (6) B[E] (1) B[DS] → B[E] (1)		
Task 6	19	S (11)	B[HT] (7)		S → B[HT] (1)
Task 7	19	S (12)	B[HT] (6)	B[E] → S (1)	
Task 8	19	S (14)	B[HT] (4)		S → B[HT] → S (1)
Task 9	19	S (11)	B[DS] (6) B[HT] → B[DS] (1)	B[HT] → S (1)	
Task 10	19	S (8)	B[HT] (2) B[DS] (6) B[HT] → B[DS] (1)	B[DS] → S (2)	
Task 11	19	S (12)	B[HT] (5)	B[HT] → S (2)	
Task 12	19	S (8)	B[HT] (6) B[E] → B[HT] (1)	B[E] → S (1)	S → B[HT] (1) S → B[HT] → S (1) S → B[RT] (1)

Note. **Abbreviations**:
D for Dictionary; S[D] : Searching the dictionary
DS for Drugs and Supplements; B[DS]: Browsing in the Drugs & Supplements section
E for Encyclopedia; B[E]: Browsing in Encyclopedia
HT for Health topic; B[HT]: Browsing by health topic
RT for Related Topics; B[RT]: Browsing by following related topics provided by MedlinePlus

TABLE 9. User experience ratings.

Dimensions		Aspects	Rating (Mean (SD))
1	Usability	Easy to learn	3.73 (.87)
		Easy to use	4.03 (.67)
	Usefulness of the content		4.07 (.58)
	Interface design		3.95 (.76)
2	Understanding the website's working mechanisms		3.58 (.49)
3	Emotional reactions	Engagement	3.37 (1.12)
		Enjoyment	3.32 (.75)
Intention to use the system in the future			3.92 (.71)

Note. Rating scale: 1 – *Strongly disagree*; 2 – Disagree; 3 – Neutral; 4 – Agree; 5 – *Strongly agree*

User Experience

Three dimensions of the participants' experience with MedlinePlus were measured: perceptions of the usability, usefulness, and interface design of the system; cognitive understanding of the system's working mechanisms; and emotional reactions to the system. The overall experience was measured by their intentions to use the system in the future. Table 9 shows the ratings.

Participants had a neutral to positive experience with MedlinePlus. They perceived the system to be easy to learn and easy to use; the content was useful; and the interface was clear. Nevertheless, they were not highly confident about their understanding of how the system works and were reserved about their engagement and enjoyment of the system.

Spearman's rho tests were performed to identify correlations between participants' perceptions of the systems and their cognitive and affective experiences with the system. It was found that participants' cognitive understanding of the system's working mechanism was positively correlated with their perceptions of the system's usability (easy to learn, $r(17) = .590$, $p < .01$; easy to use, $r(17) = .590$, $p < .01$). The engagement of the system was positively correlated with the ease of use of the system, $r(17) = .470$, $p < .05$, and the

DOI: 10.1002/asi

clarity of the interface design, $r(17) = .538$, $p < .05$. The enjoyment of the system was strongly correlated with the usefulness of the content, $r(17) = .507$, $p < .05$. It is worth noting that participants' emotional reactions to MedlinePlus were independent of their cognitive understanding of the system's working mechanisms.

Overall, participants expressed an intention to use MedlinePlus in the future. This intention was highly correlated with all the elements of their perceptions of the system: usability (easy to learn, $r(17) = .583$, $p < .01$; easy to use, $r(17) = .746$, $p < .01$), usefulness of the content, $r(17) = .785$, $p < .01$, and the interface design, $r(17) = .538$, $p < .05$. However, this intention was independent of their cognitive understanding of and emotional experience with the system.

The use of an information system is situated in the context of a task and constrained by the characteristics of the user. Thus, the UX should also be examined in relation to tasks and individual differences. Spearman's rho tests suggest that, when looking for specific information in MedlinePlus, the UX was not significantly correlated with task features, including task difficulty, mental effort required, and users' satisfaction with performance. One exception was that participants' cognitive understanding of the system's working mechanisms was positively correlated with task difficulty, $r(17) = .481$, $p < .05$, which indicates that as the task difficulty increased, so did participants' understanding of the system. Participants' perceptions of system usability were positively correlated with their spatial ability, which indicates that those with higher spatial ability felt the system was easier to learn, $r(17) = .478$, $p < .01$, and easier to use, $r(17) = .537$, $p < .01$. They were also more likely to use it in the future, $r(17) = .492$, $p < .01$.

Discussion

The first research question was how users use MedlinePlus to solve specific health information queries. To answer this question, several aspects of search behavior were examined: deployment of the two interaction strategies (i.e., searching and browsing) in completing tasks (interaction paths); query formulation and reformulation; accessing search results; and browsing strategies. Our research findings suggest that users' interaction paths were short and lacked transitions between interaction strategies, as most completed a task using either searching (42.1%–73.7% across tasks) or browsing (21.1%–47.3% across tasks). In only 8.3% of the instances (19 participants; each performed 12 tasks, 228 task instances in total) did participants transition from one strategy to the other, with the majority having only one transition. In examining users' behavior of performing exploratory tasks in MedlinePlus, Zhang et al. (2012) observed longer sequences of transitions between searching and browsing, as well as between browsing different resources. This comparison suggests that task specificity may have an influence on the deployment of searching and browsing strategies.

Between the two interaction strategies, searching was the dominant one, with at least 60% of the participants using it to perform each task. In information searching models, searching is characterized as analytical and is employed when users are able to articulate their needs (Marchionini, 1995; Sutcliffe & Ennis, 1998). Thus, this result is not surprising as the tasks in this study were specific with clearly defined goals. During the search, participants formulated short queries, with the average length being fewer than three terms. This result is consistent with previous studies on consumer health information searching (Spink et al., 2004; Zeng et al., 2004). In line with prior studies on web searches (Markey, 2007), this study also found that some participants formulated queries as questions. Most of those prior studies were based on transaction log analyses where the specificity of users' needs could not be inferred. Coupled with the results from Zhang et al. (2012) that no queries were formulated as questions for exploratory tasks in MedlinePlus, the observation method used in this study allows us to speculate further that users may be more likely to express their needs in questions when the needs are specific. It seems that to better support specific health-related tasks, systems should provide a more "natural" user interface to accommodate users' behavior of writing full sentences rather than artificial keywords (Hearst, 2011). This function could be particularly useful for health information searching, as users tend to have difficulties with medical vocabularies (Boden, 2009). Allowing them to provide longer descriptions of their needs will likely enhance the chance of finding information that matches their cognitive situations. Research on the search of question-answering sites has provided insight on improving results for longer queries (e.g., Bian, Liu, Agichtein, & Zha, 2008).

Query reformulations were not as pervasive and complex as those observed in Zhang et al. (2012) when users performed exploratory tasks. The majority of the reformulations here involved either moving along a conceptual hierarchy to make the previous query more specific or more general, or remaining at the original conceptual level by iterating queries (replacing a term with its synonym or changing the form of the query). Very few movements were made to expand the conceptual space horizontally by replacing part or all of the concepts in the previous query. Such behaviors suggest that in performing specific tasks, users' cognitive activities were focused more on experimenting with the level of specificity of particular concept(s) than exploring the relationships between concepts or new conceptual territories. This finding, to a large degree, supports the notion that information searching is also a learning process (Jansen, Booth, & Smith, 2009). To better support the retrieval of specific information, health information systems should be able to support the learning of concept hierarchies. A potentially fruitful approach is to enhance the systems with medical thesauri that provide hierarchical relationships among concepts, such as Medical Subject Headings (MeSH) (Blocks, Cunliffe, & Tudhope, 2006).

When accessing results, the most examined section on a result page was the result list, followed by the summary

DOI: 10.1002/asi

blurb of the health topic being searched. Only a few participants attempted to refine search results by information types. The most explored results were the top five results, and none of the participants went beyond the first page. It seemed that the participants expected to find results very quickly, indicated by their propensity to examine fewer than one result, on average, after each search; when they could not find answers within a couple of clicks, they gave up the current search to begin a new search or to browse a particular resource. It is worth noting that some participants found answers directly from the summary blurb on the results page. These results suggest that, in supporting users' needs for specific health information, the result page should serve as a place to provide answers or to offer salient interactions to the user rather than as an intermediary step to accessing documents.

The other interaction strategy, browsing, was also widely used. Browsing is particularly effective for ill-defined problems when users have difficulties in formulating queries (Belkin, Marchetti, & Cool, 1993; Marchionini, 1995). Because the tasks in this study have clearly defined goals, we speculate that participants may still have difficulty articulating their health-related needs even when task descriptions are provided. But on the other hand, because effective use of browsing requires users to have appropriate mental models of an information space, the wide use of browsing might also be attributed to MedlinePlus' well-constructed information architecture, which was reflected well in two respects. First, a clear identification of different resources on the homepage and in the global navigation appeared on almost every page, allowing participants to effectively select a source. Second, the use of both exact organization schemas (e.g., alphabetical list and anatomy) and inexact schemas (e.g., subject) to organize Health Topics allowed participants to access a topic of interest simply by locating it when it was known and by exploring conceptual structures when it was vague. In this study, the alphabetical list was used most, followed by the anatomic schema Body Location/System. It is worth noting that very few participants traced related topics to find information.

The second research question pertained to users' experience with MedlinePlus. Unlike most studies, we took a holistic approach and examined the UX from three aspects: perceptual experience, cognitive experience, and emotional experience. Consistent with the technology acceptance model (TAM) (Davis, 1989), the results indicated that users' intention to use MedlinePlus in the future was significantly correlated with their perceptions of the usability and usefulness of the system. Although their cognitive understanding of the system and their emotional experience were significantly correlated with perceptions of system usability and usefulness, they were not strong indicators of intended future use. A finding worth noting was that users' cognitive and emotional experiences were independent of each other; better understanding of the system's working mechanism was not necessarily associated with a better emotional experience. The results suggest that researchers intending to study the UX of information systems should measure both cognitive and emotional experience, as they seem to represent distinct aspects of the UX. Furthermore, as there are currently no standard means of measuring the UX of information systems (Kelly, 2009), studies are needed to tease out other possible dimensions of UX.

Our results also suggest that researchers should examine UX in relation to task features and individual differences. In this study, participants' cognitive understanding of the system was positively correlated with task difficulty. It is possible that difficult tasks require more exploration of the system. For individual differences, participants with higher spatial abilities felt the system easier to learn and easier to use, and they were also more likely to use it in the future. It is possible that users' spatial ability, as suggested by previous studies (e.g., Chen, 2000), has an impact on their ability to develop a mental map of a system, and subsequently affects their ability to navigate the system and the willingness to use it again. In future studies, researchers can explore how other individual differences influence users' interactions with health systems, particularly those related to health information seeking, such as tendency for searching for health information (e.g., monitor vs. blunter) and health beliefs (Case, Andrews, Johnson, & Allard, 2005).

Conclusions

Searching for specific factual health information constitutes a large part of consumer health information queries, but little is known about how users look for such information. This study attempted to fill this gap by examining users' behaviors and experiences using MedlinePlus to search for specific health tasks. The results indicate that searching for specific health information is not only an information-finding process (submitting queries and examining search results), but also a learning process (exploring conceptual hierarchies of health topics). To better support the search for specific health information, system designers could provide a more "natural" interface to encourage users to ask questions and an algorithm to provide more effective responses to natural language queries; effective conceptual hierarchies could be implemented to help users understand the problem at hand; and the search result page could be reconceptualized as a place for displaying answers or as a springboard for more fruitful interactions. Moreover, the system should have an information architecture that clearly delineates different resources and provides multiple schemas with which to locate a health topic.

The results also suggest that users' experience with a health information system is multidimensional and influenced by contextual factors (task difficulty and users' spatial abilities). It follows that to gain an in-depth understanding of users' experience with a health information system, researchers should consider measuring not only users' perceptions of the system's usability and their level of

DOI: 10.1002/asi

emotional engagement, but also their cognitive understanding of the system, as it is distinct from the emotional dimension and could be associated with the cognitive load that users experience during the search. Moreover, researchers should examine users' experience in relation to task features and individual differences.

The results of this study should be interpreted in the light of its limitations. First, the participants, a group of undergraduate students, were not typical users of MedlinePlus and the assigned tasks were not their real needs. Thus, the results should not be generalized to the public at large. Second, query reformulations were analyzed solely based on the queries submitted by the participants. More in-depth understanding of the conceptual changes involved in query reformulations could be achieved by asking users to think aloud during the search. Third, we observed the participants' searching and browsing behavior, without probing why they demonstrated such behaviors. Exploring cognitive reasons for such behaviors could be a theme for future studies. Future studies should also observe users searching for their own specific health information needs. Most studies in this vein have relied on transaction log analyses, which may be limited in revealing how users employ different interaction strategies as well as the reasons for those behaviors. Thus, user observations may be used in combination with transaction log analysis.

References

Bates, M.J. (1990). Where should the person stop and the information search interface start? Information Processing & Management, 26, 575–591.

Belkin, N.J., Brooks, H.M., & Oddy, R.N. (1982). Ask for information retrieval. Journal of Documentation, 38(2), 61–71.

Belkin, N.J., Marchetti, P.G., & Cool, C. (1993). Braque: Design of an interface to support user interaction in information retrieval. Information Processing & Management, 29(3), 325–344.

Bian, J., Liu, Y., Agichtein, E., & Zha, H. (2008). Finding the right facts in the crowd: factoid question answering over social media. In Proceeding of the 17th international conference on World Wide Web, pages 467–476.

Blocks, D., Cunliffe, D., & Tudhope, D. (2006). A reference model for user-system interaction in thesaurus-based searching. Journal of the American Society for Information Science and Technology, 57(12), 1655–1665.

Boden, C. (2009). Overcoming the linguistic divide: a barrier to consumer health information. Journal of the Canadian Health Libraries Association, 30, 75–80.

Borgman, C.L. (1986). The users mental model of an information-retrieval system—an experiment on a prototype online catalog. International Journal of Man-Machine Studies, 24(1), 47–64.

Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. Journal of Documentation, 56(1), 71–90.

Byström, K. (2002). Information and information sources in tasks of varying complexity. Journal of the American Society for Information Science, 53(7), 581–591.

Capra, R., Marchionini, G., Oh, J.S., Stutzman, F., Zhang, Y. (2007). Effects of structure and interaction style on distinct search tasks. Proceedings of ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL) (pp. 442–451).

Cartright, M., White, R.W., & Horvitz, E. (2011). Intentions and attention in exploratory health search. In the Proceedings of the SIGIR'11, 65–74.

Case, D.O., Andrews, J.E., Johnson, J.D., & Allard, S.L. (2005). Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. Journal of the Medical Library Association, 93(3), 353–362.

Chen, C. (2000). Individual differences in a spatial-semantic virtual environment. Journal of the American Society for Information Science, 51(6), 529–542.

Cole, C. (2012). Information need: A theory connecting information search to knowledge formation: Information Today, Inc.

Cool, C., & Spink, A. (2002). Issues of context in information retrieval (IR): An introduction to the special issue. Information Processing & Management, 38(5), 605–611.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Dillon, A. (2000). Spatial-semantics: How users derive shape from information space. Journal of the American Society for Information Science, 51(6), 521–528.

Eckstrom, R.B., French, J.W., Harman, H.H., & Derman, D. (1976). Kit of Factor-referenced Cognitive Tests. Princeton, N.J.: Educational Testing Service.

Eysenbach, G., & Kohler, 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, 573–577.

Fidel, R. (1985). Moves in online search. Online Review, 9(1), 61–74.

Fidel, R. (1990). Searchers' selection of search keys: III. Searching styles. Journal of the American Society for Information Science, 42(7), 515–527.

Fox, S. (2011). The Social Life of Health Information, 2011. Accessed on November 17, 2011, at http://www.pewinternet.org/~/media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf

Freund, L. (2008). Exploiting task-document relations in support of information retrieval in the workplace. Doctoral Dissertation, University of Toronto.

Hackos, J., & Redish, J. (1998). User and task analysis for interface design. New York: Wiley.

Harter, S.P., & Hert, C.A. (1997). Evaluation of information retrieval systems: Approaches, issues and methods," Annual Review of Information Science and Technology, 32, pp. 3–94.

Hearst, M. (2011). 'Natural' search user interfaces. Communications of the ACM, 54(11), 60–67.

Hong, Y., Cruz, N., Marnas, G., Early, E., & Gillis, R. (2002). A query analysis of consumer health information retrieval. In the Proceedings of AMIA 2002, 1046.

Hu, P.J., Ma, P., & Chau, P.Y.K. (1999). Evaluation of user interface designs for information retrieval systems: a computer-based experiment. Decision Support Systems, 27, 125–143.

Jansen, B.J., Booth, D., & Smith, B. (2009). Using the taxonomy of cognitive learning to model online searching. Information Processing & Management, 45(6), 643–663.

Karapanos, E., Hassenzahl, M., & Martens, J.B. (2008). User experience over time. In Proceedings of CHI2008, 3561–3566.

Keselman, A., Browne, A.C., & Kaufman, D.R. (2008). Consumer health information seeking as hypothesis testing. Journal of the American Medical Informatics Association, 15(4), 484–495.

Kelly, D. (2009). Methods for evaluating information retrieval systems with users. Foundations and Trends in IR, 3, 224pp.

Kim, K.S., & Allen, B. (2001). Cognitive and task influences on web searching behavior. Journal of American Society for Information Science and Technology, 53(2), 109–119.

Kim, S., & Soergel, D. (2005). Selecting and Measuring Task Characteristics as Independent Variables. In the Proceedings of the ASIST 2005.

Li, Y., & Belkin, N.J. (2008). A faceted approach to conceptualizing tasks in information seeking. Information Processing and Management, 44, 1822–1837.

Marchionini, G. (1989a). Information-seeking strategies of novices using a full-text electronic encyclopedia. Journal of the American Society for Information Science, 40(1), 54–66.

DOI: 10.1002/asi

Marchionini, G. (1989b). Making the transition from print to electronic encyclopedias—Adaptation of mental models. International Journal of Man-Machine Studies, 30(6), 591–618.

Marchionini, G. (1995). Information Seeking in Electronic Environments. Cambridge, England: Cambridge University Press.

Marchionini, G. (2006). Exploratory search: From finding to understanding. Communications of the ACM, 49(4), 41–46.

Marill, J.L., Miller, N., & Kitendaugh, P. (2006). The MedlinePlus public user interface: Studies of design challenges and opportunities. Journal of Medical Library Association, 94(1), 30–40.

Markey, K. (2007). Twenty-five years of end-user searching, Part I: Research findings. Journal of the American Society for Information Science, 58(8), 1071–1081.

O'Brien, H.L., & Toms, E.G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. Journal of the American Society for Information Science and Technology, 59(6), 938–955.

Qiu, L. (1993). Analytical searching vs. browsing in hypertext information retrieval systems. Canadian Journal of Information and Library Science 18(4), 1–13.

Rieh, S.Y., & Xie, H. 2006. Analysis of multiple query reformulations on the web: The interactive information retrieval context. Information Processing &Management, 42, 751–768.

Rouet, J.F., Vidal-Abarca, E., Bert-Erboul, A., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. Discourse Processes, 31(2), 163–186.

Rouet, J. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' Search Strategies in Hypertext, 15, 409–428.

Sillence, E., Briggs, P., Fishwick, L., & Harris, P. (2004). Trust and mistrust of online health sites. In the Proceedings of CHI 2004, 663–670.

Spink, A., Yang, Y., Jansen, J., Nykanen, P., Lorence, D.P., Ozmutlu, S., & Ozmutlu, H.C. (2004). A study of medical and health queries to web search engines. Health Information and Libraries Journal, 21, 44–51.

Su, L.T. (1992). Evaluation measures for interactive information retrieval. Information Processing and Management, 28, 503–516.

Su, L.T. (2003). A comprehensive and systematic model of user evaluation of web search engines: I. Theory and background, Journal of the American Society for Information Science and Technology 54(13), 1175–1192.

Sutcliffe, A.G. & Ennis, M. (1998). Towards a cognitive theory of information retrieval. Interacting with Computers, 10, 321–351.

Taylor, R.S. (1991). Information use environments. In Progress in Communication Science. (Eds.) B. Dervin, and M.J. Voigt. Norwood, NJ: Ablex: 217–255.

Toms, E., & Latter C. 2007. How consumers search for health information. Health Informatics Journal, 13(3), 223–235.

Tullis, T., & Albert, B. (2008). Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Morgan Kaufmann.

White, R.W., & Horvitz, E. (2009). Cyberchondria: studies of the escalation of medical concerns in web search. ACM Transactions on Information Systems, 27(4) Article: 23.

Vakkari, P. (2003). Task-based information searching. Annual Review of Information Science and Technology, 37, 413–464.

Wildemuth, B.M. (2004). The effects of domain knowledge on search tactic formulation. Journal of the American Society for Information Science and Technology, 55(3), 246–258.

Wildemuth, B.M., & Freund, L. (2009). Search tasks and their role in studies of search behavior. Paper presented at HCIR 2009: Bridging Human-Computer Interaction and Information Retrieval, Washington, DC, October 23, 2009.

Zeng, Q.T., Kogan, S., Ash, N., Greenes, R.A., & Boxwala, A.A. (2002). Characteristics of consumer terminology for health information retrieval. Methods of Information in Medicine, 41, 289–298.

Zeng, Q.T., Kogan, S., Plovnick, R.M., Crowell, J., Lacroix, E.M., & Greenes, R.A. (2004). Positive attitudes and failed queries: an exploration of the conundrums of health information retrieval. International Journal of Medical Informatics, 73(1), 45–55.

Zhang, J., Wolfram, D., Wang, P., Hong, Y., & Gillis, R. (2008). Visualization of health subject term analysis based on user queries. Journal of the American Society for Information Science and Technology, 59(12), 1933–1947.

Zhang, Y. (2008). The influence of mental models on undergraduate students' searching behavior on the Web. Information Processing and Management 44(3): 1330–1345.

Zhang, Y. (2010). Contextualizing consumer health information searching: an analysis of questions in a social Q&A community. In the Proceedings of the ACM International Health Informatics Symposium, Arlington, VA, 210–219.

Zhang, Y., Wang, P., Heaton, A., & Winkler, H. (2012). Health information searching behavior in MedlinePlus and the impact of tasks. In the Proceedings of the ACM International Health Informatics (IHI) Symposium 2012, Miami, FL, 641–650.

Zhang, Y. (2012). Consumer health information searching process in real life settings. In the Proceedings of the ASIST 2012, Baltimore, MD.

Zickuhr, K. (2010). Generations 2010. Pew Internet & American Life project. Accessed on October 17, 2011, at http://www.pewinternet.org/~/media//Files/Reports/2010/%20PIP_Generations_and_Tech10.pdf

DOI: 10.1002/asi