



ELSEVIER

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

The impact of task complexity on people's mental models of MedlinePlus

Yan Zhang*

School of Information, The University of Texas at Austin, 1616 Guadalupe, Austin, TX 78701, United States

ARTICLE INFO

Article history:

Received 3 December 2009

Received in revised form 2 February 2011

Accepted 20 February 2011

Available online xxxx

Keywords:

Mental models

Task complexity

MedlinePlus

Task impact

Information retrieval

Consumer health information systems

ABSTRACT

This study explored the impact of task complexity on people's mental models of MedlinePlus, an information-rich web space providing consumer health information to the general public. Thirty-eight subjects were randomly assigned to two groups, performing either simple or complex search tasks. After the search session, subjects' perceptions of MedlinePlus were elicited using a concept listing protocol where subjects listed concepts concerning MedlinePlus in the order each concept occurred to them. The analysis of the concepts suggested that task complexity impacted subjects' mental models by influencing the objects in the system that they perceived and represented, the specificity of the representations, their evaluations of and emotions about the objects, and the heuristics that they developed for what works and what does not in the system. The pragmatic aspect of mental models was represented by subjects' descriptions of the steps that they would follow to tackle a hypothetical task. The analysis showed that task complexity affected the strategies that subjects perceived themselves using to solve new tasks using the system.

© 2011 Published by Elsevier Ltd.

1. Introduction

Mental models are internal knowledge structures that serve as proxies for people to interact with the external world (Craik, 1943; Gentner & Stevens, 1983; Johnson-Laird, 1983). In the fields of human computer interaction (HCI) and information retrieval (IR), researchers generally postulate that people employ mental models to interact with various systems (e.g., Norman, 1983; Young, 1983); specifically, mental models describe "how a system works, its component parts, the processes, their interactions, and how one component influences another" (Fein, Olson, & Olson, 1993, p. 157). Therefore, knowledge about mental models can, when appropriately transformed, effectively inform system design and user instruction (Carroll & Olson, 1987; Young, 2008). Thus, there is a continuous effort in HCI and IR to better understand mental models (e.g., Brandt & Uden, 2003; Cool, Park, Belkin, Koenemann, & Ng, 1996; Doyle & Ford, 1998; Hendry & Efthimiadis, 2008; Papastergiou, 2005; Rieh, Yang, Yakel, & Markey, 2010).

In the past several decades, research on mental models, particularly in Information Science (IS), mainly focused on examining: (1) features, attributes, or characteristics of mental models (e.g., incomplete, not scientific, and parsimonious) (e.g., Papastergiou, 2005); (2) elements or content of people's mental models of IR systems (e.g., Makri et al., 2007; Westbrook, 2006); and (3) the impact of mental models on people's information searching behavior and performance (e.g., Borgman, 1986; Dimitroff, 1992).

In these studies, researchers gradually found that the features, composition, or quality of people's mental models of an IR system could be affected by various factors, including users' individual differences, such as computer experience (e.g., Thatcher & Greyling, 1998), educational status and academic background (e.g., Zhang, 1997), gender (e.g., Zhang, 2008b),

* Tel.: +1 512 471 9448.

E-mail address: yanz@school.utexas.edu

and level of familiarity with a hypertext system (e.g., Otter & Johnson, 2000), system images, particularly system feedback (e.g., Muramatsu & Pratt, 2001), and environmental or contextual factors, mainly training conditions (e.g., Borgman, 1986; Savage-Knepshield, 2001).

Understanding how mental models are affected by various factors has both theoretical and practical implications. Theoretically, it could enhance our knowledge of the nature of mental models and their roles in people's interaction with IR systems. Practically, it can inform the design of systems that are able to facilitate users in building better understanding of the systems. Thus, it is worthwhile to initiate a series of studies to investigate the relationship between mental models and factors involved in interactive information searching. This study is an attempt to explore the impact of search tasks, an important contextual factor, on people's mental models of a consumer health information system, MedlinePlus.

Search tasks are consistently reported to have significant effects on information searching behavior and performance (e.g., Byström & Järvelin, 1995; Ingwersen, 1996; Ingwersen & Järvelin, 2005; Toms et al., 2007; Vakkari, 1999; Wildemuth & Hughes, 2005). People's mental models of IR systems are constructed, developed, validated, and modified when they complete tasks (Katzeff, 1990). Thus, tasks can affect the development of mental models by constraining the range of interactions that a user has with the system. However, little is currently known about how tasks affect mental models.

MedlinePlus (<http://www.medlineplus.gov>) was selected as the platform for this study. The website is a typical consumer health information portal. The site's content is updated often, and it has a large user base. A recent Pew study reported that eighty percent of adult Internet users in the US have searched for health information on the web (Fox & Jones, 2009). Many more web-based consumer health information systems are appearing (Kim & Chang, 2007). We expect that this study could also inform the development of such systems.

2. Related literature

The existing research on relationships between tasks and mental models has primarily focused on the impact of mental models on people's performance on search tasks. Studies showed mixed results. Some found that better mental models led to faster completion of tasks, fewer errors, and more results (e.g., Dimitroff, 1992; Kerr, 1990). Some reported that mental models did not have a significant impact on recall, precision, number of documents that subjects saved (Savage-Knepshield, 2001), subjects' behavior of initiating interactions with a system, navigation, query construction, and search patterns (Zhang, 2008b). Others found that the effect of mental models on people's information searching behavior and performance was contingent on the complexity of tasks. Better mental models led to better performance on complex tasks but not on simple tasks (e.g., Borgman, 1986; Halasz & Moran, 1983).

Another camp of research has focused less on the impact of mental models on task performance, and instead has attempted to understand the nature of mental models in relation to tasks. For example, based on the research on people's use of pocket calculators, Young (1983) suggested that one type of the subjects' mental models of calculators was based on task-action mapping. A task/action model encompasses a core set of corresponding relationships between tasks and actions. In such mental models, new tasks are expressed as variants of the core tasks. Subsequently, sequences of actions corresponding to the new tasks are derived from the core actions that correspond to the core tasks. diSessa (1986) distinguished two types of mental models, structural and functional models, in terms of contextual specificity. Structural models contain information about the internal structure of the system and are independent of specific tasks; while functional models are task related and contain information about how to use a selected set of functions to perform a specific task.

In the existing body of literature, search tasks were considered to be an important contextual factor and were consistently reported to have a significant impact on people's information searching behavior and experience (Li, 2009). For example, Liu, Gwizdka, Liu, and Belkin (2010) found that subjects who performed difficult tasks spent significantly more time, issued more queries, and examined more unique content pages to complete the tasks than those who performed easy tasks. In using an experimental IR system, White, Ruthven, and Jose (2005) found that subjects preferred explicit relevance feedback for less complex tasks and implicit relevance feedback for more complex tasks. Capra, Marchionini, Oh, Stutzman, and Zhang (2007) found that subjects who performed lookup tasks were more satisfied with the trial IR system than who performed exploratory tasks.

Although some researchers have proposed that tasks might influence people's mental models of an information system (e.g., van der Velden & Arnold, 1992), little empirical research has been done to investigate that speculation. In the single study found (Savage-Knepshield, 2001), the author investigated the impact of tasks on the quality of mental models, specifically, whether combinations of two types of tasks, a standard task and an aspectual task, affected the quality of subjects' mental models of an experimental IR system. She found that subjects who performed identical tasks over two different trials (1 week apart) did not improve their mental models' accuracy and completeness, whereas subjects who performed different tasks in the two trials showed increased congruency in their mental models.

In order to improve our understanding of the formation and construction of mental models, it is not sufficient to learn the fact that different types of tasks have impact on mental models. Rather, we need to learn how tasks, specifically their various facets or characteristics, might affect, shape, or constrain people's mental models of a system.

To illustrate how tasks would impact mental models, the first challenge is measuring the mental models. The main methods used to elicit or represent people's mental models in the existing literature are: (1) eliciting verbal accounts using interviews or think-aloud protocols (e.g., Katzeff, 1990; Slone, 2002); (2) asking people to draw their mental images of a system

(e.g., Efthimiadis & Hendry, 2005; Papastergiou, 2005; Zhang, 2008b); and (3) imposing different training conditions on subjects and assuming that subjects who received different training developed different mental models of a system (e.g., Borgman, 1986; Savage-Knepshield, 2001).

There are limitations associated with these methods. Verbal accounts are affected by subjects' ability to articulate their thoughts. Meanwhile, people tend to justify or rationalize what they have done in verbal protocols rather than describe what they have actually experienced (Hoffman & Militello, 2009). Furthermore, transcribing and analyzing the verbal accounts are extremely time-consuming. Drawings are often hard to interpret; a picture has different meanings to different people. Also, drawings are not effective in conveying subjects' understanding of certain underlying mechanisms of IR systems, such as information organization and relevance ranking (Zhang, 2008a). The training method tends to impose a pre-defined structure of the system to subjects. Furthermore, this method assumes that users who received model-based instructions will develop better models of the system than those who received procedure-based instructions. This assumption is problematic, as studies have shown that some users are able to construct a model of a system in the absence of any instructions (Neumann & Ignacio, 1998; Shrager & Klahr, 1983).

In addition to measuring mental models, exploring tasks' impact on mental models also calls for the comparison of mental models. Comparing mental models requires not only effective, but also efficient methods to represent the models. Given the limitations of those popular methods, a new approach needs to be employed. Concept listing is proposed as an alternative method in this study. Concept listing is a variant of the word association method, a method based on the principle of learning by contiguity. The principle argues that objects once experienced together tend to become associated in the mind and are often recalled together in the same order of sequence or coexistence as they were initially experienced (Wettler & Rapp, 1996). In a concept listing protocol, the subject is given a prime (e.g., words, objects, situations, and so on) and asked to respond as rapidly as possible with the concepts that come to mind.

Cramer (1968) pointed out that concept listing provides an efficient means to elicit key concepts that people have about a domain. Pejtersen (1991) employed the concept listing method to study people's mental models of IR systems and found that subjects' responses clearly indicated various facets of their conceptual framework of the systems. Wang, Bales, Reiger, and Zhang (2004) used the method to explore the development of students' knowledge structure of a subject domain, information organization, over a semester. They found that the number of terms that students contributed increased over the semester, and the quality of their vocabulary improved as well. More importantly, concepts listed by some subjects formed distinct clusters on particular topics in the subject domain.

3. Method

3.1. Subjects

Thirty-eight undergraduate students (20 females and 18 males) majoring in non-medical-related subject areas, such as art history, communication studies, business, geography, mathematics, or psychology, were recruited to participate in the study through a campus-wide email list for undergraduates. The age of the subjects ranged from 18 to 22 years old ($Mean = 21$, $SD = 1.07$), and none had ever used MedlinePlus before the study.

Previous research suggested that spatial ability might affect people's mental models of information systems (Chen, 2000; Dillon, 2000). In this study, subjects' spatial ability was measured by the ETS VZ-2 paper folding test (Eckstrom, French, Harman, & Derman, 1976).

3.2. Tasks

Two types of tasks were defined in the study: simple and complex tasks. The complexity of a task was determined by three factors: (1) the clarity of the information required to answer the question (Marchionini, 2006; Rasmussen, Pejtersen, & Schmidt, 1990; White, Ruthven, & Jose, 2005); (2) the distribution of the answer (Capra et al., 2007; Rasmussen et al., 1990); and (3) the extent to which a higher level cognitive activity, such as synthesizing information, is required to complete the task (Bilal, 2001; Marchionini, 2006; Rasmussen et al., 1990). A simple task is a well-defined question where it is clear what information is required. The answer to a simple task is located on one page, and little cognitive effort is required to solve the task. A complex task is an open-ended question. The desired information is less clear, and the answer to a complex task is often located on multiple pages. High-level cognitive activities, such as information comparing, interpreting, and synthesizing, are required to tackle a complex task.

To be realistic, tasks in the study were selected from Yahoo! Answers, a question and answer website where people post their own questions and answer questions posted by peer users. Some of the tasks were directly used for the study, while others were modified to suit the scope of the MedlinePlus system. For every task, a scenario was provided.

To ensure that each task was correctly categorized as a simple or a complex task, two information professionals specializing in medical information were asked to rate the complexity of each task. Before the rating, they were given the definition of task complexity, as described above, and instructed to perform every task using MedlinePlus. Tasks used for this study were those whose complexity was agreed on by both raters. An example of a simple task is:

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.

An example of a complex task is:

Imagine that your friend recently was diagnosed with asthma and was put on two inhalers. But he thinks it is chronic bronchitis. So he wants to know what the similarities and differences between asthma and chronic bronchitis are. Also, he wants to know various means to treat or soothe asthma, such as medicine, diet, alternative medicines, exercise, etc. You want to help him by finding as much information as possible in MedlinePlus.

3.3. Measuring mental models

In the study, subjects’ mental models of MedlinePlus were elicited using a concept listing protocol, in combination with semi-structured interviews. In the concept listing protocol, subjects were asked to list concepts related to the MedlinePlus system in the order that the concept appeared in their mind. Each concept was viewed as a memory node, and the list of concepts was viewed as the result of subjects’ cognitive process of making sense of and representing the system. By this means, concept listing does not impose pre-defined structures on the subjects; rather, it allows subjects’ mental structure to emerge from the concepts and the order in which the concepts were listed.

Mental model theory suggests that people’s mental models encompass procedural knowledge of how to use a system to solve problems (e.g., Mayer, 2002). A pilot test of the concept listing protocol suggested that the method had a limited capacity to elicit the procedural knowledge. Therefore, semi-structured interviews were employed. In the interviews, subjects were presented with a hypothetical task, as shown below, and asked to describe steps that they plan to follow to solve the task.

Imagine that you are required to write a paper about hepatitis to arouse public attention to the seriousness of this disease. In the paper, you want to include, but are not limited to, aspects such as what is the difference between chronic and non-chronic hepatitis? What determines this? What are the differences between hepatitis A, B, and C, and what are the treatments for each of them, and can we prevent the development of liver cirrhosis among patients with chronic hepatitis? You decide to use MedlinePlus to collect information for your paper.

They were also asked to describe their impressions of MedlinePlus. After the interviews, subjects were asked to draw their perceptions of MedlinePlus. Due to the space limit, the drawings will not be analyzed and reported here.

3.4. Procedure

Fig. 1 shows the procedure for the study.

At the beginning of the study, subjects finished a VZ-2 test and a demographic questionnaire. Then, they spent 5 min exploring MedlinePlus freely. After the exploration, subjects were asked to describe their impressions of MedlinePlus. Then, they were randomly assigned to two groups, the simple task group and the complex task group, within gender. Subjects in the simple task group performed 12 simple tasks and those in the complex task group performed three complex tasks. Upon the completion of the tasks, subjects’ mental models of MedlinePlus were measured by: (a) a concept listing protocol where, for five minutes, subjects listed concepts concerning MedlinePlus in the order that the concepts appeared in their mind; (b) a semi-structured interview where subjects described the steps that they plan to follow to solve a hypothetical task using MedlinePlus, as well as their impressions of the system; and (c) a drawing task in which subjects drew their perceptions of MedlinePlus. After the three measurements, subjects were instructed to finish a series of other tasks, including an exit

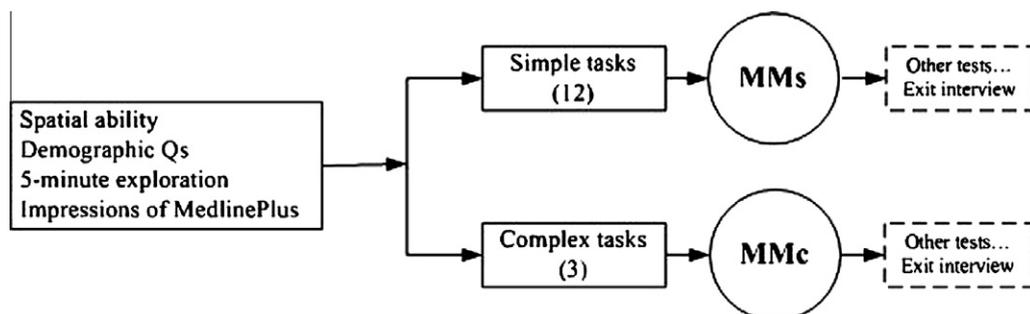


Fig. 1. Research procedure.

interview at the end of the study. These tasks did not have an impact on the theme and the results reported in this paper; thus, they are not introduced here.

3.5. Data analysis

The concepts contributed by the subjects in the concept listing protocol were analyzed using a grounded theory approach. The coding unit was a concept. When multiple concepts were involved in a concept-listing entry, the entry was assigned multiple codes. A coding schema was developed by coding a subset of the concepts. The schema consisted of five top categories: system, content, information organization, interface, and heuristics of using the system. The first four referred to major components of MedlinePlus, respectively, the overall MedlinePlus system, its content, its information organization, and its interface. The heuristics of using the system were rules of thumb that subjects developed concerning how to use the system. Each top category had a varied number of sub-categories. Each sub-category was one aspect from which the top category was represented. Details of the sub-categories will be reported in the results section. The schema was then applied to code the rest of the concepts. Here are several coding examples:

Concepts contributed by subjects	Assigned codes
Lung cancer	Content: specific
Alphabetical	Information organization: schema
Well organized	Information organization: evaluation
Tabs	Interface: element
Search	Interface: function
CDC	System: agencies involved

For example, the code for “lung cancer” meant that lung cancer was a specific topic covered by the content of MedlinePlus.

The author coded all the concepts and a second coder coded 10% of the data. The inter-coder agreement (Krippendorff's α) was 78.4% in the first round of coding (Krippendorff, 2004). A review session between the two coders revealed that the majority of the disagreements appeared at the second level of the code. For example, the second coder initially coded “NIH” as “system: similar sites.” After the discussion, she agreed with the author and coded it as “system: agencies involved.” After the review, the inter-coder reliability reached 95.7%.

Subjects' descriptions of the steps that they plan to follow to tackle a hypothetical task were transcribed and analyzed using content analysis. Subjects' descriptions of their impressions of MedlinePlus were transcribed as well. On occasion, the transcripts used to facilitate the interpretation of some concepts listed by the subjects in the concept listing protocol.

4. Results

4.1. Characteristics of the subjects

Table 1 summarizes the demographic information of the two task groups, their spatial ability as measured by the ETS VZ-2 test, and the time that each group spent completing their corresponding tasks.

T-tests show that the two groups were similar in age, computer experience, spatial ability, and time spent finishing the tasks.

4.2. Subjects' mental models of MedlinePlus

This section first provides a general overview of the composition of the subjects' mental models of MedlinePlus, followed by an overview of the two groups' shared model of the system. Both overviews will set the stage for the discussion of the differences between the two task groups' mental models.

4.2.1. General overview

The content analysis of the concepts that subjects contributed in the concept listing protocol suggested that subjects in both the simple and the complex task groups perceived and represented MedlinePlus in terms of four components:

- System: the overall MedlinePlus rather than any part of it
- Content: the content included in MedlinePlus

Table 1

Demographic information of the subjects and session length.

Group	F	M	Age	Comp. exp. (years)	VZ-2	Session length (mins)
Simple	9	10	20.95 (0.85)	9.89 (1.76)	12.99 (3.56)	20.11 (4.68)
Complex	11	8	20.37 (1.21)	9.84 (2.03)	12.42 (4.04)	19.86 (7.20)

- Information organization: how information is organized in MedlinePlus
- Interface: the interface of MedlinePlus

At the same time, subjects formed opinions and expressed their evaluations concerning the four components. Both the representations and evaluations tended to form clusters. For example, subjects represented the content in MedlinePlus with regard to its subject, information type, format, and presentation, and they evaluated the content with respect to its quality, utility, attributes, and quantity, as well as some specific sections of the content. Table 2 summarizes the aspects from which each component was represented and evaluated.

As shown in the table, the two groups represented and evaluated MedlinePlus from similar aspects: the only two aspects represented by one group but not the other were the audience and similar sites aspects of the system component. However, on a number of aspects, as shown in bold, the two groups showed qualitative differences in their representations or evaluations, which will be reported in the section of tasks' impact on mental models.

Table 2
Subjects' representations and evaluations of MedlinePlus.

	System	Content	Info. organization	Interface
Reps.	Structure Agencies involved <i>Audience</i> ^a <i>Similar sites</i> Usage of the system ^b System behavior	Subject Type Presentation Format	Information organization schemas	Elements Functions Results
Evals.	Usefulness Attributes Usability	Quality Utility Attributes Quantity Specific sections	Information organization	Look and fee Search Navigation Results

^a Italicized fonts indicate aspects represented by only one group.

^b Bold fonts indicate aspects represented differently by the two groups.

Table 3
Mental models of MedlinePlus shared by the two groups.

Component	Representations and evaluations
System	Representations <i>Structure of the system:</i> MedlinePlus pulled information from outside medical information websites <i>Agencies involved:</i> MedlinePlus is a government site and institutions such as CDC and NIH are involved in the site. Evaluations <i>Usefulness:</i> Useful <i>Usability:</i> Easy to use; user friendly; accessible (easy access to information)
Content	Representations <i>Subject:</i> e.g., doctors; care; drugs; disease; diagnosis; medical information; symptoms; illness; treatment <i>Type:</i> Glossary; news; journals; statistics; and dictionary <i>Presentation:</i> Overview <i>Format:</i> Pictures; PDF Evaluations <i>Quality:</i> Accurate; reputable; consistent; well researched <i>Utility:</i> Informative; interesting; could be easily understood, but sometimes difficult to use <i>Attributes:</i> Superficial; comprehensive; current; factual <i>Quantity:</i> Lots of information; lots of text; numerous articles on topics; overwhelming; expansive; large volume of knowledge; lots of data; vast; copious; plenty of information
Info. org.	Representations <i>Information organization schemas:</i> by category; topic; alphabetical list; anatomy Evaluations <i>Evaluation of information organization:</i> well organized
Interface	Representations <i>Elements:</i> e.g. links; headings; related topics; tabs; search bar <i>Functions:</i> e.g., search; browse; spelling check <i>Results:</i> results and response in general; relevance; relatedness Evaluations <i>Look and feel:</i> Clear; simple; decent; more diagrams; more pictures; enhancement of the visual cues; a more futuristic design <i>Search:</i> Easy to use; effective; accurate; limited ability to define the specificity of a search <i>Navigation:</i> Easy to navigate; logical layout; easy transition from one item to the next; streamlined; easy to follow links; home is easy to get to; helpful links on the main site

The pragmatic how-to aspect of the mental models was represented by subjects' descriptions of their plans to solve a hypothetical task using MedlinePlus. Three distinct strategies (A–C) were indentified based on the content analysis of the transcripts. Strategy A is a general searching strategy: subjects submit a search query to the system and then inspect the search results. Strategy C is a general browsing strategy: subjects follow links from the homepage to the page where the answer is located. Strategy B is a combination of searching and browsing, where subjects reach the health topic page pertaining to a disease or condition of interest by searching the disease or condition and then browse the links on the health topic page to find relevant information. Strategy B is more closely aligned to the capabilities of MedlinePlus and requires fewer clicks to reach the desired information. The three strategies were mentioned by subjects in both groups.

Furthermore, the analysis of the concepts that subjects contributed in the concept listing protocol revealed that subjects formed heuristics based on their interactions with MedlinePlus. Heuristics refer to rules of thumb concerning what work well, what do not, what is easy, and what is difficult to do within MedlinePlus. For example, one subject commented that “typing in a question does not work well.” Because heuristics could impact subjects' strategies of using a system, they were considered as a part of the pragmatic aspect of the subjects' mental models of MedlinePlus. The two groups developed different heuristics, which will be discussed in the section of tasks' impact on mental models.

4.2.2. Shared mental models of MedlinePlus

Besides the compositional similarities, the two groups' mental models of MedlinePlus also had many semantic similarities. Table 3 shows the representations and evaluations of the system shared by the two groups.

In terms of the pragmatic knowledge of how to use the system, most of the subjects in both groups reported that they would use multiple strategies in combination to solve the hypothetical task. In addition, the two groups shared heuristics concerning looking for information in the system: both groups developed a conflicting view, pointing out, at the same time, that it was “easy to find information on specific topics” and “specific searches [are] difficult.”

4.3. Tasks' impact on subjects' mental models of MedlinePlus

In this section, the differences between the two task groups' mental models of MedlinePlus are reported in terms of the subjects' representations and evaluations of the four components, their procedural knowledge of solving problems using MedlinePlus, and the heuristics for using the system.

4.3.1. Tasks' impact on subjects' representations and evaluations of the system component

In representing MedlinePlus as an integrated system, the two groups differed in their perceptions concerning the usage, behavior, and audience of the system, as well as similar sites to the system. In evaluating the system component, they differed in their opinions concerning the attributes and usability of the system. Table 4 illustrates the differences.

In representing the usage of MedlinePlus, the simple task group valued the practical usefulness of MedlinePlus; for example, one subject pointed out that it could be used for seeking medical help. In contrast, the complex group viewed the system from an educational perspective, with four subjects pointing out that it was a starting point for doing research and could educate households. In representing the behaviors of MedlinePlus, the simple task group focused on the general behavior of the system, such as aggregation and pop-up windows (noted by five subjects); whereas the complex task group focused more on specific instances experienced. For example, one subject noticed “some links did not work” and another commented “the search function tends to bring up several links to the same article (on different sites).”

The two groups also each had a unique aspect from which they perceived the system component: three subjects in the complex task group represented the audience (layman and doctors) for MedlinePlus, and two subjects in the simple task group represented similar sites (Google) to MedlinePlus.

Table 4

Differences between the two groups in their representations and evaluations of the overall MedlinePlus system.

		Simple task group	Complex task group
Reps.	Usage of the system	Seek medical help	Background research; educating households; starting point
	System behavior	Aggregate; pop-up windows; separate screens	Some links did not work; search bar directs to links; the search function tends to bring up several links to the same article (on different sites).
	Audience		Layman; doctors
Evals.	Similar sites	Google	
	Attributes	Information-rich; resourceful; legitimate; search-heavy; personalized	Complex; complicated
	Usability	Quick; fast; responsive; direct; straightforward; logical; not stressful to use; not intrusive; convenient Harder to use than Google; difficult; not good for browsing	

When similar expressions were used to describe one concept, only one expression is listed in the table.

In evaluating MedlinePlus as an integrated system, the two groups showed different opinions about the attributes and usability of the system. Concerning the attributes, six subjects in the simple task group recognized that MedlinePlus was information-rich, legitimate, search-heavy, and allowed personalization; whereas the complex task group focused more on the complexity of the system, with two subjects pointing this out explicitly. Concerning the usability, both groups agreed that MedlinePlus was usable, but subjects in the simple task group expressed more positive feelings toward the usability of the system, such as fast, responsive, and convenient (noted by 13 subjects). Meanwhile, three subjects in the simple task group were critical, pointing out that the system was more difficult to use and harder than Google and that it was poor for browsing.

4.3.2. Tasks' impact on subjects' representations and evaluations of the content component

In representing the content of MedlinePlus, the two groups showed differences in their perceptions of the subject, type, and presentation of information. In evaluating the content, they differed in their opinions about the quantity of information and paid attention to different sections of the content in the system, as shown in Table 5.

Subjects in the simple task group (14 out of 19) listed a greater number of specific subjects, such as heart disease and kidney problems, than subjects in the complex task group (8 out of 19), and the representations of the two groups had little overlap. An inspection of the concepts revealed that most of the specific subjects were from the task descriptions. In representing the types of information in MedlinePlus, subjects in the simple task group (5 out of 19) identified types of information that convey mostly factual information, such as fact sheets, FAQs, reports, and what-to-do articles, whereas subjects in the complex group (11 out of 19) identified types that convey mostly complex and debatable information, such as scholarly and academic articles, clinical trials, and tutorials. In representing the presentation of information in MedlinePlus, both groups listed "overview" as a way in which information was presented. However, subjects in the simple task group (7 out of 19) identified more additional forms of information presentation, such as summaries, tables, diagrams, and figures.

In evaluating the content of MedlinePlus, the two groups developed different opinions toward the quantity of information and focused on evaluating different sections of the content. In terms of the quantity of information, both groups recognized that MedlinePlus contained an overwhelmingly large amount of information, but the complex group was more critical, with six of the subjects pointing out that there were "few scholarly articles" and "limited external sources," and "exclusive MedlinePlus information is scant."

Specific sections of the content that each group evaluated tended to match with their representations of the content. As shown in the table, subjects in the simple task group (4 out of 19) evaluated summaries, descriptions, and overviews. For example, one subject commented that "drug summaries [are] easy to navigate and especially good." Subjects in the complex task group (6 out of 19) focused more on evaluating the encyclopedia and tutorials, with one subject stating that "additional content for the encyclopedia would be nice."

4.3.3. Tasks' impact on subjects' representations and evaluations of the information organization component

The differences between the two groups in representing and evaluating the information organization component are illustrated in Table 6.

Table 5

Differences between the two groups in their representations and evaluations of the content in MedlinePlus.

	Simple task group	Complex task group
Reps. Subject	E.g., health insurance; protein; Lou Gehrig; STD; AIDS; blood pressure; HIV; kidney; vitamins; heart disease; smoking; side effects; dosing information; exercise; essentials; black window; diet; nutrition; hepatitis; kidney; AZT drugs; and virus	Herbal medicine; medical ethics; organs; support groups; insulin; diabetes; liver problems
Type	Fact sheet; FAQs; reports; what-to-do articles	Scholarly articles; academic articles; clinical trials; encyclopedia; tutorials; directory
Presentation	Tables; summaries; descriptions; overviews; specifics; diagrams; figures	
Evals. Quantity		Limited; limited external sources; limited general info; exclusive Medline plus information is scant, few scholarly articles; not a lot of simple information
Specific sections	Drug summaries easy to navigate and especially good; good summary of what articles would be about; overview of diseases useful; Sometimes too general in description, more information on some topics than others (lots on AIDS, less on low blood pressure)	Good overview by medline (Encyclopedia) similar information to health issues; encyclopedia function is useful as a starting point for searches; tutorial may be very helpful; anatomy page limited, anatomy links vary in usefulness; it would be helpful if the encyclopedia included more articles based on general descriptions of organs or systems, rather than consisting almost entirely of articles on disorders; additional content for the encyclopedia would be nice

When similar expressions were used to describe one concept, only one expression is listed in the table.

In representing information organization schemas in MedlinePlus, the complex task group provided a more comprehensive view than the simple task group. In addition to the basic information organization schemas, such as alphabetical and anatomic, five subjects in the complex task group also pointed out that the information was organized hierarchically, and some of the information was organized based on the applicable demographic.

In evaluating the information organization, the simple task group commented at a very general level, pointing out that the information was well organized. Conversely, the complex task group evaluated information organization at multiple levels, from general information organization (“well organized”), to specific information organization elements, such as sub-categories (“like the sub-categories”) and listings (“listings could be broader”) (6 subjects out of 19). Meanwhile, subjects in the complex task group were more critical, with eight of them pointing out that information in MedlinePlus was not optimally organized and that “concepts were not very tied together.” They also commented that “sometimes there was too much information, too many subgroups” and “consolidation of information” was needed.

4.3.4. Tasks' impact on subjects' representations and evaluations of the interface component

The differences between the two groups in representing and evaluating the interface of MedlinePlus are illustrated in Table 7.

As shown in the table, in representing the results, the two groups showed different focuses. Subjects in the simple task group (3 out of 19) perceived features of the results in MedlinePlus, such as the domain from which the results were from, the reliability of the source, and whether the results were pre-formulated, whereas the complex task group's perceptions focused more on the organization of the results, with two of them pointing out that the search results were organized into categories.

Correspondingly, the two groups expressed different evaluations and emotions about the search results in the system. Two subjects in the simple task group thought that the results were reliable, and another one expressed curiosity about the ranking mechanisms, whereas the complex task group reflected on more aspects of the results, such as quantity (“lots of results”), relevance (“relevant”), organization (“organized”), and specificity (“specific”). Meanwhile, the complex task group was more critical, with seven subjects pointing out limitations of the results. A few examples are illustrative: one subject reported conflicting results, one reported repetitive hits, one commented that there were “too many results,” and the other commented that breaking results into categories was “not helpful.”

The two groups also developed different opinions about the navigation supported by the interface. Both groups recognized that the interface was streamlined and easy to navigate, but the complex task group was more critical, with two subjects pointing out that it was “hard to navigate to find specific [information]” and it was “[...] harder to find functions of the system.”

4.3.5. Tasks' impact on subjects' procedural knowledge of using MedlinePlus

As has been mentioned, subjects described three distinct strategies when asked to propose means to solve a hypothetical task using MedlinePlus: a general search strategy (strategy A), a browsing strategy (strategy C), and a combination of searching and browsing (strategy B). Table 8 shows the number of subjects who planned to use each strategy. If a subject planned to use multiple strategies, the subject was counted multiple times.

As shown in the table, the two groups showed significant differences in their proposed use of strategies A and B. Fisher's exact tests suggest that subjects in the complex group were more likely to adopt strategy A than those in the simple task group ($p = 0.039$), while subjects in the simple task group were more likely to adopt the strategy B ($p = 0.022$).

4.3.6. Tasks' impact on heuristics of using MedlinePlus

The two groups also formed different heuristics for using MedlinePlus after the completion of their corresponding tasks, as shown in Table 9.

Concerning looking for information in MedlinePlus in general, both groups had conflicting views; some thought it was easy, and some thought it was difficult. But two subjects in the simple task group also pointed out that the system was easy for “general search” and “guided searches,” when none in the complex task group recognized such.

As shown in the table, the two task groups also developed heuristics concerning different decision-making points in the information searching process. One subject in the simple task group formed a heuristic for accessing information: you can

Table 6

Differences between the two groups in their representations and evaluations of the information organization in MedlinePlus.

	Simple task group	Complex task group
Reps. Info. org. schemas		Hierarchal; some of the information is organized based on what demographic it applies to most
Evals. Information organization		Many disease listings; lots of browsing sub-categories; like the sub-categories Concepts not very tied together; disorganized; listings could be broader, then separated into more specific within the topic; not consolidation of info; sometimes listings are too specific; sometimes too much information; too many subgroups; could make listings more...

When similar expressions were used to describe one concept, only one expression is listed in the table.

Table 7

Differences between the two groups in their representations and evaluations of the interface of MedlinePlus.

	Simple task group	Complex task group
Reps. Results	Features of results: sometimes hits were in domain sometimes out; pre-formulated search responses; site hits were major organizations	Organization of results: search results listed in categories; search results broken down into categories
Evals. Results	Only provided hits on reliable sites; site hits were major organizations; curious about page ranking formula	Lots of responses to search; relevant; organized; specific Limited results; too many results; search results broken down into categories on left not helpful, sometimes conflicting results; the search function tends to bring up several links to the same article (on different sites)
Navigation		Hard to navigate to find specific info; easy to find components of a system, but harder to find functions of the system

When similar expressions were used to describe one concept, only one expression is listed in the table.

Table 8

Use of search strategies A, B and C for the two groups.

Strategies	Number of subjects	
	Simple	Complex
A ^a	10	16
B ^b	11	4
C ^c	5	10

Bold face indicates that the differences between groups were statistically significant ($p < .05$).

^a Strategy A: Search → read results.

^b Strategy B: Search → top result → health topic page → review content and links.

^c Strategy C: Browse: Homepage → Health issues → health topic page → review content and links.

Table 9

Differences between the two groups in their heuristics for using MedlinePlus.

Simple task group	Complex task group
Looking for information Easy for . . . Easy for general search Medline is good for guided searches	Processing information Harder to relate topics to one another
Accessing information Everything you need basically if you are willing to read multiple articles sometimes	
Use of information types Hard to find answers outside of FAQ sites	

When similar expressions were used to describe one concept, only one expression is listed in the table.

get “everything you need basically if you are willing to read multiple articles.” Another subject in the group formed a heuristic for using information types in MedlinePlus: it was “hard to find answers outside of FAQ sites.” One subject in the complex task group developed a distinctive heuristic for processing information, pointing out that it was “harder to relate topics to one another.”

5. Discussion and conclusions

Tasks reflect people's information needs. IR systems are designed to help people accomplish certain tasks to fulfill their specific information needs (e.g., Belkin, Brooks, & Oddy, 1982). Thus, tasks tend to shape the design of the functions of a system (Card, Moran, & Newell, 1983; Nielsen, 1989) and influence the ways in which the system will be used (e.g., Vakkari, 2003). Research on mental models has been exploring factors affecting mental models, with the intention to identify means of helping people develop better mental models of a system so as to use the system more effectively. However, little is known about the impact of tasks, an important contextual factor that has a significant impact on the design and use of systems, on mental models construction. This study presents an effort to explore the impact of task complexity on people's mental models of MedlinePlus.

A mental model is an abstract concept and difficult to measure (Borgman, 1986). In this study, we designed a concept listing protocol to elicit subjects' mental models of MedlinePlus. In this methodological endeavor, we learned that, compared to more traditional methods like interviews and think-aloud protocols, the concept listing protocol was efficient: it took subjects only five minutes to complete. It was also effective in eliciting perspectives from which subjects perceived the system and in representing elements of the mental models (such as objects and functions in MedlinePlus) and subjects' evaluations

of the elements. However, it has a limited capacity in eliciting procedural knowledge. Therefore, we supplemented the concept listing protocol with a semi-structured interview to elicit subjects' procedural knowledge of using MedlinePlus to solve particular tasks. In the interviews, subjects were asked to describe steps that they would take to solve a hypothetical task using MedlinePlus. They were also asked to describe their general impressions of the system. The descriptions helped the interpretation of some ambiguous concepts in the concept listing protocol.

The analysis of the mental models of the two groups, the simple and complex task groups, after they completed their corresponding tasks, revealed that their mental models of MedlinePlus had many structural and semantic similarities. Both groups represented the system in relation to four components: MedlinePlus as an integrated system, the content, the information organization, and the interface of the system. Meanwhile, they formed emotions or evaluations about each of the components. They both also developed general heuristics for using the system and were able to articulate strategies when asked to use the system to solve a new task.

However, the two groups differed in their perceptions and evaluations of some aspects of the system and strategies of using the system. This result suggests that mental models should not be studied only in relation to the user and the system. People employ mental models to interact with the external world (Johnson-Laird, 1983) and mental models are constructed as a result of the interaction between the knowledge in the mind and the knowledge in the world (Jonassen & Henning, 1999). In order to better understand people's mental models of information systems, it is necessary to consider relevant contextual factors that exist in the environment, such as tasks. In the following several paragraphs, the differences of the two groups' mental models of MedlinePlus will be discussed in relation to the features of the tasks.

First, the clarity of the information required by the two types of tasks defined in the study was different. The simple tasks were factual questions and the answers to them involved minimum ambiguity, while the complex tasks were research-oriented questions with exploratory nature. This feature could account for the differences that the two groups had in their representations of the usage of the system and the types and presentations of information in the system. Concerning the usage of the system, the simple task group developed a practical view, thinking MedlinePlus as a place for seeking medical help; whereas the complex task group perceived the system from a more educational point of view, thinking the system as a good educational site that could support research. In representing information types, the simple task group perceived information types that provide factual information useful for completing the simple tasks, such as fact sheets and what-to-do articles, whereas the complex task group represented information types that provide more in-depth information required by the complex tasks, such as scholarly articles and clinical trials. Similarly, in representing the presentations of information, the simple task group perceived more presentation forms that were likely to contain factual information, such as summaries and figures.

Second, the distribution of the answers for the two types of tasks was different. Answers to the simple tasks were located at one place, whereas answers to the complex tasks were distributed at multiple places. Finding answers from multiple places requires more exploration of the system. Therefore, it is possible that subjects in the complex task group developed more specific representations of MedlinePlus. This speculation was supported by the finding that, in representing the system's behavior, the simple task group focused on general behavior of the system, such as pop-up windows, but the complex task group focused more on specific instances, such as "some links did not work" and "the search function tends to bring up several links to the same article." It was also found that in representing information organization in MedlinePlus, the complex task group pointed out more types of information organization schemas.

Third, the cognitive demand of the tasks was different. Answers to the simple tasks were explicit and easy to recognize, therefore, did not require high-level cognitive activities, whereas answers to the complex tasks were indefinite and required subjects to relate, compare, judge, and synthesize information. The difference in cognitive demand might explain the two groups' differences in the heuristics they developed for using MedlinePlus. The simple task group developed heuristics corresponding to less cognitively demanding activities: accessing and using information types in the system, while the complex task group developed heuristics concerning a more cognitively demanding activity, processing information, pointing out that MedlinePlus was limited in helping them relate one topic to another.

The three features of the complex tasks, exploratory, distributed answers, and high cognitive demand, work together to make completing the complex tasks more difficult than completing the simple tasks. The joint effect of the three aspects could lead to some differences in the two groups' evaluations of MedlinePlus. The simple task group thought that MedlinePlus was resourceful, information was well organized, the interface was streamlined, and the results were reliable. However, the complex task group thought that the system was complex, information was scant and not effectively organized, the interface was hard to navigate, and the results were conflicting, repetitive, and not well organized.

The two groups also differed in their perceived strategies for completing new tasks using MedlinePlus. The simple task group was more likely to adopt a strategy more aligned to the capacity of MedlinePlus, for future tasks. The complex task group was more inclined to adopt a general search strategy. An inspection of the navigation path for answers to the two types of tasks suggested that the path to the answers to the simple tasks tended to reveal the information architecture of MedlinePlus, thus helping users recognize and adopt the more system-specific strategy. A comment from a subject in the simple task group supported this speculation:

The tasks are an effective way of learning how to navigate the site. That helped a lot. If I hadn't had that purpose, I would have wandered a lot more.

The complex task group's preference for the general search strategy might be attributable to their need to synthesize and integrate information from different places.

Tasks could also affect people's perceptions and representations of the specific content of a system. In the study, the simple task group contributed significantly more concepts to represent the specific content in MedlinePlus and the majority of these concepts were from task descriptions. This could be explained by the fact that, in the study, the simple tasks outnumbered the complex tasks, therefore, provided more scenarios and medical terms.

Kelly et al. (2008) demonstrated that providing users with performance feedback changed their evaluations, particularly satisfaction with an IR system. The results of this study suggested that users' perceptions and evaluations of IR systems might also be influenced by various features of a task, such as the clarity of information required by the task, distribution of its answers, the cognitive demand, the level of difficulties in completing the task, the degree to which the navigation path for the answer matches with the system's information structure, and the descriptions of the task. This finding suggests that in evaluating IR systems, researchers need to consider the potential impact of tasks. A system that functions well with one type of tasks might not work as well for other task types. A more comprehensive approach that takes into consideration different types of tasks users would carry out using a system is needed to achieve a more accurate evaluation of the system. The potential impact of tasks also calls for more research to understand the nature of self-reported evaluative measures that are widely used in studies evaluating IR systems and more research on how to validate these measurements used in.

This study has limitations. First, subjects' procedural knowledge of how to use MedlinePlus was represented by their descriptions of the steps that they would take to complete a task. In future studies, it would be beneficial to observe people's natural behavior of using the system and examine the relationship between their actual behavior and their mental models. It is also worthwhile to further explore whether behavior mediates the impact of task factors on mental models. Second, the subjects in the study were undergraduate students. This group of users is young and has a relatively low level of health information needs. Searching for health information is highly personal and often has critical consequences; therefore, it is worthwhile to explore how patients or caregivers, who are actually in need of health-related information, use the system.

In addition to complexity, tasks could be characterized by other features, such as difficulty (e.g., Liu et al., 2010), urgency (Li, 2009), familiarity to the user (Zhang, 2008b), and users' motivations for completing the task (real vs. assigned tasks) (e.g., Borlund, 2000). In future studies, it would also be worthwhile to explore the impact of these features on people's mental models of an IR system. Enhanced knowledge of the relationships between tasks and people's mental models of a system not only could improve our understanding of the nature of users' self-reported evaluations of systems, but it could also have implications for design. First, it might allow us to apply task analysis as a tool not only to inform the design of the functions of a system (Card et al., 1983), but also to help designers develop informed models of end users, which is considered critical for designers (Hammond, Jorgensen, MacLean, Barnard, & Long, 1983). Secondly, it could inform the design of out-of-the-box demos for software applications and instructional materials to help users learn the systems and develop an optimal understanding of the systems.

Acknowledgements

I would like to thank my dissertation advisor, Barbara Wildemuth, as well as my committee members, Gray Marchionini, Paul Solomon, Diane Kelly, Javed Mostafa, and Peiling Wang. I would also like to thank Andrew Dillon and the anonymous reviewers for their insightful comments on earlier drafts of this article.

References

- Belkin, N. J., Brooks, H. M., & Oddy, R. N. (1982). Ask for information retrieval. *Journal of Documentation*, 38(2), 61–71.
- Bilal, D. (2001). Children's use of the Yahoo!igans! Web search engine. II. Cognitive and physical behaviors on research tasks. *Journal of the American Society for Information Science and Technology*, 52(2), 118–136.
- 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). *Evaluation of Interactive Information Retrieval Systems*. Åbo, Finland: Åbo Akademi University Press.
- Brandt, D. S., & Uden, L. (2003). Insight into mental models of novice Internet searchers. *Communications of the ACM*, 46(7), 133–136.
- Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, 31(2), 191–213.
- Capra, R., Marchionini, G., Oh, J. S., Stutzman, F., & Zhang, Y. (2007). Effects of structure and interaction style on distinct search tasks. In *Proceedings of ACM/IEEE-CS joint conference on digital libraries (JCDL07)* (pp. 442–451).
- Card, S. K., Moran, R., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum.
- Carroll, J. M., & Olson, J. R. (1987). Mental models in human-computer interaction: Research issues about what the user of software knows. Committee on Human Factors. *Commission on behavioral and social sciences and education*. National Research Council. Washington, DC: National Academy Press.
- Chen, C. (2000). Individual differences in a spatial-semantic virtual environment. *Journal of the American Society for Information Science*, 51(6), 529–542.
- Cool, C., Park, S., Belkin, N. J., Koenemann, J., & Ng, K. B. (1996). Information seeking behavior in new searching environments. In *Proceedings of conceptions of library science* (pp. 403–416).
- Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.
- Cramer, P. (1968). *Word association*. NY: Academic Press.
- Dillon, A. (2000). Spatial-semantics: How users derive shape from information space. *Journal of the American Society for Information Science*, 51(6), 521–528.
- Dimitroff, A. (1992). Mental models theory and search outcome in a bibliographic retrieval system. *Library and Information Science Research*, 14(2), 141–156.
- diSessa, A. (1986). Models of computation. In D. A. Norman & S. W. Draper (Eds.), *User-centered system design: New perspectives in human-computer interaction*. LEA, NJ: Hillsdale.
- Doyle, J. K., & Ford, D. N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review*, 14(1), 3–29.
- Eckstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Efthimiadis, E. N., & Hendry, D. G. (2005). Search engines and how students think they work. In *Proceedings of SIGIR'05* (pp. 595–596). New York: ACM.
- Fein, R. M., Olson, G. M., & Olson, J. S. (1993). A mental model can help with learning to operate a complex device. In *INTERACT '93 and CHI '93 conference companion on human factors in computing systems* (pp. 157–158).

- Fox, S., & Jones, S. (2009). The social life of health information. Pew Internet & American Life project. <http://www.pewinternet.org/~media/Files/Reports/2009/PIP_Health_2009.pdf> [Retrieved 03.03.10]
- Gentner, D., & Stevens, A. L. (1983). *Mental models*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Halasz, F. G., & Moran, T. P. (1983). Mental models and problem-solving in using a calculator. In *Proceedings of CHI'83 human factors in computing systems* (pp. 212–216). New York: ACM.
- Hammond, N., Jorgensen, A., MacLean, A., Barnard, P., & Long, J. (1983). Design practice and interface usability: Evidence from interviews with designers. In *CHI'83 proceedings* (pp. 40–44).
- Hendry, D. G., & Efthimiadis, E. N. (2008). Conceptual models for search engines. In A. Spink & M. Zimmer (Eds.), *Web searching: interdisciplinary perspectives* (pp. 277–307). Dordrecht: Springer.
- Hoffman, R. R., & Militello, L. G. (2009). *Perspectives on cognitive task analysis: Historical origins and modern communities of practice*. NY: Psychology Press.
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), 3–50.
- Ingwersen, P., & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Netherlands: Springer.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, Massachusetts: Harvard University Press.
- Jonassen, D. H., & Henning, P. (1999). Mental models: Knowledge in the head and knowledge in the world. *Educational Technology*, 39(3), 37–42.
- Katzeff, C. (1990). System demands on mental models for a fulltext database. *International Journal of Man–Machine Studies*, 32, 483–509.
- Kelly, D., Shah, C., Sugimoto, C. R., Bailey, E. W., Clemens, R. A., Irvine, A. K., et al. (2008). Effects of performance feedback on users' evaluations of an interactive IR system. In *Proceedings of the 2nd symposium on information interaction in context (IliX)*, London, UK, pp. 75–82.
- Kerr, S. T. (1990). Wayfinding in an electronic database: the relative importance of navigational cues vs. mental models. *Information Processing & Management*, 26(4), 511–523.
- Kim, D., & Chang, H. (2007). Key functional characteristics in designing and operating health information websites for user satisfaction: An application of the extended technology acceptance model. *International Journal of Medical Informatics*, 76, 790–800.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, California: Sage Publication, Inc..
- Li, Y. (2009). Exploring the relationships between work task and search task in information search. *Journal of the American Society for Information Science and Technology*, 60(2), 275–291.
- Liu, J., Gwizdzka, J., Liu, C., & Belkin, N. J. (2010). Predicting task difficulty for different task types. In *Proceedings of ASIST 2010*.
- Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C., & Buchanan, G. (2007). A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *Journal of the American Society for Information Science and Technology*, 58(3), 433–445.
- Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4), 41–46.
- Mayer, R. E. (2002). *Multimedia learning*. NY: Cambridge University Press.
- Muramatsu, J., & Pratt, W. (2001). Transparent queries: investigating users' mental models of search engines. In *Proceedings of the twenty-fourth international ACM conference on research and development in information retrieval SIGIR 2001* (pp. 217–224). New York: ACM.
- Neumann, L. J., & Ignacio, E. N. (1998). Trial and error as a learning strategy in system use. In *ASIS '98: proceedings of the 61st ASIS annual meeting* (Vol. 35, pp. 243–252).
- Nielsen, J. (1989). The matters that really matter for hypertext usability. In *Proceedings of ACM hypertext '89* (pp. 239–248).
- Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale, NJ: Erlbaum.
- Otter, M., & Johnson, H. (2000). Lost in hyperspace: metrics and mental models. *Interacting With Computers*, 13, 1–40.
- Papastergiou, M. (2005). Students' mental models of the Internet and their didactical exploitation in informatics education. *Education and Information Technology*, 10(4), 341–360.
- Pejtersen, A. M. (1991). *Interfaces based on associative semantics for browsing in information retrieval*. Roskilde: Risø Laboratory.
- Rasmussen, J., Pejtersen, A. M., & Schmidt, K. (1990). *Taxonomy for cognitive work analysis (Rise-M-2871)*. Roskilde, Denmark: Rise National Laboratory.
- Rieh, S. Y., Yang, J. Y., Yakel, E., & Markey, K. (2010). Conceptualizing institutional repositories: Using co-discovery to uncover mental models. In *Proceedings of the Information Interaction in Context (IliX) Symposium* (pp. 165–174).
- Savage-Knepshield, P. A. (2001). Mental models: Issues in construction, Congruency, and Cognition. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.
- Shrager, J., & Klahr, D. (1983). Learning in an instructionless environment: observation and analysis. In *Proceedings of the 1983 CHI conference on human factors in computing*. NY: ACM.
- Slope, D. J. (2002). The influence of mental models and goals on search patterns during Web interaction. *Journal of the American Society for Information Science and Technology*, 53(13), 1152–1169.
- Thatcher, A., & Greyling, M. (1998). Mental models of the internet. *International Journal of Industrial Ergonomics*, 22(4–5), 299–305.
- Toms, E., MacKenzie, T., Jordan, C., O'Brien, H., Freund, L., Toze, S. et al. (2007). How task affects information search. Workshop Pre-proceedings of in initiative for the evaluation of XML retrieval (INEX), pp. 337–341.
- Vakkari, P. (1999). Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing and Management*, 35(6), 819–837.
- Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science & Technology*, 37, 413–464.
- van der Velden, J. M., & Arnold, A. G. (1992). Mental models and the evaluation of user interfaces: A case study of a library system. In M. J. Tauber & D. Ackermann (Eds.), *Mental models and human-computer interaction 2* (pp. 179–189). North-Holland, Amsterdam: Elsevier Science.
- Wang, P., Bales, S., Reiger, J., & Zhang, Y. (2004). Survey of learners' knowledge structures: Rationales, methods and instruments. In *Proceedings of the 67th annual meeting of the ASIS&T* (pp. 218–228). Medford, NJ: Information Today.
- Westbrook, L. (2006). Mental models: A theoretical overview and preliminary study. *Journal of Information Science*, 32(6), 563–579.
- Wettler, M., & Rapp, R. (1996). Computation of word associations based on the co-occurrence of words in large corporations. <<http://www.fask.uni-mainz.de/user/rapp/papers/wvlc93/latex2html/wvlc93.html>>. [Retrieved 27.02.10]
- White, R. W., Ruthven, I., & Jose, J. (2005). A study of factors affecting the utility of implicit relevance feedback. SIGIR'05, pp. 35–42.
- Wildemuth, B. M., & Hughes, A. (2005). Perspectives on the tasks in which information behaviors are embedded. In Fisher, K E., Erdelez, S., & McKechnie, L. (Eds.), *Theories of Information Behavior* (pp. 275–279). Medford, NJ: Information Today, for the American Society for Information Science & Technology.
- Young, R. M. (1983). Surrogates and mappings: Two kinds of conceptual models for interactive devices. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 35–52). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Young, I. (2008). *Mental models: Aligning design strategy with human behavior*. NY: Rosenfeld Media.
- Zhang, X. (1997). A Study of the effects of user characteristics on mental models of information retrieval systems. Unpublished doctoral dissertation, University of Toronto.
- Zhang, Y. (2008a). Undergraduate students' mental models of the Web as an information retrieval system. *Journal of the American Society for Information Science and Technology*, 59(13), 2087–2098.
- Zhang, Y. (2008b). The influence of mental models on undergraduate students' searching behavior on the Web. *Information Processing & Management*, 44(3), 1330–1334.

Yan Zhang is an assistant professor at the School of Information at the University of Texas at Austin. Her research interests focus on psychological processes involved in people interacting with information retrieval systems, consumer health information seeking behavior, and consumer health information system design.