Available online at www.sciencedirect.com
ScienceDirect
Information Processing and Management 44 (2008) 1330–1345
INFORMATION PROCESSING & MANAGEMENT
www.elsevier.com/locate/infoproman

The influence of mental models on undergraduate students' searching behavior on the Web

Yan Zhang *

School of Information and Library Science, The University of North Carolina, Chapel Hill, NC 27599, USA

Received 27 June 2007; received in revised form 6 September 2007; accepted 8 September 2007
Available online 22 October 2007

Abstract

This article explores the effects of undergraduate students' mental models of the Web on their online searching behavior. Forty-four undergraduate students, mainly freshmen and sophomores, participated in the study. Subjects' mental models of the Web were treated as equally good styles and operationalized as drawings of their perceptions about the Web. Four types of mental models of the Web were identified based on the drawings and the associated descriptions: technical view, functional view, process view, and connection view. In the study, subjects were required to finish two search tasks. Searching behavior was measured from four aspects: navigation and performance, subjects' feelings about tasks and their own performances, query construction, and search patterns. The four mental model groups showed different navigation and querying behaviors, but the differences were not significant. Subjects' satisfaction with their own performances was found to be significantly correlated with the time to complete the task. The results also showed that the familiarity of the task to subjects had a major effect on their ways to start interaction, query construction, and search patterns.

Keywords: Mental models; Information searching behavior; Query construction; Search patterns

1. Introduction

Mental models are mental mimics of elements and structures of physical objects (Craik, 1943; Johnson-Laird, 1983). Research in human factors and human–computer interaction (HCI) suggested that it is mental models that enable people to interact with complex devices, including computer systems (Conant, 1970; Gentner & Stevens, 1983). In order to differentiate mental models from similar concepts involved in HCI, such as system conceptual models and system images, Norman (1983) referred to mental models specifically as users' mental representations of the system that they work with. At a more detailed level, Fein, Olson, and Olson (1993, p. 157) defined mental models as knowledge that users have about "how a system works, its component parts, the processes, their interactions, and how one component influences another".

It has been widely acknowledged that people build mental models when they interact with information retrieval (IR) systems (Frohmann, 1992; Jacob & Shaw, 1998). Although mental models are limited, they

* Tel.: +1 919 923 7173.
E-mail address: yanz@email.unc.edu

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2007.09.002

are enduring, accessible, and have great potential to account for human behavioral patterns and errors (Doyle & Ford, 1998). In previous studies on people's mental models of IR systems, mental models were imposed by training conditions (model/no model) (Borgman, 1986; Fein et al., 1993; Sasse, 1997), or assessed in terms of completeness by comparing to a reference model (Dimitroff, 1990; Efthimiadis & Hendry, 2005; Zhang, 1997). Both of the methods are unfit for the current study. Training is inapplicable because in reality people use the Web on a daily basis without any formal training; model quality assessment is unpractical because the Web is too complex a system to build a normative model for comparison.

Faced with these constraints, we need a new approach to assess mental models. Because our subjects are a group of homogenous undergraduate students who grew up with the Web, we treated mental model as a personal style, which reflects students' understanding of the Web and impacts their interactions with it. Styles of mental models were measured by students' drawings of their perceptions about the Web. The categorization of mental models was based on the theme expressed in the drawing rather than the comprehensiveness that the drawing illustrated. Previous studies have tried to categorize people's mental models based on drawings (e.g., Papastergiou, 2005), but few took this judgment-free approach. Also, few studies have examined the relationship between mental model styles and people's information searching behavior. In the current study, we explored the effects of mental model styles, reflected by drawings, on people's Web searching behavior. Our study is also a methodological attempt to explore whether drawing is an effective method to represent people's mental models of an information system.

2. Related literature

2.1. Drawing method

As abstract mental structures, mental models are hard to operationalize (Doyle & Ford, 1998). The main methods that researchers employ to solicit users' mental models of IR systems are interview and think-aloud protocol (e.g., Brandt & Uden, 2003; Katzeff, 1990). Both tools have limitations in that they are dependent on respondents' articulation ability. Some studies took an experimental approach by imposing different training conditions to create different model groups (e.g., Borgman, 1986; Kieras & Bovair, 1984). The training method falls short when subjects' background and prior experience are not well controlled. The inherent limitations of these methods call for new ways to represent people's mental models.

Drawing is a primitive form of communication. Gray (1990) pointed out that drawings can be used to elicit and illustrate structural aspects of users' mental models. Denham (1993) used drawing as a means to examine children's conception of computers. To investigate how people conceptualize Web search engines, Efthimiadis and Hendry (2005) asked 279 students in the University of Washington to draw sketches of how an Internet search engine works. In a proposed framework for IR systems evaluation and design, Pejtersen and Fidel (1998) pointed out that questions like "Please draw a diagram or picture of it" can help determine mental models of the Internet.

Empirical research also suggested that drawing is an effective method to represent mental models (Kerr, 1990). Thatcher and Greyling (1998) used drawing to solicit people's conceptualization of the Internet. Based on the structure and content represented in the drawings, they categorized them into six categories, such as interface and utilitarian functionality, simple connectivity, simple modularity, and networking. The categories were significantly correlated with the subjects' experience with the Internet. Papastergiou (2005) collected drawing sketches from 340 high school students in Greece to represent their mental models of the Internet. Eight categories, such as non-digital entity, services and content, user's computer, and network of computers, were identified. Significant correlations were found between students' drawings and their answers to a set of Internet related questions.

2.2. Mental models and search behavior

Young (1983) pointed out that mental models should be able to explain users' performance with the system they interact with. Borgman (1986) observed two groups of undergraduate students searching a library online catalog. One group was trained by a conceptual model of the system and the other was trained by procedural knowledge of how to search the system. No difference was found between the two groups on the success rate of simple tasks, but the model group outperformed the procedure group on complex tasks. Fein et al. (1993) imposed three training conditions, rote, explicit model, and full model, on three groups of subjects. Subjects' performance was measured by time spent working on the computer, success rate on retrieval tasks, and success rate on transfer tasks. Both the model groups outperformed the rote group on all the three measurements, while no significant differences were found between the two model groups. In an investigation of the relationship between mental models and search outcome, Dimitroff (1990) found that students with more complete mental models made significantly fewer errors and found significantly more items when they were searching the University of Michigan's online catalog.

Mental models should also be able to inform understanding of specific user behaviors, such as choice of method and nature of errors (Young, 1983). To explore the effects of mental models on users' behavior of using Excel spreadsheets, Sasse (1997) asked two groups of users with different levels of mathematics and computing background (mental models) to describe and use Excel. The math and computing savvy group (comparison group) described the system at a conceptual level, whereas the other group (main group) gave a purely procedural introduction to the system. The two groups also showed different behaviors; subjects in the main group tended to trade off physical efforts against cognitive efforts by specifying their own formulae instead of using system built-ins.

By interviewing and observing five nursing PhD students using the SPSSX statistical package, Staggers and Norcio (1993) also found that novice users and expert users differed in their behavior of using the software. Novice users had limited knowledge repositories and problem solving strategies and tended to depend on notes when performing tasks. Expert users, however, were well organized and confident. They employed trial-and-error strategies to solve emerging problems. In her observation of public library users using the Web and/or a Web based library catalog, Slone (2002) found that users' mental models affected their search approaches, Web sites visited, and sources used. Users with immature mental models of the Internet relied more heavily on the online catalog and off-line sources.

Researchers also investigated the effects of mental models on search patterns. Search patterns or search tactics are a set of temporally and semantically related moves (Wildemuth, 2004). They are an important component of search quality (Debowski, 2001). Different from search performances, search patterns reflect processing differences (Borgman, 1986). In Borgman (1986), search patterns were operationalized as search-state transitions. States were users' actions and system responses. Twelve states, such as search, display, and error message, were identified. Search patterns were measured at three levels: zero- (simple frequencies of states), first- (the frequency of two-state transitions), and second-order (the frequency of three-state transitions). Significant differences were found between the conceptual group and the procedure group at all the three levels on complex tasks but not on simple tasks.

In exploring characteristics of elementary school children's mental models of a full-text encyclopedia, Marchionini (1989) employed a similar strategy to examine search patterns. Nine states, such as new term, reorder, and show titles were defined based on children's behavior. Search patterns, reflected by first-order and second-order transitions, were then examined through the state-transition matrix for each search. He found that the children used a heuristic and highly interactive search strategy, indicating their unique mental models for the system.

These studies illustrate that mental models have strong effects on how people use IR systems. They motivate new research on the effects of mental models on people's information searching behavior in the World Wide Web environment.

3. Research questions

The study examined two research questions:

1. What types of mental models of the Web do undergraduate students have?
2. What are the relationships between students' mental models and their searching behaviors?

4. Research design

4.1. Data collection

The participants were 44 undergraduate students, 22 females and 22 males, from the University of North Carolina, Chapel Hill. They received class credits for attending the study. Among them, 90.9% were freshmen and sophomores, 6.8% were juniors, and 2.3% were seniors. The majority of them had not decided their major at the time the data were collected. The participants averaged 18 years of age.

The instrument used for data collection consisted of four sequential parts: a demographic questionnaire asking students' experience with the Web, an interview to solicit students' points of view about the Web, a request to draw a picture or diagram of their perceptions about the Web and provide descriptions for the drawings, and two search tasks. The first two parts of the instrument were introduced in detail in Zhang (submitted). We focus on reporting the remaining two parts here.

For the drawing task, every participant received the same instruction: "*Please draw a diagram or picture of your perceptions about the Web.*" After drawing, participants were asked to write one or two paragraphs to describe their drawings.

The two search tasks were:

(1) I want to buy a book named Sphinx written by Robin Cook. My main consideration is price. Please help me find one copy of the book with the lowest price.
(2) Each year, the US Census Bureau reports on the projections of national population. I am interested in the most current estimate for the population of the United States in 2010. Find this information.

The first question simulates tasks that are common and typical to undergraduate students. This intends to provide a familiar context for students to perform searches on the Web so as to elicit their typical Web searching behaviors. The second question was adopted from Wang, Hawk, and Tenopir (2000), where it was a simulation of questions recorded at a reference desk.

Subjects used the same computer setup to finish the search tasks. Search processes were captured by SnagIt software. After each task, subjects were asked to rank their perceived difficulty of the task and satisfaction with their own performance on a 5-point Likert scale. Data collection sessions were one-to-one. The author self monitored all sessions, which reduced biases caused by multiple monitors.

4.2. Data analysis

The first two parts of the data (demographic and interview) have been analyzed and reported in Zhang (submitted), where subjects' mental models of the Web as an IR system were examined in terms of four main components of general IR systems: information sources, information organization, search mechanisms, and interfaces. Subjects' collective mental model of the Web as an IR system was graphically illustrated. In this paper, we focus on analyzing drawings and recorded search sessions.

Drawings were classified from bottom up into four categories based on themes that they represent. Drawings are very illustrative, while sometimes hard to interpret. The author constantly referred to the descriptions provided by the subjects along with the drawings to ensure correct interpretation. To ensure consistency and reliability, a drawing was compared to other drawings to identify similarities and differences before the categorization decision was made. In addition, the author coded the drawings twice within one month and the coding results were very close. To ensure validity, a second coder coded the same 44 drawings independently, resulting in a more finely grained 8-category classification. After the two coders discussed the conceptual level at which the drawings should be coded to make the comparison between different groups possible, the second coder combined some categories, resulting in a 4-category classification similar to the author's. The agreement rate reached 93.2%. After discussion, undecided drawings were categorized into the most appropriate categories.

Search processes were recorded. To keep the search as natural as possible, students were not required to think aloud. The recorded clips were transcribed into analytical units. Each analytical unit is a webpage. When

Table 1
Coding scheme for search patterns

Beginning moves	
New concept	Enter term (s) for a concept that was not included in previous cycle
Move to reduce the size of the set	
Add concept	Add a concept that is not represented in the previous query
Combine with AND	Combine two concepts using the AND operator
Narrow term	Replace a term with a narrower term for the same concept
Move to increase the size of the set	
Delete concept	Delete a concept from the previous search cycle
Broaden term	Replace a term with a broader term for the same concept
Combine with OR	Combine two concepts using the OR operator
Others	
Repeat concept	Repeat the same concept in two consecutive moves
Replace concept	Replace a term with a synonym for the same concept

a participant moved from one webpage to another, one record was created. The actions executed on the webpage, such as clicking, inputting queries, and modifying queries, were also transcribed and incorporated into the record. There were two states, movement and backtracking, associated with each record. When a participant moved to the current webpage by linking or typing in URLs, the record was marked as "movement"; when he/she moved to the current webpage by clicking "Back" buttons on Internet Browsers, the record was marked as "backtracking". The rationale of the differentiation is that movement is more of an indication of sense of direction and sense of control, while backtracking is often regarded as a symptom of being lost (Fidel et al., 1999). Queries were also transcribed. There were two types of queries: queries submitted to search engines, and queries submitted to a specific website (non-search-engine sites).

Search patterns were analyzed in terms of conceptual transitions during the search. Each analysis unit is a concept. A coding scheme (Table 1) was constructed based on the scheme adopted by Wildemuth (2004). In most cases, multiple conceptual changes were involved in query construction and modification, for example, the query *sphinx robin cook cheap* had three conceptual transitions: New concept (sphinx)—add concept (robin cook)—add concept (cheap).

5. Results

5.1. Categories of mental models

Four mental model styles were identified based on drawings and drawing descriptions:

(1) Technical view. Students with this view looked at the Web mainly as a composition of computers, servers, modems, and CPUs. Some of the drawings or drawings descriptions also included people as a part of these systems. Fig. 1 shows a typical example of this view of the Web.
(2) Functional view. Students with this view saw the Web as a place for shopping (books, movies, tickets, and clothes), entertainment (movies, games, and chatting), emailing, paying bills, looking for information (news, sports, weather, and maps), and doing research (libraries). Fig. 2 shows a typical example of this view of the Web.
(3) Process view, or search engine centered view. Students with this view saw search engines as the center of the Web; all information branches off from search engines. Fig. 3 is an example.
(4) Connection view. Many students viewed the Web as a global-wide connection between information, people, computers, mobile phones, and webpages. They also viewed the Web as an important communication channel. Figs. 4a and 4b are two examples.

Table 2 shows the distribution of mental model styles among the subjects. The connection view was the most popular, followed by the functional view.

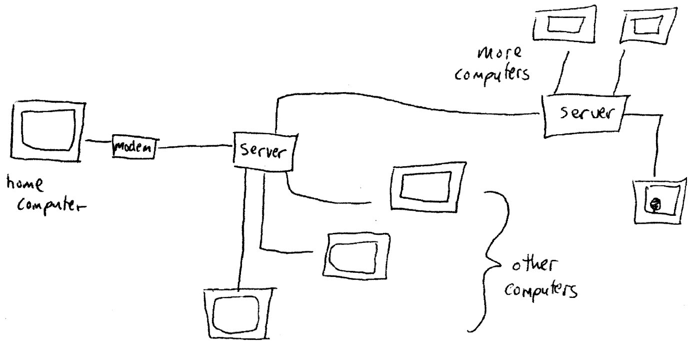

Fig. 1. Technical view of the Web.

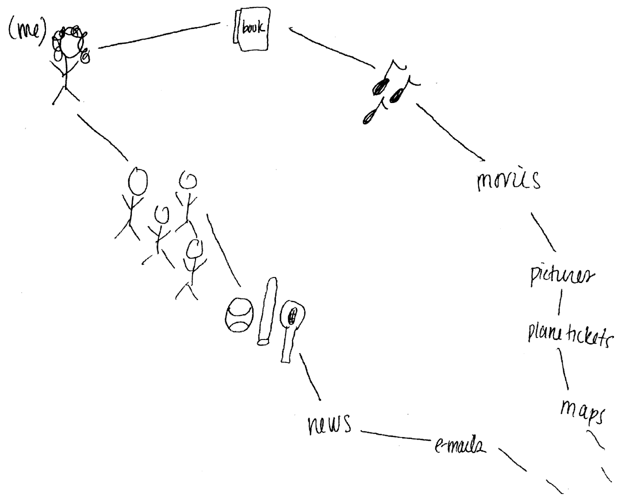

Fig. 2. Functional view of the Web.

Fisher's exact test showed that the distribution of gender among the four mental model groups was significantly different ($p = .0025$). Subjects in the technical view group were more likely to be males and subjects in the process view were more likely to be females. Pre-analyses of the data showed that gender did not have significant effects on participants' search behaviors in this study. The interaction effects of gender and mental model styles on search behaviors were not significant either. Thus, only behavioral differences in terms of mental model styles are reported in the result.

Past research reported that experience has major impacts on people's information search behavior (e.g., Elkerton & Williges, 1984; Borgman, 1989). Participants in the study on average had been using the Web for 8.2 years, ranging from 3 to 11 years (SD = 1.84). To examine the possible compounding effects of experience, a one-way ANOVA test was conducted. The test result showed that the four mental model groups did not differ significantly on their experience of using the Web ($F(3,40) = .566$, $p = .64$).

5.2. Search behaviors

Subjects' search behavior was measured from four aspects: (1) descriptive measurements, including time to finish tasks, accuracy of responses, number of movements, number of backtrackings, and ways to start interaction; (2) psychological measurements: feeling about the difficulty of the task and satisfaction with their own performance; (3) query constructions; (4) search patterns.

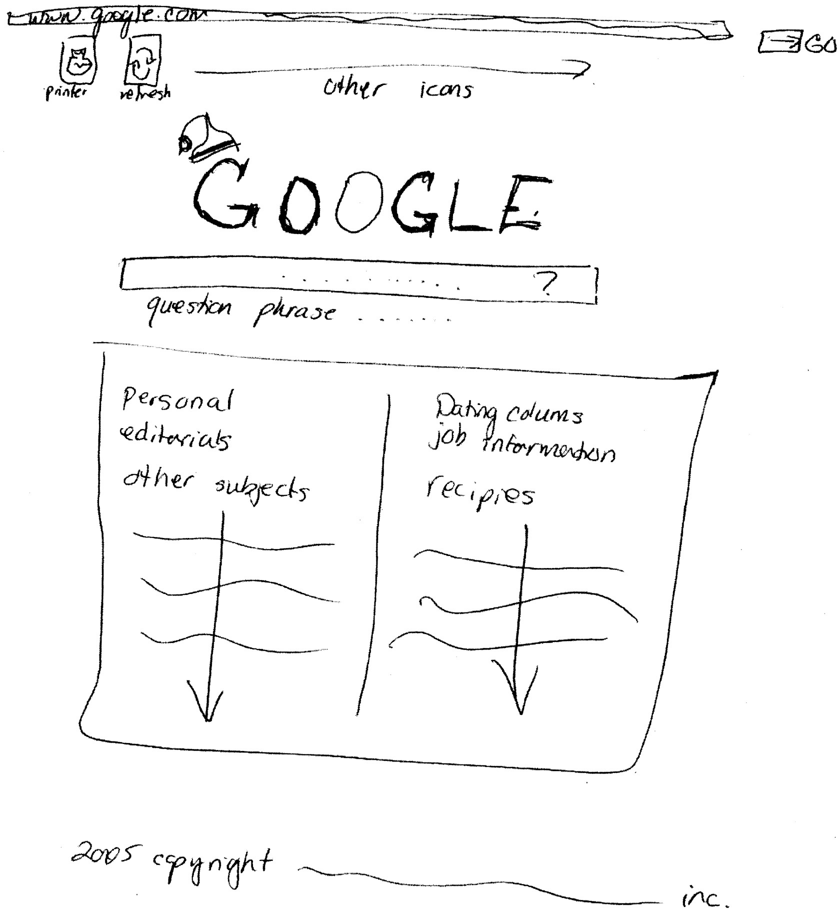

Fig. 3. Process view/Search engines centered view of the Web.

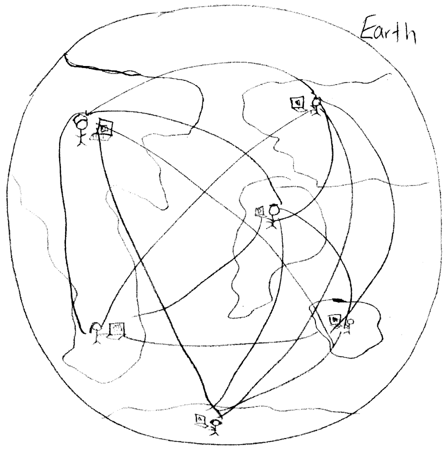

Fig. 4a. Connect people and computers around the world.

5.2.1. Descriptive measurement

5.2.1.1. Session length, accuracy of responses, number of movements and backtrackings. Session length is the time from the start to the end of the search. Accuracy evaluates the closeness of subjects' responses to the

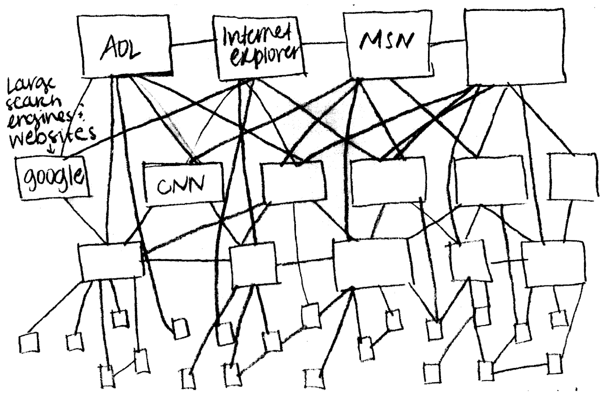

Fig. 4b. Connect information and webpages.

Table 2
Distribution of different types of mental models

Sex	Male (%)	Female (%)	Total (%)
Technical view	8 (36.4)		8 (18.2)
Functional view	6 (27.3)	5 (22.7)	11 (25.0)
Process view	1 (4.5)	7 (31.8)	8 (18.2)
Connection view	7 (31.8)	10 (45.5)	17 (38.6)
Total	22 (100)	22 (100)	44 (100)

actual answers. For task 1, the accuracy was evaluated by the price of the books that participants found. The lower the price was, the better the performance. For task 2, the accuracy was evaluated on an integer scale from 0 to 3: 0 – did not find the answer; 1– found answers on.com websites or speculated answers from graphs on the US Census Bureau website; 2 – found an answer on the Census Bureau website, but the answer was not the most recent statistics; and 3 – found the most recent statistics on the Census Bureau website.

Movement is defined as moving to a new webpage by linking or typing URLs. Clicking "Back" buttons on Internet Browsers to previously viewed pages is conceptualized as backtracking. Table 3 shows the descriptive statistics.

On average it took subjects 4.16 min to finish task 1 and 9.35 min to finish task 2. One-way ANOVA tests found no significant differences between the four groups on any of the four descriptive measurements.

5.2.1.2. Ways to start interaction. Students were offered two Web Browsers: MS Internet Explorer and Mozilla Firefox. They could pick either one to start searches. In each session the default page was the university's homepage for task 1 and the page on which they finished task 1 for task 2. The majority of the subjects

Table 3
Time to finish tasks (minute), accuracy, number of movement and backtracking actions (mean)

Mental models	*N*	Task 1				Task 2			
		Time	Accuracy	Move	Back	Time	Accuracy	Move	Back
Technical view	8	3.39	1.12	5.50	2.13	8.26	1.25	13.25	5.38
Functional view	11	3.63	1.09	3.91	1.18	8.17	1.27	12.91	4.36
Process view	8	5.04	1.06	5.88[a]	1.50[a]	10.41[b]	2.37	14.75[a]	9.63[a]
Connection view	17	4.45	2.70	6.31	1.93	10.17	1.41	15.94	9.13
Total	44	4.16	1.71	5.47	1.70	9.35	1.52	14.44	7.30

[a] One subject's data was missing for movement and backtracking.
[b] One subject's data was missing for time to finish task 2.

(70%) started their interaction with the Web using MS Internet Explorer. The four groups did not show any significant difference in their preference of Web Browsers ($\chi^2(3, n = 44) = 2.75$, $p = .43$).

For task 1, 25 (56.8%) students initiated the search process using search engines (Google: 19; Yahoo: 5; Firefox Google search box: 1). Nineteen students (43.2%) started from a specific website (Amazon: 13; Ebay: 3; Barnes and Noble: 1; Froogle: 1; Shopping Yahoo: 1). For task 2, the majority of the students (93.2%) started off from search engines (Google: 32; Yahoo: 6; Firefox Google search box: 3). Only 3 students tried to begin their exploration from the US Census Bureau website, and two of them typed incorrect URLs in the forms of "uscensusbureau.com" and "US CensusBureau". χ^2 test showed no significant differences among the four mental mode groups on their ways to start the search for both tasks (Task 1: $\chi^2(3, n = 44) = 1.78$, $p = .60$; Task 2: $\chi^2(3, n = 44) = 1.06$, $p = .79$).

5.2.2. *Psychological measurement*

After finishing each task, subjects ranked their feelings of the difficulty of the task and the satisfaction with their performance on a 5-point scale (Table 4).

No significant differences were found among the four groups on their feelings of difficulty and satisfaction. However, several significant correlations between subjects' feelings and their performances emerged. Nonparametric Spearman's rho tests showed that subjects' feeling of the task difficulty was significantly correlated with session length (Task 1: $r = .42$, $p = .005$; Task 2: $r = .40$, $p = .008$). For task 2, the feeling of satisfaction was negatively correlated with session length ($r = -.36$, $p = .02$). Furthermore, for task 2, both feelings significantly correlated with the accuracy of responses; the more accurate the response, the less difficulty subjects felt ($r = -.53$, $p = .000$), and the more satisfied with their performance ($r = .57$, $p = .000$).

5.2.3. *Query construction*

5.2.3.1. Descriptive analysis: number of queries and number of terms. In total, subjects submitted 95 queries for task 1 and 182 queries for task 2. Tables 5a and 5b show descriptive statistics of queries and terms.

Queries for task 1 and task 2 showed different patterns. First of all, for task 1, the number of queries submitted to search engines and specific websites were balanced: 47 (49.47%) in search engines and 48 (50.53%) in specific websites. However, in task 2, the great majority of the queries (84.6%) were submitted to search engines. Secondly, the number of terms per query for task 2 in both search engines and specific websites were more than those for task 1. Within each task, queries submitted in search engines on average had more terms than queries in specific websites. However, ANOVA tests did not show significant differences among the groups on the number of queries and query length.

5.2.3.2. Query facets. Queries are people's mental representations of the problem space that they want to tackle by searching IR systems. We conducted a conceptual analysis of queries with emphasis on identifying facets that subjects used to represent the search tasks. Tables 6a and 6b illustrate the analysis results.

For task 1, although the number of queries submitted to search engines and specific websites were almost the same (Table 5a), participants' use of concepts differed in the two cases. As shown in Table 6a, subjects used

Table 4
Difficulty with task and satisfaction with performance

Mental models	N	Task 1		Task 2	
		Difficulty[a]	Satisfaction[b]	Difficulty[a]	Satisfaction[b]
Technical view	8	1.25	4.75	3.38	3.38
Functional view	11	1.27	4.63	3.18	3.45
Process view	8	1.75	4.25	3.63	3.13
Connection view[c]	16	1.50	4.68	3.50	3.38
Total	43	1.44	4.60	3.42	3.35

[a] 1 – Very easy, 5 – very difficult.
[b] 1 – Very disappointed, 5 – very satisfied.
[c] One subject's data is missing.

Table 5a
Number of queries and terms – task 1

Mental models	Search engines		Specific websites	
	No. of query (mean)	No. of terms (terms/query)	No. of query (mean)	No. of terms (terms/query)
Technical view	9 (1.50)	29 (3.22)	5 (1.25)	11 (2.20)
Functional view	9 (1.50)	30 (3.33)	9 (1.29)	20 (2.22)
Process view	7 (1.40)	26 (3.71)	9 (1.29)	16 (1.78)
Connection view	22 (1.47)	92 (4.18)	25 (2.27)	47 (1.88)
Total	47 (1.47)	177 (3.77)	48 (1.66)	94 (1.96)

Table 5b
Number of queries and terms – task 2

Mental models	Search engines		Specific websites	
	No. of query (mean)	No. of terms (terms/query)	No. of query (mean)	No. of terms (terms/query)
Technical view	29 (3.62)	135 (4.66)	7 (1.40)	22 (3.14)
Functional view	42 (3.82)	192 (4.57)	1 (1.00)	1 (1.00)
Process view	24 (3.00)	136 (5.67)		
Connection view	59 (3.47)	318 (5.39)	20 (2.50)	71 (3.55)
Total	154 (3.50)	781 (5.07)	28 (2.00)	94 (3.36)

Table 6a
Concept analysis of task 1: facets, unique terms, and concept appearance

Facet	Terms	N. of terms	Search engines (%)	Specific websites (%)
Bibliographic information (title and author)	Sphinx/sphinx;	88	46.6	53.4
	Robin/robin/Robert; Cook/cook	126	63.5	36.5
Price and price related attributes	Price/price/prices/pricing; wholesale; lowest/low; cheap/Cheap; Used	29	100.0	
Product description	book/books; Novels	10	90.0	10.0
Purpose description	buy; compare/comparison	4	100.0	
Others	By/the/50 cents	14	92.9	7.1

both general and specific concepts in search engines: more than half of the specific terms (title and author) and the great majority of the general terms describing price, product, and purpose were submitted in search engines. However, in specific websites, subjects were more likely to submit specific requests (title and author).

Table 6b
Concept analysis of task 2: facets, unique terms, and concept appearance

Facet	Terms	No. of terms	Search engines (%)	Specific websites (%)
Description of purpose	Population/population/popukation; Census/census	153	85.6	14.4
Attributes of the population	projection/Projection/Projections/projections/projected/orjections; estimate/Estimate/estimated/Estimated/estimates/estimation/estimate; Predictions/prediction/predicted/predications; total	97	87.6	12.4
Organization	Census/census; Bureau/Bureau/bearu/Bereau	154	98.3	1.7
Geography	us/US/USa/Us; United/united; States/states; National/national	215	94.9	5.1
Time	2010; current; 2000; 2005; future; newest	148	79.1	20.9
Format	graph; report	2	50.0	50.0
Others	in, for, the, of, is, what, will, be, was, by	107	92.5	7.5

For task 2, subjects used the same set of terms in search engines and specific websites. Consistent with the fact that the majority of the queries were submitted to search engines, all the facets were used more in search engines than in specific websites, especially organization and geography facets (Table 6b).

Terms in both Tables 6a and 6b were original terms from subjects. It is apparent that subjects used synonyms to express concepts, such as estimation, projection, and prediction. Also, they used different forms of a term, such as prediction, Prediction, and predicted. Sometimes, there were misspellings, such as 'popukation' and 'bearu'. By and large, subjects tended to use natural language in their expressions, reflected by the use of stop words and the total absence of Boolean operators. However, other operators, specifically −, + and "", were used, albeit mostly in search engines.

5.2.3.3. Use of facets by mental models. Tables 7a and 7b illustrate the use of facets by different mental model groups. The number in parentheses is the number of subjects who used the facet.

For task 1, the title facet was used the most by subjects, followed by the author facet. Among groups, subjects with the connection view of the Web on average used more title, author, price, and product facets to search for the book than the other groups. But one-way ANOVA tests did not reveal significant differences among groups on the use of facets.

For task 2, the geography facet was used by all subjects in their searches. The other facets, target (population), attributes (projection), organization (the Census Bureau), and time (2010), were also used by the majority of the subjects. Similar to task 1, subjects with the connection view had more searches on most of the facets. Also, as in task 1, one-way ANOVA tests did not show significant differences among groups on the use of facets.

5.2.4. Search patterns

Search patterns were examined using maximum repeating patterns (MRP) analysis based on participants' conceptual movements in query constructions and modifications during the search process. MRP is an inductive approach to identify sequences of moves that repeat. It is a relatively low-cost method for locating interface problems (Siochi & Ehrich, 1991; Wildemuth, 2004). Due to the small sample size, our analysis of Web search patterns is exploratory. Figs. 5a and 5b illustrate the occurring families of search patterns for task 1 and task 2, respectively.

The first family of search patterns illustrates that in searching for task 1, participants initiated the search with a *New concept*, then added one or more new concepts to the search. The individual members of the family, that is, the unique search tactics, have different lengths. Some represent the entire search session, while

Table 7a
Use of facets by mental models (mean) – task 1

Mental models	Title	Author	Price	Product	Purpose
Technical view	1.63 (8)	1.43 (7)	1.67 (3)	1.00 (1)	1.00 (1)
Functional view	1.40 (10)	1.22 (9)	1.50 (4)	1.50 (2)	2.00 (1)
Process view	1.88 (8)	1.67 (6)	1.25 (4)	1.00 (1)	1.00 (1)
Connection view	2.71 (17)	2.13 (15)	1.86 (7)	1.67 (3)	
Total	2.05 (43)	1.70 (37)	1.61 (18)	1.43 (7)	1.33 (3)

Table 7b
Use of facets by mental models (mean) – task 2

Mental models	Target	Attributes	Organization	Geography	Time	Format
Technical view	4.29 (7)	2.71 (7)	2.14 (7)	3.38 (8)	4.29 (7)	
Functional view	3.60 (10)	2.00 (10)	1.63 (8)	3.64 (11)	3.10 (10)	
Process view	3.14 (7)	2.57 (7)	1.86 (7)	2.88 (8)	3.29 (7)	1.00 (1)
Connection view	4.06 (16)	3.08 (13)	2.25 (16)	4.65 (17)	4.00 (16)	1.00 (1)
Total	3.83 (40)	2.62 (37)	2.03 (38)	3.84 (44)	3.70 (40)	1.00 (2)

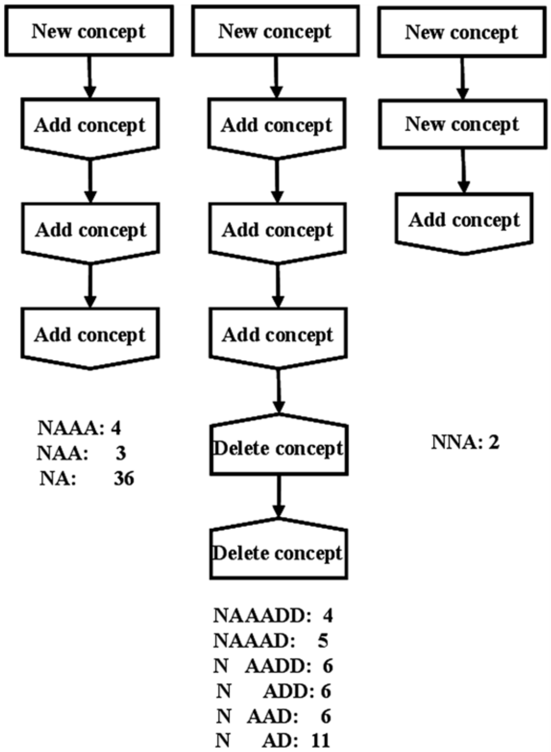

Fig. 5a. Repeating search patterns for task 1.

others are only fragments of a search. As would be expected, shorter tactics occur more frequently. It is worth noting that seven participants started with a *New concept* and ended the search there. This tactic is not listed in the figure.

The second family of search patterns for task 1 was characterized by one or more *Delete concept* modifications following an *Add concept*. The third family rarely occurred. It is listed as a distinct search pattern because in this case, participants gave up the foundation that they had laid for the search, and started a new one to tackle the task. This phenomenon is conceptually interesting and could be very common in the context of Web search given a larger sample size.

Inspection of search patterns among different mental model groups revealed that the connection view group showed more dynamic conceptual changes, but the differences were not statistically significant.

As illustrated in Fig. 5b, the families of search tactics for task 2 were similar to that found in task 1. The unique search pattern for task 2 is the third family: one or more *Add concept* movements were followed by one or more *Delete concept* movements, which were then followed by another *Add concept*. In task 2 there were only 3 participants who started with a *New concept* and ended the search there. As in task 1, the inspection of search patterns among mental model groups did not reveal significant group differences.

6. Discussion and conclusions

In existing studies, mental models were often measured on an ordinal scale in terms of completeness. By this means, mental models were implicitly judged as good or less good in quality. In a comparatively simple environment, the evaluation could be done by referring to a conceptual model of the system under research. However, the Web is a huge system. With the fast expansion and permeation into people's daily life, it is becoming a virtual counterpart of the physical world. For a system as big and complex as the Web, it is not appropriate to judge people's mental models of it in terms of completeness. The way people look at the Web is more of a reflection of their views of the virtual world represented by the Web. Thus, in this paper, people's mental

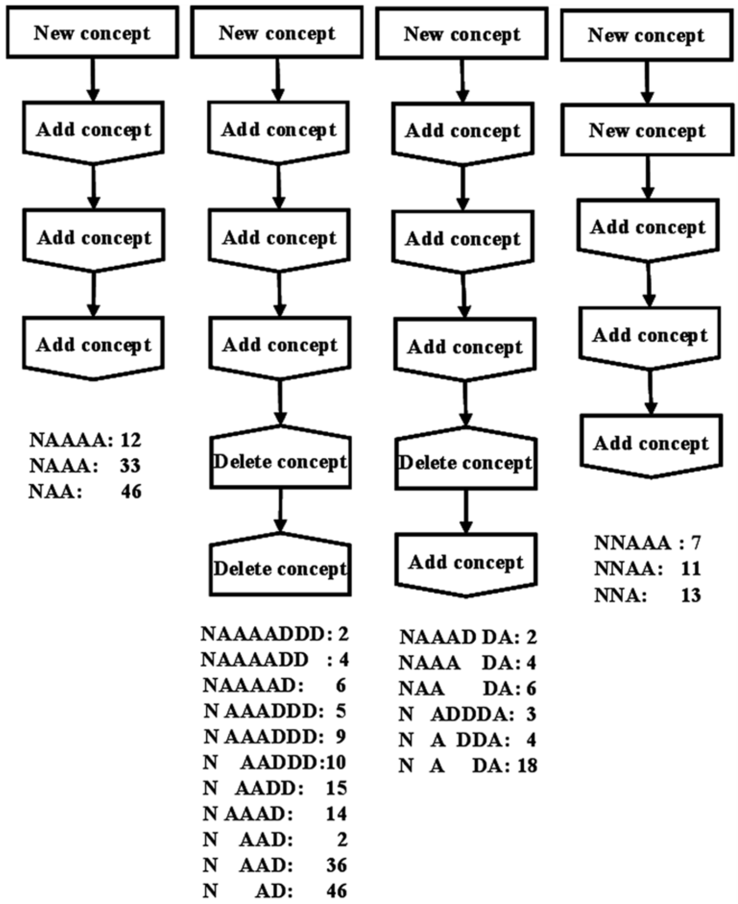

Fig. 5b. Repeating search patterns for task 2.

models of the Web were treated as a personal tendency, shaped by both people's personality traits and past experience.

In the study, students' experience with the Web, measured by how long they have been using it, did not significantly affect their mental models of the Web. The result suggested that measurement of people's past experience with a system needs to go beyond the time of usage. Other factors, such as frequency of use and usage patterns, should be given corresponding weights in the measurement.

As a primitive way to pass on knowledge and wisdom, drawings were effective in representing people's mental models of the Web in the study. Subjects with different mental model styles showed different performances and feelings during the interaction with the Web. For example, subjects with a process view of the Web spent the most time on both tasks and performed the best; nevertheless, they felt the most difficulty with the tasks, and the least satisfied with their performances. On the other hand, the technical view subjects took the least time to find a mediocre answer for task 1, but felt the least difficulty with the task and the most satisfied with their performances; the same happened to the subjects with the functional view in task 2. The correlation tests showed that subjects' feelings of the task difficulties were significantly correlated with the time they spent on the task. This suggests that time could be a strong indicator of people's feelings about tasks and their performances. Whether this relationship holds only in the Web environment needs further investigation.

Subjects' navigation behavior was investigated in terms of two actions: movement and backtracking. Movement indicates a well planned interaction, and backtracking indicates being lost (Fidel et al., 1999). The four mental model groups showed consistent differences on movement across the tasks. The connection view group on average made the most number of movements, followed by the process view and technical view groups.

On the other hand, the differences in backtracking among groups were not consistent. Future studies are needed to differentiate underlying cognitive rationales for backtrackings. For example, in some cases, subjects simply went back to the result list to check out another entry, which does not indicate being lost at all. Consistent patterns of backtracking might emerge after the differentiations are made.

How participants start their information searching process is of great value to website owners, especially commercial websites. The results suggested that the ways subjects started the interaction with the Web were affected by tasks' familiarity. Task 1, shopping for a cheapest book, is common and familiar to the subjects. Nearly half of them started from a specific website that they had already known and trusted, such as Amazon.com and Barnesandnoble.com. Conversely, task 2, looking for statistical information, is less common and familiar. Over 90% of the subjects started from search engines, specifically Google and Yahoo. The four mental model groups did not show significant differences on their ways to start interaction.

Querying is an important means for people to negotiate with the system. Large scale query studies that analyze large amounts of queries submitted to search engines (e.g., Jansen, Spink, & Saracevic, 2000) are effective in identifying subject trends and search errors. However, these studies often missed searching contexts and people's searching behaviors beyond the particular search engine. In the current study, the observation of both the context of query construction and subjects' behaviors beyond search engines were sustained. Query analysis suggested that the query attributes, specifically number of queries and query distribution, are also affected by the task. Subjects submitted more queries for task 2 than for task 1. In task 1, the number of queries submitted to search engines and to specific websites were nearly equal, whereas in task 2, the majority of the queries were submitted to search engines. The four mental model groups did not show significant differences on either attributes.

Attention was also paid to the average number of terms in a query. Past research on Web searching reported that the majority of the queries submitted to search engines were very short. Silverstein, Henzinger, Marais, and Moricz (1999) reported 2.35; Jansen et al. (2000) reported 2.21; Spink, Wolfram, Jansen, and Saracevic (2001) reported 2.16; and Wang, Berry, and Yang (2003) reported that 94.3% queries in their 1997 dataset were 1–3 words. In the current study, the average number of terms per query was 3.74 in search engines for task 1 and 5.14 for task 2, both of which are more than the means reported by previous Web studies. The result might be due to the fact that tasks in the current study are imposed questions and were explicitly described in the form of task statements. That is often not the case in real-life information seeking.

The average length of queries was 1.09 terms in specific websites for task 1 and 0.64 for task 2; both were much less than queries submitted in search engines. The difference suggests that information space might have implicit impacts on people's querying behavior. In search engines, subjects tend to be more descriptive given the huge information space that search engines connect to, while in specific websites, the boundary of the information space is more clearly defined, and subjects tend to express their needs in less, and more specific, terms.

Marchionini (1989) pointed out that a query serves as an indication of how the task was internally represented in people's minds. He conducted semantic analysis on queries, finding that subjects did not have difficulty grasping major facets such as person, place, and activity, and they typically used terms present in the task statement. Similar results were found in this study. The most used terms in queries were those that also appeared in the task statements, and our subjects did not have major difficulty in identifying facets. However, the subjects also used terms that connotate implicit aspects of task statements. For example, in task 1, some subjects used the terms 'comparison', 'used', and 'wholesale' to look for low prices. Among the four groups, the connection view group on average had more searches on most of the facets than the other groups, for both tasks. However, the differences were not significant.

Analysis of search patterns can improve system interactivity by informing new designs that support query formulation and modification. MRP analysis showed that the most common pattern of search tactic across two tasks was gradually narrowing down the retrieved set through addition of new concepts to the initial new concept. Another common search pattern was more exploratory: narrowing down the search by adding new concepts and then broadening the search by deleting concepts. The same search patterns were used by medical students when searching a microbiology database (Wildemuth, 2004). In the Web environment, participants also tended to switch to a brand new concept after the first concept failed to bring back satisfying results. In Wildemuth (2004), this pattern occurred only when participants' domain knowledge and experience with the database increased over a semester. Participants showed different search patterns for the two tasks; search patterns in task 2 were longer, more dynamic, and occurred more frequently than those in task 1, which might reflect the complexity of task 2 in comparison to task 1. The four groups did not show significant differences in search patterns.

In the Web environment, people can easily link from one Web space to another. The switch of Web space can be seen as another important dimension of search patterns. Although the search pattern analysis was limited to conceptual changes in the current study due to the small sample size, we conducted a preliminary examination of website transitions. The result showed that in task 1, subjects made more transitions between search engines and specific websites, indicating a progressive search path, whereas in task 2, subjects made more transitions from search engines to search engines, indicating a constant modification process.

The results of the study showed that although subjects with different mental model styles showed different online searching behaviors, few of the differences were statistically significant. Future studies can use other methods, such as think-aloud protocol and interview, in combination with the drawing method, to solicit people's mental models of information systems. Tests can be conducted to see if mental models identified by composite methods serve as a stronger indicator of people's information searching behavior. Consistent with findings in the current literature, task has a major effect on people's online searching behavior. In this study, search tasks affected an array of measurements, including the way that subjects started search, query constructions, and search patterns.

Acknowledgement

The author sincerely thanks Dr. Barbara Wildemuth for suggestions on the revision of the paper and Ben Birch for editorial help.

References

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.

Borgman, C. L. (1989). All users of information retrieval systems are not created equal: An exploration into individual differences. *Information Processing & Management, 25*(3), 237–251.

Brandt, D. S., & Uden, L. (2003). Insight into mental models of novice Internet searchers. *Communications of ACM, 46*(7), 133–136.

Conant, R. C. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science, 1*(2), 89–97.

Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.

Debowski, S. (2001). Wrong Way: Go back! An exploration of novice search behaviors while conducting an information search. *The Electronic Library, 19*(6), 371–382.

Denham, P. (1993). Nine- to fourteen-year-old children's conception of computers using drawings. *Behavior and Information Technology, 12*, 346–358.

Dimitroff, A. (1990). *Mental models and error behavior in an interactive bibliographic retrieval system.* Unpublished doctoral dissertation, The University of Michigan.

Doyle, J. K., & Ford, D. N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review, 14*(1), 3–29.

Efthimiadis, E. N., & Hendry, D. G. (2005). Search engines and how students think they work. *SIGIR'05* (pp. 595–596). ACM.

Elkerton, J., & Williges, R. C. (1984). Information retrieval strategies in a file-search environment. *Human Factors, 26*(2), 171–184.

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).

Fidel, R. et al. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society for Information Science, 50*(1), 24–37.

Frohmann, B. (1992). Cognitive paradigms and user needs. In *Cognitive paradigms in knowledge organization* (pp. 35–50). Madras: Sarada Ranganathan Endowment for Library Science.

Gentner, D., & Stevens, A. L. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Gray, S. H. (1990). Using protocol analyses and drawings to study mental models construction during hypertext navigation. *International Journal of Human–Computer Interaction, 2*, 359–377.

Jacob, E. K., & Shaw, D. (1998). Sociocognitive perspectives on representation. In M. Williams (Ed.). *Annual review of information science and technology* (vol. 33, pp. 131–185). Information Today, Inc.: Medford, NJ.

Jansen, B. J., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the Web. *Information Processing and Management, 36*, 207–227.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Katzeff, C. (1990). System demands on mental models for a fulltext database. *International Journal of Man-Machine Studies, 32*, 483–509.
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.
Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8*, 255–273.
Marchionini, G. (1989). Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science, 40*(1), 54–66.
Norman, D. A. (1983). Some observations on mental models. In D. R. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale, NJ: Erlbaum.
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., & Fidel, R. (1998). A framework for work centered evaluation and design: A case study of information retrieval on the Web. *Working Paper for MIRA workshop*.
Sasse, M. A. (1997). *Eliciting and describing users' models of computer systems.* Unpublished doctoral dissertation, The University of Birmingham.
Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. *SIGIR Forum, 33*(1), 6–12.
Siochi, A. C., & Ehrich, R. W. (1991). Computer analysis of user interfaces based on repetition in transcripts of user sessions. *ACM Transaction on Information Systems, 9*(4), 309–335.
Slone, 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.
Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the Web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52*(3), 226–234.
Staggers, N., & Norcio, A. F. (1993). Mental models – Concepts for human–computer interaction research. *International Journal of Man-Machine Studies, 38*(4), 587–605.
Thatcher, A., & Greyling, M. (1998). Mental models of the internet. *International Journal of Industrial Ergonomics, 22*(4–5), 299–305.
Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management, 36*, 229–251.
Wang, P., Berry, M. W., & Yang, Y. (2003). Mining longitudinal web queries: Trends and patterns. *Journal of the American Society for Information Science and Technology, 54*(8), 743–758.
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.
Young, R. M. (1983). Surrogates and mappings: Two kinds of conceptual models for interactive devices. In D. R. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 32–52). Hillsdale, NJ: Erlbaum.
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. (submitted). Undergraduate students' mental models of the Web as an information retrieval system. *Journal of the American Society for Information Science and Technology*.