Undergraduate Students' Mental Models of the Web as an Information Retrieval System

Yan Zhang
School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. E-mail: yanz@email.unc.edu

This study explored undergraduate students' mental models of the Web as an information retrieval system. Mental models play an important role in people's interaction with information systems. Better understanding of people's mental models could inspire better interface design and user instruction. Multiple data-collection methods, including questionnaire, semistructured interview, drawing, and participant observation, were used to elicit students' mental models of the Web from different perspectives, though only data from interviews and drawing descriptions are reported in this article. Content analysis of the transcripts showed that students had utilitarian rather than structural mental models of the Web. The majority of participants saw the Web as a huge information resource where everything can be found rather than an infrastructure consisting of hardware and computer applications. Students had different mental models of how information is organized on the Web, and the models varied in correctness and complexity. Students' mental models of search on the Web were illustrated from three points of view: avenues of getting information, understanding of search engines' working mechanisms, and search tactics. The research results suggest that there are mainly three sources contributing to the construction of mental models: personal observation, communication with others, and class instruction. In addition to structural and functional aspects, mental models have an emotional dimension.

Introduction

The term "mental model" was coined by Craik in his 1943 book *The Nature of Explanation*. He proposed that mental models represent external objects and phenomena internally in terms of words, numbers, and other symbols. Johnson-Laird further developed the construct of mental model into a mental model theory as an effort to explain human beings' logical fallacies concerning simple logical inferences. He argued that instead of employing formal logic, people construct mental models to make sense of and interact with the external world. In particular, Johnson-Laird saw mental models as sharing a similar relation-structure to that of the processes they imitate (Johnson-Laird, 1983).

Norman (1983) applied the notion of mental models to the field of human–computer interaction (HCI) and further distinguished it from the other models involved in the HCI process: system conceptual model (designers' model), system images (interfaces), and researchers' models of users' mental models. The term "mental model" thereafter specifically referred to the user's mental representation of systems. In line with Norman's definition, researchers in the information and library science (ILS) field generally define mental models as people's mental representation of information objects, information systems, and other information related processes.

Mental models are characterized as simple and naïve but critical constructs that people use to interact with various systems (Norman, 1983). Because the concept is abstract and difficult to measure directly, evidence of mental models was often gleaned indirectly through participant observation and interviews. The existing research has examined characteristics of mental models, such as accuracy and completeness (Dimitroff, 1990; Zhang, 1998), their developmental phases (Katzeff, 1990), transfer of mental models between systems (Marchionini, 1986), mental models' effects on people's interaction with systems (Borgman, 1986; Muramatsu & Pratt, 2001), and factors that affect mental models' sophistication (Papastergiou, 2005; Otter & Johnson, 2000; Thatcher & Greyling, 1998). However, the fundamental questions, such as what elements constitute mental models, what relationships exist between the elements, and what forms mental models take, have received less attention in the current literature.

This article presents an attempt to explore undergraduate students' mental models of the Web as an information retrieval (IR) system. The mental models are presented with respect to important components of general IR systems. A holistic view of the students' mental models of the Web also is presented. The Web is of interest because it is currently the most accessible and the most widely used information resource. In many instances, it has become the first stop or

Received September 23, 2006; revised May 9, 2008; accepted May 22, 2008

 • Published online 16 July 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20915

even the only place for finding information for many users. For younger users, furthermore, it is often the first IR system to which they are exposed. This population's mental models of the Web are likely to affect their behaviors with other types of IR systems, such as Online Public Access Catalogs (OPACs) and commercial databases. Undergraduate students were selected as a user group because they have grown up with the Web. Their mental models of the Web appear to be different from people who encountered the Web at a relatively later stage in their life. We expect that this research will inform better system design and user instruction for next generation.

Related Literature

The concept of mental model has been adopted by various HCI domains for the purpose of exploring cognitive processes underlying the observable behavior. A nonexclusive list includes text-editing programs (Douglas & Moran, 1983), calculators (Young, 1981, 1983), programming languages (Bayman & Mayer, 1983), software development (Dewier & Karnas, 1991), and manual controls (Wickens & Hollands, 2000). Researchers in the ILS field are interested in people's mental models of various IR systems, in particular, how users develop, use, and adapt their mental models to interact with systems. Research in this area includes mental models of online library catalogs (Borgman, 1984; Chen & Dhar, 1990; Dimitroff, 1990; Kerr, 1990; Slone, 2002), online news databases (Katzeff, 1990), Web search engines (Efthimiadis & Hendry, 2005; Muramatsu & Pratt, 2001), digital libraries (Makri et al., 2007), experimental IR systems (Savage-Knepshield, 2001), college Web sites (Otter & Johnson, 2000), and the Internet (Bruce, 1999; Kerr, 1990; Papastergiou, 2005; Slone, 2002; Thatcher & Greyling, 1998). Researchers also are concerned with users' mental models of information seeking when using specific media (Marchionini, 1989).

Mental models are abstract and, as a result, difficult to measure. A major challenge facing researchers involves the effective elicitation and representation of users' mental models of IR systems. Various methods have been used to represent or describe people's mental models. The main methods are:

1. *Eliciting people's verbal accounts.* There are three main forms: asking participants to describe a system or system's working mechanism (Dimitroff, 1990; Kerr, 1990; Muramatsu & Pratt, 2001; Papastergiou, 2005; Slone, 2002), asking participants for analogies or metaphors related to the system under study (Bruce, 1999; Sasse, 1997), and requiring participants to think aloud when performing search tasks (Katzeff, 1990). Transcripts of the verbal accounts were later analyzed to develop representations and evaluations of people's mental models of the particular system.
2. *Drawing.* Participants were asked to draw a picture or diagram depicting their mental image of a system. The complexity and sophistication of their mental models were evaluated through analysis of their drawings (Efthimiadis & Hendry, 2005; Otter & Johnson, 2000; Papastergiou, 2005; Thatcher & Greyling, 1998).
3. *Observing participants' errors during searches to identify gaps in people's mental models of a system* (Chen & Dhar, 1990; Huang, 1992). When the research purpose is to represent mental models, observation method is often used along with think-aloud protocols (e.g., Katzeff, 1990).

Although great efforts have been made to elicit mental models using interviews, few studies have illustrated mental models explicitly in terms of composite elements and structures grounded in the transcript. The existing studies have focused more often on evaluating mental models against a predefined framework—presumably a conceptual model of the system under study. For example, Dimitroff (1992) constructed an eight-concept scale based on system documentation and her own experience. The eight concepts covered the most basic features of an OPAC system, such as database contents, interactive nature of the system, multiple files, and use of controlled vocabulary. The completeness of mental models was measured by comparing participants' interview transcripts with the scale. In the exploration of the effects of users' characteristics on their mental models, Zhang (1998) arrived at nine essential concepts and attributes of IR systems by consulting a group of experts using the repertory grid technique. The nine concepts identified important components of an IR system, such as browsing, classification, data structure, and interface. The three attributes—format/process, targeted/untargeted, and specific to IRs/applicable to all information systems—were the dimensions on which concepts were judged. Subjects' mental models were measured by their ratings on the concepts and attributes.

Taking a data-grounded approach, Borgman (1986) developed a holistic index to evaluate users' mental models based on the interview transcripts. The index consists of three measurements: completeness of the model, accuracy of the model, and level of abstraction. "Completeness of the model" was based on the number of features mentioned. "Accuracy of the model" measured internal consistency of the models. "Level of abstraction" was a ranking based on whether the participant articulated an independently invented metaphor. In the study, she did not evaluate participants' mental models against a predefined conceptual framework as the two studies reviewed earlier; however, she also was not intent to construct participants' mental models from the interview transcripts.

Drawings are more illustrative than are verbal accounts. Studies using drawings to represent participants' mental models often categorized mental models based on the content and structure reflected in the drawing. Based on high-school students' drawings of their perceptions of the Internet, Papastergiou (2005) identified eight categories of mental models in terms of sophistication: (a) nondigital entity, (b) services and content, (c) user's computer, (d) huge remote computer, (e) connection between two computers, (f) few computers linked though a connection point, (g) computer network, and (h) network of computer networks. In analyzing drawings from a college sample that included undergraduates, postgraduates, and academic and administration staffs, Thatcher and

DOI: 10.1002/asi

Greyling (1998) identified six categories of mental models with respect to the model complexity: (a) interface and utilitarian functionality, (b) central database, (c) user to the world, (d) simple connectivity, (e) simple modularity, and (f) modularity and networking. The six categories of mental models were found to be significantly correlated with participants' experience with the Internet.

In exploring mental models and their impacts on information-seeking behaviors, attributes of mental models of information systems were identified. Papastergiou (2005) found that high-school students formed utilitarian rather than structural mental models of the Internet. Their mental models also involved misconceptions. Mental models of information systems are often simple, and sometimes rudimentary. In their investigation of students' mental models of digital libraries, Marki et al. (2007) found that students had limited understanding of how documents are organized, how to tailor queries to particular search engines, and how search results are ranked. By exploring the transfer of high-school students' mental models of information seeking from an old print-encyclopedia system to a new electronic system, Marchionini (1989) found that mental models represent both structural and functional aspects of systems. Students who transferred only content knowledge were more capable of building new mental models for the electronic system than were those who transferred all levels of detail.

In summary, in the current literature, verbal accounts are a main resource for representing mental models. However, few studies using verbal accounts to represent mental models have tried to construct people's mental models bottom-up from the data; rather, they predefined a conceptual framework and evaluated users' mental models against the framework. This approach limits the emergence of dimensions of mental models beyond the predefined framework. The other main method used to elicit and represent mental models is drawing. Drawing can neatly illustrate the structural aspect of mental models; however, the analysis of drawing is very subjective and vulnerable to misinterpretations. Thus, studies that employ multiple methods are needed to explore the fundamental elements and characteristics of mental models.

Research Questions

In this study, the following research questions are explored:

1. What experiences do undergraduate students have with the Web?
2. What are students' mental models of the Web as an IR system?
   - 2a. What do they understand about information sources on the Web?
   - 2b. What do they understand about information organization on the Web?
   - 2c. What do they understand about Web search engines' search mechanism?
3. What is undergraduate students' collective mental model of the Web?

Method

Participants

A total of 44 (22 males, 22 females) undergraduate students from the University of North Carolina at Chapel Hill participated in the study. Of the 44 participants, 25 (56.8%) were freshmen, 15 (34.1%) were sophomores, 3 (6.8%) were juniors, and 1 (2.3%) was a senior. The ages of the participant ranged from 17 to 22 years, with an average of 18 years ($SD = .83$). Since the majority of the participants were freshmen and sophomores, many of them had not decided on a major area of study. Participation was voluntary. The students were given class credits as compensation for their time spent in the study.

Materials

To illustrate mental models in a holistic manner, multiple data-collection methods were used in the study. Using multiple data-collection techniques is useful for triangulating research results and making interpretations more objective. Data were collected using four approaches. The first approach is a short questionnaire measuring students' prior experience with the Web. Previous research has noted that the construction of mental models is affected by users' prior experience with the system (Papastergiou, 2005; Thatcher & Greyling, 1998; Zhang, 1998). Knowledge about participants' experience with the Web will help us to make sense of their mental models of the Web. The current indicators for users' experience include the length and frequency of Web usage, the types of activities on the Web, and prior learning experience.

The second approach is a semistructured interview. The user–Web interaction model proposed by Wang, Hawk, and Tenopir (2000) was employed as a framework to develop the interview guideline. The model has four main components: Web space, information structure, search functions, and Web interface. The semistructured interview asks participants' views of the four components, with a focus on the first three components. All interviews were audio recorded and transcribed.

The third approach is a drawing task. Participants were asked to draw their perceptions about the Web in the form of a picture or diagram. They were asked to describe the drawings, which reduced the risk of drawing misinterpretation.

The last approach consists of two search tasks. The first search task asks participants to find the lowest price for a book on the Web, and the second search task asks participants to seek the most current estimate for the U.S. population in 2010. This approach allows the researcher to observe how students' mental models affect their search behavior and navigation through the Web space while performing actual information-seeking tasks. Participants could choose either MS Internet Explorer (IE) or Mozilla Firefox for their searches. Each

DOI: 10.1002/asi

participant performed the two tasks in the same order. After each task, students were asked to rank the difficulty of the task and their satisfaction with their own performance on a 5-point Likert scale (1-very easy, 5-very difficult; 1-very disappointed, 5-very satisfied). Screen movements during the Web search were video recorded. Students were asked to describe their search strategies in postsearch interviews.

Procedure

Participants were scheduled to take part in the study at their convenience. Upon arrival, students were welcomed and were provided a brief introduction to the purpose and components of the study. After signing an informed consent form, all participants followed the same procedure.

Participants first completed the questionnaire. Then they were interviewed based on the interview guideline (see Appendix). Upon completing the interview, participants were asked to draw a diagram or a picture of their perceptions about the Web. Finally, they were provided an IBM laptop computer to perform two search tasks. Video of participants' Web navigation was captured using SnagIt 7 software. Once the participants finished the search tasks, they were interviewed about their search strategies for each task. All participants used the same computer setup and were given the same two search tasks in the same order. There were no treatments or controls on how participants interacted with the Web.

Data Analysis

Four types of data were produced in the study: (a) demographic data acquired through the questionnaire, (b) interview transcripts, (c) drawings of users' perceptions of the Web and descriptions of the drawings, and (d) the videotaped search process. The questionnaire data were analyzed using descriptive statistics. Interview transcripts were analyzed through content analysis to identify emergent themes. The participants' drawings were grouped into categories. The videotaped search processes were analyzed to identify participants' search patterns on the Web. Due to space limitations, only the first two types of data along with drawing descriptions are reported in the article. The other data (i.e., drawings and videotaped search processes) are reported in Zhang (2008).

Interview transcripts and drawing descriptions were imported into QSR N6 software. Open coding was employed to code the data. Whenever an entity such as database, computer, network, or a statement about relationships between entities appeared, it was coded into a category. If the entities and/or relationship were illustrated in the drawings instead of the drawing descriptions, they also were coded into corresponding categories. Triangulation through multiple methods revealed a more complete picture of mental models. The identified categories then were organized based on the user–Web interaction framework proposed by Wang et al. (2000). The researcher did all coding. To ensure the validity of the coding schema, a second coder coded 20% of the transcripts. The intercoder reliability was calculated using the formula $A = M/(\sum_i^n N_i/n)$, where M is the number of coding events agreed upon by all cross-coders, N_i is the number of coding events assigned by the ith coder, and n is the number of coders (Kracker & Wang, 2002). The intercoder agreement reached 87.5%.

Results

Experience with the Web

The construction of mental models of the Web is affected by users' experience with the Web. In the study, participants' Web experience was measured by how long and how frequently they used the Web, how they learned to use the Web, and what they do on the Web.

How long and how frequently did they use the Web? Participants averaged about 8.2 years of experience using the Web ($SD = 1.84$), and they search the Web about 6.94 times per day ($SD = 7.48$). This group of students saw themselves as growing up with the Internet and regarded surfing the Web as an important part of their lives. One participant remarked during the interview that "I have been using computers for a great part of my life. I grew up with them. I think it is easier for me to access information on the Web than my dad."

How did they learn to use the Web? Thirty-eight (86.4%) participants reported that they learned to use the Web through trial-and-error, 11 (25.0%) were taught by friends, 4 (9.1%) learned by reading materials, 2 (4.5%) learned at school, and 2 (4.5%) learned by attending seminars and workshops. One (2.3%) participant indicated that he or she was taught by parents. Eleven (25.0%) students learned to use the Web by more than one means.

What do they do on the Web? Table 1 shows the aggregated results of undergraduate students' online activities by gender. It is not a surprise that these college students all use e-mail and use the Web for their course and research work. In addition to the activities listed in the table, participants also added that they play games, post photos, check the weather, search offbeat topics, and check for church news and events on the Web.

TABLE 1. Online activities by gender.

	Female (%)	Male (%)	Total (%)
E-mail	22 (100)	22 (100)	44 (100)
Course/research	22 (100)	22 (100)	44 (100)
Socialize	20 (90.9)	20 (90.9)	40 (90.9)
Entertainment	17 (77.3)	22 (100)	39 (88.6)
Reading news	17 (77.3)	18 (81.8)	35 (79.5)
Shopping	15 (68.2)	15 (68.2)	30 (68.2)
Medical information	6 (27.3)	7 (31.8)	13 (29.5)
Financial management	6 (27.3)	5 (22.7)	11 (25.0)

DOI: 10.1002/asi

Undergraduate Students' Mental Models of the Web as an IR System

Four components are important for understanding an IR system: information source, information organization schema, search mechanism, and interface (Wang et al., 2000; Zhang & Wang, 2005). To illustrate undergraduate students' mental models of the Web as an IR system, data are reported in terms of the following components:

- *The Web space.* This component measures students' understanding of information sources, information objects, and collections of information objects on the Web. An information object is regarded as the basic unit of the Web, and it can be in the form of text, an image, a video, etc.
- *The Web's structure.* This component measures students' understanding of how information objects on the Web are organized for access.
- *The search mechanism.* This component measures students' knowledge of search engines' underlying logic, rules, and mechanisms; that is, how search engines process queries and generate output.
- *The interface.* This component measures participants' understanding of general interface elements, including navigation tools and results access.

Because there were no specific interfaces involved in the study, students' mental models of the Web as an IR system were mainly illustrated by their understanding and perceptions of the first three components. Participants' perceptions of the interface elements were briefly reported when they occurred in the interview or drawing descriptions.

The Web space—Information sources and objects. The Web is viewed by most of the students (77.3%) primarily as a source of information. One of the students said that "The World Wide Web is an information super highway. It is a quick way to search for information. It replaces [the] library as the most important source of information." When talking about information on the Web, students referenced its creation, format, subject, and organization.

Creation. Students recognized that Web sites are created by different entities. There were personal Web sites created by individual people and organizational Web sites created by various organizations. Students also tended to judge the credibility of information based on the credibility of the information creators. A typical comment was: "You have to learn which Websites to go to for the right material that is correct. [You] research a topic; you should go check all the sources: where it is coming from. If it is [from] school, that is trustful."

Format. Dublin Core defines information format as "the physical or digital manifestation of the resource." Students were able to identify various information formats in the interviews or drawings, such as moving image (movie), pictures, library databases, books, specific articles, and sound (music). One student also saw instant messaging as a format of information resource that he or she can turn to for information.

Subject. Students access a wide range of information on the Web. A nonexhaustive list includes weather, news, maps, games, sports, company information, stores, products, advertisements, and dictionaries.

Information organization. Most students (86.4%) have mental models of how information is organized on the Web and were able to articulate them. But the rest of the participants (13.6%) either do not have a model or could not put their thoughts into words. Specific information-organization schemes described by the students are reported in the following section.

Information organization.

Levels of understanding. When asked how they think information is organized on the Web, students had different levels of information organization in mind. Table 2 shows the views and corresponding examples. The number in parentheses is the number of students who had the particular view.

Information-organization schemas. The information-organization schemas that students mentioned may be classified into the following categories. The number in parentheses is the number of participants mentioning the category. Due to space limitations, examples are not provided for self-explanatory categories.

- By subject/topic/category (15)
- By hierarchy (9)

TABLE 2. Understanding of information organization on the Web.

Level of perceptions about information organization	Example comments
No view (4) (No view or could not articulate it)	I never really think about that. No, I really haven't thought about it.
Micro view (15) (Individual Web site)	[The Web] is very structured. You have your links. Then it goes from broad to more specific. Just keep getting more specific. I believe. You get your homepage. There are links on that.
Macro view (15) (The whole Web)	I guess it is organized by different work it does. Organization, military, or government, commercial. Those are organized by the last three letters in the address.
Holistic view (10) (Individual Web sites & the whole Web)	In websites, it is organized by page, file, and links that get to other Websites. On a higher level, it organized in giant networked servers. Different people have various spaces in different servers. They list their things there, so they can share with everybody.

DOI: 10.1002/asi

- Organized by what you are looking for (7)
- By search engines (6)
- By keywords (4)
- By Web sites (4)
- By servers and computers (2)
- By URL domain (.com, .org, and .gov) (2)
- By dates (1)
- In alphabetical order (1)
- By how many people have looked at them (1)
- By meta tags (1)
- By closeness/relevance (1)
- By individual people (1)

  Example: I think it is organized on an individual basis. You have one person who organizes one thing. That information is passed on to everyone. You have a bunch of different individuals organizing their own stuff and putting it on the Web.

- By accessibility (1)

  Example: Information is also organized in a way where it protects certain people, certain age groups like I know little kids are not able to access something that older adults are able to access.

Students' mental models of information organization on the Web differed in complexity. Among the students who were able to articulate their mental models, 68.4% had only one schema in mind while the others held multiple organizational schemas simultaneously, contingent on particular scenarios. For example, one student noted that:

> [Information organization on the Web] is a little more complex. It is kind of hard to understand exactly how it is in there. A lot of times, you got articles that come with dates and [are] organized on dates. A lot of stuffs are organized alphabetically depending on what you are looking at. Most of them are organized by dates or by how many people have looked at them. I know if you use a search engine, it will come up with the mostly used one.

Evaluation of information organization on the Web. Students had different attitudes toward information organization on the Web. Eight participants (18.2%) were very positive, suggesting that information on the Web is well laid out, structured, and user friendly. For example, one participant said that the Web is "Easily formatted; it is easy to find [information]; it is organized very well."

Five participants (11.4%) expressed negative attitudes, suggesting that information organization on the Web is complex, random, and hard to understand: "It is not organized that well [. . .] A lot of things, they are just a bunch of random webpages that [are] usually unhelpful at all."

Searching mechanism.

Avenues for getting information. Students had various ways to get information on the Web. The majority (95.5%) mentioned that search engines, especially Google and Yahoo, provide a primary avenue for them to access information. Over half of the participants (56.8%) had a list of specific Web sites in mind to which they refer for something specific. For example, participants went to Web sites such as CNN for news, weather sites such as weather.com to check weather, UNC Webmail to check e-mail, the UNC Web site for university-related information, the UNC library site for articles and papers, Amazon.com for shopping, and dictionary and Wikipedia for research work.

The extent to which students are aware of their needs has a major impact on how they look for information. If students knew exactly what they wanted, it is likely that they went to the specific Web sites on their list; if they were uncertain about what they wanted, they turned to search engines. One student commented that:

> I go online, depending on what type of information I am looking for. Just random information, I usually use Google [. . .] But if I know specifically what I want, like if I want a book note, I go to a website for book notes. And if I want some stuff on UNC, I search their website.

Another student stated that:

> When I am searching for sort of superficial information, I go straight to Google. Usually that will give me the most relevant ones to click on for that general topic. But if I am searching for something pretty specific, like an article or certain book [. . .] I pretty much go straight to school through the UNC website, the library link. If I am looking for a movie, I go to the movie website, if I am looking for new music, I can go to the music website.

Four students (9.1%) said that they also took recommendations for information sources from other people, especially friends. One student recalled that: "Being in college, I don't even have a facebook account until I got here. A couple of weeks in the school, everybody asked me about do you have one, do you have one?" Five students (11.4%) mentioned that e-mail, instant messaging, and some commercial toolbars were other sources where the found information.

How search engines work. Students use search engines as a primary source for locating information. Students' understanding of search engines' working mechanism influences their search strategies. Based on the analysis of the transcripts, five aspects of the students' mental model of search engines were identified: (a) components of search engines, (b) search process, (c) ranking mechanism, (d) supported functions, and (e) attributes of search engines.

Components of search engines. Nielson (2005) noted that searching is a prominent part of the Web experience; Web users have developed a firm mental model for how it is supposed to work. In terms of the search interface, users expect three components: a box to type in words, a button they can click to run the search, and a result list. In our interviews with the students, their mental models of search engine interfaces conformed to Nielson's model. In addition to the interface, students mentioned other components of search engines, such

DOI: 10.1002/asi

TABLE 3. Ranking mechanisms of search engines.

Ranking mechanism	Example comments
Word frequency (11)	The first one probably is the one that has the most frequent appearance of key words.
The number of hits/popularity (7)	The most popular ones will come first. If I am searching for American history for 1950s, there maybe the most popular sites that people click on. Those [sites] will come up first.
Relevance (5)	[Search engines] try to put the most relevant one onto the top, as you go on, it will become less relevant.
No. of keywords contained (3)	If you type in three keywords, if [the source] has all three keywords, it will be ranked higher than those [that] have two keywords.
Where the keywords appear (2)	If [the keyword] is in the title, it will be ranked higher than if it is just in the body of the webpage.
If keywords are together (2)	Search engines take out the words to see if those two words are found together versus they are found separately in an article; the one they are together might have higher relevancy.
Specific to general (1)	Most specific would be the top one [. . .] from more specific to general.
Importance of sources (1)	I do feel that they may list the most important at the top and people's comments at the bottom. Something like a company, a major site like encropeida.com or the dictionary [will] come first. And then random people's comments, personal websites.
Sponsored or not (1)	I know that search engines like Google is sponsored by a lot of websites. These sponsored websites come up with higher level of relevance when you search for stuff.

as Web crawlers, algorithms, and the search engines' encoded databases.

Search process. Students possessed different mental models of how search engines produce search results. The most common understanding of the process was through keyword matching, which 40.9% of the participants explicitly recognized in the interview. A typical example was: "[Search engines] use the keywords that you typed in. It just searches for websites or webpages that have those same words, basically."

Two students had a more complex model of search engines' search process, which involved processes beyond the keyword matching. For example:

> When web pages are created, [search engines] have giant crawlers that go cross the Internet to find documents or webpages, and then they categorize them by all sorts of things, by the words that are contained within webpages, by the title of the webpages, by the authors of the webpages, and then, they are assigned certain status, I guess. So when you search for something on the Internet, you enter key words, and it looks through the catalogs of keywords, and titles, and authors, all that . . ., that kind of combined content score, so that's what you see the top entries.

Three students held a naïve model of how the system returns search results, thinking that an actual person was performing the search for them: "I think there are people who don't have anything to do, sit at home, and type in this little box and they got anything containing this word as specific keyword that is in your search."

Additionally, one student indicated that search engines have presearched everything: "Google is set up so that it has already presearched everything, like it knows what's already on the Web."

Ranking. Students have various mental models of how search engines rank the results. Some models are vague: "You type in something, whichever sources have those words in there, that's what comes first. Like you type in Tom Sawyer, then everything have Tom Sawyer will come up first. I don't know what they do after that." However, in most cases, students were able to articulate their conceptions of the criteria that search engines use to rank the search results. Table 3 lists ranking mechanisms that students mentioned. The number in parentheses is the number of students who mentioned the particular mechanism.

Note that students did not simply view the results ranking as "one mechanism rules all." Instead, they often thought that ranking was based on multiple criteria.

Functions supported by search engines. Table 4 summarizes the functions mentioned by the students in the interview.

Search engines provide many operators for end users. For example, Google provides the operator ":pdf" for users to retrieve PDF files only; however, users do not keep all the

TABLE 4. Mentioning of search engine functions.

Functions		No. of participants (%)
Phrase search	""	21(47.7)
	+	6(13.6)
	( )	4(9.1)
Boolean operators		4(9.1)
Advanced search		2(4.5)
Wildcards		1(2.3)
Eliminate duplicate results		1(2.3)
Highlights of the cashed page		1(2.3)

DOI: 10.1002/asi

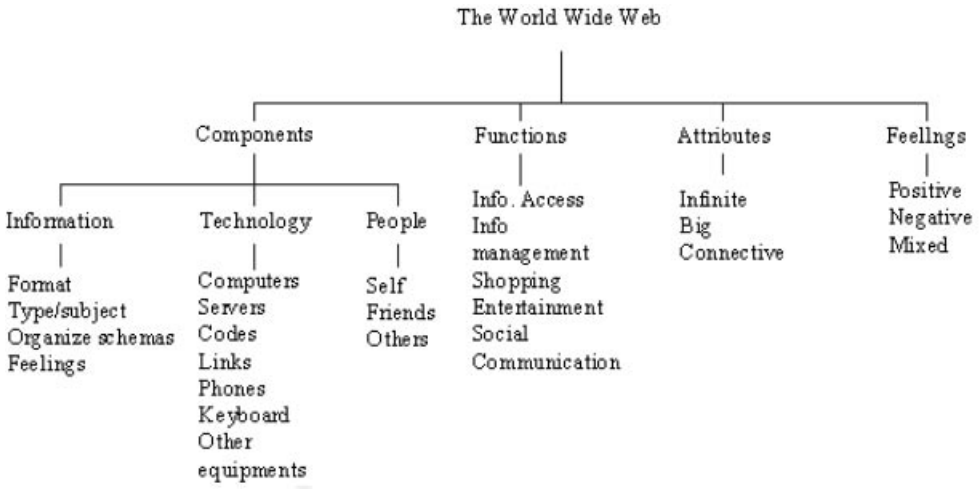

FIG. 1. Undergraduate students' collective mental model of the Web.

operators in mind. During the interview, only the six functions listed in the table were reported by the students. Among the functions, phrase searching was mentioned the most, followed by Boolean operators. The fact reflected that undergraduate students had simple and incomplete mental models of functions supported by search engines.

Attributes of search engines. During the interviews, students mentioned several search engines' attributes. One student recognized that search engines are different from each other in terms of scope: "Different search engines sometimes have different things. [If you] search for pictures. Yahoo, Google, they came up with different lists, not exactly the same." Two students realized that search engines overlap with each other: "A lot of search engines seem to overlap. Some search engines use the resources of other search engines." One student found out that search engines do not deal with current events: "It is hard to find things for me about current events for projects. You should not just use Web search engines."

It is widely known by students that search engines are often sponsored by certain Web sites: "Google makes so much money because companies pay Google a certain amount of money to have them to be the suggested sites that you visited."

Search tactics. Students used various tactics when searching on the Web. The following is a nonexhaustive list of general tactics that students used, followed by the number of students who mentioned the tactic.

- Be precise and specific: Use more words instead of one. (20)
- Try several different words choices. (5)
- Start specific, then broader. (4)
- Start general, then more specific. (4)
- Leave stop words/common words. (4)
- Skim reading and compare. (4)
- Search a couple of different search engines. (3)
- Pick up the main words. (2)
- Do not use sentences. (1)

Undergraduate Students' Collective Mental Models of the Web

The Web is a huge system with infrastructures and functions that go beyond the scope of traditional IR systems. Correspondingly, students' perceptions of the Web also went beyond the four main components of general IR systems outlined earlier. By combining data from interviews, elements depicted in drawings, and drawings descriptions, a broader picture of undergraduate students' "collective" mental model of the Web emerged (see Figure 1). This "collective" representation of the mental models of the participants is intended to provide a holistic view of students' understandings of the Web. Meanwhile, it serves as a tangible structure based on which different user groups' mental models of the Web could be compared.

As shown in Figure 1, the undergraduate students' collective mental model consists of four aspects: (a) components (i.e., what the Web consists of), (b) functions of the Web, (c) characteristics and attributes of the Web, and (d) feelings about the Web. In the previous section, students' mental models of the Web as an IR system have been discussed in terms of information source, information organization, and search mechanism. In the following sections, other aspects of students' mental models of the Web are briefly reported.

Components of the Web. The students' view of the Web, which emerged from the transcripts, drawings, and drawing descriptions, consists of people, technology, and information. People are users of the Web, various technologies constitute the infrastructure of the Web, and information is the content of the Web.

People. The concept of a user is complicated with multiple dimensions (cognitive, affective, and physical) and is influenced by dynamic situational factors (Wang et al., 2000). The "user" concept was simply viewed by students as people who conduct a series of activities of uploading, using, and

DOI: 10.1002/asi

sharing information. The Web is viewed as a connection between "everyone in the world who has a computer."

Technology. Technology is an integral part of students' perceptions of the Web, and it was often interwoven with information in their descriptions. They thought that the Web consists of various technologies and technical devices such as servers, computers, programming codes, service providers (ISPs), databases, files, zeroes and ones, files, search engines, links, and other equipment such as keyboard and phones.

Students sketched the Web in terms of its structure at different levels of complexity. The Web has been described from a flat and single-layered technical perspective. For example, one student described the Web as: "Lots of computers, all contain information and enable [the information] to be viewed by all the other computers."

The Web also has been described as a hierarchy of different technical devices: "Individual computers connect to the Internet through some sort of servers. The Internet itself is comprised of many servers and computers, large and small, around the globe."

Students also described the connection made possible by Web technologies from both macro and micro levels. One example of the macro view of the connection is: "Individual users are connected through their ISPs. ISPs are connected at the POP sites, and the POP sites are connected so, in turn, every one is connected somehow." An example of the micro view of the connection is:

> You start out on the web with the search engine or website that you have as the homepage and then there are a million link[s] that can be clicked on to take you to a new website and then on that website there are even more links and it is just a never ending process.

Information. Information and information organization have been elaborated in two subsections, *The Web Space* and *Information Organization*, in the second part of the result (Undergraduate students' mental models of the Web as an IR system).

Functions. It is very common for students to have a utilitarian view of the Web. The Web has been seen as a connection between people, companies, and organizations, as a tool for quick and easy communication with others, and as an easy-to-access information resource that is more preferable than the library. For example, one student said that: "A lot of people use it to communicate with each other among long distances [. . .] It is an easy way to store information and search it. It will be hard if you had to go to the library or some other catalog places." Another student commented that: "It is an access to any information that you need at any time. It is not just for entertainment or academics. You can do about anything on there. If you are bored, you can find something like games, jokes, funny pictures."

Attributes. When describing the Web, students inevitably expressed the characteristics of the Web as they saw it. In their mind, the Web is big, improving, growing, and its growth will not be ending. One exemplar comment is: "[The Web] keeps improving. There are countless websites. Billions, I am sure."

Feelings. Students also expressed their affective feelings about the Web. They were amazed by the convenience made possible by the Web, but also were careful about the validity and credibility of information on the Web. Table 5 lists examples of students' feelings about the Web. The number in parentheses is the number of students who expressed the feelings.

Discussion and Conclusion

Mental model is an abstract concept that is difficult to explicate (Borgman, 1986). There are methodological challenges to researching the concept. In the existing literature, mental models were often elicited by a single method: interviews, drawing, or participant observation. In this study, we took a holistic approach, using semistructured interview, drawing, and drawing description in combination to explore mental models. In this methodological endeavor, we learned that drawing is an effective method to illustrate concrete

TABLE 5. Feelings about the Web.

Feelings	Example comments
Positive (9) Quick, easy, amazing, phenomenal, convenient, fun	It is a way of communication for people to use email or instant messaging or post on an Internet site. It is a way to research for people or assignment [. . .] and a way to find out what is happening in the world, find out the news and breaking events. It is quick and easy.
	[The Web] is a place where everybody can access information about pretty much anything on the world. It is pretty easy to use. I like it. It is really convenient. It is fun.
Negative (6) Not trustful, invasive, be careful	I think sometimes, the Web can be invasive because it has a lot of things that are very private issues. People have blogs, things like that; give up too much personal information.
	Half the information I found on the web is more like that they are not true [. . .] Majority information probably not trustful on the Web.
Mixed feelings (2)	Can access different things but also could be [. . .] I have mixed feeling about it.

DOI: 10.1002/asi

elements and relationships between the elements in people's mental models of an information system. Drawing, however, has difficulties in revealing abstract aspects of mental models, such as students' understanding of information organization on the Web (e.g., information is organized by what you are looking for), search mechanisms of Web search engines (e.g., search engines have presearched everything), and relevance ranking of search results (e.g., results are ranked based on the frequency of visit). To shed light on these aspects of mental models, an interview is a more effective method. Thus, in this sense, drawing and interview methods complement each other. Note that it is useful to ask participants to describe their drawings because the descriptions to a large degree help reduce misinterpretations on the researchers' part.

In this study, undergraduate students' mental models of the Web as an IR system were discussed in terms of their perceptions of three important components of general IR systems: information source, information organization, and search mechanism. This generation of undergraduate students grew up with the Web (They had about 8 years experience using the Web.) The Web has permeated many aspects of their lives, and they used it with a high frequency on a daily basis. When these students talked about information on the Web, they thought about information creation, format, and subject. They recognized that information on the Web is created by individuals, corporations, and various organizations. Some of them were particularly concerned with who created the information because they use the authorship as a criterion to judge the validity of the information. Students mentioned general information formats that have been defined in Dublin Core Metadata Schema while they also looked at instant messaging as a format of information that they can refer to when needed. Students tended to feel that they can find information on any subject on the Web for academic research, for entertainment, or for other purposes.

Students' mental models of information organization on the Web are of particular interest to researchers in the ILS field. Gaining a deeper understanding of users' mental models of information organization may shed light on presenting information in various information landscapes to fit people's expectations and creating information structures that best support people's cognitive functions. Most undergraduate students had at least a vague model in their mind of how information is organized on the Web, and were able to articulate that model. The articulated models varied in complexity and sophistication; some of them were not necessarily correct. Among the articulated models of information organization schemas, the most popular one was by subject. Meanwhile, there were a few students who had never considered the issue of information organization or could not articulate their thoughts in a model-based way. Instructional materials might be needed to help these students gain a clearer picture of how information is organized on the Web. By constructing more sophisticated mental models, students' tolerance of complex information environments on the Web could be improved, and they could be more comfortable with navigating through the Web to achieve their information-seeking goals.

Students' mental models of searching mechanisms were discussed from three aspects: avenues for getting information, how search engines work, and search tactics. Extending the notion of mental models beyond information objects and information systems, Marchionini (1989) used mental models to refer to users' mental representations of information seeking and problem solving, reflected by the relationship between goals and strategies to achieve these goals. Students had different methods for locating information on the Web depending on the goals that they wanted to achieve. They started with search engines, particularly Google and Yahoo, when they wanted general information or were uncertain about their information needs. They kept certain Web site addresses in mind and went to those sites directly when they had specific needs. For example, students went to the UNC library's Web site to look for research-related information and went to weather.com to check weather. Corresponding to different search tasks and purposes, students employed different tactics to solve search problems. The most common search strategy for the participants was "use more words instead of one."

Students' mental models of search engines were discussed from five aspects: search engines' components, search process, search results ranking mechanism, supported functions, and attributes of search engines. In terms of components of search engines, the result in the current study echoed Nielson's (2005) opinion that general users have developed a firm mental model for Web searches, which consists of a text box, a search button, and a result list. Some participants in the study went beyond the stereotype by mentioning other components such as crawlers and databases. The majority of the students understood the search process as a process of keyword matching. Only two students considered the search process on a broader spectrum by mentioning the crawling and indexing processes. It is apparent that students formed mental models of search engines mainly based on system cues and feedback. Thus, their mental models of search engines' search process were sometimes naïve and sometimes incorrect. A typical example is that several students regarded that there were people sitting behind "a curtain," searching everything, and getting results back to them.

Students had various understandings of how search engines rank search results. These criteria are listed in Table 3. Note that students' mental models of search engines' ranking mechanisms were not necessarily simple. Some students possessed systematic mental models that encompass multiple parameters affecting search results ranking. Search engines offer a handful of functions for users to use; however, only a few of the functions were recognized by the students. And only the phrase search function was reported to be used by over half of the participants. A few students mentioned attributes of search engines. They recognized that search engines overlap with each other while they also are different from each other in terms of scope.

DOI: 10.1002/asi

The construction of mental models is affected by various factors. Based on the interview transcripts, three main sources contributing to the construction of undergraduate students' mental models of the Web were identified: personal observation, classroom instruction, and communication between people. Students formed mental models of the information organization and search engines' search mechanisms and attributes based primarily on system feedback. They tended to ignore the invisible part of the Web, such as metadata, the descriptive and representative information attached to an object that provides access to the object (Wang et al., 2000). Students learned search skills in classrooms. They were taught in high-school classrooms of various search techniques such as Boolean operations (although they did not use them very often); they were taught that search engines would ignore common words (stop words). Students also expanded their spaces on the Web by communicating with each other.

The Web is a huge system that goes beyond the scope of common IR systems. Students' mental models of the Web also were not limited to the Web as an IR system. A more comprehensive picture of undergraduate students' mental model of the Web is illustrated in Figure 1. Four main components were identified: components of the Web, functions, attributes, and people's feelings toward the Web. Students acknowledged that the Web consisted of people, information, and technology. The Web offered many functions: It is a versatile information resource, a quick way for communication, and a wonderful place for entertainment. Overall, the students' mental models of the Web were functional rather than structural and technical. This finding echoed Papastergiou's (2005) conclusions about high-school students' mental models of the Internet.

The interview transcripts revealed that emotion was an inseparable ingredient of students' perceptions about the Web. Students inevitably expressed their feelings when talking about various aspects of the Web. Research in information seeking has found that emotions affect judgment and play a critical role in the overall effectiveness of searchers. This emergent theme suggests that emotions should be integrated into mental model representations as a dimension. Bilal (2002) called for teachers of information literacy skills to incorporate the affective aspects of information seeking into their instruction models. We agree that instructors should take into consideration students' emotions attached to the Web to enhance the effectiveness of instruction.

This study is an initial effort to illustrate the building blocks of people's mental models of the Web. Future research is suggested in each of these three directions: (a) to explore other user groups' mental models of the Web for comparison with the current undergraduate-students group, (b) to explore the effects of students' mental models on their information search behaviors and performance, and (c) to test how to incorporate findings about users' mental models and cognitive behaviors to the design of information-retrieval tools.

Acknowledgment

I thank Dr. Paul Solomon and Dr. Peiling Wang for their comments on drafts of the article. Great appreciation also goes to Sanghee Oh for double coding and Steve Bales for editorial help.

References

Bayman, P., & Mayer, R.E. (1983). A diagnosis of beginning programmers' misconceptions of BASIC programming statements. Communications of the Association for Computing Machinery (ACM), 26, 677–679.

Bilal, D. (2002). Children's use of the Yahooligans! Web search engine: III. Cognitive and physical behaviors on fully self-generated search tasks. Journal of the American Society for Information Science and Technology, 53(13), 1170–1183.

Borgman, C.L. (1984). The user's mental model of an information retrieval system: Effects on performance. Unpublished dissertation, Stanford University, Palo Alto, CA.

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

Bruce, H. (1999). Perceptions of the Internet: What people think when they search the Internet for information. Internet Research, 9(3), 187.

Chen, H., & Dhar, V. (1990). User misconceptions of information retrieval systems. International Journal of Man–Machine Studies, 32, 673–692.

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

Dewier, A., & Karnas, G. (1991). Level of experience in text-editing: The role of mental representations on performance. In M.J. Tauber & D. Ackermann (Eds.), Mental model and human–computer interaction 2 (pp. 191–201). Amsterdam: North-Holland.

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

Dimitroff, A. (1992). Mental models theory and search outcome in a bibliographic retrieval-system. Library & Information Science Research, 14(2), 141–156.

Douglas, S.A., & Moran, T.P. (1983). Learning text editor semantics by analogy. In Proceeding of CHI'83 Human Factors in Computing Systems Conference (pp. 207–211). New York: ACM.

Efthimiadis, E.N., & Hendry, D.G. (2005). Search engines and how students think they work. Proceedings of the 28th annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR'05 (pp. 595–596). New York: ACM.

Huang, M. (1992). Pausing behavior of end-users in online searching. Unpublished dissertation, University of Maryland, College Park.

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.

Kracker, J., & Wang, P. (2002). Research anxiety and students' perceptions of research: An experiment. Part II. Content analysis of their writing on two experiments. Journal of the American Society for Information Science and Technology, 54(3), 295–307.

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

DOI: 10.1002/asi

Muramatsu, J., & Pratt, W. (2001). Transparent queries: Investigating users' mental models of search engines. Proceedings of the 24th annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '01 (pp. 217–224). New York: ACM.

Nielson, J. (2005). Mental models for search are getting firmer. Retrieved June 15, 2006, from http://www.useit.com/alertbox/20050509.html

Norman, D.A. (1983). Some observations on mental models. In D. Gentner & A. 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.

Sasse, M.A. (1997). Eliciting and describing users' models of computer systems. Unpublished doctoral dissertation, University of Birmingham, Birmingham, AL.

Savage-Knepshield, P.A. (2001). Mental models: Issues in construction, congruency, and cognition. Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

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.

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.

Wickens, C.D., & Hollands, J.G. (2000). Engineering psychology and human performance (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Young, R.M. (1981). The machine inside the machine: Users' models of pocket calculators. International Journal of Man–Machine Studies, 15, 51–85.

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. 35–53). Hillsdale, NJ: Erlbaum.

Zhang, X. (1998). A study of the effects of user characteristics on mental models of information retrieval systems. Unpublished dissertation, University of Toronto.

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

Zhang, Y., & Wang, P. (2005, October–November). Measuring mental models: Rational and instruments. In Proceedings of ASIST 2005. Charlotte, NC.

Appendix

1. Can you describe the World Wide Web to me?
2. How do you think that information on the Web is organized?
   a. How do you get information on the Web?
   b. Is it easy for you to find information on the Web?
   c. Do you have any difficulties when you look for information on the Web?
3. a. Please describe the way Web search engines work.
   b. What search features/strategies do you usually use?
4. a. What metaphor would you use to describe the Web?
   b. Why? Describe your reasons.

DOI: 10.1002/asi