Do doctoral students need instructions on using Google? -- An exploratory study

Ping Li
pli@qc.cuny.edu
Graduate School of Library and Information Studies
Queens College, City University of New York

Jamshid Beheshti
jamshid.beheshti@mcgill.ca
School of Information Studies, McGill University

Abstract
This paper reports on the findings of an exploratory study on doctoral students' mental models of a search engine and their information seeking behaviour. Semi-structured interview and direct observation techniques were used for the research purpose. The findings show that doctoral students can be differentiated in the dimension of completeness of their mental models of the search engine and any training programs should incorporate the technical aspects of the search engine.

1. Introduction

This paper reports on the findings of an exploratory study on doctoral students' mental models of a search engine and their information seeking behaviour. The study goes beyond previous studies by situating users' interactions with the web search engine Google within the framework of mental model research.

Doctoral students are a special group of information users because they are potential future scholars, conducting "advanced study in a special field" (Merriam-Webster Inc. 1984, 1051). Their web information seeking behaviour has not been studied before. Cole suggests that it is advantageous to study doctoral students rather than professors or other scholars because this group of users is generally more sympathetic and willing to be the subjects of research. They usually have the interest and ability to talk about the process of acquiring information for their theses (Cole 1998, 33 - 54).

This research specifically aims at those doctoral students who have completed their courses and have started the research for their degree thesis. This group of students usually needs to read and synthesis a large body of literature as the background to their dissertation research. It is probable that they often use web search engines in addition to the highly structured information retrieval systems.

Web search engines, though often deal with data with very little consistency of structure (Hock 2001, 20), are able to provide information on the web that encompasses as many subjects as a library (Matherly 2000).

Google, the retrieval system under the study, is "the favorite search engine for the majority of search engine users" (Hock 2001, 86), and, as reported in 2005 by OneStat.com (http://www.onestat.com) – "the number one provider of real-time website analysis software in the world" (OneStat.com 2002-2006a), Google has been the number one search engine in the world and had a global usage share of 56.9% in a two month period (OneStat.com 2002-2006b).

2. Conceptual Framework

The conceptual framework for this study is based on a modified version of the mental model completeness scale proposed by Borgman (1984), Saxon (1997), and Dimitroff (1990).

The term "mental model" has been defined in various ways, yet the different definitions share the common broad implication, that is, mental model refers to "any thought process in which there are defined inputs and outputs to a believable process which operates on the inputs to produce outputs" (Carroll and Olson 1987, 12). It is "representations in the mind of real or imaginary situations" (Johnson-Laird and Byrne 2000), which has a structure that captures or resembles the situation it stands for (Hampson and Morris 1996; Johnson-Laird 1995). In this study, mental model is defined as the model that a doctoral student has built in his/her mind representing the structure and internal relationships of the web search engine Google, based on his/her perception. This definition combines the different ways of expressing the common idea of the concept and makes it fit for the purpose of the research.

In the information retrieval literature several studies focused on a particular dimension of mental model – completeness (e.g. Borgman 1984; Dimitroff 1990; Saxon 1997). For the purpose of this study, a modified version of the mental model completeness scale was developed from Borgman's (1984), Dimitroff's (1990) and Saxon's (1997) scales. This modified scale fits in with the nature of the particular type of information system – web search engine and the particular group of users – doctoral students in this study. With it, doctoral students' mental models of Google are measured in the dimension of completeness, which refers to their perception of all components and structures of this web search engine. The following table illustrates the components that make up this modified version of the scale:

Model Element	Definition and Delimitation
Existence and Nature	**Choose all that apply.**
1. Specific/General Info	1. Perception of Google's ability to retrieve specific or general information.
2. Limited Info	2. Perception of Google's inability to retrieve all information.
3. Indexed Pages/Sites	3. Perception of the indexing of web pages/sites for retrieval.
4. Info Authority	4. Perception of different levels of authority of the information retrieved via Google.
Search Features	**Choose all that apply.**
5. Search Assistance	5. Perception of search assistance availability.
6. Matching	6. Perception of some matching process.
7. Search Types	7. Perception of different types of searches.
8. Restricting/Broadening	8. Perception of mechanisms for restricting or broadening searches.
9. Ranking/Sorting	9. Perception of the ranking or sorting process of search results.
Interactivity Level	**Choose only one category.**
A. Magic Finding	A. Effortless retrieval. Expresses the mere sense that information "is obtained."
B. Stimulus/Response	B. Search is system-directed, user-responsive. Limited system/user partnership. Explicit stimulus-response sequence in evidence.
C. Negotiated Dialog	C. Search is user-directed, system-assisted. Sense of user reformulating search terms; system's difficulty in matching.

With this modified scale, a participant's mental model completeness level is determined by how many components of the first two sections of the scale are described and which level of interactivity between the participant and Google is revealed.

3. Methodology

Sixteen English-speaking doctoral students were recruited from two universities in Canada to participate in the study. Efforts were made so that participants were heterogeneous in the areas of discipline, computer experience, web search experience, and formal training they had received in using computers, in searching the web or other databases.

Individual semi-structured interviews were conducted first to solicit the participants' demographic information, their disciplines, the formal training they had received in using computers, their web searching experiences, the training they had received in web searching as well as other online information searching provided by librarians, and their mental models of Google. All the conversations were captured using a digital audio recorder. Direct observation technique was used to observe the participants' interactions with Google, which helped reveal their actual search practices. Their activities on the screen were captured with the software "Macromedia Captivate". The authors watched the recordings together with the participants right after they finished the searches. Students were asked to talk through the recordings and the researchers discussed their search strategies with them. The combination of the screen recordings, audio recordings, and direct observations provided the researchers with a rich data set for analysis.

Content analysis technique was employed to analyze the data. The modified version of the mental model completeness scale functioned as the coding frame for the participants' mental model completeness levels and each participant's search behaviour was identified, encompassing such aspects as the lengths of queries, Boolean queries, phrase searching, query reformulation, and number of results viewed.

4. Results

Based on the interview data, the participants' background and online information searching experience (including web searching experience) were identified. Of the sixteen participants, seven are females and nine are males. Six participants are from the science and engineering disciplines and ten are from the humanities and social sciences. Nine of them received formal training in computer usage while the other seven have never received any training in using computers. Thirteen participants received online information search training provided by librarians, and only three never received any training. The sixteen participants' web search experiences span from four to twelve years, and their Google usage span from one to seven years.

The sixteen participants' web search behaviour was identified through the analysis of the direct observation data. Most of these doctoral students started with keyword searching. The lengths of the queries were usually one, two or three keywords. Two participants exhibited different behaviour than the rest. They seemed to enter keywords with no apparent patterns or search strategies. Eleven participants searched using phrases (with quotation marks), while nine participants used the advanced search options, and 12 used Boolean operators. Ten participants made some efforts to reformulate their searches when they were unhappy with the results, while the others just looked through the results for anything relevant. Most of the participants browsed through the first one or two pages of the search results before they either reformulated the search strategies or just ended the search sessions. Only two participants went beyond the first two pages of results and looked through the results one by one for relevant information.

Based on the analysis of both the semi-structured interview and direct observation data, the participants' perceptions of the mental model components of the first two sections of the modified scale and their interactivity levels with Google were inferred, illustrated by the following table:

Participants' Mental Model Completeness Levels

Model Element		Perception															
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
Existence and Nature	Specific/General Info	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Limited Info	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Indexed Pages/Sites	N	N	N	N	N	N	Y	N	N	N	Y	N	N	N	N	N
	Info Authority	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Search Features	Search Assistance	N	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	N
	Matching	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Search Types	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N
	Restricting/Broadening	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
	Ranking/Sorting	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
Interactivity Level	Magic Finding																X
	Stimulus/Response	X	X		X				X					X			
	Negotiated Dialog			X		X	X	X		X	X	X	X		X	X	

5. Discussion

Based on the above results, we propose that doctoral students' mental models of Google may be measured relatively in the dimension of completeness. Not surprisingly, Google, the most popular search tool on the web (Goldsborough 2007), was used by all the doctoral students to conduct actual searches for their thesis. Their mental model completeness levels of Google are generally good considering that among the sixteen participants, two had "ideal" mental models, and fourteen had the perception of at least seven out of the total nine components of the first two sections of the modified scale and their interactivity level with Google is either "Negotiated Dialog" or "Stimulus/Response" (except participant 16). Some mental model components, for example, "Specific/General Info", "Limited Info", "Info Authority" and "Matching" are perceived by all the participants, indicating that these features of Google have, to some degree, become the basic knowledge to the doctoral students. But on the other hand, this group of users seem to have a poor perception of the technical aspects of the search engine, in particular how it indexes and ranks the web pages. This finding may raise the question: is it important for doctoral students to learn about the inner workings of the search engine?

By comparing the participants' perceptions of the mental model components of the first two sections of the scale and their interactivity levels with Google, one can see that the participants with the perception of more mental model components have the Negotiated Dialog (the most ideal) type of interactivity with Google while those with the perception of fewer mental model components tend to have the Stimulus/Response type of interactivity.

One finding from the data analysis is that those participants who have the perception of the mental model components relating to the more technical aspects of Google have comparatively more complete mental models of this search engine. The two students (participant 7 and 11) who seemed to understand the indexing and ranking mechanism of Google both have the perception of all the 9 components of the first two sections of the scale and the interactivity level between them and the web search engine is the most ideal Negotiated Dialog.

The participants' web information search behaviour identified through the direct observation technique supports the above findings. Those who used Negotiated Dialogue and who tended to have more complete mental models did make use of more search features of Google and were more active in reformulating their search strategies so as to obtain better search results. The two participants who have ideal mental models of Google tended to be very confident and active in directing the searches. This shows that doctoral students' interactivity levels with Google can be the practical application and manifestation of their perception of the nature and search features of Google.

This research echoes previous studies in the finding that users' mental models of a device or a system do have an impact on users' performance or behaviour with the system (e.g. Norman 1983; Chen and Dhar 1990; Muramatsu and Pratt 2001). According to Dimitroff (1990), end-user searchers, who have a more complete mental model, are better searchers. They make fewer errors and find more items. On the other hand, previous research also shows that training programs are effective in helping users to develop better mental models of the device or system, which can then improve the users' behaviour or performance (e.g. Norman 1983; Bayman and Mayer 1984; Briggs 1988).

For future training programs, it may be necessary to incorporate technical instructions so as to assist doctoral students to develop more complete mental models of Google, which in turn may affect their web information search behaviour and performance.

6. Conclusion

This study is of exploratory nature and its findings still need further research for validation and may not be generalizable. Yet they can at least, to some degree, shed some light on the doctoral students' perceptions of Google in general.

Google is a very popular web search tool among doctoral students, and it should be important for any training and user instruction programs to incorporate the features of this search tool, with which doctoral students have difficulty understanding. In particular, more emphasis should be put on the instruction of its technical aspects, particularly the indexing and ranking processes. With an understanding of these features, doctoral students can have a better idea of how Google works. As a result, they can know better how to formulate search strategies to get more satisfactory results. Also they can better understand the importance of the active role a user should play in conducting web information searches, and thus they will more conscientiously direct their searches to achieve better search performance.

References

Bayman, Piraye, and Richard E. Mayer. 1984. Instructional manipulation of users' mental models for electronic calculators. *International Journal of Man-Machine Studies* 20: 189-199.

Borgman, Christine L. 1984. *The user's mental model of an information retrieval system: effects on performance*. Ph.D. diss., Stanford University.

Briggs, Pamela. 1988. What we know and what we need to know: the user model vs. the user's model in human-computer interaction. *Behaviour and Information Technology* 7: 431-442.

Carroll, John M., and Judith R. Olson, eds. 1987. *Mental models in human-computer interaction.* Washington, D.C.: National Academy Press.

Chen, Hsinchun, and Vasant Dhar. 1990. User misconceptions of information retrieval systems. *International Journal of Man-Machine Studies* 32: 673-692.

Cole, Charles. 1998. Information acquisition in history Ph.D. students: inferencing and the formation of knowledge structures. *Library Quarterly* 68, no. 1: 33-54.

Dimitroff, Alexandra. 1990. *Mental models and error behavior in an interactive bibliographic retrieval system*. Ph.D. diss., University of Michigan.

Goldsborough, Reid. 2007. New developments in web searching. *Teacher Librarian* 35, no. 1: 50.

Hampson, Peter J., and Peter E. Morris. 1996. *Understanding cognition.* Oxford UK & Cambridge USA: Blackwell.

Hock, Randolph. 2001. *The extreme searcher's guide to web search engines: a handbook for the serious searcher.* Medford, NJ: CyberAge Books.

Johnson-Laird, Phil. 1995. Mental models and probabilistic thinking. In *Cognition on Cognition,* ed. Jacques Mehler and Susana Franck. Amsterdam: Elsevier Science Publishers.

Johnson-Laird, Phil, and Ruth Byrne. 2000. *Mental Model Website: A Gentle Introduction.* Accessed 9 October, 2005. Available from http://www.tcd.ie/Psychology/Ruth_Byrne/mental_models/.

Matherly, Mark. 2000. *Tips on Using Search Engines.* Accessed 1 April, 2008. Available from http://www.netpath.net/search_tips.htm

Merriam-Webster Inc. 1984. *Webster's ninth new collegiate dictionary.* Springfield, Massachusetts: Merriam-Webster Inc., Publishers.

Muramatsu, Jack, and Wanda Pratt. 2001. Transparent queries: investigating users' mental models of search engines. *SIGIR'01*: 217-224.

Norman, Donald A. 1983. Some observations on mental models. In *Mental Models*, 7-14, ed. Dedre Gentner and Albert L. Stevens. Hillsdale, NJ: Laurence Erlbaum Associates.

OneStat.com. 2002 – 2006a. *Company Information.* Accessed 11 April, 2008. Available from http://www.onestat.com/html/aboutus.html

OneStat.com. 2002 – 2006b. *Global Usage Share MSN Search Has Slightly Increased According to OneStat.com.* Accessed 11 April, 2008. Available from http://www.onestat.com/html/msn_search_usage_share_20050926.html

Saxon, Scharlotte. 1997. *Seventh grade students and electronic information retrieval systems: an exploratory study of mental model formation, completeness and change*. Ph.D. diss., University of North Carolina.