Is Searching the Internet Really Different? Search Process Models for Two Electronic Environments

Elisabeth Logan and Lori L. Driscoll-Eagan
School of Information Studies
Florida State University
logan@lis.fsu.edu
lld2113@garnet.acns.fsu.edu

A pilot study devised to explore factors contributing to searcher persistence in varying electronic environments was conducted in the usability lab at the School of Information Studies at Florida State University in the summer and fall of 1997. Subjects were students who claimed some knowledge of WWW searching and students who had completed a basic course in traditional electronic searching. Detailed Search Process Models were created and analyzed for both Dialog and WWW searches and found to be similar. These preliminary findings may have identified a generalizable Search Process Model for specific query searches independent of the searching environment.

Introduction

Since the first online systems in the early 1970s until fairly recently, seekers of information have had two options: learn the command languages of the electronic information providers or pay or persuade a knowledgeable searcher to find the desired information. Command languages and protocols of these systems are typically highly structured and require training and experience as well as support from additional printed sources before searching can be considered effective. Nevertheless, these traditional "online" command languages have been employed effectively by information professionals with experience in their use. Now, however, the Internet, specifically the WWW with browsers like Netscape and Internet Explorer, has opened electronic information searching to casual and untrained users. Browsing the Internet has become a way of life for groups of people previously unskilled and perhaps even uninterested in accessing electronic information.

For decades research has explored and reported many aspects of command-based electronic searching, Bates (1979); Fidel (1984); Borgman (1989); Bellardo (1985); Saracevic et al. (1988); Logan (1990); Logan and Woelfl (1986), to mention only a few. In addition, Marchionini (1995); Hildreth (1982); Cove and Walsh (1988); Kwasnik (1992); Bates (1989); Borgman (1995); Twidale and Nichols (1996); and Hawkins (1996), among others have sought to understand and document aspects of browsing behavior.

In this exploratory study, both Internet and Dialog searches for information on a selected topic are investigated using quantitative and qualitative methodologies. Although the traditional command language searches on Dialog were different in many ways from the Internet searches, we observed that detailed Search Process Models created for each type of search showed general similarities.

The Study

A pilot study devised to explore factors contributing to searcher persistence in varying electronic environments was conducted at the School of Information Studies at Florida State University during the summer and fall of 1997. Questionnaires, audio and video recordings as well as computer logs were made of each search using facilities in the Usability Lab. Searchers were asked to employ think-aloud protocols as they did their searching to provide additional material for analysis. Detailed Search Process Models were created and analyzed for both Dialog and Internet searches. Although there were differences in search output and methods, the detailed Search Process Models for the two environments were surprisingly similar. These preliminary findings may have identified a Search Process Model generalizeable to searches conducted for specific information regardless of the searching environment.

Methodology

Seven masters students in Information Studies at Florida State University were asked to search for a specific query. Some searched Dialog, some searched the WWW, and some searched on both searching environments. Questionnaires requesting demographic as well

as computer and online searching experience information were administered prior to each search. Searches were carried out at the convenience of the subjects in the Usability Lab at the School of Information Studies at Florida State University. The laboratory consists of two adjoining rooms separated by a one-way window. The evaluation room holds a computer and printer as well as a small camera and microphone. There is also a desk and chair for a facilitator. In the adjoining observation room is the equipment for monitoring the activities of the subject as well as a desk for the observer/operator. Usability equipment permits audio, video, and search log capture of online activity. Since subjects were also asked to think aloud as they worked, it was possible to collect information about searcher motivation and responses while online.

Subjects were introduced to the room, oriented, and given search instructions. Blue Sheets were available for Dialog searches. A facilitator remained in the room to answer questions of procedure and to encourage subjects to think-aloud as they worked. No time limits were set for the searches since we were primarily interested in recording and monitoring user persistence. As subjects indicated their searches were completed, each was given a post-search questionnaire and interviewed briefly to give each an opportunity to clarify and expand upon questionnaire answers if they chose. NUD*IST software was used to analyze the data from these observations.

Results—Quantitative

Demographic data obtained from the pre-search questionnaires indicates that the reported levels of both computer and searching skills are uniformly high.

Subjects included students who claimed some knowledge of Internet searching and students who had completed a basic course in traditional electronic searching, so the results were not unanticipated; however, the uniform nature of the reported skills is noteworthy. Results from the questionnaire are summarized below.

Reported Basic Skill Levels

Dialog Searchers: All but one had taken a basic online course; four indicated advanced skills.

Internet Searchers: All indicated familiarity with WWW searching; four stated being "very familiar" with the web environment.

Computer Technology: All noted they were either "familiar" or "very familiar" with computers; one reported a level of "expert" knowledge.

Education: All had more than 18 hours of graduate information courses; some had as many as 31 hours.

Professional Goals: All subjects anticipated using the WWW in future employment; four anticipated using commercial databases and command-based searching.

Search Results

Search results, although specifically not what the study was designed to evaluate, are of interest when compared with searchers' self-evaluation of search success summarized in Table 2.

WWW Searchers: Four of five were able to locate some useful hits; one of the group found a good collection of hits; one found some good hits that could have been followed up, but were not; one searcher found many false drops.

Dialog Searchers: Three searchers located more than one good hit; one found good potential hits but lost them through user error; one searcher found one good hit, but did not pursue more.

Search Times

Here again, subjects were specifically asked not to concern themselves with time, but times varied widely both between and within the two environments. Search times on WWW ranged from 24 minutes to 74 minutes; Dialog searches lasted from 19 minutes to 54 minutes.

Post-Search Questionnaire

All participants completed a questionnaire as they finished searching. Five of the questions were designed to identify points in the

search process where the subjects could identify and assess their responses. Since we were particularly interested in persistence and abandonment, several of the questions address these issues. Table 1 summarizes responses. Since the population is small, actual numbers are not useful at this stage and are not reported. However, it is interesting to see that issues of quantity are mentioned in three of the five categories and quality is mentioned in four. Issues of quantity and quality of information are distributed fairly evenly across both environments and are associated with general satisfaction, frustration, feeling lost, and search termination. "Quantity of information" is identified in Internet but not Dialog searching as being associated with both frustration and general satisfaction. "Technology" is noted in Dialog searches as contributing to general satisfaction as well as to frustration and feeling lost, but not mentioned in any connection with Internet searches. It appears that information quantity can be both blessing and curse of the WWW. Dialog users found that for them technology could serve a similar dual role—both aiding and undermining their searches.

The sixth question asks for searchers' overall responses to the search process and satisfaction with their results. Table 2 summarizes responses to this question.

Searcher responses are uniform across both environments and indicate, at least for this group, there is little difference in satisfaction between Internet and Dialog searches. Over half the respondents indicated for both environments that they were able to locate useful information, and that although they were able to find useful information there was probably more available. Note too that not a single searcher in either environment felt he or she had been unable to locate useful information or that such a failing might have been due to user error. These results are particularly interesting when compared with the more objective assessment of subjects' actual search results. This seems to reinforce reports from older studies where in many cases searchers were notoriously unable to evaluate their own search results effectively.

Results—Qualitative

Search Process Model

From video and audio transcripts of seven searchers engaged in ten separate searches, search process scripts were generated for each search. These were compared and a Search Process Model derived that applies to both Internet and Dialog environments (Table 3). In our analyses, only one stage, Place-marking, was environment specific. Although logically it could have been used in both environments, here it was observed only in the Internet searches. Although each search began with the Creation stage and ended with either Completion or Abandonment, the intervening stages were not necessarily in any order nor were they used only once in a given search. We also differentiated between Revision, revising a search statement, and Iteration, an ongoing interplay between keywords, indexes, or search engines. Interestingly, many of these observed stages have been identified in earlier studies of traditional online searching processes and, although named and described somewhat differently in earlier studies of command-based searching, here they appear to be true of both Internet and Dialog searches by our subjects (Bates 1979; Fidel 1985).

Search Approach

We also observed that there were two specifically different overall approaches to the searching process that occurred in both searching environments. We identified and characterized the two as follows:

1. Methodical - searcher follows paths, links, or search steps in a logical progression; often looks carefully at each result before moving to the next stage.

2. Casual - searcher looks at many sites doing quick checks to see what if possible is available; may go back later and check some sites.

These two were not necessarily limited to a single searcher or search session; we observed that in some cases they were mixed within the same search session and were observed in both Dialog and Internet environments. Here also, these categories are reminiscent of many

early descriptions of online searchers. Fidel contrasts operational and conceptual moves (1985); Marchionini contrasts browsers with analytical searchers (1995), Logan and Woelfl contrast active experimentation with reflective observation searching styles (1986). Of interest in this study is evidence that both types of searches occurred in both environments and were sometimes used by the same searcher. Whether this mix of styles can be demonstrated in a larger population over a variety of searches and environments remains to be seen.

Decision Point Model

From search logs and audio and video recordings of the search sessions, we further identified a set of Decision Points at which searchers made a series of alternate choices. We noted that as a search session progresses, there are points in the search process where decisions must be made in order for the searcher to proceed. Actions at these decision points may be taken as the searcher understands the process and what is happening, or as the searcher just forges ahead and tries anything without seeming to understand the processes or problems that have occurred. The majority of decision points identified in this study are similar for both Internet and Dialog search environments although the responses within the two environments may differ. These are shown in Table 4. Three decisions in this study, however, are associated with individual environments: two are Internet specific ("new insights" and "links not found") and one is specific to Dialog ("can't remember commands").

Conclusions/Discussion

In light of the prevalent view of much of the profession which assumes that searching for information using commercial database vendors is substantially different from searching the Internet with browsers like Netscape and Internet Explorer, the findings of this exploratory study are somewhat surprising. Although the Search Process Model derived from both Dialog and Internet searches in the study certainly represents a "big picture" view of the search process, it does appear that at least in overview, searches in the two different environments follow the same basic script. In our investigations of browsing versus searching behavior, there appear to be more similarities than differences.

The Decision Point Model derived from examination of searching behavior and processes in both environments indicates further similarities between the two. Six of the nine decision points identified are common to both, although responses at the individual points are not necessarily similar. Of the three environment specific points identified in the study, two are Internet specific and one is peculiar to Dialog. In fact, one of the Internet specific points, "new insights", could be associated with either searching environment. Of the other two, "links not found" is probably an Internet specific point and "can't remember commands" is just as likely a command-based search problem. In the Decision Point Model as in the previously mentioned Search Process Model, it appears that Internet searching may incorporate much that has been traditionally associated with command - based searching and that both are closer cousins than we have been willing to recognize.

Looking at subjects' evaluation of their searching both in response to specific questions and in their overall search evaluation, issues of the quantity and quality of information appear frequently and in several different contexts. In this study, searchers of both Internet and Dialog environments seem to be aware that both information quantity and quality play an important role in the retrieval process. Searchers in both environments also indicated an awareness that there was probably additional information on their topic that had not been retrieved.

Also worth noting is a difference between the subjects' evaluation of their searches and those of more objective observers. Although not a single searcher indicated errors that resulted in an unsuccessful search, "user error" was listed as a source of frustration, and from a more objective review of search results, certainly was a cause of search failure. Evidently regardless of searching environment, searchers are still not very reliable judges of their own success.

Although there appear to be striking similarities between searching in command-based and Internet environments in this study, there is much that needs further investigation. First, the small number of subjects in this pilot study makes these findings sources for further exploration rather than confirmed findings. Second, the search in this study was for a specific topic assigned by the investigators and as such may have limited the "browsing behavior" that might otherwise have produced observable differences between the two envi-

ronments. Third, the population itself limits the ability to generalize from these results. From the demographic data, one can see that this group of subjects has a similar profile. All are master's students with at least 18 credit hours in a program in information studies and all but one are at least conversant with command-based searching. Perhaps an even more important observation is that all these subjects are familiar with the overall search process and that whether they are searching the Internet or Dialog, they bring with them an appreciation of information retrieval in a larger context.

Further investigations are called for to address different populations, different search environments, different question types, and different motivational scenarios. By compiling observations from a larger and more diverse population as well as from searching scenarios with additional variables, it may be possible to verify some of these early observations. Additional studies are currently underway which incorporate some of the above.

Table 1. Post Search Questionnaire

What contributed to your:	D	I
1. General Satisfaction with the search		
Special features of environment	x	x
Quality of Information	x	x*
Quantity of Information	0	x
Entire Search	x	0
Technology	x	0
2. Why did you continue searching?		
To broaden the search	x	x
To narrow the search	x	x
Did not effect my search	0	x
After verification	x	0
No response	x	x
3. What caused you frustration?		
Term Selection	0	x*
Quality of Information	x	x
Quantity of Information	0	x
Technology	x	0
User Error	x	0
No frustration in search	x	0

	D	I
4. What caused you to feel lost?		
Special Features	x	x
Quality of Information	x	x
Term Selection	0	x
Technology	x	0
Beginning the query	0	x
Not lost	x	x
5.What caused you to quit searching?		
Quality of Information	x	x
Quantity of Information	x	x
Repeat of Previous Information	0	x
Tired	x	x
Wanted to Renegotiate	x	0

*** Indicates selected by more than half of respondents**

Table 2. Overall Search Evaluation

Question	Dialog	Internet
Searched well and I am satisfied with my search	X	X
Searched well, but could have done better	X	X
I was able to find useful information	X *	X *
I was unable to find any useful information	0	0
I found useful information, but I'm sure there is more	X *	X *
My search was unsuccessful because I made errors	0	0
Unsuccessful due to not much information available	X	X
I found good information, but it was due to luck	X	X

* indicates selected by half or more respondents

Table 3. Search Process Model

Stages	Description	Observed
1. Creation	Select terms; sites; engines	Dialog; Internet
2. Revision	Revise strategy to narrow, broaden, or refocus	Dialog; Internet
3. Place-marking	Transcribe or mark site for further investigation	Internet
4. Iteration	Change search with interplay of engines, keywords, or indexes	Dialog; Internet
5. Extraction	Use information from search to refine search	Dialog; Internet
6. Abandonment	Stop search with nothing or little found	Dialog; Internet
7. Completion	Finish search with information found	Dialog; Internet

Table 4. Decision Points - Both Environments

Points	Response	Environment
1. No relevant results	New terms, new engine, back up level, stops to plan,quits, re-reads query	Internet
	Ignores result; changes database	Dialog
2. Too much information	refine search at same location; specific techniques to limit, add operators, scans quickly, tries new search, quits	Internet
	uses limiters (language)	Dialog
3. Marginally relevant data	re-read query, try other terms	Internet
	re-read query, logoff hold, check Blue Sheets change databases,	Dialog
4. User error	correct error and return to search; new strategy; quit; change database	Internet
	correct error and reselects terms	Dialog
5. Partial success (size of return)	re-read query and evaluate results; select terms from hits, change strategy; quits with verbal justification	Internet
	check terms;adds databases; tries codes	Dialog
6. Appropriate results (successful search)	quits with satisfaction; no attempt to find more	Internet
	prints hits to determine relevance; selects items to print; Bluesheets for format	Dialog

BIBLIOGRAPHY

Bates, M. 1979. "Information search tactics." *Journal of the American Society for Information Science* 30: 205-214.

Bates, M. 1989. "The design of browsing and berrypicking techniques for the online search interface." *Online Review* 13: 407-424.

Bellardo, T. 1985. "An investigation of online searcher traits and their relationship to search outcome." *Journal of the American Society for Information Science* 36: 241-250.

Borgman, C.L. 1989. "All users of information systems are not created equal: An exploration into individual differences." *Information Processing and Management* 25: 237-252.

Borgman, C.L., S.G. Hirsh, and V.A. Walter. 1995. "Children's searching behavior on browsing and keyword online catalogs: the Science Library Catalog Project." *Journal of the American Society for Information Science* 46: 663-684.

Cove, J.F., and B.C. Walsh. 1988. "Online text retrieval via browsing." *Information Processing and Management* 24: 31-37.

Fidel, R. 1984. "Online searching styles: A case-study-based model of searching behavior." *Journal of the American Society for Information Science* 35: 211-221.

Hawkins, D.T. 1996. "Hunting, grazing, browsing: a model for online information retrieval." *Online* 20: 71-73.

Hildreth, C. 1982. "The concept and mechanics of browsing in an online library catalog." In *Proceedings of the Third National Online Meeting* March 30-April 1, 1982, 181-193. New York City.

Kwasnik, B.H. 1992. "A descriptive study of the functional components of browsing." In *Proceedings of the IFIP TC2aG2.7 Working Conference on Engineering for Human-Computer Interaction*, August 10-14, 1992, 191-202. Ellivuori, Finland.

Logan, E.L. 1990. "Cognitive style and online behavior of novice searchers." *Information Processing and Management* 26: 503-510.

Logan, E.L., and N.W. Woelfl. 1986. "Individual differences in online searching behavior of novice searchers." *Proceedings of the American Society for Information Science* 23: 163-165.

Marchionini, G. 1995. "Browsing strategies." In *Information Seeking in Electronic Environments: Cambridge Series on Human-Computer Interaction 9.* London: Cambridge University Press: 100-138.

Marchionini, G., and B. Schneiderman. 1988. "Finding facts vs. browsing knowledge in hypertext systems." *IEEE Computer* 21: 70-80.

Saracevic, T., P. Kantor, A.Y. Chamis, and D. Trivison. 1988. "A study of information-seeking and retrieving. I. Background and methodology." *Journal of the American Society for Information Science* 39: 161-176.

Twidale, M.B., and D.M. Nichols. 1996. "Collaborative browsing and visualization of the search process." In *Aslib Proceedings* 48: 177-182.