5-1-1998

An Exploratory Study of User's Interaction with World Wide Web Resources: Information Skills, Cognitive Styles, Affective States, and Searching Behaviors

Peiling Wang
University of Tennessee - Knoxville

Carol Tenopir
University of Tennessee - Knoxville, ctenopir@utk.edu

AN EXPLORATORY STUDY OF USERS' INTERACTION WITH WORLD WIDE WEB RESOURCES: INFORMATION SKILLS, COGNITIVE STYLES, AFFECTIVE STATES, AND SEARCHING BEHAVIORS

Peiling Wang and Carol Tenopir
University of Tennessee

Keywords: World Wide Web, Searching Behavior, Web Searching

Abstract: Understanding user behavior in retrieving factual information from the World Wide Web will assist web system designers, web instructors, and reference librarians. In this study, 24 graduate students were asked to search the Web to find the answers to two factual questions. A process-tracing technique, including transaction logs and concurrent verbalization of thoughts during the search, provided data on their search processes, including cognitive, sensorimotor, and affective behaviors. Pre-tests measured cognitive styles (Embedded Figures Test) and previous computer and search experience. Anxiety levels were monitored before and after searching by administering the State-Trait Anxiety Inventory. Differences in cognitive style are significantly related to the time spent on each question and the number of sites visited. Generally, those who found an answer in a shorter time and expressed confidence about their results felt less anxiety. Search procedures used in the course of each search session varied, but many users found search screens confusing, had trouble with the sensorimotor domain, or did not know always know where they were searching. Users expressed many kinds of emotions as they searched, which seemed to be an important part of the search process. Once an answer was found, even if it was incorrect, most users felt confident in their ability.

1. INTRODUCTION

End users have been searching online databases for more than a decade. With the growing popularity of the Internet, more end users are connected to more multimedia resources available on the World Wide Web (Web). For libraries, the Web has been a valuable addition to or substitute for their reference resources. For end users, the Web is certainly an alternative to and sometimes a replacement for their previous channels of information. The Web is being used for factual information gathering, in addition to recreation and surfing.

Searching for factual information on the Web differs from surfing on the Web in that the interaction is purposeful, need-oriented, and well-focused. The success of the interaction can be measured by: 1) effectiveness (whether or not the needed information is found); 2) efficiency (what strategies and paths the searcher used during the interaction); and, 3) satisfaction (how the searcher feels at the end of the search).

Not all users interact with the Web in the same way. User-Web interaction is a

communication process between the user and the system; it is also a decision-making process in which users decide on the next action at each point. Differences in information skills, cognitive styles, and affective states are likely to affect how individuals carry out the searching processes.

This study examined how well users can perform simple factual searches on the Web, what factors might affect users' interactions with the Web, and how the Web can be improved to incorporate individual differences.

2. LITERATURE

Studies of end user behavior with a variety of electronic information retrieval (IR) systems have been conducted by many researchers in the last two decades (Ref. 1). Of particular relevance to this research are those studies which analyze the search process, feelings while searching, and search results of end users on a variety of IR systems.

Studies of online public access catalogs (OPAC) by Borgman conclude that these systems designed for end users are not easy to use and they require knowledge of the systems at three levels: conceptual, semantic, and syntactic (Refs. 2-3). Individual differences have been shown to influence users' interaction with and success with online retrieval systems (Ref. 4). Nahl and Tenopir have demonstrated the importance of the affective and sensorimotor domains for novices in addition to the cognitive elements of online searching (Ref. 5).

The Web has features that can be found in OPACs and traditional online IR systems, yet there are many differences as well. It is a multimedia environment, presumed to be extremely user-friendly, and relies on a combination of search engines, directories, and links to facilitate information retrieval. A growing body of research is examining user behavior in interacting with the Internet (and, in particular, the Web) to find information, but this research is still in the early stages (Ref. 6). Much additional research is needed to make the IR function of the Web truly user-friendly. In a recent National Research Council report, the goal of "Every-Citizen Interfaces" (ECI) to the nation's information infrastructure is defined and specific research areas for improving interface design are identified (Ref. 7). Without understanding all aspects of how users behave when interacting with the Web (cognitive, affective, and sensorimotor), the ECI cannot be achieved.

3. RESEARCH FOCUS: VARIABLES AND FACTORS

This study focused on three factors that might affect users' interaction with the Web: 1) computer and information retrieval experience; 2) cognitive style in terms of an individual's information processing style; and 3) affective states in terms of the anxiety level before and after a search.

Computer and information retrieval experience can be measured by pretest questions about how long and at what level of experience subjects have with Windows, IR systems, and the Web.

Cognitive style can be measured by a standard test; in this study we used the Embedded Figures Test (EFT). In the EFT an individual is shown a simple figure and then is asked

to find this simple figure embedded in a complex figure. The test identifies how a person processes information by providing a score in terms of field-dependent and field-independent (Ref. 10). For high field-dependent individuals (those who have a more difficult time identifying a simple figure embedded in a complex one), the Web was expected to be a more difficult environment.

Affective states were measured by the State-Trait Anxiety Inventory (STAI), which consists of two forms: S-anxiety and T-anxiety. S-anxiety indicates an individual's feelings at that moment; T-anxiety indicates an individual's general tendency of feelings (Ref. 11). Affective states may influence a person's performance of a task, while self-satisfaction after performing a task may also affect affective states, especially S-anxiety.

In reference services, it is critical to provide correct and accurate answers to questions seeking specific information. For such factual information needs, the outcome can be objectively measured as correct or incorrect. This study observed users' interaction with the Web in finding factual information. The searches were observed in a natural setting and the only control was the search questions as all the participants were given the same questions to search.

The following three main research questions were explored:

1. How successful are end users in finding information on the Web? In other words, how effective and efficient are they in searching the Web to find information that will answer a factual (ready reference) question? How do searchers feel about their searches afterwards?
2. What are the individual differences that result in differences in the searching process? In other words, how do the cognitive, affective, and physical domains affect search behavior?
3. What are the search strategies Web users apply in searching for factual information?

Both quantitative and qualitative variables were considered in this study. Quantitative data were quantified directly, by measuring such things as online time, number of sites visited, scores on cognitive and anxiety tests, and self-reported previous experience. Qualitative data were coded based on content analysis and some of them were quantified as well. Specifically, the following variables were observed and analyzed quantitatively.

3.1 Dependent Variables

Dependent variables are the measurement of success of an interaction. They included:

- effectiveness (Y1): whether or not the needed information is found;
- efficiency (Y2--time and Y3--paths): which processes occurred during the interaction;
- satisfaction (Y4—confidence level in correct answer and Y5—change of anxiety level): how the searcher feels at the end of the search.

3.2 Independent Variables

Independent variables are the measurement of individual differences. They included:

- computer and information experience (X1—years of computer use, X2—years

of database retrieval, X3—years of Web use, X4—frequency of Web access);
- cognitive style (X5—Embedded Figure Test);
- affective state (X6—pretest of state anxiety, X7—trait anxiety).

3.3 Qualitative Data

Qualitative data are non-numerical in nature and are coded into categories based on content analysis. Categories are derived from data collected. Since the study focused on search behavior from a holistic viewpoint, three classes of search behavior were analyzed qualitatively:

- cognitive behavior—related to users' information processing, problem solving and decision making,
- affective behavior—related to users' feelings and emotional states, and
- physical behavior—related to sensorimotor skills and control of devices.

4. METHODOLOGY

The setting, which was designed to be as natural as possible for the participants who are accustomed to using the school 's Computer Laboratory, consisted of a PC with Windows 95 and Netscape 3.01 Gold. The input device was a standard keyboard with mouse. For data collecting purposes, a VCR and a tape recorder were also connected to the computer to record the screen output and searchers' verbalization during the interaction. The recording equipment was visible to the user, but was not on the same desk as the computer.

4.1 Participants

Participants were recruited from graduate students enrolled in the University of Tennessee Master's program in information sciences. Twenty-four volunteers participated in this study during Fall, 1997, all of whom completed the session. Of the 24 participants, 14 were entry-level students who had just started the core courses (which teach basic concepts of information organization and access). The remaining ten participants were advanced-level students who had finished the core information access and retrieval course (which includes intensive online searching training using the DIALOG system and Web searching experience).

Table 1 summarizes the computer use, traditional IR systems experience, and Web experience levels of all 24 participants. In addition, it summarizes the Embedded Figure Test and State-Trait Anxiety Inventory scores, which will be discussed later.

Table 1. Participants' computing experience, cognitive style, and affective state (N=24)

Variables	Computer Use (Mon.) X1	IR systems Use (Mon.) X2	WWW Use		EFT Score* X5	STAI**	
			Month X3	Times X4		S-Anxiety X6	T-Anxiety X7
Min. Value	43	2	3	1	16.30	20	24
Max. Value	132	144	60	25	172.90	63	63
Mean	41	51.5	25.3	15.0	46.80 &	35.42 #	38.58 #
SD	33.8	34.8	13.0	11.1	32.10 &	10.17 #	9.28 #

* EFT: Embedded Figure Test monitored individually; the score is measured by total response time in seconds divided total number of cards.
** STAI: State-Trait Anxiety Inventory by Speilberger, 1983.
& EFT for male: mean 45.80 SD 17.65; for female: mean 47.10 SD 35.31
# S-Anxiety for male: mean 31.60 SD 9.53; for female: 36.42 SD 10.34
T-Anxiety for male: mean 38.80 SD 8.64; for female: 38.53 SD 9.67

4.2 Search Questions

Participants were given two factual search questions to answer on the Web. The searchers were allowed to do whatever they decided to do, and were not guided or assisted on any aspect of the search process. The following questions were supplied to the participants:

1 This summer, a faculty member at the University of Tennessee, Knoxville submitted a full grant proposal to the National Science Foundation through the FastLane, the agency's new electronic proposal system. It was announced in the Research Good News. Can you find the news for me? [You may use this sheet for scrap paper]

2 Each year, the U.S. Census Bureau reports on the projections of national population. I am interested in the most current estimate for the population of the United States in 2000. [You may use this sheet for scrap paper]

Based on the results reported by Hsieh-Yee that the order effect was not significant in her study of online search tactics, all participants were given the questions in the same order (Ref. 8).

4.3 Procedure

The participants were scheduled to perform the searches at their convenience. After signing a consent form, each searcher followed the same procedure. The searcher:

- filled in pretest questionnaires to measure level of information skills and affective states (State-Anxiety and Trait-Anxiety).
- took an individually administered Embedded Figure Test (EFT) about cognitive styles (12 questions with a 3-minute cut-off time on each).
- was given the first question sheet and asked to find the answer. (The Web browser was set at home state the university's homepage) and the searcher was reminded to

think aloud during the search.)

- filled in two questionnaires when the searcher completed the question (or decided to stop). These questionnaires measured satisfaction of the search result and comments and affective state (State-Anxiety).
- was given the second question sheet and asked to find the answer. (The Web browser was on the page where the first search ended and the searcher was again reminded to think aloud during the search.)
- filled in two questionnaires when the searcher completed the question or decided to stop. These questionnaires measured satisfaction of the search result and comments and affective state (State-Anxiety).

5. PRELIMINARY RESULTS

Because of the complex mix of variables, the results reported here should be considered preliminary. Further analysis is ongoing and will be reported in future publications.

5.1 Whether Or Not A *Correct Answer* was Found and How Did The Searcher *Feel* About The Result: Tables 2 and 3?

Question 1: About half of the participants failed to find a correct answer. Three of them gave up after some effort. Of the nine who found an answer that was incorrect, three were very sure they had found the right information.
Question 2: All participants claimed to have found the correct answer. For those who found the right information (N=22), half of them were highly confident about their search, the other half were moderately confident. The two who found an incorrect answer also felt moderately confident about their results.

Table 2. Search result (Y1) for question 1

Post-search confidence level (Y4)	Correct answer N=12	Incorrect answer N=9	No answer N=3
4 (high)	12	3	
3 (moderately high)		4	
2 (moderately low)		2	
1 (low)			3

Table 3. Search result (Y1) for question 2

Post-search confidence level (Y4)	Correct answer N=22	Incorrect answer N=2
4 (high)	11	
3 (moderately high)	11	2

5.2 The *Time Spent* On Interacting with The System And The *Number Of Sites* Visited For Each Search Question: Table 4

Time spent on each question varied among the participants. In general, about 15 minutes was spent on each question, although the standard deviation for question one is much larger than for question two. The number of sites visited has a similar distribution for question one and two. The Pearson coefficient is .93 and .89 respectively, indicating a high correlation between the search time and visited sites.

Table 4. Search Time (Y2) and Sites (Y3)

	Time in minutes		Number of Sites visited	
	Q1	Q2	Q1	Q2
Min	3	3	4	10
Max	57	39	106	77
Mean	14.50	15.04	28.13	29.67
SD	13.77	9.79	24.16	17.46

5.3 The *Change of Affective State*: Table 5

The change of State Anxiety levels caused by Web searching was measured after each search by using the same STAI Form Y1 (S-Anxiety) that was administered before the search. The value of change was calculated by the formula:

Δ S-Anxiety = Post-search S-Anxiety -- Pre-search S-Anxiety

for question 2,
Pre-search S-Anxiety = Post-search S-Anxiety after question 1

Table 5. Change of State Anxiety Level (Y5)

	Number of participants	
	After Q1	After Q2
No change	2	1
Less stressed	13	10
More stressed	9	13
Range of change	-17 to +18	-18 to +13
Mean	-.38	-.33
SD	8.85	6.49

5.4 Verbalization related to Cognitive, Affective and Physical Domains

Information processing. Users selectively processed information presented to them. Often they missed the information that was wanted, as shown by this participant's thinking aloud:

"Here it is. I did miss it before. Nation's total population projected to reach 275 million by the year 2000. I found it I just didn't read that very closely last time. I skimmed this article."

Subject knowledge. To start a search, the user must have some understanding of the subject matter being searched. Participants might be confident or less confident about their knowledge level:

"I feel confident that it will either already been in a press release since the year 2000 is fairly close or in a report by the Bureau."
"I don't know what proposal means."

Knowledge of Web search engines. How a system accepts a query may not always be obvious to users. Most participants seemed confused about which space they were searching. They input queries to whatever the search input slot was available on a page

and had trouble differentiating a subject directory on the Web and a search engine. For example, the Excite homepage places a subject directory (more like a classification) right under a form fill-out search slot for query input. In one case, the participant entered the query "Research Good News" in the slot and clicked on the link labeled Education under the subject directory. She thought she was narrowing the search in this category. However, the search query was ignored by the system and the user was led to the Education Page, the next level of directory.

Feelings. As users go through the search process, they revealed different types of feelings. Expressing their emotions seems to be an important part of the search process, as also described by Nahl and Tenopir (Ref. 5). Whether or not the computer understands it, the user needs to get it out, as illustrated by these thinking-aloud comments:

"I don't want ... What in the world?" [puzzled]
"Why I want education and they give me careers?" [disappointed]
"I am not sure what the NetSearch will do." [uncertainty]
"This is getting me where I was before." (Breathes heavily) [lost & frustrated]
"Here, let's try it again." [persistent]
"Well, Okay... then, why do we have it [a link returned a "document contains no data"] there? (in a sarcastic tone) [sarcastic]
"Huh, ... we're still waiting. I hate to wait for stuff like this ... Hmm. I get bored." [impatient]
"Ah...ah....ah! There it is." [excited]

Physical domain. Users are not always able to control the input devices. Typos (especially in URLs), clicking the wrong button, and double mouse clicks when a single click was intended all occur. Also, the mouse is not for left-handed users. "I am left-handed, so these mice things ... I know buttons are often a problem for me."

Information-seeking strategies. The Web searchers in this study either prefer to trace links from known pages or to search various search engines. With link-tracing the searcher started from a page with which they were familiar and followed links to find an answer. This type of user tends to look at the page thoroughly by scrolling the screen all the way down. In one interesting case, when the participant failed to find the information via several Web search engines he went to his own Web page and from there to his professor's homepage.

When the searcher had no clue as to where to start, he or she often used familiar Web searching engines to search and to modify searches. These users tended to input searches and look at the results briefly. They seldom went through more than two-thirds of the pages retrieved in a search. They often changed search formulas or tried the search in a different search engine.

6. DISCUSSION AND CONCLUSIONS

As mentioned above, the population we studied was graduate level students in a library and information sciences program. This may be considered a group of early adopters of Web searching or at least people who are motivated to use the Web regularly as required by their future profession. This group will be frequent users for their daily information services and provide instructions to other users.

Studying this user group can provide insight into differences in interacting with online IR systems and with Web systems, however, and are not atypical. In our sample, more than half were entry-level students whose behavior was probably more like typical academic users at the time of data collection. The mean EFT scores for all college students are 45.5 for male (S.D. 28.5) and 66.9 for female (S.D. 33.6), which is reported by Witkin et al. in 1954. Our student participants had similar scores (Table 1). The mean scores for STAI for college students are: S-Anxiety for male 36.47 (S.D. 10.02) and for female 38.76 (S.D. 11.95), T-Anxiety for male 38.30 (S.D. 9.18) and for female 40.40 (S.D. 10.15). Again, our participants are similar.

If we assume that our participants are at least in some ways similar to typical college students, our results should sound some note of concern for Web designers and Web instructors. From the preliminary results, we can see that half of the searchers failed to find an answer or a correct answer. On average, 15 minutes were spent for each question and about 29 sites were visited. In an extreme case, a participant spent as much as 57 minutes and visited 106 sites, without finding an answer. A more serious problem is that some searchers actually believed that they had found the right answer, even though the answer was incorrect. (This actually is not too far off from the classic studies that showed correct answers are given at the reference desk on an average of only about 50% of the time.) (Ref. 9.)

Individual differences are obvious from our results. These participants are quite diverse in terms of information skills, cognitive styles, affective states, and physical skills. Their search outcomes are also diverse. The revealed differences and their impact on search process and search outcomes merit further understanding.

Some basic design features on the Web are confusing to novices. (For example, placing a search engine together with a directory proved confusing to users. The system should be programmed to understand that when a user inputs a query before clicking on a category she may mean to search the query within this category.)

Individual preferences for search strategies in finding information may prevent a user from finding information effectively and efficiently. In our two questions, if the user begins in a related web space (such as the university's home page for question one), link-tracing will be the most effective and efficient way to find the answer. Those who used search engines for question one were taken to other Web pages and got lost in cyberspace. For the second question it was most efficient to go directly to the Census Bureau homepage. If a searcher does not know or cannot guess the Census page URL, it was more efficient to use a smaller search engine or a directory such as Yahoo, that indexes government resources.

7. REFERENCES

1 Hewins, Elizabeth T. "Information Needs and Use Studies," *Annual Review of Information Science and Technology* 25 (1990): 145-172.

2 Borgman, Christine L. "Why Are Online Catalogs Hard to Use? Lessons Learned From Information Retrieval Studies," *Journal of the American Society for Information Science* 37 (6, 1986): 387-400.

3 Borgman, Christine L. "Why Are Online Catalogs *Still* Hard to Use?" *Journal of the American Society for Information Science* 47 (7, 1996): 493-503.

4 Borgman, Christine L. "All Users of Information Retrieval Systems Are Not Created Equal: An Exploration Into Individual Differences," *Information Processing and Management* 25 (3, 1989): 237-251.

5 Nahl, Diane and Tenopir, Carol. "Affective and Cognitive Searching Behavior of Novice End-Users of a Full-Text Database," *Journal of the American Society for Information Science* 47 (April 1996): 276-286.

6 Tenopir, Carol, guest editor. "Perspectives on...Internet Issues," *Journal of the American Society for Information Science* 49 (1998): forthcoming.

7 *More Than Screen Deep: Toward Every-Citizen Interfaces to the Nation's Information Infrastructure.* Toward an Every-Citizen Interface to the Nation's Information Interface Infrastructure Steering Committee, Computer Science and Telecommunications Board, Commission on Physical Sciences, Mathematics, and Applications, National Research Council, 1996. URL: www.nap.edu.

8 Hsieh-Yee, I. "Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experienced Searchers," *Journal of the American Society for Information Science* 44 (3, 1993): 161-174.

9 Crews, Kenneth, "The Accuracy of Reference Service: Variables for Research and Implementation," *Library & Information Science Research* 10 (July 1988): 331-335.

10 Spielberger, C. D. (1983). State-Trait Anxiety Inventory (Form Y). Palo Alto, CA: Mind Garden.

11 Witkin, H. A. and et al. (1971). *A Manual for the Embedded Figures Tests.* Palo Alto, CA: Consulting Psychologists Press.

ACKNOWLEDGEMENT

This study is partially supported by a Faculty Development Research grant, Graduate School, University of Tennessee, Knoxville. We wish to thank Elizabeth Layman for her help with this project.