

Interaction on the Web: Scholars' Judgment of Information Quality and Cognitive Authority

Soo Young Rieh, Nicholas J. Belkin

School of Communication, Information and Library Studies
Rutgers University, New Brunswick, New Jersey

Abstract

This paper reports on a study of scholars' judgments of information quality and cognitive authority when interacting with information in the World Wide Web. This is the second in a series of studies on information quality and cognitive authority in the Web. The first study obtained data through retrospective accounts, and this study collected the data based on actual searching behaviors. Multiple data collection methods were utilized, combining verbal protocols during the searches, search logs, and post-search interviews. The scholars, recruited from diverse discipline areas, conducted four searches regarding four different tasks: research project, travel, medicine, and computer price. It was found that the subjects made two kinds of judgments: before they looked at some web pages (predictive judgment) and after they looked at it (evaluative judgment). Their judgments were analyzed in terms of facets indicating information quality, cognitive authority, topical interests, aesthetic aspects, and affective aspects. The factors influencing each judgment of quality and authority were identified, including major categories in characteristics of information objects, characteristics of sources, knowledge, and situation.

INTRODUCTION

When searching for "useful" information, people often base their actions on the concepts of quality and authority. In traditional information environments, there are well-established mechanisms for indicating or controlling for those factors, such as authors, publishers, and document type (Wilson, 1983) and refereeing processes (Janes & Rosenfeld, 1996). In other words, there are accumulated standards for publications according to which people judge the goodness of information. However, as the WWW lacks such mechanisms, people are likely to have difficulty in identifying "good" information based on authoritative sources. In addition, people who have relatively little experience in Web use might have even more difficulties in making judgment of which information is "useful" enough for their information problems.

This study concentrates on scholars' judgment because scholars are more likely to be concerned with information quality and cognitive authority than other populations. As a scholar's work, by nature, is heavily involved in interaction with information (e.g., finding, assessing, using, creating, and disseminating information), most scholars should feel competent to judge quality and authority based on their own evaluation criteria in the printed environment with which they are familiar. However, the same rules and criteria that have been used in printed collections might not apply exactly to the Web environment. In this study, faculty members and doctoral students constitute the scholar group. Doctoral students might have less experience in assessing quality and authority in the printed environment than faculty members, but they might be as concerned with these issues as faculty members are.

It is important to note that we examine the problem of information quality and authority from the perspective of users who are interacting with information in the Web environment. There have been several authors who were concerned with the problem of evaluation of information resources in the Web (e.g., Cooke, 1999; Kjartansdottir & Widenius, 1995; Pratt et al., 1996; Tate & Alexander, 1996). The common objective of these studies was to develop guidelines and critical evaluation skills related to bibliographic instruction. Unlike these studies in which little consideration has been given to real users, we consider users as active seekers and interpreters of information, who evaluate the usefulness, goodness, trustworthiness of information with respect to their goals, problems, intentions, and situations.

This is the second in a series of studies on information quality and cognitive authority in the Web. The first study (Rieh & Belkin, 1998) investigated this problem by conducting semi-structured interviews with 14 scholars. The results of the first study indicated that the issues of quality and authority were indeed important to people who

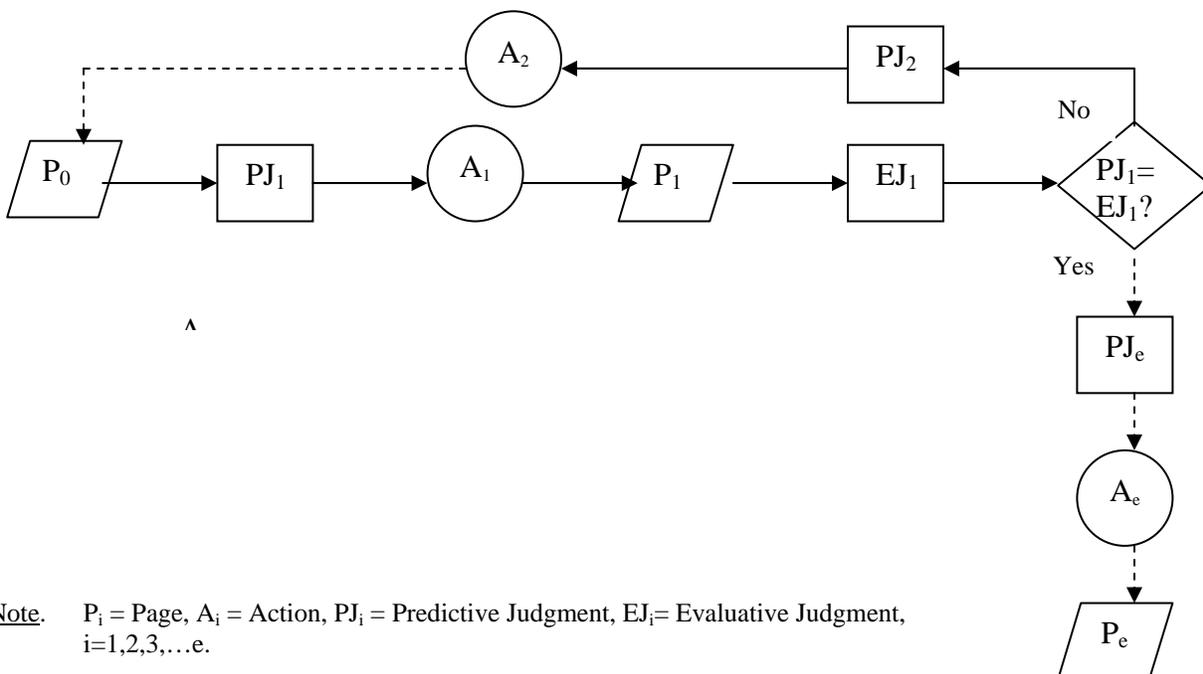
search in the Web. The subjects had substantial doubts about general quality of information sources, making comments indicating the necessity to expend more effort on quality and authority assessment in the Web than in other information systems. They assessed information quality based on source credibility and authority, giving considerable attention to institutional authority such as types of institutions and reputation of institutions, and individual authority including affiliation of author and creator. A general limitation of the first study was that we had no data regarding actual searching behavior, as we obtained the data through the interviews only. So, in the second study we utilized more diverse data collection methods, capturing search logs and verbal protocols during the searches as well as post-search interviews. We also attempted to recruit the subjects from more diverse scholar groups in terms of discipline areas.

CONCEPTUAL FRAMEWORK AND RESEARCH QUESTIONS

Studies of decision making suggest a framework for understanding the nature of judgment of information in the course of information retrieval interaction processes. The selection of one alternative out of a set of alternatives have been a central topic for decision theory for a long time (Huber, 1989). According to Huber, the decision process starts with an initial situation in which the decision maker is confronted with a set of alternatives to choose from. There is a desired goal situation (goal state) in which exactly one alternative has been chosen. The decision maker transforms the initial state into goal state by applying one or more operators. An operator can be defined as an activity to alter states.

In this process, the decision is made internally, but the choice is actual behavior that can be directly observed (Rachlin, 1989). Since a researcher cannot see directly into the mind of another person, that person's actual choice behavior constitutes a clue to what the person's decision process might be. It appears that two kinds of judgment are involved in this type of choice behavior: evaluative judgment and predictive judgment (Hogarth, 1987). According to Hogarth, evaluative judgment denotes the value judgments by which people express preferences, while predictive judgment refers to what they expect to happen. In other words, some decisions depend on a person's preferences (evaluative judgment), and other decisions are based on what a person anticipates might happen in the future (predictive judgment). Taking this model of decision-making, people's judgment and decision process on the Web can be characterized in the following way (see Figure 1).

Figure 1. Judgment Processes on the Web



When opening a Web browser, people start the search process with a default page of the Web browser (P_0). They take the first action (A_1) based on their own reasons. These reasons are associated with their predictive judgment (PJ_1). They decide to go to P_1 because they expect that it is going to be a good place to look or a useful site for the task. The predictive judgment is made before P_1 is seen. The reasons for choosing P_1 can be based on their knowledge, experience, recommendation from other people, or other characteristics of information objects and sources that might be visible on P_0 . These are the criteria for predictive judgments. People sometimes combine several criteria to make a decision (e.g., choice) toward P_1 . Once they reach P_1 , they make an evaluative judgment (EJ_1) about the page: for example, how good the information is, how useful the information is, whether there is a link which seems to be interesting, and whether they can trust the information. The reasons underlying such a judgment can be based on some characteristics of information objects, characteristics of sources, the person's own knowledge, or other situational factors. These are the criteria for evaluative judgments. If people find that evaluative judgments of the page do not match the expectations of their predictive judgments, then they might go back to a previous page or might decide to start with a new page. When making such a decision, another predictive judgment is made, and the action A_2 takes them to the next page. If people find that this page is what they expected, they would continue to use the page by selecting a link on the page, for instance. Iterating all these actions a number of times, they would reach P_e where their predictive judgments and evaluative judgments match, and they would decide to use the information from P_e .

Among the several facets influencing people's judgments in the information seeking process, this study specifically investigates the dimensions of information quality and cognitive authority. The definitions and dimensions of information quality and cognitive authority are described in detail in Rieh and Belkin (1998). Thus, only brief definitions will be discussed here. Taylor's (1986) notion of quality in his value-added model and Wilson's (1983) concept of cognitive authority provide a useful basis for the conceptual framework for this study. Taylor defines quality as "a user criterion which has to do with excellence or in some cases truthfulness in labeling" (p. 62), and characterizes five values included in the judgment of quality: accuracy, comprehensiveness, currency, reliability, and validity. The term "cognitive authority" was coined by Wilson (1983) where he differentiates cognitive authority from the other kind of authority, that of the person who is in a position to tell others what to do. He defines cognitive authority as "influence on one's thoughts that one would consciously recognize as proper" (p. 15). According to Wilson, the situation is one in which we may be faced with a number of different people all claiming to be knowledgeable but all having different things to say on the same subject. Then how can we choose among them, or how can we defend our choice once made? He discusses various indirect tests for recognizing the cognitive authority of a text including examining author, publisher, document type, and content of text.

To understand the decision process with respect to information quality and authority during searching on the Web in detail, the following research questions are addressed in this study.

1. To what extent do scholars make judgments of information quality and cognitive authority when they interact with information in the Web?
2. What is the nature of judgments of information quality and cognitive authority performed by the scholars?
3. What are the characteristics and factors that influence scholars' judgments of information quality and cognitive authority?

METHODOLOGY

Multiple data collection methods were utilized, combining verbal protocols during searches, search logs, and post-search interviews. The participants were 15 scholars from diverse disciplines, recruited at Rutgers University. The experiments took place at the Information Interaction Lab at School of Communication, Information, and Library Studies in Rutgers University from December 1998 to June 1999. Each subject performed four searches on the four topics described below. The tasks were given to the subjects in a random order. Before conducting their searches, participants completed a background questionnaire which asked about their demographic characteristics and their experiences with the Web. The subjects were given 15 minutes to complete each search. Two kinds of Web browser, Netscape 4.0 and Internet Explorer 4.0 were installed on a PC so that the subjects could choose the one with which they were more familiar. The PC used was equipped with a 100 MHz processor, 32 MB memory and 5.3 GB hard disk under Windows 95, using a 15" color monitor. The entire search interaction was logged using Lotus ScreenCam software. The subjects were instructed to "think aloud" about what they were doing, and why, as they searched. These verbal protocols and the computer monitor were captured on videotape. After the subjects

completed all four searches, the experimenter took a seat next to the subject so they could look at the computer monitor together. While the search logs saved in the ScreenCam were played back, a semi-structured interview was conducted about their search activities. The interviews were recorded on both videotapes and audiotapes.

The “generic” tasks given the subjects outlined the kinds of topics they would search on, while not restricting the specific information problems. The dimensions of quality and authority were embedded in the tasks by using such words as “good papers,” “useful information,” “credible information,” and “best price.” The tasks were given in a written form, and were associated with research, travel, medical, and product information as follows:

1. For the research project in which you are currently engaged, you would like to find some good papers which are new to you, which you think will be useful (*research* task)
2. You are planning for the next conference which you are going to attend, and would like to find useful information about hotels, restaurants, and features of interest in that city (*travel* task)
3. A friend of yours has just been diagnosed as having schistosomiasis, and you want to find credible information about the disease itself, and the best methods of treatment (*medicine* task).
4. You've decided that you want to buy a new computer to use at home, and now you need to find the best price for it (*computer* task).

The principal advantage of this methodological approach was that it made it possible to collect both concurrent and retrospective verbal data as it obtained the think-aloud data during the searches and the interview data after the searches. Also, by using generic rather than specific tasks, it was possible to gain some measure of verisimilitude, while maintaining the possibility of comparison between subjects. In addition, the features of ScreenCam software allowed us to play back the search logs including screen activities and cursor movements, which helped the subjects remember the actions that they had taken during the searches. This software also enabled us to pause and play the screen any time so that the subjects could answer the questions in length while the interviewer was holding the screen.

The tapes of the interviews and think-aloud tapes for the 15 subjects were transcribed. Content analysis was used as a technique to inductively identify and categorize the type and facets of judgment and the criteria mentioned by the subjects. Taking three kinds of data (search logs, interview transcripts, and think-aloud transcripts) together, the coding scheme was developed. The basic unit of analysis was a web page viewed by the subjects, and the data from the log, interview and thinking aloud protocol with respect to that page. Following Pharo's (1999) definition, the term web page is chosen to denote the “constellation of text, images, videos, and sounds that emerges as the result of activating a URL in a browser (p. 208).” For each page, the following items were coded: evaluative judgment, criteria for evaluative judgment, action on the page, predictive judgment, and criteria for predictive judgment. Pages and actions were coded from the search logs, and the other categories were coded from the interview and think-aloud data. After the coding of raw data was completed, the categories of the schemes were identified as described in the next section.

RESULTS AND FINDINGS

In the following sections we report on our analyses based on 1321 web pages from 60 searches by 15 subjects.

Characteristics of the Subjects

The subject group included 6 faculty members and 9 doctoral students. Among the faculty members, four held the status of professor while two were assistant professors. The subjects were distributed across diverse discipline areas. Faculty members were affiliated with departments of communication, library and information science, linguistics, sociology, chemistry, and computer science. Doctoral students were in departments of communication (4), library and information science (2), biomedical engineering, computer engineering, and organizational psychology. The average age for the doctoral student subjects was 32.2 while that for the faculty subjects was 45.3. Five subjects were female and 10 were male. All were experienced web users: 9 of them (60%) said that they use the Web every day or more than once a day, and all of them responded that they use search engines and directory services from “somewhat” to “a great deal” ($M=3.7$, $SD=0.8$ on a five point scale where 0 is “no experience,” 3 is “somewhat,” and 5 “a great deal”).

Predictive Judgment

A predictive judgment is defined as people's making predictions that reflect what they expect to happen by opening a new page. There were 442 responses coded as predictive judgments in answer to the interview questions of "why did you select this page to look at" and "what aspects of the page were important to you in making decisions about what to do next." We found that the subjects were indeed concerned about information quality and cognitive authority to a substantial extent (N=156, 51.1%) when they made decisions what to select on the Web before they looked at some web pages. The response was coded as "Information quality" when the subjects mentioned that the information, the web page, or the web site selected "is going to be" or "would be" *good, best, accurate, correct, current*. Cognitive authority was used as a facet when they said the information or the source that they had selected was *trustworthy, credible, reliable, reputable, respectable*. The facet of topical interest, which has been traditionally the representative of relevance judgment, accounted for 43.2% of the total responses regarding the judgment that the subjects made as predictions. Topical interest was coded when the subject mentioned that the information was going to be interesting because of what the information was about with respect to the task and the topic that he/she was looking for. 94.3% of all the responses regarding predictive judgments fell into these three categories. There were only 25 responses which were associated with the subjects' affective aspects or general expectations. Table 1 shows some examples for each coding category. In the following examples, task, subject number, and line number of interview transcripts are presented.

Table 1. Coding Examples of Predictive Judgment

Code	Examples (direct quotes)
Information Quality	<ol style="list-style-type: none"> 1. [Selected UnCover] "I want to come up with a couple of articles that I know are going to be good" (research, S010: L241). 2. [Selected Infoseek from Netscape NetSearch] "Infoseek would be a good search engine" (travel, S005: L90). 3. [Selected WHO Division of Control of Tropical Diseases – Schistosomiasis Control] "I saw the World Health Organization, and thought that information in there is correct" (medicine, S004: L331-332). 4. [Typed in www.zdnet.com] "I'm going to reduce a lot of uncertainty by just going with the one or two brands that I know are going to be good, like Gateway, Dell, Compaq, you know" (computer, S010: L105-107).
Cognitive Authority	<ol style="list-style-type: none"> 1. [Clicked on Libraries] "I think the data I get are more reliable than I might get from Yahoo or Altavista someplace" (research, S014: L224-225). 2. [Selected Microsoft expedia.com – Make a Hotel Reservation] "That's a Microsoft Network...they are credible to me" (travel, S005: L202-203). 3. [Selected World Health Organization from travelhealth.com site] "The World Health Organization is a trustworthy source to me so I decided to see how much information is there" (medicine, S003: L549-550). 4. [Typed in jandr.com] "I would log on from my own knowledge which I trust" (computer, S012: L305).
Topical Interests	<ol style="list-style-type: none"> 1. [Selected Computation and Language E-Print Archive] "Because much of what I do combines computations and linguistics, there's a listing of an archive because actually one I heard of" (research, S013: L219-221). 2. [Selected www.Aachen-Tourist.De] "Maybe that one will give information about hotels and restaurants" (travel, S016: L258-259). 3. [Selected Fact Sheet] "It would be in perfect resource for the question I was attempting to answer was, what is this disease and how do people get it and how do they prevent, how do they cure it?" (medicine, S008: L408-410) 4. [Selected Yahoo!] "I thought that might fit this search task better" (computer, S014: L509).

Evaluative Judgment

An evaluative judgment refers to people's making a value judgment by which they express preferences about the web page with respect to use. The total number of evaluative judgments was 534. When the subjects made evaluative judgments, they were concerned with information quality and cognitive authority to a great extent (N=354, 66.3%). This is not a surprising finding since we asked questions of the subjects directly such as "do you believe that the information is good, accurate, current, or correct?" "do you think that this is useful information for your information problem?" and "can you trust this information?" However, it is still worthy of note that our subjects mentioned topical interests only 71 times (13.3%) as a facet of evaluative judgment of information. One of the most significant findings here is that the subjects characterized the nature of information quality and cognitive authority with various dimensions. Some of the sub-facets with respect to information quality and cognitive authority correspond to Taylor's (1986) "values" of quality in which he characterized quality as accuracy, comprehensiveness, currency, reliability, and validity. We found new facets of information quality and authority such as "trustworthy," "scholarly," "official," "important," which were not identified in previous studies about dimensions of information quality (see Rieh & Belkin, 1998). Also it is interesting to note that the scholars used very diverse terms to characterize the perception of "goodness" and "usefulness," as shown in the Table 2. Other facets which emerged regarding evaluative judgment were aesthetic aspects, affective aspects, and general expectations. The keywords mentioned by the subjects, which determined the coding, are listed in Table 2.

Table 2. Facets of Evaluative Judgment of Information

Facets	Values	Keywords (direct quote) mentioned by the subjects
Information Quality	Good	Good job, bad, better, excellent, fine, nice, great, best, perfect, wonderful, incredible, cool, the state of art, well kept site, well developed site
	Accurate	Accurate, correct, right, precise
	Current	Current, recent, up-to-date, out-of-date, old, timely
	Useful	Useful, useless, hard to use, informative, helpful, doesn't help, can't understand, it's not going to be of much use, didn't make good use
	Important	Important
Cognitive Authority	Trustworthy	I trust it, trustworthy, believe in, confidence that this is true, seems real, faith in the quality
	Credible	Credible
	Reliable	Reliable, reliably done
	Scholarly	Scholarly, serious, academic, professional, biological, superficial, deep thing
	Official	Official
Topical Interests	Authoritative	Authoritative
		I find it interesting, the subject matter was of interest, wasn't interesting, approach that I'm interested in, close to, not relevant to what I was looking for, wasn't about it
Aesthetic Aspects		I liked the look, I didn't like the color, I liked the logo, weird looking, like visually, nice appearance, looks too ugly, bad design
Affective Aspects		Disappointed, surprised, frustrated, I got impatient with, those satisfy me, get confused, encouraging, puzzling, I was happy, weird, amazed, bizarre, I'm pleased, disturbed, annoying
General Expectation		Didn't meet my expectation, medium expectation, I was not what I expected, worth of my time, waste of my time, this stuff is insufficient, I wasn't what I was thinking of
Don't know		I don't know, I'm not sure, I don't care

Criteria for Judgment of Information Quality and Cognitive Authority

Criteria were coded when the subjects mentioned characteristics of information objects, sources and other factors that affected their judgments of information quality and cognitive authority. Most of the responses were from the question of “why do you think so?” and “what makes you to think so?” as follow-up questions of their judgments. Six major categories and 16 subcategories of criteria which applied to both evaluative judgments and predictive judgments emerged from the interviews and think-aloud data (Table 3).

1. Characteristics of Information Objects

The following categories were used to identify the characteristics of information objects (e.g., a web site, a database, an information system, an individual item) mentioned by the subject.

- Types of Information Objects

This category was coded when the subject was concerned whether the type of an information object is a book, bibliography, dissertation, journal article, personal homepage, or others. Not surprisingly, most of the responses occurred when the scholars searched for information about their research projects. An example of criteria for evaluative judgment is: “I wasn’t expecting to have a huge bibliography like this (research, S006: L667-668).”

While making predictive judgment, one scholar said, “I was pretty sure that it was a conference proceedings. So I thought it would be relevant (research, S003: L135-136).”

Table 3. Criteria for Judgment of Information Quality and Cognitive Authority

Major Categories of Criteria	Sub-Categories of Criteria
Characteristics of Information Objects	Types of Information Objects Title Content Organization/Structure Presentation Graphics Functionality
Characteristics of Sources	URL Organization Type Type of Source Reputation of Source One – Collective Source Author/Creator Credentials
Knowledge	Types of Knowledge: Domain Knowledge System Knowledge Mode of Acquisition: First-Hand Experience Second-Hand Knowledge
Situation	
Ranking in Search Output	
General Assumption	

- Title

Some scholars in our study mentioned that they based their judgment on cues from title, name, or label of the page. Most of the examples were found in the criteria for predictive judgment including: “When this label, special collections, caught my eye, I thought that it might have, things like the linguist list (research, S013: L130-131);”

“Because the title indicates that it is an overview or review of this issue which is something that could be very useful to me (research, S014: L242-243).”

- Content

A lot of scholars participating in this study said that what was “in” the page was the key for information quality and authority. The keywords for coding as content include: various information, a lot of information, complete, has comparison, summary, introduction, had research, and long list. Regarding evaluative judgments, examples are: “It’s an archive dedicated directly to papers relating to linguistic theory that I do most of my work in (research, S013: L72-73);” “It gives me more detailed features (computer, S009: L387).” In the case of predictive judgments, content is predicted based on some words and descriptions: “I think I went into it because I noticed it pertains to some of the lobbying efforts like the ALA and so on. But I thought that would be worth looking at. (research, S003: L178-180).” It is worth noting that content only rarely has to do explicitly with topical interests.

- Organization/Structure

To some scholars, how the page is organized or structured was a criterion they considered for making judgments. For example: “But this is much better organized and they do a good job (travel, S016: L124);” “I like visually, I like things, that are very clear, right up front. And in table format. That way I can print this out, use it for other stuff (computer, S001: L599-600).” The scholars also mentioned organization/structure when deciding to select a portal site as seen in the following examples: “So I went back to Yahoo, which seems to be the easiest one to browse (travel, S009: L472);” “Yahoo, I know it has a hierarchical organization...Searching Alta Vista for Korea would be nonsense because it would give me pages and pages on Korea; I wouldn’t get the right one (travel, S012: L532-535).”

- Presentation

The scholars mentioned how the page is written. This criterion was mentioned only when the scholars made evaluative judgments. The examples include: “No typos, good syntax, no misspelled words, these usual kind of stuff (research, S005: L543);” “Because it’s written in very scientific way (medicine, S006: L352-353);” “I noticed that they aren’t really giving me useful information about the systems. They are more like typical advertising pitch (computer, S013: L379).”

- Graphics

Some scholars were interested in how the page is designed in terms of graphic arts. The subjects showed different reactions toward graphics depending on the tasks. In the cases of travel and computer price information, the subjects liked to have some graphics. Here are some examples: “I liked the background. So, it looks elegant and I like the look. The colors is green and brown, I also like the logo and everything, so it was like kind of what I expected. (travel, S006: L71-73);” “Well, I don’t think I like this page very much. There was no picture of computer (computer, S008: L240-241).” However, with the tasks of research projects and medicine, the subjects did not like to see the graphics as shown in the following examples: “They had trick animation. It was a waste of time. So I got frustrated with this one real fast (research, S005: L553-554);” “It’s far too busy, cute, lots of pictures but impossible to find things (medicine, S016: L519-520).”

- Functionality

The scholars made comments about whether a function in the page is (isn’t) working as it is intended. For example: “I reached the conclusion that this search engine is not working properly (research, S013: L217-218);” “It doesn’t help that much. I can’t type in something that I am looking for (medicine, S001: Think-aloud);” “That’s not very nice. I can enter only one word (medicine, S016: Think-aloud).”

2. Characteristics of Sources

The following categories were used to identify the characteristics of sources mentioned by the subject.

- URL Organization Type

Sometimes scholars’ judgment was made based on the domain suffix – whether it is a .org site, .gov site, .com site, or .edu site. The examples include: “Absolutely [I trust it] because it’s an ‘.org’ (research, S016: L1017);” “Now that I look up, I see ‘travelhealth.com’ so it’s a public thing (medicine, S016: L460-461).”

- Type of Source

The Type of Source was used as a criterion when the scholars mentioned the kinds of source or organization from which the information originates. S008 explicitly mentioned why he decided not to go to web sites for finding information about research projects: “For web site information, there is no reason to believe that it has any validity to it, it could just be something that’s published, something that someone personally just published or posted on their Web site. And I guess the difference being that, for your own academic oriented research, you tend to want to read things that have been approved and are kind of endorsed by a journal with some type of reputation, some type of expertise that you know, has been evaluated, which I know is different than the other tasks because I kind of have a different level of trust than those Web sites (research, S008: L516-523).” With respect to the travel task, S016 said “I would trust less, one of the services which gives you hotel, I would say it’s very clear that service is not really complete because they make some deals with the hotels and the hotels have to pay the money to be listed, etc., so it won’t be complete. Something put up by the city itself, I expect to be more complete (travel, S016: L276-279).”

- Reputation of Source

A number of scholars mentioned that a source of information is reputable, famous, or well-known. This category was mentioned more frequently with the tasks of medicine and computer than other tasks. The examples include: “For this task, I know that I did not trust the information until, I did not completely buy into the information until it came from the Center for Disease Control, some place that I knew, that was very, very reputable in an area like this; (medicine, S008: L381-384);” “It’s Gateway. And they put a lot of effort into updating their prices and their computer configuration all the time so it’s in part, just my knowledge and reputation of that manufacturer (computer, S013: L339-341).” Some scholars said that they trusted the information because they found the information from “reputable search engines” as shown in the following example: “I suppose part of it is trust. I trust it in some ways because of the way I’ve got to it, through the Excite home page. I mean, it’s, I would consider it one of the more reputable or established search engines (computer, S003: L258-262).”

- One – Collective Source

To some scholars, what mattered was whether the information is based on a single person’s opinion or on that of group of people including an organization, a company, or an institute. Examples can be found in the following comments with regard to the travel task: “And [it was] good because it’s an independent thing. It was not only one, one restaurant page but this guide has links to several restaurants, so I thought, that’s good, that’s a third party which is going to give me an evaluation of different restaurants. That was what I needed (travel, S006: L135-138);” “It seems to be adequate...Because it had a lot of listings and it seemed to have different input from different sources (travel, S008: L115-119).”

- Author/Creator’s Credentials

This criterion was coded when the scholars mentioned an author’s or a creator’s affiliation or credentials. It was mentioned often by the subjects when they searched for information with the tasks of research and medicine. Here are some examples: “I would trust that because...it’s someone, some scholar, someone, some professional ideas on this subject matter, which is of interest to me (research, S003: L172-174);” “So this is the health advisory panel for this information that this company provides. So I’m looking at it and I see that in fact, there’s whole list. It’s real doctors and they’re at respected medical schools, largely, University of California, San Francisco. So now my trust level goes up somewhat. It’s not just an editor of some newsletter that I’ve never heard of. I’m not an expert in the medical field so I don’t know for certain but names are, but given my judgment, that makes me have more faith in the quality of the information, given that a legitimate medical advisory board (medical, S013, L745-752).”

3. Knowledge

Knowledge was given when the scholars mentioned their own personal knowledge, experience, familiarity with a site, source, or information itself. Knowledge was categorized in two facets: type, and mode of acquisition. First, based on the types of knowledge, we identified domain knowledge and system knowledge. Domain knowledge refers to knowledge of a topic area, including sources of information in the domain. System knowledge refers to the knowledge of system functionalities and system structures in a portal site, a web site, or any other information systems available on the Web. Second, according to the mode of acquisition, we classified knowledge as first-hand experience and second-hand knowledge. This classification is based on Wilson’s concepts (1983) in which he argues that we know of the world on the basis of either own personal experience or what we have second hand (we have read or heard of) from others. Each instance of knowledge was coded as either domain or system; if how the knowledge was acquired was mentioned, that code was also added.

- Domain Knowledge

Regarding the task of medicine, S010 decided to find the NLM site because: “Well, I know the National Library of Medicine, that they are going to have all of the medical databases. I also know that they’re the most prestigious medical resource in the U.S. I know that Medline is one of their services, one of many services I don’t know about. So Medline was a possibility but I also know that the National Library of Medicine is the major place to go. And that Medline would be a possibility under that. And that’s pretty much all that I knew (medical, S010: L418-423).”

- System Knowledge

With respect to the task of finding the best price for computer, S005 decided to go to onsale.com site because he knew about this web system. The following example was coded as both system knowledge and first-hand experience: “I had prior experience with buying something, buying computer stuff from them before. So the task said, find a good price and I knew that the prices were very good there so in this, e-commerce site, I have, I am very familiar with how to get around the site and I knew from experience that I had gotten things at a good price and I knew how the things work. A lot I had familiarity with, so I would go here just to see if they had computers that are, would be attractive to me in some way (computer, S005: 710-715).”

- First-Hand Experience

The following example represents the reason of choosing a particular portal site based on system knowledge acquired through the subject’s own experience: “Well that’s just prior experience. I got the sense that it was, that I would get more quote, serious, or sophisticated sources from Altavista. And from Infoseek I would get sort of mid range stuff. And when I was looking for, I just figured Infoseek would be as good as any that, a good search engine for that because travel arrangement and local attractions and restaurants and stuff is what I was looking for. So that’s just a force of habit (travel, S005: L87-92).”

- Second-Hand Knowledge

The following case is an example of choosing a site on the basis of second-hand system knowledge: “And I know a couple of people who spent a lot of time studying search engines and I’ve been told by some people with judgment I think is pretty good that Excite as a search is particularly good because it doesn’t, some kind of searches don’t just do simple key word matching. They try to do semantics similarity type of things and update their indexes [or something]. So it’s in part, there’s a certain amount of belief that Excite searches are likely to be good but that’s in part because I’ve been told by people who spend more time thinking about Web searches than I do (computer, S013: L289-295).”

4. Situation

This criterion refers to situational factors including the given task, time constraints, and duplication of information found. The examples coded as Situation include: “I went down a little further but I realized when I looked at this bar here, that it was going to be a very long list and, if I were at home doing it, I still would look at the size of that list and I would decide that I didn’t have time to look at all of them (research, S013: L198-201);” “This is the same thing that I already found in Yahoo (medicine, S009: think-aloud).”

5. Ranking in Search Output

Some scholars mentioned that they chose the particular page simply because it was ranked in first, or second, or was generally highly placed in the search results. It was mentioned only when they made predictive judgments. The examples were found as follows: “It was at the top of the list of search engines (travel, S014: L33);” “I think it was the first, I think it was the highest ranked pages in that (medicine, S005: L337).”

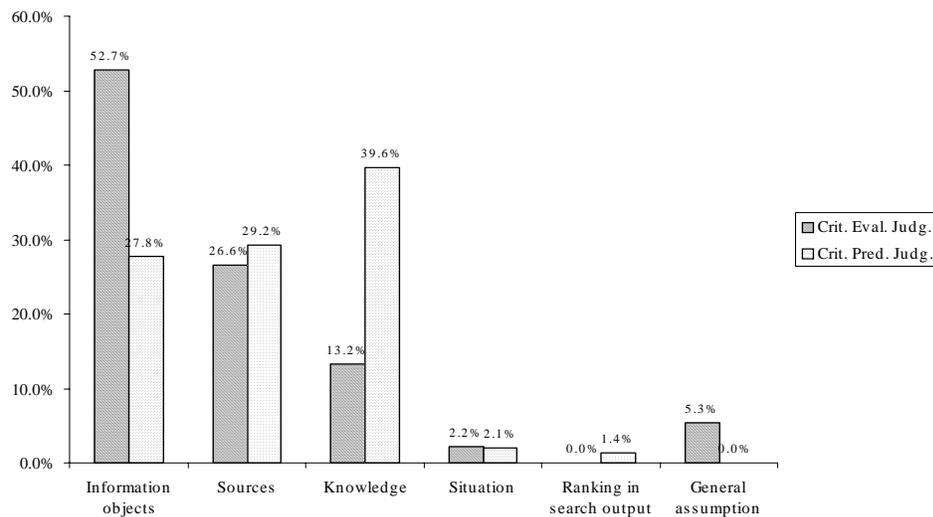
6. General Assumption

There were cases in which the scholars made judgments of quality or authority based on general assumptions that they had rather than any particular reasons. For example: “I didn’t run across anything that made me think it was false or not credible (travel, S005: L164);” “I’ve no reason to distrust this information. It’s just information about bands and people and things like that (travel, S007: L185-186).”

Comparison of Criteria between Evaluative Judgment and Predictive Judgment

We note that the criteria for quality and authority were used differently when deciding to select a page before looking at it (predictive judgment), than when deciding to use a page after looking at it (evaluative judgment). Figure 2 presents how the scholars in our study used the different kinds of criteria when they made predictive judgments and evaluative judgments, in terms of major categories. Apparently, Knowledge (N=114, 39.6%), especially System Knowledge (N=66, 22.9%), played an important role for predictive judgments while Characteristics of Information Objects (N=188, 52.7%), especially Content (N=108, 30.3%), was a critical category for evaluative judgments. Characteristics of Source was important for both evaluative judgments (N=95, 26.6%) and predictive judgments (N=84, 29.2%). In particular, Reputation of Sources was more frequently mentioned than any other aspects of Sources for both evaluative judgments (N=36, 10.1%) and predictive judgments (N=39, 13.5%). Ranking in Search Output did not occur as a criterion for evaluative judgments, and General Assumption was never mentioned as a criterion for predictive judgments.

Figure 2. Comparison of Criteria for Judgments of Quality and Authority



Note: Criteria for evaluative judgments N=357; Criteria for predictive judgments N=288

DISCUSSION AND CONCLUSION

A number of interesting findings emerged from the analyses of interview and think-aloud data. First, we found that people make two kinds of judgment, predictive judgment and evaluative judgment during the interaction with Web pages. The Web interaction environment permitted a unique situation in which we could observe people's behaviors with respect to prediction while activating a new page and evaluation while looking at the Web in a continuous process. The scholars made predictions about the next Web page, and if the Web page matched their expectation, then they continued to use the page. If it was not good or useful enough, according to their prediction or expectation, they would go back to the previous page or would try a new page. It appears that our empirical findings with respect to judgment and decision process correspond to our framework which we adopted from Hogarth's (1987) concepts described in the section on our conceptual framework. This may be analogous to the situation in bibliographic retrieval systems where people make predictions of relevance based on document surrogates, and evaluations of usefulness based on the documents themselves. There has been little research which has examined

people's decision process of judgment of information quality and authority in these two stages of information seeking behavior. Recently, Wang and Soergel (1998, 1999) have investigated three stages of document use: selecting, reading, and citing. They collected the data about relevance criteria and decision rules in 1992 during the selection processes and in 1995 on the reading and citing decisions, and compared them among three different stages. However, unlike our study, which focuses on judgment of quality and authority aspects in the Web environment, Wang and Soergel addressed relevance judgment and criteria in general with the context of traditional information interaction environments.

The second significant finding of this study is that people indeed make such judgments based on information quality and cognitive quality to a great extent. Although topicality (topical interests) was an important facet of judgment, especially when the subjects make predictive judgments (43.2%), the facets related to information quality and cognitive authority constitute 66.3% and 51.5% of the total responses for evaluative judgment and predictive judgment respectively. The facets which emerged from the subjects' utterances with respect to information quality and cognitive authority (see Table 2) were more diverse than those identified in the previous research (e.g., Taylor, 1986). For example, the facets such as "scholarly," "official," and "authoritative," identified by the subjects in their own terms, were not found in the literature. Considering that our subjects were all scholars, it is interesting to note that they made comments that the information looked "scholarly," "academic," and "biological" (for the medicine task) with respect to the judgment of cognitive authority. When the subjects made judgments, such as whether the information is good, accurate, current, useful, and important, or whether the source is trustworthy, credible, reliable, scholarly, official, and authoritative, they conducted "tests" for recognizing those facets of quality and authority (Wilson, 1983). In other words, there were specific criteria which served as bases or reasons for people's decisions. These criteria can be considered as factors that influence their judgments, with respect to quality and authority.

One such factor is the person's own knowledge. The results of our study indicated that the scholars used their knowledge a great deal when they interacted with information on the Web. The subjects mentioned their knowledge as a criterion more often when they made predictive judgments (39.6%) than evaluative judgments (13.2%). This suggests that the scholars used their knowledge to decide "where to go" or "what to select" rather than to make judgment of "whether this is good or useful." In other words, we might argue that the scholars tend to rely on their own knowledge, accumulated by either first-hand experience (42.1% of total knowledge) or second-hand knowledge (13.2% of total knowledge), as a way of identifying *authoritative sources* that they have known. They used both types of knowledge, domain knowledge and system knowledge to predict information quality and cognitive authority. Sometimes (22.9%) the scholars in our study knew about systems and sites available in the Web, and decided to go to that particular site. This site could be a portal site, an archive site, a medical site, or some article database available on the Web. They also decided to look at web pages based on knowledge that they learned with respect to subject areas (16.7%). Such cases were not limited to the task of research projects. For example, for the task of medicine, S008 expected to find the "American Medical Association" site, and S010 wanted to go to the "National Libraries of Medicine." For the task of travel, six of subjects tried to find the "official site" run by the city. Regarding the task of computer price, a lot of subjects (S001, S006, S007, S008, S011) went to big computer manufacturer sites (e.g., Dell, Gateway, IBM), a computer retailer site such as CompUSA (S006), or a PC Magazine site (S010). The common strategy of these decisions was that the scholars knew about these sources either as domain knowledge or system knowledge, and wanted to go to these sites directly. In other words, before they went to these sites (in fact, in most of cases, they never had been these sites before), they knew that the information provided by these sites was going to be good, useful, and trustworthy because they had knowledge about the sources.

We found that Characteristics of Information Objects is the criterion category which the scholars used most in order to make evaluative judgments on Web pages (N=188, 52.7% of all responses). This category was ranked first mainly because of the criterion of "Content" (N=108, 30.3% of all responses). Content should be differentiated from "topicality" which often refers to what the page is "about," as it refers to various aspects of what is "in" the page. For example, the scholars believed that the information is good and useful if the page had "various information," "a lot of information," "complete," "comparison," "summary," "introduction," "research," or "long list." These results indicate that the scholars considered content, rather than graphics (N=23, 6.4%), presentation (N=12, 3.4%), organization/structure (N=18, 4.6%), as the key criterion for evaluative judgments of information quality and cognitive authority. This finding is different from that of the previous studies which investigated Web searching behaviors with different user populations. Although these authors (e.g., Fidel et al., 1999; Scull, Milewski, and Millen, 1999) did not investigate the issue of quality and authority in the Web directly, some of their

findings are related to this issue. For example, Fidel et al. (1999) found that “[high school students] often relied on information that was displayed in a graphic form. They examined the graphics of a site to determine its relevance and quality (p. 35).” Also, Scull, Milewski, and Millen (1999), in a comparison study based on the behaviors of “novice college students and Web savvy librarians,” reported that “link reliability” and “site navigability” were criteria that “significantly impacted” librarians’ credibility assessments. However, our user population, scholars, consistently mentioned content most frequently as their criteria for all four tasks.

Further analyses are underway. One of the questions that we are examining is to what extent the scholars used different criteria for different kinds of tasks. The scholars were familiar with some tasks such as research projects, however, they were not as familiar with the medicine task. How familiarity with the task influences the factors of judgments is of interest. We are also interested in comparing the criteria used to the demographic characteristics of the subjects in terms of academic status (doctoral students vs. faculty members) and discipline areas (natural scientists vs. social scientists). Furthermore, we are also looking at the relationships between the types of page viewed and evaluative judgment, and between the types of action taken and predictive judgment. A general limitation of this study is that we presumed that people’s relevance criteria and decision rules accumulated in the traditional information systems might not be directly applicable to this new interaction environment, the Web, and information quality and cognitive authority might be more important dimensions of relevance judgment. However, we don’t have actual comparison data between two different interaction environments. Therefore, in the future, we will extend this study by comparing directly people’s judgments of information quality and authority in the Web environment and in the printed environment, focusing on the perceptions of information accuracy and credibility in different information interaction environments.

REFERENCES

- Cooke, A. (1999). *Authoritative guide to evaluating information on the Internet*. New York: Neal-Schuman.
- Fidel, R. et al. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society for Information Science*, 50(1), 24-37.
- Hogarth, R. M. (1987). *Judgment and choice: The psychology of decision*. 2nd Ed. New York: John Wiley & Sons.
- Huber, O. (1989). Information-processing operators in decision making. In H. Montgomery & O. Svenson (Eds.), *Process and structure in human decision making* (pp. 3-21). New York: John Wiley & Sons.
- Janes, J. W. & Rosenfeld, L. B. (1996). Networked information retrieval and organization: Issues and questions. *Journal of the American Society for Information Science*, 47(9), 711-715.
- Kjartandottir, A. & Widenius, M. (1995). The quality of business information on the Internet: Evaluation criteria applicable to Internet resources. *Swedish Library Research*, 3-4, 43-50.
- Pharo, N. (1999). Web information search strategies: A model for classifying web interaction? *Proceedings of the 3rd International Conference on the Conceptions of the Library and Information Science*, 207-218.
- Pratt, G. F., Flannery, P., & Perkins, C. L. D. (1996). Guidelines for Internet resource selection. *College and Research Libraries News*, 57(3), 134-135.
- Rachlin, H. (1989). *Judgment, decision, and choice: A cognitive/behavioral synthesis*. New York: W. J. Freeman and Company.
- Rieh, S. Y. & Belkin, N. J. (1998). Understanding judgment of information quality and cognitive authority in the WWW. *Proceedings of the 61st ASIS Annual Meeting*, 35, 279-289.
- Scull, C., Milewski, A., & Millen, D. (1999). Envisioning the Web: User expectations about the cyber-experience. *Proceedings of the 62nd ASIS Annual Meeting*, 36, 17-24.

Tate, M. & Alexander, J. (1996). Teaching critical evaluation skills for World Wide Web resources, *Computers in Libraries*, 16(10), 49-55.

Taylor, R. S. (1986). *Value-added processes in information systems*. Norwood, NJ: Ablex Publishing.

Wang, P. & Soergel, D. (1999). A cognitive model of document use during a research project. Study II: Decision at the reading and citing stages. *Journal of the American Society for Information Science*, 50(2), 98-114.

Wang, P. & Soergel, D. (1998). A cognitive model of document use during a research project. Study I: Document selection. *Journal of the American Society for Information Science*, 49(2), 115-133.

Wilson, P. (1983). *Second-hand knowledge: An inquiry into cognitive authority*. Westport, CT: Greenwood Press.