DOCUMENT RESUME

ED 405 001 IR 056 280

AUTHOR	White, Marilyn Domas; Wang, Peiling
TITLE	Document Selection and Relevance Assessments during a Research Project.
INSTITUTION	Maryland Univ., College Park. Coll. of Library and Information Services.
REPORT NO	CLIS-TR-97-02
PUB DATE	Jan 97
NOTE	102p.; Research funded by the Council on Library Resources.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Citations (References); *Decision Making; Higher Education; Information Sources; Interviews; Models; Needs Assessment; *Relevance (Information Retrieval); *Research Methodology; *Selection; *User Needs (Information)
IDENTIFIERS	*Search Behavior

ABSTRACT

User's information needs and the tasks they face change over the stages of a research project. In previous research by Peiling Wang, a cognitive model of users' document selection behavior for their research projects was developed. This study looks at the general applicability of Wang's model to subsequent decision-making about items selected initially. Data is gathered from a sample of the same researchers on use patterns and reasons for use and citation of relevant items in the original set. The sample consists of 15 faculty and graduate student agricultural economists. Post decisional data related to criteria and decision rules were collected via interviews. The project identifies additional criteria participants used to decide to read or cite documents; finds that not all criteria used to select documents initially, as determined by the 1994 model, continue to be applied; and determines that the decision-making rules identified in the original model are generally applicable to subsequent decision points in the research process. In connection with citation behavior, it also expands Wang's original model by identifying the contributions of the cited documents and metalevel documentation concerns as decision-making factors. Appendices include materials related to clearance and consent; interview guidelines; coding schema; sample coded interview; and definitions of criteria. Fourteen figures and tables accompany the text. (Contains 50 references.) (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 405 001

COLLEGE OF LIBRARY AND INFORMATION SERVICES

UNIVERSITY OF MARYLAND AT COLLEGE PARK

Document Selection and Relevance Assessments During a Research Project

by

Marilyn Domas White, Associate Professor
College of Library and Information Services
University of Maryland
College Park, MD 20742

and

Peiling Wang, Assistant Professor
School of Information Sciences
University of Tennessee
Knoxville, TN 37996-4330

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Marilyn D. White
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CLIS TECHNICAL REPORT NO. 97-02
January 1997

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

056280

BEST COPY AVAILABLE

COLLEGE OF LIBRARY AND INFORMATION SERVICES

UNIVERSITY OF MARYLAND AT COLLEGE PARK

Document Selection and Relevance Assessments During a Research Project

by

Marilyn Domas White, Associate Professor
College of Library and Information Services
University of Maryland
College Park, MD 20742

and

Peiling Wang, Assistant Professor
School of Information Sciences
University of Tennessee
Knoxville, TN 37996-4330

CLIS TECHNICAL REPORT NO. 97-02
January 1997

The research was funded by the Council on Library Resources, Grant no. 90. An earlier, slightly different version of this report was submitted to the Council on Library Resources on December 12, 1996.

Copyright © 1996
by Marilyn Domas White and Peiling Wang

The College of Library and Information Services, University of Maryland, College Park, supports the production and distribution of its Technical Report Series to disseminate research findings in a timely manner. The author(s) of each report is(are) solely responsible for its content, editorial style, and all other substantive aspects.

TABLE OF CONTENTS

Acknowledgments v
Abstract vii

1. BACKGROUND AND RESEARCH QUESTIONS 1
   - Background 1
   - Research Questions 1
   - Related Literature 2

2. METHODOLOGY 5
   - Participants 5
   - Interviews 5
   - Data Analysis 7
     - Reliability of Data 8
     - Validity of Data 8
   - Wang's Document Selection Model 8
   - Additional Definitions and Units of Analysis 11
   - Plan for Analysis 11

3. OVERALL PATTERNS OF USE 13

4. CRITERIA AND DECISION RULES 17
   - Criteria Used Across Stages 17
   - Additional Criteria 17
   - Decision Rules 20
   - Case Studies of Document Use by One Participant for One Item 22
     - Case 1 22
     - Case 2 24
     - Case 3 26

5. CITING BEHAVIOR 29
   - Background and Research Questions 29
   - Model of Citing Behavior 30
   - Contributions of Documents 31
   - Criteria Used in Decision-making 35
   - Metalevel Documentation Concerns 40
   - Conclusions 45

6. SUMMARY AND CONCLUSIONS 47
   - Summary 47
   - Suggestions for Future Research 48
   - Dissemination of Results 49

Bibliography 51

Appendices 55

A. MATERIALS RELATED TO CLEARANCE AND CONSENT 55
Application for Review of Research Project Using Human Subjects 57
Research Consent Agreement 59

B. INTERVIEW GUIDELINES 61

C. CODING SCHEMA 65
Coding Schema Used for Chapter 4 67
Revised Coding Schema for Criteria 73

D. SAMPLE CODED INTERVIEW 79

E. DEFINITIONS OF CRITERIA 97

LIST OF FIGURES AND TABLES

FIGURES

1. Document Use at Different Stages . . . 3
2. Document Selection Model . . . 9
4.1. Judgments Related to One Document (Case 1) . . . 23
4.2. Judgments Related to One Document (Case 2) . . . 25
4.3. Judgments Related to One Document (Case 3) . . . 27

TABLES

2.1. Characteristics of Participants . . . 6
3.1. Extent of Document Use by Participants . . . 14
4.1. Criteria Used Across Stages . . . 18
4.2. Decision Rules Across Stages . . . 21
5.1. Contributions of Cited Documents . . . 32
5.2. Criteria Grouped by Type . . . 36
5.3. Criteria Applied by the Nature of the Citing Decision . . . 39
5.4. Specific Coding for Selected Criteria . . . 41
5.5. Metalevel Documentation Concerns . . . 42

ACKNOWLEDGMENTS

The authors are grateful to the participants in this study who have reported graciously and frankly at various times on their manner of doing research and selecting related literature. We especially acknowledge those who commented subsequently on the analysis to validate the results.

The funding provided by the Council on Library Resources helped immeasurably in data analysis, providing funds for transcribing the interviews and data entry. We are very appreciative for their support.

In addition, the authors wish to acknowledge the helpful comments of individual referees on this research as reports have been submitted to conferences and journals.

ABSTRACT

Users' information needs and the tasks they face change over the stages of a research project, but most relevance studies have focused on judgments made at one time. In previous research, Wang (1994) developed a cognitive model of users' document selection behavior for their research projects.[1] It was based on analysis of users' judgments and decisions about bibliographic output from an online search. This study looks at the general applicability of Wang's model to subsequent decision making about items selected initially. This follow-up study gathers data from a sample of the same researchers on use patterns and reasons for use and citation of relevant items in the original set. The sample consists of 15 faculty and graduate student agricultural economists. Post decisional data related to criteria and decision rules were gathered via interviews.

The project identifies additional criteria participants used to decide to read or cite documents; finds that not all criteria used to select documents initially, as determined by the 1994 model, continue to be applied; and determines that the decision-making rules identified in the original model are generally applicable to subsequent decision points in the research process. In connection with citation behavior, it also expands Wang's original model by identifying the contributions of the cited documents and metalevel documentation concerns as decision-making factors.

[1]Peiling Wang. (1994) A cognitive model of document selection of real users of information retrieval systems. Ph.D. dissertation, University of Maryland, College Park.

1. BACKGROUND AND RESEARCH QUESTIONS

Background

Most of the recent user-criteria relevance research has focused on judgments users make at one time only, usually during an early search stage (Schamber 1994). The researchers ask users to react to output from a computerized search and to make judgments about each item's relevance in relation to the reason for the search. Each user generally bases his decisions on factors in the bibliographic record and his anticipated information needs based on tasks in the research project as he sees it at that time.

Users make relevance judgments at several other times during a project, however. Comparatively little is known about the correspondence between the decisions made at different stages or, at a general level, if the factors, criteria, and decision-making rules remain the same throughout the project. Because judgments of usefulness are so closely related to task requirements, it seems reasonable to suspect that these aspects vary across the process of completing the project, which encompasses many tasks. Dervin's sense-making theory also buttresses this idea since Dervin points out the dynamic nature as well as the uniqueness of sense-making in a situation (Dervin 1983). Perceptions of the relevance of the same item may change over time, and an item once regarded as irrelevant may be seen as relevant as the person's perception of the task evolves.

In an earlier study based on evaluation of relevance by judges, Rees and Schultz found that mean relevance ratings attached to items dropped significantly between the problem definition stage and the second or experimental stage and less so between the experimental stage and the final or data analysis stage, although judgments about some documents did not decrease over time (Rees and Schultz 1967). They also found that where a person was in the completion of a task was a more important factor than document representations, which varied in this experiment from title to bibliographic citation to full-text. White also found that information-seeking behavior changed during a research project but did not consider relevance (White 1975).

Research Questions

This research studies the criteria and decision-making rules used by users in making relevance judgments during a research project.[1] It builds on Wang's dissertation which identified these aspects at a previous stage in a project for these same users (Wang 1994). In that research

[1]The notions of both topical relevance and pertinence are incorporated into the new idea of "relevance."

Wang examined how users selected documents for the task at hand from bibliographic surrogates retrieved in an online search. She derived a cognitive model identifying criteria and decision rules which characterized the users' behavior in the early stage of a research project.

This study looks at the general applicability of Wang's model to subsequent decision-making about document use. The project is concerned with documenting decisions about items considered relevant in the original search: if they were used, why and how they were used or not used, the criteria and decision rules that the users employed in making choices. In addition, the project also addresses similar decisions on items used and cited which were not identified in the original search.

The research seeks to answer the following major questions:

1. How are items considered relevant at an early stage of research used subsequently in the same project?

2. Do users apply the same criteria and rules whenever they make subsequent decisions about use of an item?

Figure 1 shows graphically the decisions a researcher makes for a document during a research project, as conceptualized for this research. Wang's original research covered the decisions in the first box; the project funded by Council on Library Resources, presented in this report, emphasizes the decisions in the second box.

Related Literature

The literature relevant for this project appears in several different, but related, areas: relevance, decision making, and citation behavior.

Relevance in information retrieval has been used frequently as a criterion to evaluate the effectiveness of a system or a single search. Saracevic synthesized research of the 1960's and 1970's and summarized factors considered for relevance judgment (Saracevic 1970; Saracevic 1975). Most of the earlier research was based on experimental studies involving the use of judges to assess relevance. More recent research, which has generally emphasized qualitative studies of real users engaged in actual problems, has been summarized by Schamber (Schamber 1994). In an earlier review, Schamber and others called for a cognitive review of relevance and a user-oriented approach (Schamber, Eisenberg, and Nilan 1990). There is a clear shift of research in this field from a system-centered view to a user-oriented view. Four recent dissertations, including Wang's, have examined the actual users' criteria in making relevance judgments of the documents for the task at hand (Park 1992; Barry 1993; Wang 1994). Schamber, Park, and Barry emphasized criteria; Wang also included decision rules.

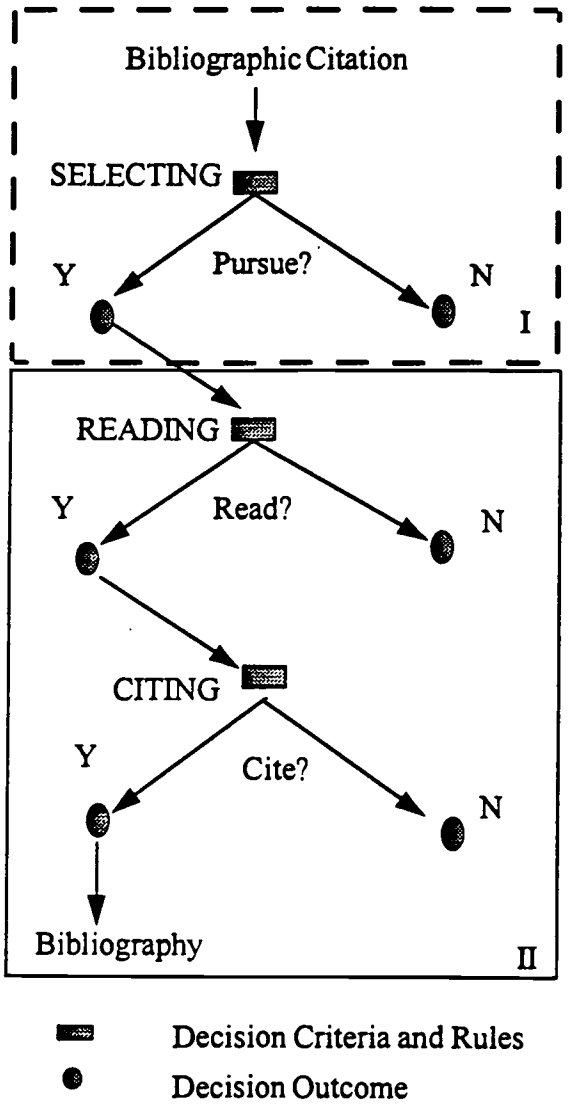

Figure 1. Document Use at Different Stages

The psychology literature related to decision making is useful, especially in its discussion of decision rules and factors. Using literature either in surrogate or full-text format involves decision-making by the user. In psychology, research in decision-making has shifted from studying final decisions to focusing on cognitive processes and the structure of decision-making (Svenson 1979; Wright 1984; Fischoff 1988). Decision processes are frequently studied by process-tracing techniques, *e.g.* think aloud reports (Ericsson and Simon 1993). Studies of decision making in psychology show that human beings process information and apply criteria and rules in making decisions (Svenson 1979; Simon 1985; Montgomery and Svenson 1989).

Acknowledging related work is a widespread and essential practice in scholarly communication. This practice and the artifacts it creates have been studied in several disciplines for a variety of reasons (Snyder, Cronin, and Davenport 1995). In library and information science, the ensuing literature is covered by several excellent literature reviews and syntheses (Smith 1981; Cronin 1984; White and McCain 1989; Borgman 1990; Liu 1993). Of particular relevance in the analysis of citing behavior in this project are studies presenting a typology of user motivations or criteria (Chubin and Moitra 1975; Hodges 1978; Moravcsik and Murugesan 1978; Spiegel-Rosing 1977; Frost 1979; Peritz 1983; Brooks 1985; Brooks 1986; Vinkler 1987; Cano 1989; Snyder and Bonzi 1989; Bonzi and Snyder 1989; Hooten 1991) and studies looking at citing behavior related to negative decisions (Vinkler 1987; MacRoberts and MacRoberts 1988). This project is interested in the behavior underlying the practice, especially the reasons why researchers choose to cite or not to cite particular works as they document their research activity.

2. METHODOLOGY

Because many of the faculty and graduate students who had served as subjects in Wang's original research were still available in 1995 and had completed their projects, the researchers were in the relatively unique position of being able to follow a set of items judged relevant at one stage to see how they were evaluated at subsequent stages.

In the longitudinal study, data were collected at two stages of research projects: (1) early in the project, when the participants were involved in surveying existing literature on a topic; and (2) at the conclusion of the projects, when the written products were completed or near completion. In the second stage, information was gathered about reading/consulting decisions and about citing decisions.

Participants

In Summer 1992, 25 agricultural economics faculty and graduate students provided concurrent verbal reports about their decisions as they reviewed output from a search of online databases (Wang 1994). All participants in the 1992 study were invited to participate in the second study. Initially the participants were questioned about the status of their projects by phone or personal contact. Those who had dropped or modified their initial research significantly were excluded. Of the original 25 participants, 10 were excluded for this reason. The other 15 participants agreed to participate in the follow-up interviews. For the study of citation behavior reported in Chapter 5, 3 of the 15 participants (N=12) were excluded because they were not able to provide data about citing decisions: two participants still had research in progress and the third was a joint author whose co-author had made the citing decisions. Table 2.1 shows the characteristics of the participants (See Appendix A for description of project and human subjects permission form).

For the sake of clarity and brevity, hereafter each study is referred to as either the 1992 or 1995 study respectively, the faculty and students in the studies as participants. The term researchers refers to the authors of the 1995 study, a generic "researcher," or others studying document use behavior, as evident from the context.

Interviews

In the 1995 project, data were gathered through a structured interview with each participant (See Appendix B for the interview schedule). Essentially the interview followed up on the decisions each user had made earlier in the 1992 study to see what further use had been made of the items he originally considered relevant. After determining the relationship between

Table 2.1. Characteristics of Participants

(1) No.	(2) Status	(3) Experience (Years)	(4) Task	(5) Items Selected in 1992	(6) Total Cites in Product
2	Faculty	31	Report	22	6
3	Faculty	24	Wk. Paper	32	118
4	Faculty	22	Book	14	30
6	Faculty	19	Book chap.	8	17
7	Faculty	17	Wk. Paper	14	22
8	Faculty	14	Article	33	42
10	Faculty	5	Grant Prop.	19	25
11	Faculty	4	Article	5	17
12	PhD Student	5	Dissertation	13	125
13	PhD Student	5	Dissertation	23	118
16	PhD Student	4	Proposal	27	n.a.
17	PhD Student	2	Article	8	n.a.
20	PhD Student	1	Proposal	53	29
21	PhD Student	1	Proposal	27	28
25	MS Student	1	Thesis	7	52

Column notes:
1) No. is the participant identifier assigned in the first phase of this study. Related information about the participants is included in Wang (1994).
3) Number refers to years after Ph.D. for faculty, years in program for students.
4) Task refers to the generation of the product indicated. Variety in products partially explains the range of total cites. Products are: Article: Journal article; Grant Prop.: Grant proposal; Proposal: Dissertation proposal; Report: Technical report (2); Thesis: Master's thesis (25); Wk. Paper: Working paper (3, 7).
6) At the time of data collection, the final products for Participants 16 and 17 were still in progress. n.a. = not applicable.

the written product and the original query, the interview solicited information about the user's decision-making regarding use or non-use of items in three categories: (a) items judged useful in the 1992 study that were cited; (b) relevant items from the 1992 study that were not cited; and (c) items in the bibliography of the particiant's written product which were not from the 1992 search. In category (c), the information was obtained for at least a sample of eight items, although some respondents provided information for all items in their bibliographies.

The open-ended questions were designed to elicit extent of use, reasons for use or non-use, and decision-making rules in making the judgments. During the interview the participants could refer to items as presented in the original output, their papers, actual publications, anything

that could stimulate their memories regarding the decision-making. One participant checked a referee's comments, for example.

Because information about decision-making rules is often interwoven throughout the discussion and may need additional questioning to elicit completely the factors considered, both researchers were present during each interview. One conducted the interview; the other took field notes and asked follow-up questions when necessary. The interviews were audio-taped and transcribed for analysis.

The data in the 1995 data-gathering phase, unlike the data in the 1992 study, were gathered post decision. This timing was necessary because the participants' decisions were dispersed throughout their research, which, for some participants, extended for the entire 2½-year period between studies. Some problems for the completeness of the data, especially for recall of the reasons for using or not using or citing or not citing particular documents, may have been created by the period of time between the actual decision and data collection. Data gathering for the first stage occurred in Summer 1992; for the follow-up research, the interviews occurred in early Spring 1995.

Data Analysis

Since this study is primarily qualitative, data analysis focuses on the content of the participants' verbalized thoughts. Where appropriate, evidence is given by quoting the verbal protocols of the subjects; the quotes were chosen for their representativeness or typicality. Although the verbal data were also quantified by counting occurrences of codings for each category, the numbers serve only to indicate how many cases were found. It cannot be assumed that these are the only cases which occurred during the decision processes, because some thoughts may have passed by without being verbalized (Ericsson and Simon 1993).

In the interviews, each participant freely elected what he chose to remark on and commented in his own language without reference to categories provided by the researchers. The respondents varied in referring to document characteristics, frames of reference, criteria, generalized rules, scales, and scale values, making the data rich but, at times, inconsistent. The interviewers achieved completeness in several respects: each item in the stimulus sets was commented on;[2] for each item in the set of items selected but not cited, the respondent made a yes or no decision about subsequent use with clarification of that use.

The interview data were organized and analyzed initially using *NUD*IST* (Non-numerical Unstructured Data Indexing, Searching and Theorizing)(*Q.S.R. NUD*IST 3.0 for Microsoft*

[2]Sometimes the only comment the participant could make was that he could not remember. In a few instances, an item was overlooked because documents were discussed in an order different from that in the original set of items provided to the participant. Occasionally a comment was lost because of technical problems.

Windows, 1985-1994). *NUD*IST* is a qualitative data analysis program for organizing, indexing, and analyzing textual data. This package has powerful functions in text searching and coding and in category (coding schema) building. The data were coded for: extent of use, reasons for reading or not reading an item and citing or not citing materials; rules used to make the decisions; corollary use of items, *e.g.* as leads to other literature. The criteria and decision rules that had been identified on the basis of the 1992 study served as the initial basis for coding, but the schema was augmented as new criteria became evident. The data went through several iterations of analysis (See Appendices C.1 and C.2 for coding categories used for the interview data; also Appendix D for a sample interview coded with the criteria noted in C.1). Subsequently, for statistical analysis, the qualitative database was augmented by a numerical database.

Reliability of Data

One of the researchers did the initial coding. To assess the reliability of the coding, a sample of interviews were recoded by the second researcher, which led to dividing some of the original criteria into more specific criteria, and recoding of all relevant items according to the new schema. Taking into consideration the modifications, the overlap in coding for broad criteria for the items coded by both researchers was 86 percent. For exploratory studies, the acceptable level of intercoders' consistency is 80 percent (Krippendorff 1980), so the data may be considered reliable.

Validity of Data

Several steps were taken to insure the validity of the data in this study. First, the data were obtained directly from the participants, and they described their behavior in their own language without recourse to pre-established categories. Recall of decisions was stimulated by reference to the 1992 bibliographic output or the bibliography in their product, but the participants were able to refer to the documents themselves or other items to reinforce their recall. Most were interviewed in their offices or work settings so these materials were readily accessible. The initial content analysis used a set of criteria defined in the 1992 study but, as the analysis evolved, it was supplemented through inductive content analysis to identify other criteria and coding categories. For the most part, the categories appear across some, but not all, participants; some also appear in previous research related to relevance and motivations for citing. The categories were checked with several participants and some adjustments were made.

Wang's Document Selection Model

Figure 2 shows Wang's Document Selection Model in detail. According to this model, document selection consists of three processes:

1. The user processes information from the citation to arrive at values for each of a number of criteria.

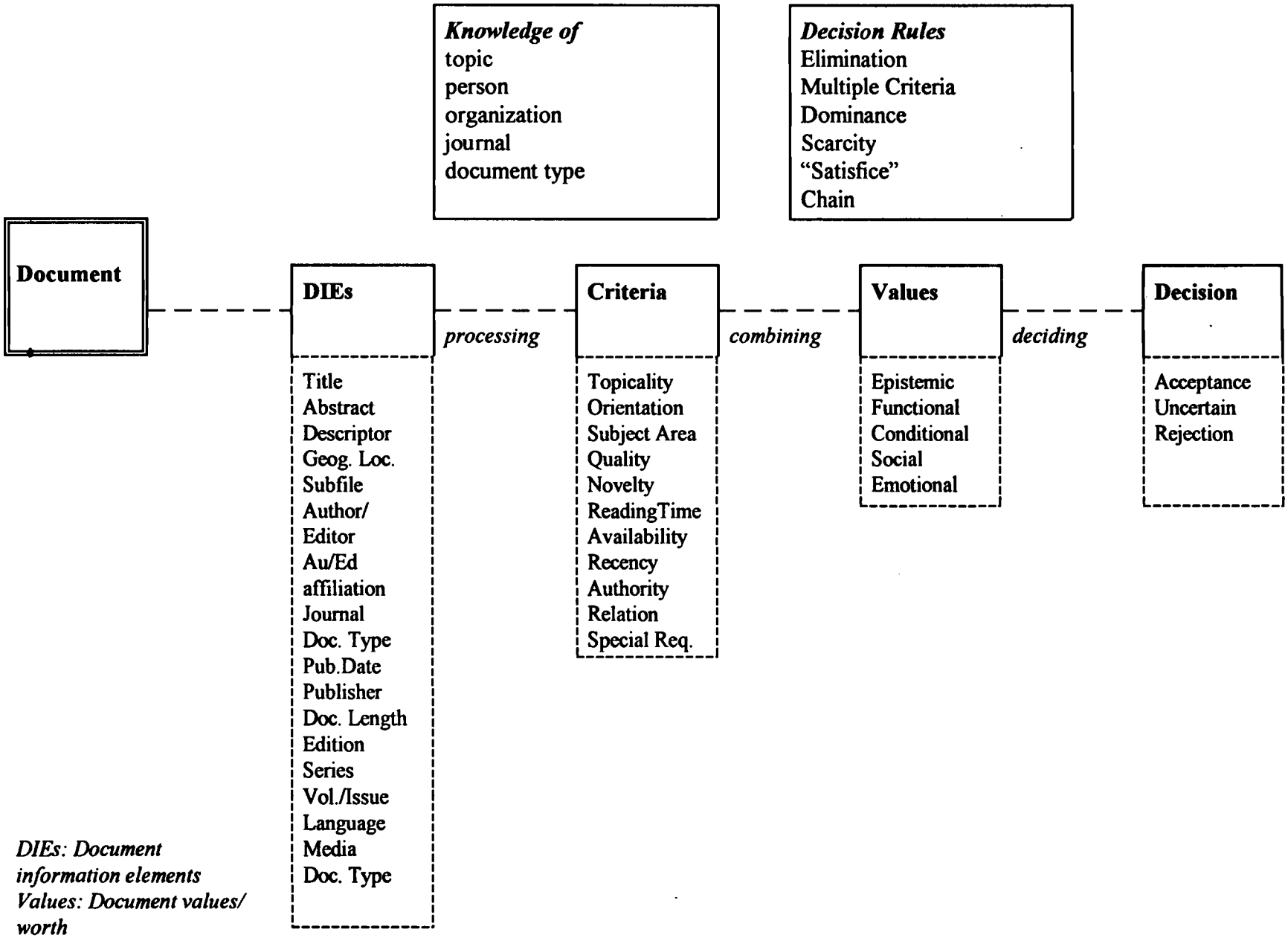

Figure 2. Document Selection Model

2. The user applies and combines criterion values to assess document value or worth in five categories.

3. The user weighs the document values and makes a decision (Wang 1994, 36).

In making these decisions, the user applies a range of decision rules either alone or together. The arrows indicate influence but the pattern of influence does not always correspond to the sequence of events. Although the 1995 project addresses the question of the general applicability of this model to subsequent stages, it does not consider all components of the model, omitting values and document information elements (DIEs). Only the components considered in the 1995 study are clarified below.[3]

Three options exist initially for the decision: accept, reject, or postpone judgment because of uncertainty. All three options are available throughout the initial scanning of the citations but, at some point, the user changes the postponed to accept or reject. Postponed documents are grouped and processed after the initial scanning of all documents, or decisions are made subsequent to seeing the document but before the user has reviewed all others.

Wang identified 11 criteria at the first stage, which are defined in Chapter 4. A *criterion* is defined as a filter a participant applies to a document to determine its usefulness to his information need. Wang divides the criteria into *document criteria* (1-9) and *document origin criteria* (10-11). With document criteria, the document is useful in itself for content, style, or whatever. Document origin criteria, which sometimes override the document criteria, relate to the "user's situation in the real world and his/her relationship with the origin of the document (Wang 1994, 46)."

In making the judgments the users applied several different rules:

1. Elimination rule
2. Multiple-criteria rule
3. Dominance rule
4. Scarcity rule
5. Satisfice rule
6. Chain rule

These rules are explained in Chapter 4. The first two apply to individual documents; the third through fifth are "trade-off principles about a set of documents (Wang 1994, 58)." The last denotes a specific relationship between or among documents.

---

[3]For a discussion of values and DIEs included in Wang's model, see her dissertation (Wang 1994, 42-46).

Additional Definitions and Units of Analysis

The primary unit of analysis is the *document use decision, i.e.* a decision made by a participant regarding use of a document. Possible decisions in the 1995 study include: decision to read or not to read; decision to cite or not to cite. In the study reported in Chapter 5, a subset of 314 document use decisions is used in this paper: those related to *uncited* items, *i.e.*, documents read, but not cited; and *cited* items: documents read and cited, and new cited items in the participants' bibliographies. Excluded from the analysis were items from the original search that were not read. With the exception of Tables 2.1 and 3.1 which provide summary data by participant, individual participant data are not provided.

In the criteria analyses, the unit is usually the criterion coded broadly. The broad criteria were also coded specifically, as indicated through selective examples in Table 5.4. In a document decision unit, a participant may use one or more criteria. Definitions of the individual criteria are included in the appropriate sections.

Plan for Analysis

The analysis presented in the next few chapters differs in detail. In Chapter 3, the emphasis is on presenting statistically the general pattern of use over time for the items in the original search output. Chapter 4 identifies the changes in the application of decision rules and criteria over the three stages: selection, use, and citing. Also included in this chapter are several case studies of document decisions for one document over time. Chapter 5 focuses on the final stage, the citation stage, and identifies more specifically the criteria applied at this stage. It also distinguishes other factors included at that stage in the decision-making process. Chapter 6 summarizes the findings, suggests future research, and indicates the dissemination of the findings of this report.

The analysis is more complete for Chapter 5; an article based on this chapter has already been accepted for publication. Data and results for Chapters 4 are more complete than the data in published conference papers, but a more thorough analysis is currently in process for submission as a journal article.

3. OVERALL PATTERNS OF USE

Table 3.1 summarizes document use at each stage of the project included in the large project to provide a perspective on citation behavior.

Several important findings are evident in Table 3.1. From the original search output, the participants identified as potentially useful items far more items than they actually consulted or read during the research process. In addition, by the time they completed their projects, they chose to cite even fewer, selectively choosing from among those they consulted.

The original searches ranged from 21 to 77 documents (Wang 1994, 107). They averaged 49 documents (S.D. 15). The participants identified as potentially useful an average of 41 percent (S.D. 21) of the items in their search outputs (Column 4). As they progressed through their projects, they elected to follow up only on some of these items. On average, the participants used, *i.e.* scanned, read completely, or consulted for relevant information about 52.5 percent (S.D. 22.9) of the items they had considered useful initially (Column 5). The searches were not the only source of information for their projects. In the interviews, they mentioned consulting colleagues and other interested parties on campus and elsewhere, scanning new literature as it appeared, referring to texts or other materials in their own collections, being notified of missing literature by referees or, in the case of the students, by their advisors.

The average percentage of items cited from the search was 19.5 percent (Column 6), but the distribution is skewed. The range was from 0 percent to 93 percent, with 6 of the 15 participants citing none of the items originally considered potentially useful. This pattern is not the result of a tendency to cite relatively few documents. As Column 8 indicates, three participants produced relatively large products with more than 118 items in their bibliographies. The shorter bibliographies averaged 27 items. In only one case did the items originally identified constitute a large portion of the final bibliography (Column 6, second percentage figure). Excluding that project as an outlier, the average percentage of items judged relevant in the original search and subsequently cited was only 2.3 percent in all projects, 5 percent in the projects citing at least some items.

Table 3.1. Extent of Document Use by Participants

(1) No.*	(2) Status	(3) Product	Based on 1992 Search ** (4) Selected	(5) Read	(6) Cited	Based on 1995 Product (7) New cites	(8) Total cites
2	Faculty	Technical report	22 (44.0%)	7 (31.8%)	0 (0; 0)	n.a.	6
3	Faculty	Working paper	32 (58.1)	18 (56.3)	7 (38.9; 6.0)	10	118
4	Faculty	Book	14 (31.8)	13 (92.9)	12 (92.3; 40.0)	18	30
6	Faculty	Book chapter	8 (20.0)	4 (50.0)	0 (0; 0)	8	17
7	Faculty	Working paper	14 (25.0)	7 (50.0)	1 (14.3; 4.5)	8	22
8	Faculty	Article	33 (63.5)	21 (63.6)	1 (4.8; 2.4)	8	42
10	Faculty	Grant proposal	19 (38.0)	3 (15.8)	0 (0; 0)	22	25
11	Faculty	Article	5 (13.5)	1 (20.0)	0 (0; 0)	17	17
12	Doctoral student	Dissertation	13 (38.2)	13 (100.0)	4 (30.8; 3.2)	8	125
13	Doctoral student	Dissertation	23 (39.7)	13 (56.5)	4 (30.8; 3.4)	8	118
16	Doctoral student	Thesis proposal	27 (42.2)	12 (44.4)	n.a.	n.a.	n.a.
17	Doctoral student	Article	8 (12.3)	5 (62.5)	n.a.	n.a.	n.a.
20	Doctoral student	Thesis proposal	53 (68.8)	26 (49.1)	0 (0; 0)	29	29
21	Doctoral student	Thesis proposal	27 (87.1)	10 (37.0)	1 (10.0; 3.6)	28	28
25	Master's student	Master's thesis	7 (33.3)	4 (57.1)	0 (0; 0)	16	52

Table 3.1. Extent of Document Use by Participants (Continued)

* Number is the participant identifier assigned in Wang (1994). Individual data for these participants for the early stage are provided in Wang's dissertation (Wang 1994, 107).
** Percentage is computed by dividing number of items at this task by the number of items from the previous task. The second percentage in Column 6 is the number of cited items as a percentage of the total cites in the participant's written product.
Participants 2, 16, and 17 are omitted from the citation study. Participants 16 and 17 have research in progress; Participant 2's co-author made the citing decisions.

4. CRITERIA AND DECISION RULES

This chapter addresses the question: Do users apply the same criteria and rules whenever they make subsequent decisions about use of an item? In this chapter, the emphasis is on identifying and characterizing the criteria and decision rules used at different times during a research project. No data are presented on the frequency of occurrence or the extent of use of each criterion or decision rule by the participants.

Criteria Used Across Stages

In Table 4.1, Column 2 shows criteria identified in previous research as affecting the relevance judgments based on a bibliographic surrogate; Columns 3 and 4 show criteria the users exhibited in two tasks at the later stage: scanning or reading a document, and citing the document in the written product.

Several important findings emerge after this analysis. New criteria were identified that seem to be used only during the last two tasks (See bottom half of Table 4.1). The document selection model needs to be revised and expanded to describe decision-making for reading and citing. While the users, as a group, use many of the original criteria in deciding to obtain a physical copy of the document and actually scan or read it, they also considered six totally new criteria in making these decisions (See Column 3, bottom half of Table 4.1). Some of the new criteria identified in connection with the reading decision carry over to citing: actual quality, classic/founder, well-known/standard reference, prolific author, and judge. But two new criteria emerge as well that appear only in making judgments to cite a source: norm and credential (Table 4.1, Column 4). Six criteria from the selection model were applicable to both subsequent tasks: topicality, novelty, recency, orientation, authority, and relation. The model also has to be adjusted to reflect that some of the criteria were not considered in later tasks.

Additional Criteria

In all, a total of eight new criteria emerged based on this analysis (See Appendix E for definitions and examples of all criteria). These are defined and illustrated with quotations as follows:[4]

---

[4]In the quotations the titles of specific articles, books, or papers and their authors (but not journals) are usually referred to by X; also ambiguous terms in the comments are clarified briefly within square brackets. These conventions hold across all participant comments cited in this paper and in Appendix D, except in relation to the three case studies in Chapter 4.

Table 4.1. Criteria Used Across Stages

(1) Criteria	Early Stage (2) Selecting	Later Stage (3) Reading	Later Stage (4) Citing
Original Criteria *			
Topicality	✓	✓	✓
Novelty	✓	✓	✓
Subject Area	✓	✓	
Expected Quality	✓	✓	
Recency	✓	✓	✓
Orientation	✓	✓	✓
Availability	✓	✓	
Time	✓	✓	
Special Requisite	✓		
Authority	✓	✓	✓
Relationship	✓	✓	✓
New Criteria			
Cognitive Requisite		✓	
Actual Quality		✓	✓
Classic/Founder		✓	✓
Well-known/Standard Reference		✓	✓
Prolific Author		✓	✓
Judge		✓	✓
Norm			✓
Credential			✓

* See Appendix E for definitions of all criteria identified in the 1992 or 1995 studies

Cognitive Requisite

Whether the user has the knowledge to understand a document.

"This is an example of the paper from agronomy side. It is in a sort of multidisciplinary journal. I needed it in order to get some idea of what are the facts and influences upon soil quality. If it had been too scientific [vs. economics], I probably would not have read it, because it would have taken too much for me to understand."

Actual Quality

The actual quality is judged by reading the whole document. (Note: at the document selecting stage users can only guess the expected quality by journal, author, or document type.)

"This one I thought I might be able to use ... It didn't at the end come up with anything. It is just has more discussion of the modeling issues ..."

Classic/Founder

The document is recognized in the field as the first substantial work on a topic or technique.

"This is the classic on water productions and there really hasn't been much else."

"It's a famous paper that's been around for a long time. And people enhanced their technique, but that's still an important paper. ... I should, have to, cite the founders of the technique."

Well-known/Standard Reference

The concepts in the document are well-known to the field; or the document is used as text book.

"In general, they are well known enough. They don't need to be cited or, if there is something that needs to be cited, there is a more specific source."

Prolific Author

When an author wrote many documents on a topic, users may take this situation into account when reading or citing.

"I didn't cite this. I did cite several papers that X wrote. ... He has written twenty or thirty of this thing. ... There may be something new. ... It does make you feel every single item is less valuable. This is readable and broad."

Judge

Sense of the person who will read or approve the finished product. This person can be a funding agent, journal referee, advisor for a prospectus or a dissertation, etc.

"If I rank it intellectually, it will not rank very high. But, if I think strategically, it's important to include it, because people on evaluation committee likely read and contribute to it and possibly edit this journal."

"I don't particularly think he [advisor] will appreciate my citing a paper [in a regional journal], especially if I can find papers in more national or leading journals."

Norm

Perceived expectation or practice in the field for the finished product. For example, in a funding proposal, specifics are rarely necessary; in a dissertation, a complete literature review is expected.

"This prospectus is short anyway, and I am trying to keep cites down to some degree. I didn't feel the need at this point to go beyond the national journals."

Credential

Usually refers to paper authored by user and is included primarily to support his own appearance of expertise.

"If he weren't a co-PI, we might not put the standard reference in. ... We want to show that we have the expertise in production modelling more generally."

Decision Rules

Table 4.2 shows that all rules were used in each task and no new rules were added. That no new rules appear is not surprising. At this level of abstraction, decision rules have been identified in studies of a wide range of choice behavior, not simply document-use behavior. It is unrealistic to expect that document use is so unique that decision makers would employ a new rule in connection with it that has not been applied before to other choice activities.

Wang's original model identified and defined the six decision-making rules, which are illustrated with examples from the 1995 study:

Chain rule

A user wants to get documents which are linked together through a range of relationships, *e.g.*, all articles in a special issue of a journal, or papers citing the same article.

"The reason why we cited the second paper, and it's not a paper, it's a book, is because he collected a number of studies that had done empirically what he suggested theoretically. We thought it was interesting to look at. The first one is a classic, the second one is a collection of readings, and there's a number of people that contributed things that followed up on his work from ten years before. I mean this paper by X revolutionized the way of looking at inequality measures."

Dominance rule

In a set of similar documents, the user selects one or more that excel in at least one aspect but are no worse than other documents in the other aspects

"... the reason why I used their book is that they have a book about water marketing theory and practice. I'm planning at the beginning to take, the area on the Middle East is really premature,

Table 4.2. Decision Rules Across Stages

	Early Stage	Later Stage	
Decision Rule	Selecting	Reading	Citing
Chain	✓	✓	✓
Dominance	✓	✓	✓
Elimination	✓	✓	✓
Multi-criteria	✓	✓	✓
Satisfice	✓	✓	✓
Scarcity	✓	✓	✓

dealing with water markets. That's why we need government intervention in the level of negotiation and cooperation. So I don't really talk much about water markets, but I just say why water markets will not be efficient at this point. But this is probably one of the best references because they do have in their book, they do talk about policy issues, about water markets, about the value of water, and not just about the theory of water markets, but also how it works in practice. So it's a good reference. But this is definitely not the only reference; there are many books about water markets. I chose it because it's theory and practice together, so it's a good reference. It's not just a case study, and it's not a book that just talks about the theory, it's a combination of both. And it's pretty recent."

Elimination rule

A users applies a salient criterion to reject a document.

"The first one I looked through, I was very disappointed in the level of economics."

Multiple-criteria rule

A user accepts or rejects a documents by applying more than one criterion, usually when confidence in the degree of salience of a criterion is low.

"... precisely on the subject. It had, I am not entirely sure now, at that point of the time, very up-to-date information as to what the legal standing of particular GATT, well, particular trade restrictions that the U.S. government had placed on Mexico -- what stage those decisions, however, what those decisions were. ..."

Satisfice rule

When the user considers that enough documents have been identified or used, he may terminate the selection or use process. Sometimes, if the search involves several facets or terms, he may stop looking for items about a particular facet.

"Just, this is another application of endogenous coalition formation where he was using it for another area. Now that I look at it, maybe I should not be citing it; maybe it's enough to cite X in my prospectus."

Scarcity rule

When the user wants more documents, but only a few are retrieved, he uses an inclusive approach, emphasizing getting those that are even tangentially relevant, instead of applying selection criteria specifically, which may exclude a few of the already scarce documents.

"... one of my chapters was on land allocation and this was one of the few papers which actually has been written on the subject in recent years, the last four or five years. So I wanted to see what they do."

Deducing or inferring decision rules from user reports is difficult for a variety of reasons. Special efforts were made in the 1995 study to identify these rules. Both researchers were present during each interview so that one researcher could focus on listening for evidence of decision rules and followup, as necessary, with probing questions. Asking the users explicitly was not possible since many could not easily relate their decision-making to applying rules.

Case Studies of Document Use by One Participant for One Item

Each of the three case studies presented in this section shows the decisions made over time by an individual participant about an individual document. Each case study represents the thinking of a different participant. The case studies generally indicate that decision rules and sometimes criteria varied over time as participants progressed through their projects.

Case 1

Figure 4.1 shows the decisions made by one participant about an article that appeared in the *American Journal of Agricultural Economics*. The search output the participant reacted to initially is boxed in the figure and includes bibliographic citation, abstract, and subject indexing. This participant continued to judge this item useful but changed the criteria and decision rules he applied at each stage. Initially he judged it relevant because it was on his subject (Topicality), and he expected the quality to be high, based on his knowledge of the author, "Adams is in Oregon State University. Very good author." He used the multi-criteria rule in selecting it.

In the later stage, the participant's comments were postdecisional. The participant had already read this item, related it to other literature he used or considered, and had decided to actually acknowledge it in his bibliography. Although the immediate stimulus in the interview was again the bibliographic reference, it is clear from the participant's comments that he was referring to his memory of the article itself and his activities in connection with it and others he considered similar. He was aware of the scarcity of material on the topic covered in this article which led him to read it: "I looked for other, for anything I could find on these topics ..." He was attracted to

Figure 4.1. Judgments Related to One Document (Case 1)

Bibliographic Citation From Original Search

Dialog/File 139/262461
TITLE: The On-Farm Costs of Reducing Groundwater Pollution
AUTHOR(S): Johnson, Scott L.; Adams, Richard M.; Perry, Gregory M.
AUTHOR(S) AFFILIATION: S Coast Air Quality Management District; OR State U; OR State U
JOURNAL NAME: American Journal of Agricultural Economics,
JOURNAL VOLUME & ISSUE: 73 4,
PAGES: 1063-73
PUBLICATION DATE: November 1991
DOCUMENT TYPE: Journal Article
ABSTRACT INDICATOR: Abstract
ABSTRACT: Agricultural chemicals are a source of groundwater pollution in some areas. Regulatory options to reduce such nonpoint pollution imply costs to producers. By integrating plant simulation, hydrologic, and economic models of farm-level processes, this study evaluates on-farm costs of strategies to reduce nitrate groundwater pollution. The empirical focus on intensively managed, irrigated farms in the Columbia Basin of Oregon. Results suggest that changes in timing and application rates of nitrogen and water reduce nitrate pollution with little loss in profits. Once such practices are adopted, further reductions in nitrates can be achieved only at increasing costs to producers.
GEOGRAPHIC LOCATION DESCRIPTOR(S): U.S.
DESCRIPTOR(S) (1991 forward only): Renewable Resources and Conservation; Environmental Management: Water; Air (Q250)
DESCRIPTOR(S): Conservation and Pollution (7220); Natural Resources--General (7210)

Comments on Selection (Concurrent)

The first one, "On-Farm Costs of Reducing Groundwater Pollution" That title sounds exactly what I am looking for. ***(Topical/relevant)***
The title is very informative. And the authors are [reads silently], In fact, I know one of these people. Adams is in Oregon State University. Very good author. ***(Expected quality/high)***
That's interesting. That is in the *American Journal of Agricultural Economics*, which is a journal I read. I have missed this article. ... OK, the abstract says, "Agricultural chemicals are a source of [reads silently]" What I am interested in is what it would cost to engage in a kind of production, but not so much pollution. And this is about. ***(Topical/relevant)***
"... The empirical focus on intensively managed, irrigated farms in the Columbia Basin of Oregon. Results suggest that changes in timing and application rates of nitrogen and water reduce nitrate pollution with little loss in profits. ..." Eh, interesting. Ok, this is the very good one to have.
Selecting Criteria: Topicality; Expected Quality
Rules: Multi-criteria Rule

Figure 4.1. Judgments Related to One Document (Case 1) (Continued)

Comments on Reading and Citing (Postdecisional)

I think it didn't have specific results, but in this topic there is even less qualitative work that has been done. ***(Scarcity in literature)***

As I recall, I didn't cite anything qualitative from this study, which is on the cost of reducing ground water pollution. I don't remember they had anything qualitative to report. ***(Orientation/not qualitative)***

But, again this is the best thing I could find to cite about the work that has been done. ***(Dominance)***

... I looked for other, for anything I could find on these topics and there is a certain amount of discussion by interest groups and other people. They were just sort of stating their side of the issue without really providing much analysis. This is more analytical. ***(Orientation/analytical)***

Well, this analytical stuff generally offers arguments that look better. ... This paper wasn't so much about the political arguments and was more about consequences. So, for this purpose, better. ***(Dominance)***

Reading Criteria: Topicality; Orientation
Rules: Scarcity Rule

Citing Criteria: Topicality; Orientation
Rules: Dominance Rule

Note: The item within the box is the output from the original search for one item, as provided by the information system. It was the stimulus for the subsequent comments, reflecting the decision-making of the participant. Criteria and decision rules are bolded and italicized in parentheses within the text. The latter are summarized, along with the decision-making rule, after each set of comments.

the article both for reading and citing because it addressed a topic he considered important for his paper (Topicality) and because it was "more analytical" (Orientation/Analytical). In commenting on other articles similar to this, he said "there is a certain amount of discussion by interest groups and other people. They were just sort of stating their side of the issue without really providing much analysis." The dominance rule is operative in citing: "But, again this is the best thing I could find to cite about the work that has been done."

Case 2

In Figure 4.2 the participant considers an article from *Metroeconomica*. The search output in this case consisted of a citation and descriptors but no abstract. The participant determined its usefulness initially by looking at the descriptors and title, but he clearly missed having the abstract,"I just want to see how to ... try to see if there is an abstract." In judging it useful from the search output in 1992, he consider it to be on his subject. Although he could not infer microeconomic theory from the title, which concept he seemed to use as a measure of subject correspondence, he could determine it from the descriptors, "Based on the descriptors, 'theory of production, economics of uncertainty ...,' I would be interested in looking at it." He

Figure 4.2. Judgments Related to One Document (Case 2)

Bibliographic Citation From Original Search

Dialog/File 139/167491
TITLE: Downside Risk and the Competitive Firm
AUTHOR(S): Honda, Yuzo
JOURNAL NAME: Metroeconomica, JOURNAL VOLUME & ISSUE: 37 2,
PAGES: 231-40
PUBLICATION DATE: June 1985
DESCRIPTOR(S): Microeconomic Theory--theory of production (0223); economics of uncertainty and information--Theory of Uncertainty and Information (0261);

Comments on Selection (Concurrent)
"Downside Risk and the Competitive Firm," Honda, Yuzo, *Metroeconomica*. I haven't heard Microeconomic Theory. There is no abstract. Based on the descriptors, "... theory of production; economics of uncertainty ..." I would be interested in looking at it. ***(Topicality/match-close)***
It seems that would be a theoretical paper relates the concept of downside risks to behavior of competitive firms. I just want to see how to... try to see if there is an abstract. The first one I will try to look at is this one. I haven't seen anyone who looked at it that formally. In general, the first thing I will look at will be the ones that have to do with downside risk --this one that is theoretical and behavioral. ***(Orientation/theoretical-conceptual)***

Selecting Criteria: Topicality; Orientation
Rules: Multi-criteria Rule

Comments on Reading and Citing (Postdecisional)
#2. Honda, in *Metroeconomica*, was essential in terms of understanding what was going on with the model. From what I read in his paper, he had a specific output and he did all the comparative statistics associated to a model very similar to the one I was applying, with the difference that my model generalized his but had a lot of extra terms. It gave me a lot of intuition and also backup on some of the results I was trying to show. ***(Topicality/match-close; Orientation/theoretical-conceptual)***

Reading Criteria: Topicality; Orientation
Rules: Multi-criteria Rule

Citing Criteria: Topicality; Orientation
Rules: Multi-criteria Rule

Note: The item within the box is the output from the original search for one item, as provided by the information system. It was the stimulus for the subsequent comments, reflecting the decision-making of the participant. Criteria and decision rules are bolded and italicized in parentheses within the text. The latter are summarized, along with the decision-making rule, after each set of comments.

was also influenced by its theoretical orientation: "The first one I will try to look at is this one. I haven't seen anyone who looked at it that formally." He used the multi-criteria rule in selecting it.

In the later stage, the participant's comments show how well he remembered the *Metroeconomica* article. He could distinguish Honda's model from his own very clearly and considered the paper important, "... essential in terms of understanding what was going on with the model." Again, the participant had already read this item, related it to other literature he used or considered, and had decided to actually acknowledge it in his bibliography. His reasoning did not change through the process, but the level of detail he provided for his reasoning was much greater. The multi-criteria rule was evident both for citing and reading.

Case 3

In Figure 4.3, the participant considered an article in the *Review of Economic Studies.* He changed his judgment about its usefulness in later stages. Originally he considered it relevant because it was on his subject (Topicality), and he expected the quality to be high, based, as in the first case, on his knowledge of the authors, "I know the kinds of paper these guys write. They are pretty good ..." He also anticipated the orientation of the article, "... they focus a lot on information issues ..." He used the multi-criteria rule in selecting it.

In 1995, the participant indicated that he had aquired the article and skimmed it but had not used it substantively in his work. The article was sufficiently related to merit his consideration because it was "more on the level of analysis that I was doing," and had the word "dynamic" in its title. But he noted that the authors' interpretation of "dynamic" was not the same as his. He took a long term view of that concept, whereas the authors took a short-term view. He indicated that, once he determined that, he eliminated the article from consideration (Elimination rule): "I don't recall that I read this deeply enough to see whether there was possibly something else in it that I could have used."

These case studies show that criteria and decision-making do not necessarily remain constant over time for the same item. Often in later stages, when the user has actually looked at the full text of the document or used it significantly in his work, his judgments are based on a more detailed perception of the document and, at times, far more refined criteria.

Figure 4.3. Judgments Related to One Document (Case 3)

Bibliographic Citation From Original Search

Dialog/File 139/205782
TITLE: Commitment and Fairness in a Dynamic Regulatory Relationship
AUTHOR(S): Baron, David P.; Besanko, David
JOURNAL NAME: Review of Economic Studies, JOURNAL VOLUME & ISSUE: 54 3,
PAGES: 413-36
PUBLICATION DATE: July 1987
DOCUMENT TYPE: Journal Article
ABSTRACT INDICATOR: Abstract
ABSTRACT: This paper considers a multiperiod model of a regulated firm that has (stationary) private information, which may be revealed through performance. A "Fairness" arrangement is proposed in which the firm agrees not to quit if in future periods the regulator allows it to earn a nonnegative profit given the type it revealed in earlier periods. The properties of such arrangements are studied, and an example is presented in which both the firm and the regulator prefer a fairness arrangement to a policy feasible without commitment.
DESCRIPTOR(S): Economics of Regulation (6190) ; Welfare Theory--Allocative Efficiency including Theory of Cost-Benefit (0242) Cost/Benefit; Theory of Uncertainty and Information (0261) ; Microeconomics--Theory of Production (0223)

Comments on Selection (Concurrent)

The next item is called "Commitment and Fairness in a Dynamic Regulatory Relationship" by David Baron and David Besanko in *Review of Economic Studies*. Both of these authors write a lot on regulation and dynamic regulation. I haven't read all their papers. OK, but let's see. "This paper considers a multiperiod model of a regulated firm that has (stationary) private information, which may be revealed through performance. A fairness arrangement is proposed in which the firm agrees not to quit if in future periods the regulator allows it to earn a nonnegative profit given the type it revealed in earlier periods. ..." This is close. ***(Topicality/match-close)***
However, I know the kinds of paper these guys write. ***(Expected quality/good)***
They are pretty good, but they focus a lot on information issues and on the property of that equilibrium. ***(Orientation/factual-descriptive)***
That I guess, I am interested in equilibrium, although the equilibrium concept is different from what they usually use. Private information is just a bit too much problem for what I want to work out. It makes a problem a lot harder. ***(Topicality/facet not match)***
That's another one I might want to look at. That means this actually looks very good but for a related thing I am working on.

Selecting Criteria: Topicality; Expected Quality; Orientation
Rules: Multi-criteria Rule

Figure 4.3. Judgments Related to One Document (Case 3) (Continued)

Comments on Reading and Citing (Postdecisional)
For number 2, number 2 is more the level of analysis that I was doing. But, despite the fact that it has the word dynamic in the title, I know from the authors that their work, even they use the word "dynamic," it's not truly dynamic. The paper I was interested in was something truly dynamic. Dynamic for economist means lots of time periods. So, you look over a very long time. Some people when they say dynamic mean today and tomorrow. So, dynamic in my paper as I can see is very long time. I know these authors when they say dynamic they mean a short time. I don't recall that I read this deeply enough to see whether there was possibly something else in it that I could have used. ***(Topicality/concept-interpreted-different)***

Reading Criteria: Topicality
Rules: EliminationRule

Note: The item within the box is the output from the original search for one item, as provided by the information system. It was the stimulus for the subsequent comments, reflecting the decision-making of the participant. Criteria and decision rules are bolded and italicized in parentheses within the text. The latter are summarized, along with the decision-making rule, after each set of comments.

5. CITING BEHAVIOR

In this chapter, the focus is on a particular type of document use behavior--acknowledging the contribution of another's work to a current project. Understanding this behavior began by coding the participants' comments with the criteria developed in the 1992 study and the additions noted in the task-based analysis presented in Chapter 4. During data analysis, some ambiguities in the original criteria categories were clarified, and the criteria were refined (See Appendix C, second part), and the data recoded. In addition, the analysis identifies and characterizes two other types of variables: contributions and metalevel documentation concerns. For criteria and contributions, numbers are presented to indicate the frequency of occurrence of the criteria and contributions and the extent of use of these variables across participants.

Background and Research Questions

This research differs methodologically from most other research on citing behavior in two ways. First, research clarifying citation practices or motivation is often document-based, *i.e.*, it deduces researchers' behavior by analyzing the content or context of the citation (For example, Chubin and Moitra 1975; Moravcsik and Murugesan 1978; Cano 1989; Small 1982). Underlying this approach are critical assumptions that the researchers' behavior is accurately and completely revealed in the content or context of the citation and that the citations themselves reveal all or most of the document use during a research project. These assumptions themselves need to be tested. In this study, the participants themselves were questioned directly about their decisions. Several other studies have used this citer-oriented approach but have generally provided the citers with a checklist of possible reasons (Hodges 1978; Brooks 1985; Brooks 1986; Vinkler 1987; Snyder and Bonzi 1989; Bonzi and Snyder 1991; Liu 1993).[5] In this study, the questions were focused, but no pre-set categories were provided; the participants formulated their own reasons. This study is qualitative, based on interviews, exploring reasons researchers express for citing documents.

Another significant difference stems from the fact that most studies of citing motivation look only at positive decisions, not at decisions resulting in *not citing* materials that were actually used in the course of a research project. Because this analysis of citing behavior is based on a larger study of document use during research projects, both types of decisions can be studied.

---

[5]Although Bonzi and Snyder provided their respondents with pre-detetermined categories, they based their typology of citing motivations on typologies in other studies and interviews with some researchers to elicit other reasons and to check the reliability of their typology (Snyder and Bonzi 1989; Bonzi and Snyder 1991).

Indeed, the emphasis in this paper is on comparing and characterizing the behavior resulting in both positive and negative decisions. Although MacRoberts and MacRoberts suggested the usefulness of this approach in 1988, only Vinkler and MacRoberts and MacRoberts have considered motivations for not citing (Vinkler 1987; MacRoberts and MacRoberts 1988).

The research questions originally stated in the Chapter 1 were expanded or specified to include the following in studying citation behavior:

1. Are all materials that are used during the course of a research project actually acknowledged in the documentation phase?

2. What reasons do the participants acknowledge in making citing decisions?

3. If all materials are not cited, what are the reasons for acknowledging some uses and not others?

4. What broad rules or concerns, if any, play an overarching role in making citation decisions?

In the criteria analyses, the unit is usually the criterion coded broadly. A total of 616 criteria were applied during the 314 document use decisions. The broad criteria were also coded specifically, as indicated through selective examples in Table 5.4. In a document decision unit, a participant may use one or more criteria. Definitions of the individual criteria are included in the appropriate section.

Model of Citing Behavior

The participants seemed to consider several different types of factors in making decisions about citing. The following model is generalized from the full range of participant comments and establishes a framework for the subsequent analyses. The capitalized, bolded items are those addressed in this chapter:

> A researcher identifies potentially useful documents from a broad range of sources. She follows up on some of these items by obtaining copies of them and consulting or reading them and using them during the research project. At the end of the project, the researcher formally acknowledges for archival and peer review purposes some of the documents she consulted based on her judgments about the ***CONTRIBUTION(S)*** they made to the paper. In making these decisions, besides topical relevance, the researcher applies one or more ***CRITERIA*** to each document as she chooses which to cite. She may also consider ***METALEVEL DOCUMENTATION CONCERNS*** which directly or indirectly affect her decisions about individual

documents. In making these decisions, the researcher applies one or more decision rules (identified in Chapter 4).

Contributions of Documents

Only the items actually cited were analyzed for contributions. With a few exceptions, the contributions closely parallel the major tasks in a research project and the structure of a paper documenting that research (See Table 5.1). The contributions are defined and illustrated with quotations from the interviews as follows:

Analogies/Contrasts/Comparisons

Concepts in the documents served as an analogy or otherwise provided a basis for comparison and contrast with the participant's research.

"My book looks at the interaction of modernization, modernized agriculture and environment. ... I have to say something about what's happening with the traditional agriculture. ... This [document] is a good thing to cite."

Corroboration

Concepts, arguments, or results in the documents supported the participant's discussion or findings.

"And there are a couple of points in my prospectus where I talk about that issue and how that, how measures of productivity, even conventional, even measures that are out there right now sort of, to some degree, indirectly reflect soil quality. And this [paper] is, papers by X and others are, provide some evidence that deal, at least indirectly reflect some soil quality reflected in productivity measures. But only indirectly."

Data

Actual data in the original documents were used. The data were not always statistical data.

"All these [reports from state agricultural agencies] were key data sources for me."

Identification of originator

The person or article was associated with the first use of that concept, theory, or analytical approach. This contribution differs from the Tangential/Ceremonial motivation in that the participant considered this type of citation important in documenting earlier attention to the acknowledged idea, in part to indicate the long-term importance of the concept, theory, or analytical approach he was using.

Table 5.1. Contributions of Cited Documents

Contribution	Number	Number of Participants Commenting
Analogies, contrasts, comparisons	15	5
Corroboration	14	6
Data, example, case	33	8
Identification of originator	9	6
Justification	33	9
Methodology	29	8
Tangential, ceremonial	6	4
Theory, concept, definition, argument	61	12
Verification	13	5

Notes:
Number refers to the document decisions units (cited documents only) in which that criteria was invoked.
Number of Participants Commenting refers to those mentioning it at least once. N of participants = 12.

"I cited this particular book because this guy, [author], was the originator of this asset fixity in overproduction trap idea back in the '50s and, if you're going to talk about that particular issue, you have to cite this guy. I cited this book instead of his original article because in this book he lays out his model in mathematical detail which he didn't originally."

Justification

The document suggested or justified an idea. This reasoning was often an indication that the document was the stimulus for the research.

"I often don't read it very carefully. I just say, OK, here is the sentence from the book that justifies what I am going to do."

Methodology

The participants applied the methodology presented in the documents, sometimes with modifications.

"This one is important for the tests that this paper performed. ... I did use one of the tests."

Tangential/Ceremonial

The participants felt the need to cite these papers, not because the papers directly contributed to their research, but because they felt a responsibility to alert readers to other literature, to consider the expectations of potential referees, or to acknowledge personal influence of the author.

"The reason we cited this particular paper wasn't because it contributed anything to our particular research. It was just tangentially related and it reflects a different approach to the construction of inequality indexes that we thought people should be made aware of. So it didn't do anything directly for our research."

"... we didn't think his [a potential referee's] work was all that valuable [to our paper] in some ways but at the same time we didn't want to alienate the person by not citing his work."

Theory, Concepts, Definitions

The document provided a theoretical framework, concepts, or definitions.

"Well, it gives the theory of external costs and benefits in agricultural terms. I believed at the time that I had to cite this chapter of this book because it provided a general framework from where to begin the analysis."

Verification

The participants had checked these documents to determine if something had or had not been done. This action was often done early in the project to verify that the idea had not already appeared in the literature.

"This [document] was the originating paper that set me off on the issue ... That's the one that gave me the basic results that I am trying to test as a hypothesis in the model."

Several researchers have established typologies of motivations for citing, usually based on content analysis (Chubin and Moitra 1975; Moravcsik and Murugesan 1978; Spiegel-Rosing 1977; Frost 1979; Peritz 1983; Cano 1989; Hooten 1991) or devised for direct questioning studies (Hodges 1978; Brooks 1985; Brooks 1986; Vinkler 1987; Snyder and Bonzi 1989; Bonzi and Snyder 1989). Among those soliciting information directly from the participant, only Hodges did not provide pre-established categories. Peritz's classification of citation roles is particularly

relevant because it is a typology of citation roles for social sciences empirical research and because it is based on Hodges' exploratory work. He tested it by applying it to citations in several social science fields. Finding some difficulties in applying Hodges' original categories within a content analytical perspective, Peritz established eight motivations for citing: Setting the stage; Background; Methodological, subdivided into Design and Method of analysis; Comparative; Argumental, Speculative, Hypothetical; Documentary; Historical; Casual (Peritz 1983, 48). His own typology more closely follows the typical structure of a research paper. He found a close association between his categories and a citation's location in an article.

That there is an underlying similarity between the typology devised from the participants' comments and Peritz's categories is not surprising. The participants are documenting task-related activity, and the basic tasks in the research addressed in both of these studies are quite similar. On the other hand, however, the qualitative data show some subtleties that are not apparent in Peritz's typology because only the participant himself could have known motivation at this level of detail and because the rationale was often "cleaned up" or simply not described at that level in the article itself. In the Justification category, for example, a cited document suggested or justified an idea. That stimulation notion is not obvious in Peritz's categories. In the Verification category, the participants indicated that some of their citations stemmed from a need to determine that something they wanted to do had not been done. That is, again, a slightly different and more subtle perspective than is apparent in Peritz's Setting the Stage category. He includes only the Comparative category, which he thinks subsumes Hodges' Corroborative category, and possibly her Oppositional and Corrective citations, whereas in this study, the participants seemed to distinguish between corroborative cites and cites that draw analogies or contrast with their work.

The category of citing to document a theoretical framework in this study is probably subsumed into either Peritz's Setting the Stage or Methodology categories. Economics, one of the more theory-rich social sciences, was not included in the social science disciplines in Peritz's study. He may not have seen enough references to documents for this purpose to suggest a more refined category. In this study, the category showed up across all subjects.[6]

The range of contributions of the cited papers noted by the respondents provide some support for Cronin's idea of tiered citation, *i.e.* that citers cite to different levels of a work (Cronin 1994). His tentative typology consists of five levels: oeuvre, motif, opus, chunk, and quantum, such as formula or phrase, method, result. In the contributions identified in this study, it is apparent in the illustrating quotations that the respondents were not always referring to the cited document in its entirety but instead to segments or aspects of it that they had found useful. This level of granularity is reflected in several of the contributions, *e.g.* data, methodology.

---

[6]Other criteria may also have been apparent across all participants, but some participants provided information about citations only on the basis of their sample of cited works, not on all documents cited.

Criteria Used in Decision-making

As noted earlier, a criterion is a filter a participant applies to a document to determine its usefulness to his information need. A range of 1 to 5 criteria were applied in a document decision unit, with an average of 1.9 (S.D. 1.0) per document decision unit. Document decisions resulting in citing averaged 2.1 (S.D. 1.0) criteria; those resulting in decisions not to cite averaged only 1.6 (S.D. .8).

The participants mentioned 28 criteria that they used in decision-making related to citing decisions (Table 5.2). Many of the criteria also appear in other phases of document use, although the same criteria were not always applied to the same document at each phase.[7] The criteria can be divided into three types: internal, self-related, and external. The *internal* are related to the document either holistically or to specific elements of the document, such as the author. *Self-related* criteria look at the document within the context of the reader's intellectual or physical capabilities. The *external* are related only tangentially to the document itself and more to how the participant's product will be received by journals or funding agencies, referees and other persons in formal roles as judges, and peers and colleagues. The emphasis shifts to the research product, even though the decision is still made about a specific document. With the self-related criteria, the important non-document frame of reference is the participant himself. With external criteria, frames of reference relating to the external environment of research production play a significant role in the citing decisions.

The distinction between internal and external, as noted above, is comparable to Vinkler's categories of professional and connectional motivations (Vinkler 1987). Cronin also mentions "extrinsic factors," which include several of those identified in this study (Cronin 1984, 31). For Vinkler, professional motivations refer primarily to the "relation of the citer to the object of his research;" connectional motivations refer to the "relation of the citer ... to persons involved in the research machinery and authors of articles to be cited" (Vinkler 1987, 50). "Authors" in Vinkler's terms are included in this study's internal criteria, regarding the author of a publication as an important bibliographic element in a document. The criterion Relation also considers authors in a relationship between a document's author and the participant. Relation may reflect judging a document's quality or paying homage. With the former interpretation, the criterion is internal; with the latter, it is external. It is shown in both columns in Table 5.2.

In analyzing the internal criteria, it is apparent, however, that elements in the bibliographic record or document trigger responses that are based in part on relating the document or element to several different frames of reference. These frames include the project itself and the project-related task or tasks the participant faces, either immediately or in the future; the subject literature; the discipline itself; and the document delivery system. But the most important frames of reference are essentially internal or based on the document or project.

---

[7]These categories modify preliminary results reported in Wang and White, 1996.

Table 5.2. Criteria Grouped by Type

Internal Criteria	Self-related Criteria	External Criteria
Actual quality	Availability	Credential
Audience	Cognitive requisite	Judge
Authority	Novelty	Norm
Classic	Time, effort	Relation *
Content		Target journal
Depth		
Discipline		
Expected quality		
Journal spectrum		
Orientation1		
Orientation2		
Peer reviewed		
Prolific author		
Publicity		
Recency		
Relation *		
Reputation		
Standard reference		
Topicality		

* See explanation in text for placing Relation in both groups.

The criteria previously identified in preliminary analyses have been modified for this paper. Many of these criteria appeared in connection with document use decisions earlier in the larger project on document use over time (See previous chapter), but several of the criteria also appear in other typologies of criteria users used in making relevance judgments (Schamber, 1991; Barry, 1993; Barry and Schamber, 1995) (See Appendix E for definitions and examples of all criteria).

The nine new criteria are defined and illustrated with quotations related to citing decisions:

Audience

The level at which the publication is written, usually indicated by reference to intended audience.

"The next one is actually in the *Economist*, which is a main sort of public magazine and again I need it to get quotes from the popular press that sort of say, this issue is important. This [situation] is what is happening in the real world. I am not really going to it for anything that is of substance to the paper."

Content

The nature of the materials included, *e.g.*, data, methodology, theory.

"The book by [authors] is just a classic mathematical reference on inequality, not necessarily on inequality measurement, but on inequality. And so a lot of the mathematical theory that is relevant to analyzing inequality would have been first studied by these three guys."

Depth

The extent of treatment about the subject within a document.

"The first one--I did look at that. That came from the Farm Bureau Federation. That was the one that, in the end, was too tendentious to use. It didn't really have any analytical information, and it had skeptical ideas about pesticide policy. And, they may be right, but they just didn't have any scientific contribution."

Journal spectrum

An assessment of the publishing journal, usually in terms of major or minor.

"The next one is by [authors]. This [article] is a short journal article in one of the main economic journals, *Review of Economic Studies*."

Orientation1

A distinction between an emphasis on a theoretical or empirical presentation.

"I did look at this [document] by [author], but it is an empirical piece that looks at the role that international inequality basically has in development. It's not relevant to what we were working on."

Orientation2

A methodological distinction between a qualitative or quantitative approach.

"This one I used. I actually use some of their qualitative estimates from this one. They looked at a scenario where you would take something like 25% reduction of use of all fertilizers or chemicals..."

Peer-reviewed

Evidence of critical evaluation by a knowledgeable judge, *e.g.*, an article published in a refereed journal. A document cited by another researcher is also considered peer-reviewed.

"The journal paper is much more credible. Conference papers, everybody gets papers at a conference, the quality of a conference paper is usually much lower than journal papers. In the conference paper you are limited to 10 pages and, if you haven't been to the conference and sat through the presentation, you don't know what objections or worries others might face. And

usually [the journal article's] a much more expanded version and a lot more carefully done and refereed a lot more vigorously. In a conference, you can't sit there and do the math or the algebra, but you know in a refereed journal somebody hopefully has looked through it and accepted it."

Publicity

Indication that the publication has received some extraordinary acknowledgment.

"X is definitely a book that is used in many theory classes. It's a very famous book."

Target journal

Consideration of the characteristics, level, demands of the journal in which the research product will be published.

"We cited the *Journal of Economic Theory* because that was the level at which we were trying to write the paper. In economics there are all different kinds of levels of journals, and the theoretical level that we were aiming at is most closely matched by the *Journal of Economic Theory*, *Review of Economic Studies*, and *Econometrica*. The paper that we actually wrote was ultimately submitted to *Econometrica*. So, when we picked out references, we tried to stay in that group. It is a little bit of gamesmanship in a way, to be citing the right people."

As the analysis for Table 3.1 shows, the respondents did not cite every item they acknowledged using during the project. If all materials are not cited, what are the reasons for acknowledging some uses and not others? As Table 5.3 indicates, at the broad level of coding criteria, many of the same criteria are apparent in both positive and negative decisions. However, in making a negative decision, *i.e.*, in deciding not to cite, the participants did not comment on Classic, Credential, Cognitive Requisite, Judge, Orientation, and Standard Reference. In the first two, the decisions are inherently uni-directional toward citing, *e.g.*, if the document is judged a classic, it is cited. A document's not being a classic does not prohibit being cited but usually is just not mentioned. Credential as a criterion says that a participant cites to prove his authority; he does not "not cite" to do that. The other four are more flexible but were only used to support actual decisions to cite in this study.

In deciding to cite, the participants made no reference to Availability; Time, Effort; Depth; and Anticipated Quality. Time, Effort and Availability are clearly related. Presumably the items were already physically accessible to the participant if he had used them, so the added cost and effort of citing were minimal. Anticipated Quality was not a suitable criterion once a participant could judge actual quality. A judgment about Depth could have resulted in both negative and positive decisions but resulted only in judgments about a document's deficiency in this study, thus resulting in negative decisions.

These findings actually support the generalizability of the criteria, showing that, by applying them, the participant can reach either a positive or a negative decision, depending on the

Table 5.3. Criteria Applied by the Nature of the Citing Decision

	Uncited Documents		Cited Documents	
Criteria	Number	Number of Participants Commenting	Number	Number of Participants Commenting
Actual quality	5	4	7	5
Audience	3	2	1	1
Authority	3	3	8	5
Availability	3	2		
Classic/founder			37	8
Cognitive requisite			2	1
Content	23	8	97	12
Credential			1	1
Depth	5	4		
Discipline	1	1	6	5
Expected quality	2	2		
Journal spectrum	3	1	7	4
Judge			7	5
Norm	5	1	2	1
Novelty	1	1	1	1
Orientation1	8	2	18	10
Orientation2			2	1
Peer reviewed	1	1	4	2
Prolific author	3	1	1	1
Publicity	1	1	16	5
Recency	2	2	7	5
Relation	6	4	19	5
Reputation	3	2	9	6
Standard reference			16	7
Target journal	1	1	2	2
Time, effort	2	2		
Topicality	95	12	143	12
Total	183		433	

Notes:
Number refers to the number of document decision units in which that criteria was invoked. N=314.
Number of Participants Commenting refers to the number of participants who mentioned the criterion at least once. N of participants = 12.
N = 616 criteria applied within the 314 decision units in this study.

frame(s) of reference he employs. The decisions were also coded more specifically and, at the specific level, it is possible to see more of the reasoning behind the decision and to identify values more directly related to the direction of the decision. A sample of the criteria with more specific coding is given in Table 5.4 to provide some insight into the application of these criteria to specific items.

The reputation of the journal, organization, or person, *i.e.*, the author, was mentioned in relatively few document decisions; however, when it was judged high, the items were usually cited. Similarly documents judged current or older were usually cited, not those that were "too old." In terms of the audience, in three decisions, an orientation to a popular audience was considered a negative, but one participant chose to cite an article written for popular consumption. On the other hand, empirical research was not always regarded as a positive factor, although most participants mentioning it indicated it contributed to their decision to cite. Although topicality was a criterion in 238 document use decisions, documents that matched all facets, were close matches, or were considered related were cited whereas documents in which the subject did not match to some degree, or the subject was unclear were usually not cited.

Metalevel Documentation Concerns

In comments about individual items the participants sometimes alluded to or explained metalevel-level documentation concerns. For some categories of material or types of authors, for example, they seem to follow a basic rule, almost independent of the extent or importance of use of a particular item in the immediate project. It should be noted that, although the evidence of these concerns exists sometimes across several participants, their decisions stemming from these concerns may have varied. The metalevel documentation concerns, then, are not the equivalent of a fixed code of citing but rather reflect a general set of guidelines that are applied flexibly to specific items and/or in specific situations as needed.

In the interviews, the participants were usually not asked to generalize over their citing behavior but to discuss document-specific decisions as a technique for getting more specific and valid data. Some metalevel comments arose in reference to specific items. In addition, this section is based in part on spontaneous, summarizing comments about decision rules or responses to probing questions asked when patterns or general decision rules were evident.

The emphasis in this section is on identifying the metalevel documentation concerns, *i.e.*, identifying the variables. Numbers are not given because they would be misleading since questions about these concerns were not asked systematically across participants. Nevertheless, the concerns that are identified are intriguing and suggest additional research. The findings presented in this section should be regarded as very tentative.

Table 5.5 summarizes the object or concern of the participant and paraphrases that concern as an overarching question. The first set of concerns/questions (Table 5.5, Nos. 1, 2) is very general and suggest a philosophical orientation. The second set (Table 5.5, Nos. 3, 4) relates

Table 5.4. Specific Coding for Selected Criteria

Criteria	Uncited Documents		Cited Documents	
	Number	Number of Participants Commenting	Number	Number of Participants Commenting
Audience/Popular	3	2	1	1
Journal spectrum/Major	2	1	2	1
Journal spectrum/Minor	1	1		
Journal spectrum/First class			4	2
Journal spectrum/Noneconomics			1	1
Orientation1/Empirical	8	2	18	10
Recency/Current			6	4
Recency/Older			1	1
Recency/Too old	2	2		
Reputation/Journal/High	1	1	5	3
Reputation/Journal/Low	1	1	1	1
Reputation/Organization/High			3	2
Reputation/Person/Low	1	1		
Topicality/Facet match	6	5	54	9
Topicality/Close match	5	4	58	22
Topicality/Related	6	4	15	8
Topicality/Analogy	1	1	1	1
Topicality/Context, bibliography	14	6	15	7
Topicality/Concept difference	1	1		
Topicality/Facet changed	7	4		
Topicality/Other interest	2	2		
Topicality/Not relevant	13	4		
Topicality/Facet not match	21	5		
Topicality/Too specific	8	4		
Topicality/Too general	8	5		
Topicality/Not clear	1	2		
Topicality/Other task	3	2		

Notes:
Number refers to the number of document decision units in which the criterion was invoked. Total N=314, not all of which are represented in this table.
Number of Participants Commenting refers to the number of participants who mentioned the criterion at least once. N of participants = 12.

Table 5.5. Metalevel Documentation Concerns

No.	Concern	Translation into Question
1.	Completeness	How complete do I need or want to be in citing relevant literature?
2.	Researcher vs. Researcher/teacher role	What do I see as my role as a citer: only documenting the research or documenting research plus teaching others about this broad area of research?
3.	Secondary citing	Do I have to see the original to cite it?
4.	Self-citing	Should I cite myself?
5.	Judges' expectations	What are the referees/outside judges going to expect?
6.	Synthesizing sources	Should I cite sources that contain information that is already synthesized and may be considered common knowledge?
7.	Journals: Coreness	How should I weigh core and non-core journals in citing articles?

very specifically to the act of citing. The third (Table 5.5, Number 5, possibly 4) indicates the influence of outside judges; the fourth (Table 5.5, Nos. 6, 7) indicates a sensitivity to typologies grounded in the nature of the literature. More complete explanations with illustrations from the interviews follow.

Importance of completeness

Participants occasionally expressed a willingness to risk missing some items, although they also seemed to apply different standards for completeness based on the journal's importance.

"For example, if I have done all of this [cited all of these articles] and I got to the end and someone said 'oh, wow, there is an article in blah, blah journal that's very close to what you did,' the chance is that it would be in one of the journals that really isn't important to this field ... not visible enough."

Researcher vs. Researcher/teacher role

Some participants document broadly, leading to related materials, thus assuming a researcher/teacher role, even as an author of research publications; whereas others in the researcher role limit their citations only to materials playing a substantial role in their own work.

"The rules are laid out by the rules of research, and you reference things as they are directly related to research, or that establish results that you are going to build on or that establish results closely related to yours ..."

Willingness to do "secondary citing"

Indirectly some participants acknowledged the possibility that they or others may cite without actually seeing a paper by simply drawing on references to the cited article in another's paper. None acknowledged the possibility of the previous paper's inaccuracy, limitation, or bias in representing the cited paper.

"I wouldn't have cited it unless I'd seen it, or if one of us [co-authors] had seen it. For example, I've never seen the X's one, but he [co-author] assures me that he's seen it."

Appropriateness of self-citing

Not many documents authored by the subjects were included in the original search or in the sample of cited, personally-located items discussed in the interview. In the few instances, the participants indicated they occasionally cited themselves to provide evidence of their own expertise, especially when the product is a grant proposal. Such citing constitutes redundancy in the proposal since typically a proposal contains the resumes of the project personnel. It may not appear for other types of publications. In a proposal, a funding agent needs to judge the capability of the authors to carry out the work successfully, whereas in an article, the work has already been done, and the referee or reader can draw inferences about the author's expertise from the quality of the research.

"If he weren't a co-PI [principal investigator], we might not have put the standard reference [textbook] in. ... We want to show that we have the expertise in production modeling more generally. ... We have to establish our credentials through them [self citations]."

Expectations of referees and/or editors

Participants were aware of referees and editors as filters for publication and modified their citing according to their perceptions of gatekeeper expectations. Sometimes they even speculated that a particular person would referee and cited his articles accordingly even if they had not actually used them during the research.

"We didn't want to be told we had neglected to cite certain people. So there are people in here, for example, X is one of these people we anticipated being a referee ..."

The appropriateness of citing textbooks/standard references

Some felt that citing a standard source allowed them to avoid developing an explanation of a concept but still provided a reference point for any reader that needed that explanation. They seemed to be trading off others' perceiving them as naive in not recognizing that something was common knowledge and a responsibility to the readership for directing them to related material that was important to the exposition, but not the paper's major contribution.

"So, I had to so some education of myself to get myself in a position to write the paper. ... You don't cite your textbook, you don't cite the stuff that other people have already known about."

Relative importance of articles in peripheral or non-core journals

Several participants alluded to levels of journals and selected citations in part based on whether the journals in which they appeared matched their own publication expectations. They may have read articles in non-core journals but cited them only if the findings were not available in a higher-ranked source. Relating journal of publication to quality was relatively common, with the expectation that the referees/editors of some journals were more rigorous in identifying articles making new, substantive contributions to the field.

"If we were trying to write at the level of the journals we're talking about, which are usually perceived as being the three best theoretical journals, it would be a mistake to be citing something like this [an item from a peripheral journal] unless it really had a point to make that nobody else had made. You don't cite much lower journals unless the journal contains something that was overlooked by everybody else."

References to some of these concerns are scattered in the literature. Prabha's study of document use and citing by business administration faculty supports the secondary citing question (Prabha 1983). He found that 96 percent of the documents the faculty cited in a recent paper had been consulted before citing, 63 percent especially for the paper. According to him, "the practice of citing 'second-hand' may be less prevalent than previously suggested (Prabha 1983, 203). His 'second-hand' is comparable to the secondary citing mentioned in this section.

His study also asked researchers about other works they had used, but not cited, during their research. Fifteen of the 19 respondents had used from 1 to over 21 documents but not included them in their bibliographies. He did not solicit rationales for their omission. MacRoberts and MacRoberts found that authors cite only about a third of their influences (MacRoberts and MacRoberts 1986; MacRoberts and Mac Roberts 1988). Wilson does not address citing but he indicates two reasons for researchers' not using relevant information: information overload and deliberate use based on policy or principle (Wilson 1995). Both of these may be operative among the participants in this study. In Vinkler's study, the most frequently mentioned reasons for not citing a relevant work were, in order: the paper was not important enough; the author or work is publicly known, *i.e.*, very well known; at the time of the work, the paper was unknown (Vinkler 1987, 59-60). The findings of other researchers and the comments related to completeness identified in this paper indicate that completeness is not a critical concern among researchers. The participants in this study did seem to indicate some relationship between completeness and coverage of literature in core journals. Completeness for core material was important, but they were less concerned about missing articles in non-core journals.

Bonzi studied both the incidence and reasons for self-citing in several disciplines, including economics. She found the reasons for self-citing were comparable usually to reasons for citing other work, but her respondents also mentioned "Establish writer's authority in the field," which

was the basis for the self-citing concern found in this research. This reason accounted for relatively few citations among the self-citations in her study (Bonzi and Snyder 1991, 248).

Conclusions

As Harter points out in his study of psychological relevance, studies of citation behavior have emphasized the purposes for citing and have not looked at the reasons the item was found relevant initially. His definition of citing as "a statement of an historical relevance relationship, captured for all time in a published article" clearly aligns relevance or usefulness decisions and citing decisions (Harter 1992, 614). The analysis in this paper shows that citing behavior is complex, multidimensional behavior. It involves an interweaving of judgments about documents' contributions to the research project, criteria applied to the documents which are comparable to criteria found in studies of relevance judgments at an earlier stage of document use, and metalevel documentation concerns.

The findings buttress concerns already raised in the literature about the normative theory of citing and the use of citations as a basis for evaluating scholars, journals, and academic departments or for showing previous influences on scholarship (Cronin 1984; Brooks 1985; MacRoberts and MacRoberts 1988; Liu 1993). One citation may represent several different motivations. Citing, as done by the economists in this study, actually underrepresents their total use of literature and seems to reflect a change from private, personal literature use to archival, peer review, literature guide considerations. The underrepresentation is not based solely on not citing poor quality, irrelevant, or tangential documents, but there may be a systematic bias to excluding some types of materials.

6. SUMMARY AND CONCLUSIONS

Summary

This study is one of relatively few studies to look at relevance judgments over time for the same set of documents by the same set of researchers. As the quotations in the study illustrate, the qualitative methods used in this study yielded rich data, allowing a researcher to see multiple motivations intermingling in document use decisions.

In addition, this study is one of only a few to look at the reasons for citing and for not citing documents known to be used during a research project and one of few citation behavior studies to interview researchers about citing behavior without specifying specific citation motivations for the respondents. Citation studies based on content or context analysis at best provide only incomplete pictures of citing behavior. Researchers divulge far more information in their personal comments than in their public documents. In fact, as this research shows, their public documents are sometimes constrained by concerns about how the research and documentation will be perceived by their peers.

The study points out that although, at the broad level, similar criteria are applied, the researchers clearly differentiate among documents and make use and citing decisions on that basis, applying one of six decision rules. Certain values of those criteria, as shown in the specific coding segment in Chapter 5, are more likely to result in negative or positive decisions. Even then, the specific situation may make a document judged inappropriate in one situation suitable for another.

To summarize briefly, the major findings of this study are:

- From their original searches, the participants identified as relevant more documents than they actually used. They also used more documents than they actually cited. During the research process, the original search output was winnowed, but other documents also entered into the pool of relevant documents, identified by other means.

- A wide range of criteria was identified which were used in deciding which documents to use. Some of these criteria have also been identified in other studies related to relevance decision-making or citing behavior.

- The participants mentioned an average of about two criteria in connection with citing decisions; this statistic is not yet available for reading/not reading decisions. But this statistic should be interpreted cautiously because of problems with postdecisional data.

- In applying the criteria, the participants made different value judgments; the document use decision could then be negative or positive. In so doing, they seemed to draw on frames of reference which allowed them to put the document into a perspective that suited their purpose. Some of these frames of reference were the research project, the immediate or future research tasks, the literature of the field, the discipline itself, document delivery systems, external judges, and self, *i.e.*, the individual's intellectual capabilities and motivation. These frames of reference are only deduced in this study.

- Many of the same criteria were applied at each stage of the process, but others appeared only in connection with the last two tasks, when the user often had access to the full-text of the document.

- The same decision rules appeared throughout the research process.

- For a particular document, a participant often changed the criteria and decision rule he used in making document use decisions.

- The document's contribution to the research seemed to be a major factor in the decision to cite a document. Although similar to motivations for citing found through content analysis and in Hodges' work, the contributions were sometimes subtle and could not always be inferred from contextual analysis of the finished product.

- The participants applied numerous other criteria in evaluating the documents for citation, often several criteria for the same document. At a broad level, these factors were relatively consistent across the documents. Only 10 of 28 factors appeared only for negative or positive decisions; otherwise, the criteria were not directional and were mentioned in connection with both negative and positive decisions.

- Some decisions appear to be influenced by external frames of reference. In these cases, the decision-making emphasis shifted away from the document being used to the document being generated and how it and the research it embodies would be judged by others.

- Perhaps the most significant finding in the citing study is that there is some evidence of overarching concerns which influence decision-making related to citing almost independent of considerations of the use of a document during the research project.

Suggestions for Future Research

The notion that frameworks for judgment about information change as the user becomes more conversant with the overall task is one that has re-appeared at various times in studies of user behavior. Additional studies need to be done to characterize those changes vis a vis certain

types of tasks and to determine if research or task phase can be used to refine interaction with information systems, perhaps through highly individualized interfaces.

The metalevel documentation concerns related to citing identified in the Chapter 5 are very intriguing and suggest an interesting area for additional research. The findings here are only tentative and need to be verified through a more focused, consistent approach. How widespread are these concerns? Can researchers articulate effectively a "style of documentation" for themselves? If they can, are these styles discipline-dependent or dependent on individual researcher characteristics? How are these styles learned or developed? How are they related to the rhetoric employed in the paper, which may be prevalent in a particular type of written product? White is currently extending the research in this regard, interviewing a sample of researchers from a range of fields.

In related work Peiling Wang is analyzing and collecting data on the knowledge structures of the researchers at different phases in the project, drawing on vocabulary used for describing their information needs related to document decisions.

Dissemination of Results

Preliminary results of this research have already been disseminated through papers presented at two annual conferences of the American Society for Information Science (ASIS):

Wang, Peiling, and Marilyn Domas White. 1995. Document use during a research project: A longitudinal study. In *Proceedings of the 58th ASIS annual meeting*, ed. Tom Kinney. Medford, NJ: Learned Information, Inc.

Wang, Peiling, and Marilyn Domas White. 1996. A qualitative study of scholars' citation behavior. In *Proceedings of the 59th ASIS annual meeting.* Medford, NJ: Learned Information, Inc.

A journal article, which is based on Chapter 5 but also includes some corroborative bibliometric analysis, has been accepted for publication:

White, Marilyn Domas, and Peiling Wang. 1997. A qualitative study of citing behavior: Contributions, criteria, and metalevel documentation concerns. *Library Quarterly.* (Forthcoming)

Another article, presenting a more holistic view of the results, based in part on Chapter 4, is in progress:

Wang, Peiling, and Marilyn Domas White. 1997. Document selection and reference assessments during a research project. (In progress)

An earlier version of this report was provided to the Council on Library Resources, the funding agency.

BIBLIOGRAPHY

Barry, Carol L. 1993. The identification of user criteria of relevance and document characteristics: Beyond the topical approach to information retrieval. Ph.D. dissertation, Syracuse University.

Barry, Carol L., and Linda Schamber. 1995. User-defined relevance criteria: A comparison of two studies. In *Proceedings of the 58th ASIS annual meeting*, ed. Tom Kinney. Medford, NJ: Learned Information, Inc.

Bonzi, Susan, and Herbert W. Snyder. 1991. Motivations for citation: A comparison of self citation and citation to others. *Scientometrics* 21:245-253.

Borgman, Christine L. 1990. *Scholarly communication and bibliometrics*. Newberry Park, CA: Sage.

Brooks, Terrence A. 1986. Evidence of complex citer motivations. *Journal of the American Society for Information Science* 37:34-36.

_____. 1985. Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science* 36:223-229.

Cano, V. 1989. Citation behavior: Classification, utility, and location. *Journal of the American Society for Information Science* 40:284-290.

Chubin, Daryl E., and Soumyo D. Moitra. 1975. Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science* 5:423-441.

Cronin, Blaise. 1984. *The citation process: The role and significance of citations in scientific communication.* London: Taylor Graham.

Dervin, Brenda. 1983. Information as a user construct: The relevance of perceived information to synthesis and interpretation. In *Knowledge structures and use: Implications for synthesis and interpretation*, ed. Spencer A. Ward and Linda J. Reed. Philadelphia, PA: Temple University Press.

Ericsson, Karl Anders, and Herbert A. Simon. 1993. *Protocol analysis: Verbal reports as data.* Cambridge, MA: The MIT Press.

Fischhoff, Baruch. 1988. Judgment and decision making. In *The psychology of human thought*, ed. Robert J. Sternberg and Edward E. Smith. Cambridge: Cambridge University Press.

Frost, Carolyn O. 1979. Use of citations in literary research: A preliminary classification of citation functions. *Library Quarterly* 49:339-414.

Harter, Stephen P. 1992. Psychological relevance and information science. *Journal of the American Society for Information Science* 43:602-615.

Hodges, Theodora L. 1978. Citation indexing: Its potential for bibliographic control. Ph.D. dissertation, University of California, Berkeley.

Hooten, Patricia A. 1991. Frequency and functional use of cited documents in information science. *Journal of the American Society for Information Science* 42:397-404.

Krippendorff, Klaus. 1980. *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

Liu, Mengxiong. 1993. The complexities of citation practice: A review of citation studies. *Journal of Documentation* 49:370-408.

_____. 1993. A study of citing motivation of Chinese scientists. *Journal of Information Science* 19:13-23.

MacRoberts, M. H., and B. R. Mac Roberts. 1988. Author motivation for not citing influences: A methodological note. *Journal of the American Society for Information Science* 39:432-433.

_____. 1986. Quantitative measures of communication in science: A study of the formal level. *Social Studies of Science* 16:151-172.

Montgomery, Henry, and Ola Svenson, eds. 1989. *Process and structure in human decision making*. Chichester: Wiley.

Moravcsik, M. J., and P. Murugesan. 1978. Variation of the nature of citation measures with journals and scientific specialties. *Journal of the American Society for Information Science* 29:141-147.

Park, Taemin Kim. 1992. *The nature of relevance in information retrieval: An empirical study*. Ph.D. Dissertation, Indiana University, Bloomington, IN.

Peritz, Bluma C. 1983. A classification of citation roles for the social sciences and related fields. *Scientometrics* 5:303-312.

Prabha, Chandra. 1983. Some aspects of citation behavior: A pilot study in business administration. *Journal of the American Society for Information Science* 34:202-206.

*Q.S.R. NUD*IST 3.0 for Microsoft Windows*. Power Version. [computer software]. (1985-1994). Melbourne: Qualitative Solutions and Research PTY, LTD.

Rees, Alan M., and Douglas G. Schultz. 1967. *A field experimental approach to the study of relevance assessments in relation to document searching: Final report.* Cleveland: Western Reserve University.

Saracevic, Tefko. 1970. The concept of 'relevance' in information science: A historical review. In *Introduction to Information Science,* ed. Tefko Saracevic. New York: R.R. Bowker.

_____. 1975. Relevance: A review of and a framework for thinking on the notion in information science. *Journal of the American Society for Information Science* 26:321-343.

Schamber, Linda, Michael B. Eisenberg, and Michael S. Nilan. 1990. A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing and Management.* 26:755-776.

Schamber, Linda. 1994. Relevance and information behavior. In *Annual review of information science and technology,* ed. Martha E. Williams. Medford, NJ: Learned Information, Inc.

_____. 1991. Users' criteria for evaluation in multimedia information seeking and use situations. Ph.D. dissertation, Syracuse University.

Simon, Hebert A. 1985. Information-processing theory of human problem solving. In *Issues in cognitive modeling,* ed. A. M. Aitkenhead and J. M. Slack. London: Lawrence Erlbaum Associates.

Small, Henry G. 1982. Citation context analysis. In *Progress in communication sciences,* v. 3, ed. Brenda J. Dervin and Melvin J. Voigt. Norwood: Ablex.

Smith, Linda C. 1981. Citation analysis. *Library Trends* 30:83-106.

Snyder, Herbert W., and Susan Bonzi. 1989. An enquiry into the behavior of author self citation. In *Proceedings of the 52d ASIS annual meeting,* ed. Jeffrey Katzer and Gregory B. Newby. Medford, NJ: Learned Information, Inc.

Spiegel-Rosing, Ina. 1977. Bibliometric and content analysis. *Social Studies of Science* 7:79-113.

Svenson, Ola. 1979. Process descriptions of decision making. *Organizational Behavior and Human Decision Processes* 23:86-112.

Vinkler, P. 1987. A quasi-quantitative citation model. *Scientometrics* 12:47-72.

Wang, Peiling, and Dagobert Soergel. 1993. Beyond topical relevance: Document selection behavior of real users of information retrieval systems. *Proceedings of the 56th ASIS annual meeting*, ed. Susan Bonzi. Medford, NJ: Learned Information, Inc.

Wang, Peiling. 1994. A cognitive model of document selection of real users of information retrieval systems. Ph.D. dissertation, University of Maryland, College Park.

Wang, Peiling, and Marilyn Domas White. 1997. Document selection and relevance assessments during a research project. (In progress)

_____. 1995. Document use during a research project: A longitudinal study. In *Proceedings of the 58th ASIS annual meeting*, ed. Tom Kinney. Medford, NJ: Learned Information, Inc.

_____. 1996. A qualitative study of scholars' citation behavior. In *Proceedings of the 59th ASIS annual meeting*, ed. Steve Hardin. Medford, NJ: Learned Information, Inc.

White, Howard D., and Kathleen W. McCain. 1989. Bibliometrics. In *Annual review of information science and technology*, ed. Martha E. Williams. Medford, NJ: Learned Information, Inc.

White, Marilyn Domas. 1975. The communications behavior of academic economists in research phases. *Library Quarterly* 45:337-354.

White, Marilyn Domas, and Peiling Wang. 1997. A qualitative study of citing behavior: Contributions, criteria, and metalevel documentation concerns. *Library Quarterly.* (Forthcoming)

Wilson, Patrick. 1995. Unused relevant information in research and development. *Journal of the American Society for Information Science* 46:45-51.

Wright, George. 1984. *Behavioral decision theory: An introduction.* Beverly Hills: Sage Publications.

Appendix A. MATERIALS RELATED TO CLEARANCE AND CONSENT

Application for Review of Research Project Using Human Subjects

Marilyn Domas White
College of Library and Information Sciences
University of Maryland at College Park

Peiling Wang
Department of Agricultural and Resources Economics
University of Maryland at College Park

This is to request approval of using the human subjects for this project.

1. Purpose

This is a longitudinal study of document relevance judgments made by real academic researchers in actual situations. Users' information needs change over the stages of a research project, but most recent relevance studies have focused on judgments made at one time. In previous research, Wang [1] developed a cognitive model of users' document selection behavior for their research projects. It was based on analysis of users' judgments about bibliographic output from an online search. This follow-up study gathers data from the same researchers on use patterns and reasons for use and citation of items in the original set plus additional documents identified later. The purpose of this research is to shed light on the nature of relevance judgments of the same document at subsequent stages of the research project. The results will expand Wang's model and suggest factors to be considered in designing future information retrieval systems that can incorporate users' long-term information behavior.

2. Participants

Participants are faculty and graduate students from the Dept. of Agricultural and Resource Economics who have participated previous research, because this is a follow-up study of the previous study.

3. Procedures

The participant will be interviewed on how the previous selected documents were used. The interview will be audio-taped.

---

[1]Peiling Wang. 1994. *A Cognitive Model of Document Selection of Real Users of Information Retrieval Systems*. Ph.D. Dissertation. University of Maryland.

4. Risks?

There is no risk to subjects in this research.

5. Protection to the subjects

Original data will not be released to any one else without the participant's permission. A pseudo name will be given to each participant. Data analysis will not use any real name.

6. Participants' consents

The attached is a consent form. The participants will be asked to sign the form before the research.

Research Consent Agreement

Title of the Study:
Relevance Assessments During a Research Project: A Longitudinal Study of Document Use.

Description of the Study:
This study follows previous study[1] to collect further data from the same users on how the document were used in the following stages.

Researchers:
Marilyn Domas White
College of Library and Information Sciences
Peiling Wang
Department of Agricultural and Resources Economics

I have freely volunteered to participate in this study and been informed in advance as to what my task(s) would be and what procedures would be followed. I have been given the opportunity to ask questions, and have had my questions answered to my satisfaction.

I am aware that I have the right to withdraw consent and discontinue participation at any time, without prejudice.

My signature below may be taken as affirmation of all of the above, prior to participation. And, I agree to be contacted by third party after participating to verify whether I have been treated in accordance with the "APA Ethical Principle 9."

Date:
Participant's Name:

Signature:

---

[1]Peiling Wang. 1994. *A Cognitive Model of Document Selection of Real Users of Information Retrieval Systems*. Ph.D. Dissertation. University of Maryland.

Appendix B. INTERVIEW GUIDELINES

RELEVANCE ASSESSMENTS DURING A RESEARCH PROJECT: INTERVIEW GUIDELINES

Note: Each participant was provided with the following materials at the start of the interview. At specific times during the interview, he was asked to refer to these materials, which included:

- *Initial query*: the query statement provided in the 1992 study.
- *Relevant/cited item bibliography*: bibliographic records for the items considered relevant in the 1992 study and cited in the pt's final product. The bibliographic record usually contained the citation, abstract, and descriptors and identifiers.
- *Relevant/not cited item bibliography*: bibliographic records for the items considered relevant in the 1992 study and not cited. The bibliographic record was the same as above.
- *Cited items from participant's bibliography*: the bibliography from the pt's final product, with a random sample of eight items marked with a check. The record consisted only of a standard bibliographic citation. No abstract or indexing was included.

Purpose of project:

We are trying to identify the decision factors and rules that influence how and why researchers use information during a project. Please be as complete as possible in indicating why you did or did not use something and how you arrived at that decision. In answering your questions, please free to refer to your paper, to the printout of the items you considered relevant in the first stage of this project, or to the actual item if you have it nearby.

Note: Refer to initial query.

1. Please read over the topic of the original search. How does the paper we are talking about today fit into the original project?

Note: Refer to relevant/cited item bibliography.

2. For each item considered relevant and cited:

2.1 What contributed to the decision to use and cite this item?

Note: Probe, if necessary, to get beyond subject match.

2.2 Any other factors besides containing the information you needed?

Note: Refer to Relevant item/not cited bibliography.

3. For each item considered relevant, but not cited:

These are the items you considered useful shortly after I did the search but decided not to cite in the paper. In this case, we are interested in the decision-making that led up to your deciding not to cite it. We have some particular questions, but we would just like you to talk about each one. Please feel free to mention other items you used if you want. We need some specific information about each one, but we will followup as necessary to make sure we get that information.

3.1 Did you obtain a copy of the item?

If yes:
3.1.1 Did you read or scan it?
3.1.2 Was it useful to you at all in this project?
If yes:
3.1.2.1 How was it useful?
3.1.2.2 Why did you decide not to cite it?

If no:
3.1.2 Why did you decide not get a copy of the item?

NOTE: Refer to checked, cited items in participant's bibliography.

For each checked, cited item:
4. We have identified a sample of eight items that you cited that were not in the original search. They are checked in red on the [paper, article, proposal] bibliography. Could we go through these individually to ask you three questions about each one.

4.1 How did you discover this document?
4.2 What contributed to the decision to use and cite this item?
4.3 Did you know about this item before the search?

Appendix C. CODING SCHEMA

CODING SCHEMA USED FOR CHAPTER 4

(6)	/criteria
(6 1)	/criteria/topicality
(6 1 1)	/criteria/topicality/not clear
(6 1 2)	/criteria/topicality/not relevant(useful)
(6 1 3)	/criteria/topicality/facet not match
(6 1 4)	/criteria/topicality/related
(6 1 5)	/criteria/topicality/facet match
(6 1 6)	/criteria/topicality/match-close
(6 1 7)	/criteria/topicality/too specific (narrow)
(6 1 8)	/criteria/topicality/too general (broad)
(6 1 9)	/criteria/topicality/useful for other task
(6 1 10)	/criteria/topicality/match other interests
(6 1 11)	/criteria/topicality/context or bib.
(6 1 12)	/criteria/topicality/concept-interpreted-different
(6 1 13)	/criteria/topicality/analogy
(6 1 14)	/criteria/topicality/facet(focus) changed
(6 2)	/criteria/novelty
(6 2 1)	/criteria/novelty/known item
(6 2 2)	/criteria/novelty/might have
(6 2 3)	/criteria/novelty/known information
(6 2 4)	/criteria/novelty/dups in P.O.
(6 2 5)	/criteria/novelty/diff. edition
(6 3)	/criteria/subject field
(6 3 1)	/criteria/subject field/not sure
(6 3 2)	/criteria/subject field/remote-not close
(6 3 3)	/criteria/subject field/related
(6 3 4)	/criteria/subject field/same
(6 4)	/criteria/peer reviewed
(6 4 1)	/criteria/peer reviewed/no
(6 4 2)	/criteria/peer reviewed/yes
(6 4 3)	/criteria/peer reviewed/cited by other paper
(6 5)	/criteria/quality
(6 5 1)	/criteria/quality/low
(6 5 2)	/criteria/quality/uncertain
(6 5 3)	/criteria/quality/average-fair
(6 5 4)	/criteria/quality/good-high
(6 6)	/criteria/recency
(6 6 1)	/criteria/recency/current
(6 6 2)	/criteria/recency/older
(6 6 3)	/criteria/recency/too old

(6 6 4) /criteria/recency/became old
(6 7) /criteria/orientation
(6 7 1) /criteria/orientation/not clear
(6 7 2) /criteria/orientation/news-popular
(6 7 3) /criteria/orientation/empirical(cases)
(6 7 4) /criteria/orientation/factual-descriptive
(6 7 5) /criteria/orientation/application
(6 7 6) /criteria/orientation/methodology-technique
(6 7 7) /criteria/orientation/theoretical-conceptual
(6 7 8) /criteria/orientation/historical
(6 7 9) /criteria/orientation/technology-engineering
(6 7 10) /criteria/orientation/academic
(6 7 11) /criteria/orientation/critique-discussion
(6 7 12) /criteria/orientation/not theoretical
(6 7 13) /criteria/orientation/review-summary
(6 7 14) /criteria/orientation/policy-political
(6 7 15) /criteria/orientation/multidisciplinary
(6 7 16) /criteria/orientation/agenda type
(6 7 17) /criteria/orientation/data related
(6 7 18) /criteria/orientation/qualitative(analytical)
(6 8) /criteria/depth
(6 8 1) /criteria/depth/not
(6 8 2) /criteria/depth/yes
(6 9) /criteria/availability
(6 9 1) /criteria/availability/departmental
(6 9 2) /criteria/availability/campus
(6 9 3) /criteria/availability/ILL
(6 9 4) /criteria/availability/outside US
(6 9 5) /criteria/availability/difficult
(6 10) /criteria/time(cost, effort)
(6 10 1) /criteria/time(cost, effort)/unaffordable
(6 10 2) /criteria/time(cost, effort)/too much(not worthwhile)
(6 10 3) /criteria/time(cost, effort)/affordable(worthwhile)
(6 11) /criteria/special requisite
(6 11 1) /criteria/special requisite/foreign language
(6 11 2) /criteria/special requisite/micro-format
(6 12) /criteria/authority
(6 12 1) /criteria/authority/no
(6 12 2) /criteria/authority/yes
(6 13) /criteria/relation
(6 13 1) /criteria/relation/advisor
(6 13 2) /criteria/relation/dept. professor
(6 13 3) /criteria/relation/committee member

(6 13 4) /criteria/relation/classmate(fellowsl)
(6 13 5) /criteria/relation/colleague
(6 13 6) /criteria/relation/affiliated organization
(6 13 7) /criteria/relation/advisor's collaborator
(6 13 8) /criteria/relation/collaborator
(6 14) /criteria/publicity
(6 14 2) /criteria/publicity/well(wide) known
(6 15) /criteria/spectrum of JN
(6 15 1) /criteria/spectrum of JN/first class(top)
(6 15 2) /criteria/spectrum of JN/national(international)
(6 15 3) /criteria/spectrum of JN/central
(6 15 4) /criteria/spectrum of JN/major(main stream)
(6 15 5) /criteria/spectrum of JN/minor
(6 15 6) /criteria/spectrum of JN/periphery
(6 15 7) /criteria/spectrum of JN/regional(local
(6 15 8) /criteria/spectrum of JN/non-economics
(6 20) /criteria/cognitive-requisite
(6 20 1) /criteria/cognitive-requisite/not high
(6 20 2) /criteria/cognitive-requisite/too high
(6 21) /criteria/actual quality
(6 21 1) /criteria/actual quality/low
(6 21 3) /criteria/actual quality/average-fair
(6 21 4) /criteria/actual quality/good-high
(6 22) /criteria/classic(founder)
(6 22 2) /criteria/classic(founder)/yes
(6 22 3) /criteria/classic(founder)/follow to classic
(6 23) /criteria/standard reference(textbook)
(6 23 2) /criteria/standard reference(textbook)/yes
(6 24) /criteria/reputation(reliability)
(6 24 1) /criteria/reputation(reliability)/person
(6 24 1 1) /criteria/reputation(reliability)/person/low
(6 24 1 2) /criteria/reputation(reliability)/person/high
(6 24 2) /criteria/reputation(reliability)/Journal
(6 24 2 1) /criteria/reputation(reliability)/Journal/low
(6 24 2 2) /criteria/reputation(reliability)/Journal/high
(6 24 3) /criteria/reputation(reliability)/Organization
(6 24 3 1) /criteria/reputation(reliability)/Organization/low
(6 24 3 2) /criteria/reputation(reliability)/Organization/high
(6 24 4) /criteria/reputation(reliability)/Newspaper(magazine)
(6 24 4 1) /criteria/reputation(reliability)/Newspaper(magazine)/low
(6 24 4 2) /criteria/reputation(reliability)/Newspaper(magazine)/high
(6 25) /criteria/prolific author
(6 25 2) /criteria/prolific author/yes

(6 26) /criteria/judge
(6 26 1) /criteria/judge/advisor
(6 26 2) /criteria/judge/committee member
(6 26 3) /criteria/judge/journal referee
(6 26 4) /criteria/judge/funding agent
(6 27) /criteria/norm
(6 27 1) /criteria/norm/for perspectus
(6 27 2) /criteria/norm/for dissertation
(6 27 3) /criteria/norm/for funding
(6 27 4) /criteria/norm/for conference-paper
(6 27 5) /criteria/norm/for working-paper
(6 27 6) /criteria/norm/for journal paper
(6 27 7) /criteria/norm/for review article
(6 28) /criteria/credential
(6 28 2) /criteria/credential/concerned
(6 29) /criteria/target journal
(6 29 2) /criteria/target journal/conscious(level up)
(6 30) /criteria/opinions
(6 30 1) /criteria/opinions/interest group
(6 30 2) /criteria/opinions/public

(7) /DM strategies
(7 1) /DM strategies/elimination
(7 2) /DM strategies/multi-criteria
(7 3) /DM strategies/dominance
(7 4) /DM strategies/scarcity
(7 5) /DM strategies/satisfice
(7 6) /DM strategies/chain
(7 6 1) /DM strategies/chain/comments-criticism
(7 6 2) /DM strategies/chain/collective volume
(7 6 3) /DM strategies/chain/dups in P.O.
(7 6 4) /DM strategies/chain/diff. edition
(7 6 5) /DM strategies/chain/by same AU
(7 6 6) /DM strategies/chain/dedicated JN issue
(7 6 7) /DM strategies/chain/bearing similarity
(7 6 8) /DM strategies/chain/original-followup

(8) /decisions
(8 1) /decisions/first selecting
(8 2) /decisions/second selecting
(8 3) /decisions/final acceptance
(8 3 1) /decisions/final acceptance/no
(8 3 2) /decisions/final acceptance/yes

(8 4)	/decisions/reading
(8 4 1)	/decisions/reading/no
(8 4 1 2)	/decisions/reading/no/given to RA
(8 4 2)	/decisions/reading/yes
(8 4 3)	/decisions/reading/previous
(8 4 4)	/decisions/reading/will read
(8 4 5)	/decisions/reading/may read
(8 5)	/decisions/citing
(8 5 1)	/decisions/citing/no
(8 5 2)	/decisions/citing/yes
(8 5 3)	/decisions/citing/likely
(8 5 4)	/decisions/citing/maybe
(8 5 5)	/decisions/citing/further research
(8 5 6)	/decisions/citing/other paper
(8 5 7)	/decisions/citing/definitely will
(8 5 8)	/decisions/citing/should not cite

REVISED CODING SCHEMA FOR CRITERIA

Note: The revised codes were developed during the analysis for the citation data.

(6)	criteria
(6 1)	criteria/topicality
(6 1 1)	criteria/topicality/not clear
(6 1 2)	criteria/topicality/not relevant
(6 1 3)	criteria/topicality/facet not match
(6 1 4)	criteria/topicality/related
(6 1 5)	criteria/topicality/facet match
(6 1 6)	criteria/topicality/close match
(6 1 7)	criteria/topicality/too specific
(6 1 8)	criteria/topicality/too general
(6 1 9)	criteria/topicality/other task
(6 1 10)	criteria/topicality/other interests
(6 1 11)	criteria/topicality/context, bibliography
(6 1 12)	criteria/topicality/concept different
(6 1 13)	criteria/topicality/analogy
(6 1 14)	criteria/topicality/facet changed
(6 2)	criteria/novelty
(6 2 1)	criteria/novelty/known item
(6 2 2)	criteria/novelty/may have
(6 2 3)	criteria/novelty/known information
(6 2 4)	criteria/novelty/dups in P.O.
(6 2 5)	criteria/novelty/diff. edition
(6 3)	criteria/discipline
(6 3 1)	criteria/discipline/not sure
(6 3 2)	criteria/discipline/remote
(6 3 3)	criteria/discipline/related
(6 3 4)	criteria/discipline/same
(6 3 5)	criteria/discipline/multidisciplinary
(6 4)	criteria/peer reviewed
(6 4 1)	criteria/peer reviewed/no
(6 4 2)	criteria/peer reviewed/yes
(6 4 3)	criteria/peer reviewed/cited by other paper

(6 5)	criteria/expected quality
(6 5 1)	criteria/expected quality/low
(6 5 2)	criteria/expected quality/uncertain
(6 5 3)	criteria/expected quality/fair
(6 5 4)	criteria/expected quality/good
(6 6)	criteria/recency
(6 6 1)	criteria/recency/current
(6 6 2)	criteria/recency/older
(6 6 3)	criteria/recency/too old
(6 8)	criteria/depth
(6 8 1)	criteria/depth/cursory
(6 8 2)	criteria/depth/substantial
(6 9)	criteria/availability
(6 9 1)	criteria/availability/department
(6 9 2)	criteria/availability/campus
(6 9 3)	criteria/availability/ILL
(6 9 4)	criteria/availability/outside US
(6 9 5)	criteria/availability/difficult
(6 10)	criteria/time, effort
(6 10 1)	criteria/time, effort/unaffordable
(6 10 2)	criteria/time, effort/not worthwhile
(6 10 3)	criteria/time, effort/worthwhile
(6 11)	criteria/special requisite
(6 11 1)	criteria/special requisite/foreign language
(6 11 2)	criteria/special requisite/micro-format
(6 12)	criteria/authority
(6 12 1)	criteria/authority/no
(6 12 2)	criteria/authority/yes
(6 13)	criteria/relation
(6 13 1)	criteria/relation/advisor
(6 13 2)	criteria/relation/dept. professor
(6 13 3)	criteria/relation/committee member
(6 13 4)	criteria/relation/classmate
(6 13 5)	criteria/relation/colleague
(6 13 6)	criteria/relation/affiliated organization

(6 13 7)	criteria/relation/advisor's collaborator
(6 13 8)	criteria/relation/collaborator
(6 14)	criteria/publicity
(6 14 2)	criteria/publicity/well-known
(6 15)	criteria/journal spectrum
(6 15 1)	criteria/journal spectrum/first class
(6 15 2)	criteria/journal spectrum/national, international
(6 15 3)	criteria/journal spectrum/central
(6 15 4)	criteria/journal spectrum/major
(6 15 5)	criteria/journal spectrum/minor
(6 15 6)	criteria/journal spectrum/peripheral
(6 15 7)	criteria/journal spectrum/regional, local
(6 15 8)	criteria/journal spectrum/non-economics
(6 20)	criteria/cognitive requisite
(6 20 1)	criteria/cognitive requisite/not high
(6 20 2)	criteria/cognitive requisite/too high
(6 21)	criteria/actual quality
(6 21 1)	criteria/actual quality/low
(6 21 3)	criteria/actual quality/average
(6 21 4)	criteria/actual quality/good
(6 22)	criteria/classic, founder
(6 22 2)	criteria/classic, founder/yes
(6 22 3)	criteria/classic, founder/follow to classic
(6 23)	criteria/standard reference
(6 23 2)	criteria/standard reference/yes
(6 24)	criteria/reputation
(6 24 1)	criteria/reputation/person
(6 24 1 1)	criteria/reputation/person/low
(6 24 1 2)	criteria/reputation/person/high
(6 24 2)	criteria/reputation/journal
(6 24 2 1)	criteria/reputation/journal/low
(6 24 2 2)	criteria/reputation/journal/high
(6 24 3)	criteria/reputation/organization
(6 24 3 1)	criteria/reputation/organization/low
(6 24 3 2)	criteria/reputation/organization/high

(6 25) criteria/prolific author
(6 25 2) criteria/prolific author/yes

(6 26) criteria/judge
(6 26 1) criteria/judge/advisor
(6 26 2) criteria/judge/committee member
(6 26 3) criteria/judge/journal referee
(6 26 4) criteria/judge/funding agent

(6 27) criteria/norm
(6 27 1) criteria/norm/for proposal
(6 27 2) criteria/norm/for dissertation
(6 27 3) criteria/norm/for funding
(6 27 4) criteria/norm/for conference-paper
(6 27 5) criteria/norm/for working-paper
(6 27 6) criteria/norm/for journal paper
(6 27 7) criteria/norm/for review article

(6 28) criteria/credential
(6 28 2) criteria/credential/concerned

(6 29) criteria/target journal
(6 29) criteria/target journal
(6 29 2) criteria/target journal/conscious(level up)

(6 40) criteria/audience
(6 40 1) criteria/audience/popular
(6 40 2) criteria/audience/scholarly

(6 41) criteria/orientation1
(6 41 1) criteria/orientation1/theoretical
(6 41 2) criteria/orientation1/empirical

(6 42) criteria/content
(6 42 1) criteria/content/factual,descriptive
(6 42 2) criteria/content/application
(6 42 3) criteria/content/methodology
(6 42 4) criteria/content/theory, models
(6 42 5) criteria/content/historical
(6 42 6) criteria/content/technology/engineering
(6 42 7) criteria/content/critique, discussion
(6 42 8) criteria/content/review,summary
(6 42 9) criteria/content/policy

(6 42 10)	criteria/content/agenda
(6 42 11)	criteria/content/data
(6 43)	criteria/orientation2
(6 43 1)	criteria/orientation2/qualitative
(6 43 2)	criteria/orientation2/quantitative

Appendix D. SAMPLE CODED INTERVIEW

SAMPLE CODED INTERVIEW

Note: This interview is coded with the first schema listed in Appendix C. The bolded lines demark specific question areas within the interview: 1) the fit of the finished project to the original search; 2) comments related to read/cited items; 3) comments related to all other items selected as useful in original search; 4) a sample or all citations in product bibliography not in original search. An Arabic number in the text preceded by #, e.g. #1, refers to a specific document; these numbers correspond to item numbers marked on the materials given to the client. The client often refers to a specific item only by number. (See, for example, no. 5.) The numbers in the margin to the right demark text units; in most cases, a text unit consists of the exchange related to one document. Coding related to a specific document follows the text unit. Names and titles mentioned in the text are indicated by an X or by [title].

PROJECT: FOLLOWUP, User Wang, 2:42 pm, Jan 26, 1996.

++
+++ ON-LINE DOCUMENT: P06-95r
+++ Document Header:
*Verbal Protocol for Participant #006 (Subject no. 8)
Journal article: 30 selected; 21 read; 1 read/cited;
8 new cites; total of 42 cites in bibliography.

+++ Retrieval for this document: 69 units out of 69 = 100%
White: You have gone through the first part of this project and have completed the product. Now we'll be able to followup and see how you use materials that you once considered relevant. So what we'll be doing is following somewhat the same kind of procedure that Peiling did before, except that we're going to focus only on the items that you considered relevant initially. And we're going to look first of all at the ones that you've cited, and then the ones that you did not cite. What we're interested in is the same kinds of factors that she was interested in before -- the decision-making process that you went through so that we can draw the decision rules from that, and also the factors that you considered in the decision making itself. Peiling is going to ask most of the questions and I will step in if a point needs to be clarified. 1

1. How does the paper fit into the original topic/project. 2

Wang: The first thing I want to do is to give you the topic sheet. And now, please look at it and tell us how this description fits your final project. Participant #006: It fits really closely; this is what we worked on. We worked on the construction of inequality indexes, income inequality indexes or income distribution indexes. So the information that was in here fits really closely with what I would have considered to be the

existing type of literature at the time. 3
2. Cited Items (one known item). 4
Wang: The next thing I will give you will be a cited article I found in the search, and we would like to know what this particular document contributed to your paper. Participant #006: The reason we cited this particular paper wasn't because it contributed anything to our particular research. It was just tangentially related and it reflects a different approach to the construction of inequality indexes that we thought people should be made aware of. So it didn't do anything directly for our research. Actually, most of the literature that we cited is going to be like that. We took what we thought was a fairly different approach to constructing these indexes. This is the X piece. [Wang: Was there any other item for same information?] Definitely. In picking the onces that we cited, definitely. We cited Journal of Economic Theory because that was the level at which we were trying to write the paper. In economics there are all different kind of levels of journals, and the theoretical level that we were aiming at is most closely matched by Journal of Economic Theory, Review of Economic Studies, and Econometrica. The paper that we actually wrote was ultimately submitted to Econometrica. So when we picked out references, we tried to stay in that group. It is a little bit of gamesmanship in a way, to be citing the right people. [Wang: What about the author?] X? There is nothing specific about him. I know him, and know that he is at the University of British Columbia, but there is nothing specific about him. He is now that the University of British Columbia; he used to be at (?) [White: Was he the only one who has used this approach that you know of?] No, this is very common, this axiomatic approach. What they do is very different from us. They specify a series of properties they think these indexes should have, and then they mathematically figure out which class of indexes would fall in that class. 5

(6 1 4) /criteria/topicality/related
(6 7 6) /criteria/orientation/methodology-technique
(6 29 2) /criteria/target journal/conscious(level up)
(7 3) /DM strategies/dominance
(8 4 2) /decisions/reading/yes
(8 5 2) /decisions/citing/yes

[White: You could have chosen a number of articles that use that approach. Did you actually do that?] We cited actually a whole array of articles. There is a long history of these type of indexes, so I know we cited X, probably cited X. I can't remember if we cited X or not. And a number of those older ones and we kind of gradually marched forward, to X and X, and people like that. [White: If there are a number of people, he's more current? (Right.) If there were a number of articles that you could have chosen to represent the

(?), you elected to get him?] There will be relatively few people working in this area, so among the more current articles that are at this level and are in economics journals as opposed to kind of mathematical type of social science journals, he would be a logical choice. 6
(6 3 4) /criteria/subject field/same
(6 6 1) /criteria/recency/current
(7 2) /DM strategies/multi-criteria

3. Go through the selected documents one by one (32 including 2 known items). 7

Wang: Now we will move on to the third part of our interview: that will be the ones you selected at that time but didn't show up in the articles which you cited. What we want to know is, one by one, can you tell us whether you actually obtained the document; if you got it, have you read it; if you read it, whether you got any information from it, or if you decided not to cite it, how did you make the decision... Participant #006: #1. I didn't get this. The reason I didn't get it is because it is focused on the traditional approaches to measuring inequality and poverty, and it appears to be totally empirical. We were interested in doing a piece of theoretical research and the information that we would have gleaned from this would have been limited for our purposes. We were looking more towards a mathematical backdrop to the theory. [White: Did you discover that after you had selected it originally?] I don't recall. 8
(6 7 3) /criteria/orientation/empirical(cases)
(6 7 6) /criteria/orientation/methodology-technique
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no

#2. I did look at this. It wasn't very informative. This is not really a professional type of journal; this is kind of a mix between a professional and a popular type of journal, so it wasn't very helpful. 9
(6 5 2) /criteria/quality/uncertain
(6 7 2) /criteria/orientation/news-popular
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no

#3. By X. I remember trying to find this one, but I couldn't find it. When I kind of experienced difficulty in getting it I gave up because it is from a very minor, minor journal, so it is unlikely that it is going to be very good. Although the individual is good, and I probably picked this one on the basis of the individual's name. 10
(5 1 2) /knowledge of/person/yes
(6 5 1) /criteria/quality/low
(6 9 5) /criteria/availability/difficult
(7 2) /DM strategies/multi-criteria
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no

#4. I don't recall seeing this, and I don't know why. But again, we wouldn't have cited this and we probably wouldn't have gone into this because this is

BEST COPY AVAILABLE

much more empirical than what we ultimately ended up doing. 11
(6 7 3) /criteria/orientation/empirical(cases)
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#5. I don't recall No. 5 at all. 12
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#6. I do recall No. 6, it just wasn't very helpful. This really could be important at an empirical level, but conceptually it's not very important. It's very clear that the point that the individual is making is obviously true. At an empirical level it could be very important, but at a conceptual level it's not. [Wang: Any comment about the author?] I'd never heard of this person before. 13
(6 7 3) /criteria/orientation/empirical(cases)
(6 7 12) /criteria/orientation/not theoretical
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#7. I don't recall. No. 7 would have been very, very hard to get because it's an LSMS study working paper, and they're very hard to get from the World Bank at times, and again, this would have been mainly empirical, and I don't know who X is. 14
(5 1 1) /knowledge of/person/no
(6 7 3) /criteria/orientation/empirical(cases)
(6 9 5) /criteria/availability/difficult
(7 2) /DM strategies/multi-criteria
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#8. I did look at. The reason I looked at it is because of X, one of the authors. It just didn't relate very much to what we were doing. Again, it was much more empirical than what we were doing. 15
(5 1 2) /knowledge of/person/yes
(6 7 3) /criteria/orientation/empirical(cases)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#9. A bad paper. It wasn't worth citing. [Wang: When you say "not worth citing," can you elaborate?] You wouldn't cite this paper. First of all, it's not in a very good journal, so if we were trying to write at the level of the journals we're talking about, which are usually perceived as being the three best theoretical journals, it would be a mistake to be citing something like this unless it really had a point to make that nobody else had made. You don't cite much lower journals unless the journal contains something that was overlooked by everybody else, and that's not the case here. This is just a bad paper. Now I'm sorry I looked at it. 16
(6 21 1) /criteria/actual quality/low
(6 24 1 1) /criteria/reputation(reliability)/person/low
(6 29 2) /criteria/target journal/conscious(level up)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#10. I did look at. It's a little difficult to

summarize in a few words, but there is a difference between poverty measures and income inequality measures. The difference is fairly technical in nature, and we ended up not saying too much about the poverty measures. We were focusing on inequality measures, and that's why we didn't cite this, and it would be the reason we didn't cite, for example, the 17
(6 1 14) /criteria/topicality/facet(focus) changed
(6 7 6) /criteria/orientation/methodology-technique
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no

#11. X paper that follows. [White: Did that evolve as you worked on the project?] Actually, when I started the project I wasn't really aware that there was a difference because the basic mathematical axioms, with the exception of one, are the same. It wasn't until we started looking at things fairly carefully that we recognized that there was a different axiom involved in poverty measurement than in inequality measurement. The basic difference is that poverty measurement defines an absolute threshold level for what they consider to be poverty, and the inequality measure doesn't do anything like that. 18
(6 1 14) /criteria/topicality/facet(focus) changed
(6 7 6) /criteria/orientation/methodology-technique
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no

#12. I didn't cite No. 12 by X, although I did cite a number of other things by him and I did look at this paper. I don't really have a firm reason for why I didn't cite this one. I cited one by X, X and X, is that it? It's by the same author, and my recollection is - yes, X - and my recollection is that's in Economica. Economic Journal? Okay. The reason is is probably the case is because it is more oriented towards poverty measurement as opposed to inequality measurement. Literally there were thousands of articles I could have cited on either inequality or poverty measurement. I did cite X because I thought he was important, but you can't possibly cite everything everybody has done. 19
(6 25 2) /criteria/prolific author/yes
(7 5) /DM strategies/satisfice
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no

#13. I don't remember seeing this paper by X. This is a hard journal to get ahold of, if it's even a journal. This may have been a book. Sometimes some of these books are very, very hard to get ahold of. They require doing things like inter-library loan, and that's typically not worth doing. [Wang: How do you think a paper published in a book compares with one published in a journal? Would you prefer to choose the one in a journal?] Yes. The ones in the journal are usually higher quality. It's interesting, because sometimes the ones in the books

are better because people have more liberty to say what they want to say, so they can be more creative. But really, what generally happens is books end up as a repository of the lower-quality papers. There are obvious exceptions to that. {One of the papers that we did cite, by X, there's a number of them, but the 1981 paper is a classic example of a paper that was in a book that is incredibly widely cited. In his case, people tend to only be able to find what he publishes in books because he doesn't publish [in journals]. People kind of have to drag the stuff out of him.} {belongs to 45} 20
(1) /quote
(6 9 3) /criteria/availability/ILL
(6 9 5) /criteria/availability/difficult
(6 10 2) /criteria/time(cost, effort)/too much(not worthwhile)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#14. I looked at this, but again, this is just too specific. This is on ranking Lorenz curves, and this is just too specific to what we were doing. This is telling whether or not the Lorenz curves are different for the different regions in the US, and that's just not that helpful. 21
(6 1 7) /criteria/topicality/too specific (narrow)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#15. I don't recall seeing this. It would have been very hard sometimes to get No. 15, simply because of the journal. The Italian journals are very episodic. 22
(6 9 5) /criteria/availability/difficult
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#16. I did look at the paper by X, but this only uses inequality in the sense that they introduce a Lorenz curve into the analysis, so it didn't really help with what I was looking at. 23
(6 1 3) /criteria/topicality/facet not match
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#17. I don't recall looking at this, and I think probably the reason I wouldn't have looked at it is because it is by a very minor person, X. X is a little better-known. [Wang: Do you know them in person?] I know them by reputation, I don't know them personally. 24
(5 1 2) /knowledge of/person/yes
(6 5 1) /criteria/quality/low
(6 24 1 1) /criteria/reputation(reliability)/person/low
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#18. I did look at the paper in the Journal of Economic Theory by X and X. It's just not relevant to what we were working on. 25
(6 1 2) /criteria/topicality/not relevant(useful)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#19. Journal of Econometrics. I did look at this, but

I don't have any recollection of what's in it. I'm looking at it now, but when I first looked at it, I didn't have any recognition. 26
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#20. I did look at this by X, but this is an empirical piece that looks at the role that international inequality basically has in development. It's not relevant to what we were working on. 27
(6 7 3) /criteria/orientation/empirical(cases)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#21. I did look at the one by X and X. The reason we looked at these measures is because these two individuals are colleagues of my co-author at X University. This is an empirical approach to actually measuring inequality, and we wouldn't have used it in a computational approach. So that wouldn't have been too relevant for what we were looking at. 28
(6 7 3) /criteria/orientation/empirical(cases)
(6 13 5) /criteria/relation/colleague
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#22. I don't recall this. 29
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#23. I don't recall this, and I think probably upon second thought, I wouldn't have been interested in it. 30
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#24. This guy X, I just haven't been impressed with his work at all. He's the one we talked about earlier. 31
(5 1 2) /knowledge of/person/yes
(6 24 1 1) /criteria/reputation(reliability)/person/low
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#25. I don't recall seeing this. But again, having looked at this more closely, I guess what would have happened, quite frankly, is that we're interested in understanding basically what everybody understands about ... Let me back up. I think we're interested in understanding the stylized facts about inequality at an empirical level that any working economist would be interested in. But we're not particularly interested in finding out just what causes inequality and what causes the differences in the various measures of inequality in the US. That's not relevant to what we were interested in. 32
(6 1 14) /criteria/topicality/facet(focus) changed
(6 7 3) /criteria/orientation/empirical(cases)
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#26. I've seen the paper by X. It's not relevant to what we were working on. 33
(6 1 2) /criteria/topicality/not relevant(useful)

(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#27. I've seen the paper by X. It's an interesting paper, but it just is not really relevant to what we were working on. I did look at the effect of inequality on fertility and ?, and it's very interesting, but it's just not relevant to what we were working on, so it wouldn't be appropriate to cite this. 34
(6 1 2) /criteria/topicality/not relevant(useful)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#28. I've seen the paper by X and X. I actually have copies of this paper. We cited other things by X and X, and we just didn't want to cite too much by these individuals. This is a lesser piece of work. The two that we cited in this area are the two main pieces of work. There is one in International Economic Review, and the other is in the Journal of Economic Theory. So they are more important pieces of work. [Wang: This one is later than the other two?] Yes, it's in a lesser journal. The other papers are much more basic to inequality measurement. 35
(6 24 2 1) /criteria/reputation(reliability)/Journal/low
(6 25 2) /criteria/prolific author/yes
(7 3) /DM strategies/dominance
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#29. I've seen the paper by X. This again is just one of these things that we could have cited. I cited other things by X. But this just didn't seem to be very appropriate to cite because, again, it's just a list of assumptions that will lead to one of these particular measures. We could have cited a number of things like this. 36
(6 25 2) /criteria/prolific author/yes
(7 3) /DM strategies/dominance
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#30. I don't recall this paper; I'm sure I've seen it, but I don't recall it. 37
(8 4 1) /decisions/reading/no
(8 5 1) /decisions/citing/no
#31. I've seen this paper, but again this is on econometric estimation of the Lorenz curve, which is something that we weren't interested in doing. [Wang: Did you cite these authors?] I don't remember citing X, I don't know who X is (known item pre-search). 38
(6 1 3) /criteria/topicality/facet not match
(6 7 3) /criteria/orientation/empirical(cases)
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
#32. I've seen this book of papers here, and this takes a very different approach to the problem we were looking at, so it wouldn't have been relevant to cite. Again, it is something that is very interesting, but it wouldn't have been relevant to

cite in the literature that we were working on. 39
(6 1 3) /criteria/topicality/facet not match
(8 4 2) /decisions/reading/yes
(8 5 1) /decisions/citing/no
Question by Wang (unintelligible) No. 18
Participant #006: Like I said, I'm very interested in these type of things, but this just isn't appropriate to cite. [Question by Wang unintelligible] Well no, the point of the paper that we actually ended up writing was to kind of merge two different literatures. This particular version of what these people are doing, or this particular version of the literature that we would be merging would be so far off kind of the mainstream that it wouldn't be worth doing. I'm interested in stochastic dominance in inequality. Stochastic dominance is important in the measurement of uncertainty, which I also happen to know a lot about. But inequality measures and uncertainty measures are very similar types of things, so that's why I was interested in seeing it. I'm sure what attracted me was the nexus between stochastic dominance and inequality measures. They are both areas that I am working on. I ranked that number 1 of the ones I saw. This is important for my overall research, but it's just not something that's relevant to this particular paper. 40
(6 1 13) /criteria/topicality/analogy
4. Reference for the paper. (8 + self select) 41
Participant #006: #1. The first one that you checked off, by X, "[Title]" the {previous paper that we cited by X, the one right above it, is considered one of the modern classics in the literature on the measurement of inequality}.{to 69} The reason why we cited the second paper, and it's not a paper, it's a book, is because he collected a number of studies that had done empirically what he suggested theoretically. We thought it was interesting to look at. {The first one is a classic}, the second one is a collection of readings, and there's a number of people that contributed things that followed up on his work from ten years before. I mean, { this paper by X revolutionized the way of looking at inequality measures.} So for example, the stuff by X and X that we cited later on is just a continuation of what X had done. {50} 42
(6 7 13) /criteria/orientation/review-summary
(6 22 3) /criteria/classic(founder)/follow to classic
(7 6 8) /DM strategies/chain/original-followup
(8 5 1) /decisions/citing/no
#2. The paper by X and X, the reason we cited that is because we looked at three basic actions that our inequality index would have to have, three alternative definite actions that it could satisfy. One of them is this notion of intermediate inequality, and they are one of the few people that have worked in that area, so that's why that one got cited. {50} 43
(6 1 5) /criteria/topicality/facet match

BEST COPY AVAILABLE

(8 5 1) /decisions/citing/no
#3. X and X, that's a classic work on aggregation, and it would be necessary to cite that. Because what we're doing is merging the aggregation and inequality literature. {50} 44
(6 1 5) /criteria/topicality/facet match
(6 14 2) /criteria/publicity/well(wide) known
(6 22 2) /criteria/classic(founder)/yes
(8 5 1) /decisions/citing/no
#4. This paper by X on "[Title]" is one I referred to earlier as a classic paper on integrability of demand systems, and so it would be necessary for us to cite that as well. [see 20 for coding] {50} 45
(6 14 2) /criteria/publicity/well(wide) known
(6 22 2) /criteria/classic(founder)/yes
(6 24 2 1) /criteria/reputation(reliability)/Journal/low
(8 5 1) /decisions/citing/no
#5. The book by X, X, and X is just a classic mathematical reference on inequality, not necessarily on inequality measurement, but on inequality. And so a lot of the mathematical theory that is relevant to analyzing inequality would have been first studied by these three guys. I even have a copy of this on my shelf. {50} 46
(6 7 6) /criteria/orientation/methodology-technique
(6 7 7) /criteria/orientation/theoretical-conceptual
(6 14 2) /criteria/publicity/well(wide) known
(6 22 2) /criteria/classic(founder)/yes
(8 5 1) /decisions/citing/no
#6. X, X, and X, they're the first people to do empirically some of the things that we're suggesting. They're not doing what we're suggesting, but they're doing something that's close to what we're suggesting. {50} 47
(6 1 5) /criteria/topicality/facet match
(6 7 3) /criteria/orientation/empirical(cases)
(6 14 2) /criteria/publicity/well(wide) known
(6 22 2) /criteria/classic(founder)/yes
(8 5 1) /decisions/citing/no
#7. The X paper that you've checked is the first paper that looks at non-linear aggregation of demand systems, and that is basically one of the ideas that we're advocating. {50} 48
(6 1 5) /criteria/topicality/facet match
(6 14 2) /criteria/publicity/well(wide) known
(6 22 2) /criteria/classic(founder)/yes
(8 5 1) /decisions/citing/no
#8. The X and X piece, which is the last one, the reason we cited that is they had a fundamental mathematical result that was important to resolving one of our theorems. {50} 49
(6 1 5) /criteria/topicality/facet match
(6 7 3) /criteria/orientation/empirical(cases)
(8 5 1) /decisions/citing/no
Wang: Can you recall where you got these papers?
Participant #006: Okay. {#8} The X and X, I didn't get that, my co-author was aware

of that. {#7} X, I've known about this for a long time; virtually everybody in the department would know about this X paper. Not as well known as {#3}the X and X paper, but the X and X paper that you checked earlier simply applies the theory that was developed in the {#7} X paper in 1975. I've known about that for a long time; I have copies of that in my file, of course. {#6} X, X and X is a very well-known empirical study. {#5} X, X, and X, again, it's a very well-known book in mathematics. {#4} The X paper I've already talked about. Anybody who's in demand analysis would know that paper. {#3} X and X, again, it's a classic. {#2} X and X, I think I got that...I don't know how I got that. I'm pretty sure I read X's paper in the Journal of Economic Theory, and then I saw the reference to his work with X. {#1} The X book, I just knew about that. 50

Wang: If you can select a few more which represent another set of criteria (unintelligible) other than the ones you just told us about, specific to... Participant #006: I can think of two basic criteria: works that are relevant to either the basic theory of inequality measurement, or the basic theory of demand systems aggregation, which are the two areas of literature that we're trying to merge. Apart from that, works that are touching on aspects of the problem that we're talking about, that have been something like we're talking about doing but either have done it in an ad hoc fashion, which typically turns out to be incorrect when you try to think about it in an axiomatic fashion, or works that have tried to do something empirical along the lines that we're suggesting. So let me go through the first ones, the ones that are the basic theory pieces. 51

The first one, "[Title]," by X; the two papers by X and X; the paper by X, the paper by X and X; the paper by X, X, and X; the paper by X, "[Title];" the paper "[Title];" X, X, and X; the two papers by X; the paper by X; the first paper by X; the paper by X; the book by X and X. The book by X and X contains a basic modern mathematical theory of inequalities. The two papers by X; the book by X; the book by X. And that would be it; they would have to be cited. 52

Now the ones that are empirically related to what we've done. The paper by X, et al; the paper that actually motivated a lot of what we're doing is the paper by X. It was actually in thinking about X that we realized that he had done something that was wrong. The paper by X and X; the paper by X, X, and X as well; the paper by X. The paper by X, I just noticed, came out in Review of Economic Studies. I just noticed that that was wrong. [White: In all of these, you clearly did adhere closely to the idea of the quality of the journals.] That's

right. We were trying to write for a particular audience, and if you go through the references, most of those papers will be at that level. 53
White: {#6} The X one is one of the few that you cited that's in a book, and yet you mentioned you did that because it is empirically close to what you have done. That goes counter to your, or is that an example of a very good article in a book Participant #006: Yes, it is. This is an extremely well-known paper. It's one of these papers that is very frequently cited but a lot of people haven't seen it. Now, you see it cited all the time but when you talk to people, a lot of them haven't seen it. It's hard to get a copy of it. I have a copy of it. I wouldn't have cited it unless I'd seen it, or if one of us had seen it. I haven't seen, for example, the thing by X, "[Title]," and one of the other ones in here. The X one. I've never seen the X one. But [co-author] assures me that he's seen those. 54
White: The age of the literature you've cited is really extraordinary. You have some things that go back to 1920. Participant #006: There are things that are earlier than that. X should be there at the turn of the century, and there's stuff from the 1880s. The interest and measure in inequality is very, very old. 55
White: So that the literature does have a long lifespan. Participant #006: Right. The work by X is considered, is just basic in this area, one of the primary mathematical properties that all of these indexes possess is something called X's [named concept]. That just means that if you transfer money from a richer individual to a poorer individual, inequality should go down. In other words you should be better off by doing that. 56
Wang: Did the paper go to Econometrica? Participant #006: Well, we sent it to Econometrica. The acceptance rate in economics journals is incredibly low. At a place like Econometrica it would be eight or nine percent. That's routine. Even regular journals it's 25 percent. I can share something with you. I just got something from the editor of the Economic Journal, which is another well-known journal, to review, and he was saying, "We receive a large number of papers submitted to the journal. Because of constraints on space we have to reject well over 90 percent of them. A journal like Econometrica, for example, these levels of journal, Journal of Economic Theory and whatnot, out of something like 80 or 90 publications I have, I have maybe five or six at that level. It's very hard to get in those. 57
White: One that you mentioned is (?) The X paper was published in a relatively good journal? Participant #006: No, well, yes, it was. But that can happen. There is a lot of randomness that's involved in this. I mean, it's not uncommon for a paper to be rejected in a lesser journal and then to be

accepted in a much better journal. For example, the Journal of Economic Theory, I have something that's just come out in the last month or so, and that was rejected at Economic Journal before that. It just happens. There's a fair amount of randomness that's involved. It really depends on who the referees are and whether the referees are competent. Sometimes the referees just aren't competent, especially when you're doing something that's very different than what anybody's done before. There's a lot of room for error. If it's just a standard straight-out piece, you usually get treated pretty fairly, but if you're thinking in a very different fashion than people have thought before, that's very, very hard. It's hard to get accepted, and what ends up happening is usually things that are really new will end up sometimes being accepted in lesser places or come out in books. So those are the exceptions where you will see things that are really better in books or in lesser journals. One of my co-authors in some of my other work is a classic example. He has a paper in a very obscure journal that has become a classic. X. He has a paper in a very obscure journal; he published it in 1982 and the best journals started publishing things along the lines of this in the late 1980s and then somebody realized that what they were doing was a special case of what he had already done. So it's not uncommon for that to happen. 58

(6 15 6) /criteria/spectrum of JN/periphery

White: It would also be possible, because it was in an obscure journal, to just get lost. Participant #006: Right. That happens. There is a lot of examples of that in economics. I imagine it's pretty true in other fields, but there is definitely a lot of examples of that in economics, for something that's so advanced just to be overlooked. The work of X is a good example. The paper that we referred to, "[Title]," I know virtually all of the really good people in this field, and nobody understands completely what he has done. To have him try to explain it is pointless, because nobody can understand what he is talking about. He's one of these people that everybody's sure he's right, but nobody is exactly sure why it's right. 59

(6 12 2) /criteria/authority/yes

White: You mentioned there is an enormous body of literature about inequality in demand systems. Do you feel that you have a good grasp of that, or are you missing something important? How concerned are you about it? Participant #006: I think I have a good grasp of the literature, of what's out there. It would be nice to understand some of the things that X has done that people basically have to take at face value. The math is right, but it's not really clear just exactly why he did some of the things that he did, and why it works. You can follow it step by step, but there is not a clear, intuitive reason as to why what he has done is correct. That's important

in economics. It's not the same in mathematics because there should be behavioral reasons for why these things are correct. In some of his stuff it's just not clear why that's the case. You can follow the math, but it's very hard to understand what it means. 60
White: For these things that are appearing in the more obscure journals, do you find that you just sort of, are you truly trying to read as broadly as you possibly can just to find those, or does it really not matter to you, that if they're good somebody else will pick it up? Participant #006: Well, I should say I read as broadly as I possibly can, but the fact is that I work on so many things it is very hard for me to do anything other than just basically skim a subset of journals that I rely upon. I will typically always look at the Journal of Economic Theory, I will look at International Economic Review, American Economic Review, Economic Journal, things like that, to see things that I'm interested in. Then I'll skim that, but I just don't have the time to sit down and look at everything. 61
White: So for some of these, like the Italian journal, the chances are if you don't read that article it is not going to be catastrophic. Participant #006: That's right. There are exceptions to that. Again, there is a paper we didn't cite by X, a paper he published in an Italian journal, that was very obscure for a number of years but turned out to be very important. 62
Wang: Do you think somebody will pick it up? When people read your article they will tell you there is something that you didn't cite. Participant #006: Actually, that is one of the reasons we cited a number of the things in here; we didn't want to be told we had neglected to cite certain people. So there are people in here, for example, X is one of these people, we anticipated being a referee so we didn't want to, we didn't think his work was all that valuable in some ways but at the same time we didn't want to alienate the person by not citing his work. 63
(6 26 3) /criteria/judge/journal referee
White: So there are a number of those that you did more for just paying lip service to? Participant #006: The paper by X, that's for sure. I can go through and tell you the ones we did that on. The papers by X. The first paper by X, and sort of the second paper, because the second paper has a legitimate role. The paper by X and X. The paper by X. One of the papers by X, the one in Quarterly Journal of Economics. The paper by X, and the paper by X. 64
Wang: Quite a few. When you write a paper you have your audience in your mind, and you also have the journal referees in your mind. And also the reputation of the journal itself, right? Participant #006: Right, very clearly. You learn this by

experience. I've had a number of times when I've been told, well, you didn't cite such-and-such. You should have cited this, or whatever. When people say that, you look bad, and it looks like you don't know the literature. There is no reason to cite the paper by X, the first one, because it doesn't really fit, but at the same time you want people to be aware of the fact that you're aware of that. That you've read that. 65

Wang: You are teaching your students that for their own career... Participant #006: This is something they have to learn themselves. 66

White: Actually, they are learning it very well! It's interesting to see the difference in the way they are approaching the literature after they're going through a project than when they did it initially. Participant #006: One of the advantages of research assistantships is that they get to spend a lot of time with professors on research and sometimes students who don't have the RAs don't have that experience. 67

White: It's part of the socialization. It should happen for every student in the doctoral program; they should get some sense of the relative quality of the literature. Participant #006: This program is really oriented toward training students to publish with good professors. We have a group of people in the department, they have shortcomings in a lot of ways, but that's not one of them. They are very prolific. 68

Peiling's note during coding:{51} The user's classification of the cited literature:(1) Theories related to (a) (b), (2) Empirical evidence to their proposed theory.{20, 45, 54, 58} Although papers in a book are considered low quality, not peer reviewed, but more freedom to new ideas. Some authors do not publish when solicited; lesser competitive journals may also be a forum for new ideas. {63, 64, 65} 6 papers cited because of the sense of possible referees. 69

Appendix E. DEFINITIONS OF CRITERIA

DEFINITIONS OF CRITERIA

Note: This appendix includes the criteria identified in the 1992 study, the criteria identified for this report's Chapter 4's task/phase-based analysis, and the additional criteria and modifications to the previous criteria made in the analysis for Chapter 5. The latter are related to the 1995 study. The date and number in parenthesis following the criterion indicate the first appearance of that formulation of the criterion.

Actual Quality (1995, ch. 3)

The actual quality is judged by reading the document.

"Another important paper, I think. ... he lists them pretty neatly, nicely, and that's where that paper comes in. ... The 1992 is a better test, though some of the specifics are more clearly given in the 1990 paper because this is sort of a review paper. They both came in handy."

Audience (1995, ch. 5)

The level at which the publication is written, usually indicated by reference to intended audience.

"The next one is actually in the *Economist*, which is a main sort of public magazine and again I need it to get quotes from the popular press that sort of say, this issue is important. This [situation] is what is happening in the real world. I am not really going to it for anything that is of substance to the paper."

Authority (1992)

This criterion acknowledges a relationship between the document's author and readers in the field in which readers generally acknowledge the individual's accuracy, breadth of knowledge, creativity, for example. An author has authority in the field, and his/her ideas and opinions are important to others in the field.

"The paper that we referred to, [title]--I know virtually all of the really good people in this field, and nobody understands completely what he has done. To have him try to explain it is pointless, because nobody can understand what he is talking about. He's one of these people that everybody's sure he's right, but nobody is exactly sure why it's right."

Availability (1992)

Physical accessibility of the document.

"Actually, there are two reasons. One, it's difficult to get a dissertation ..."

Classic/Founder (1995, ch. 3)

The document is recognized in the field as the first substantial work on a topic or technique.

"It's a famous paper that's been around for a long time. And people enhanced their technique, but that's still an important paper. ... I should, have to, cite the founders of the technique."

Cognitive Requisite (1995, ch. 3)

A judgment about whether the user has the required knowledge or background to understand a document.

"This is an example of the paper from agronomy side. It is in a sort of multidisciplinary journal. I needed it in order to get some idea of what are the facts and influences upon soil quality. If it had been too scientific [vs. economics], I probably would not have read it, because it would have taken too much for me to understand."

Content (1995, ch. 5)

The nature of the materials included, *e.g.*, data, methodology, theory.

"The book by [authors] is just a classic mathematical reference on inequality, not necessarily on inequality measurement, but on inequality. And so a lot of the mathematical theory that is relevant to analyzing inequality would have been first studied by these three guys."

Credential (1995, ch. 3)

This criterion usually applies to a paper authored by person citing or using a document and is included primarily to support his own appearance of expertise.

"If he weren't a co-PI, we might not put the standard reference in. ... We want to show that we have the expertise in production modelling more generally."

Depth (1995, ch. 5)

The extent of treatment about the subject within a document.

"The first one--I did look at that. That came from the Farm Bureau Federation. That was the one that, in the end, was too tendentious to use. It didn't really have any analytical information, and it had skeptical ideas about pesticide policy. And, they may be right, but they just didn't have any scientific contribution."

Discipline [Formerly called Subject area] (1992)

The broader subject area or branch of knowledge to which the document belongs. Agricultural economics topics are also studied by other disciplines, such as political science,

engineering, environment, environment management, finance, etc. A user may want to include or exclude documents from different disciplines.

"I would not look at this. It's very ecological oriented."

Expected Quality (1992)

The estimate of the goodness of a document. Expected quality may really indicate a positive or negative attitude that the user has toward a document before consulting the full document.

"It's put out by the National Academy of Sciences. This is good."

Journal Spectrum (1995, ch. 5)

An assessment of the publishing journal, usually in terms of major or minor.

"The next one is by [authors]. This [article] is a short journal article in one of the main economic journals, *Review of Economic Studies*."

Judge (1995, ch. 3)

Anticipation of the perception of the person who will read and/or evaluate the product [*i.e.* the product being written, not the document being considered by the respondent].

"If I ranked it [journal] intellectually, it would not rank very high. But, if I think strategically, it's important to include it, because people on evaluation committee likely read and contribute to it and possibly edit this journal."

Norm (1995, ch. 3)

Perceived expectation or practice in the field for the finished product. For example, in a funding proposal, specifics are rarely necessary; in a dissertation, a complete literature review is expected.

"This prospectus is short anyway, and I am trying to keep cites down to some degree. I didn't feel the need at this point to go beyond the national journals."

Novelty (1992)

Whether the document or its content is new to the user or whether it has been seen before and is a known item; or its content or information is known to the user.

"[Title] Well, this paper actually was presented, I think, at one of the annual meetings and I wanted to see if they had done anything new in this, and they didn't. Again, I think this is more of a general interest. Actually, this is also revisional price expectations, that's probably why I kept it. There was nothing, other than complicated estimation procedures, not much in terms of expectations."

Orientation (1992)

At which intellectual level the document is written and for which audience it is intended. Is it an academic or scholarly paper? Is it a theoretical or an empirial study? Is is a methodology paper or a review of the topic? Is it for academic or for the practitioner? Note
The category was split into Audience; Content; Orientation1 (Empirical, theoretical); and Orientation2 (Qualitative, quantitative).

"I need it to get quotes from the popular press that sort of say, this is an important issue."

Orientation1 (1995, ch. 5)

A distinction between an emphasis on a theoretical or empirical presentation.

"I did look at this [document] by [author], but it is an empirical piece that looks at the role that international inequality basically has in development. It's not relevant to what we were working on."

Orientation2 (1995, ch. 5)

A methodological distinction between a qualitative or quantitative approach.

"This one I used. I actually use some of their qualitative estimates from this one. They looked at a scenario where you would take something like 25% reduction of use of all fertilizers or chemicals..."

Peer-Reviewed (1995, ch. 5)

Evidence of critical evaluation by a knowledgeable judge, *e.g.*, an article published in a refereed journal. A document cited by another researcher is also considered peer-reviewed.

"The journal paper is much more credible. Conference papers, everybody gets papers at a conference, the quality of a conference paper is usually much lower than journal papers. In the conference paper you are limited to 10 pages and, if you haven't been to the conference and sat through the presentation, you don't know what objections or worries others might face. And usually [the journal article's] a much more expanded version and a lot more carefully done and refereed a lot more vigorously. In a conference, you can't sit there and do the math or the algebra, but you know in a refereed journal somebody hopefully has looked through it and accepted it."

Prolific Author (1995, ch. 3)

The fact that an author has published many articles in a topic is considered, even though he may not be regarded as a founder or have written a classic work.

"I've seen the paper by X and X. I actually have copies of this paper. We cited other things by X and X, and we just didn't want to cite too much by these individuals. This is a lesser piece of

work. The two that we cited in this area are the two main pieces of work. There is one in *International Economic Review*, and the other is in the *Journal of Economic Theory*. So there are more important pieces of work. ... The other papers are much more basic to inequality measurement."

Publicity (1995, ch. 5)

Indication that the publication has received some extraordinary acknowledgment.

"X is definitely a book that is used in many theory classes. It's a very famous book."

Recency (1992)

The comparative newness of a document with regard to the respondent's topic; however, recency is relative and does not directly correlate with date of publication. This criterion depends on topics and can be overridden by other factors.

"It's a little old, well, not much old."

Relation [Formerly called Relation/Origin] (1992)

"Recognition of an author or an organization brings in a relationship between the user and the source of the document. The document becomes useful to this particular user because of his/her particular situation or position (Wang 1992, 136)."

"I agree with the findings of the paper, and besides some of the people [authors] here are from X, where I got my Ph.D. So this is one of the reasons. The other is that the fellow X is now a professor here in the department. I know there are a lot of papers on pesticides."

Special Requisite (1992)

Requires a special skill, such as language, or equipment, *e.g.* microfilm reader.

"It's good. No, it's in German. The title is extremely interesting."

Standard Reference (1995, ch. 3)

The concepts in the document are well-known to the field, or the document is used as a textbook or standard reference.

"The first one is X and X. It's one of the main textbooks in my field, environmental policy, especially with an emphasis on regulation. And, because it is one of the main textbooks, it's one that I would often go to first to see what is sort of the main models in this field, what is the way that most people have done on this up to now, so, it is always a very good starting point for just about anything I do in policy regulation. Now, it's often the case that I can be pretty sure even in this paper that all I used it for in my actually research is to say that this is what we know up until now."

Target Journal (1995, ch. 5)

Consideration of the characteristics, level, demands of the journal in which the research product will be published.

"We cited the *Journal of Economic Theory* because that was the level at which we were trying to write the paper. In economics there are all different kinds of levels of journals, and the theoretical level that we were aiming at is most closely matched by the *Journal of Economic Theory*, *Review of Economic Studies*, and *Econometrica*. The paper that we actually wrote was ultimately submitted to *Econometrica*. So, when we picked out references, we tried to stay in that group. It is a little bit of gamesmanship in a way, to be citing the right people."

Time, Effort [Formerly Time or Reading Time] (1992)

The respondent's sense of whether or not she can afford the time or effort required to read a document.

"It has 266 pages. So, it's a book. ... It would be useful reading if I had unlimited time."

Topicality (1992)

What the document is about as the respondent sees it with respect to his tasks at hand, whether or not the topic of a document is related to the topic of the respondent's project.

"I think part of this is relevant, like consumption function."

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Document Selection and Relevance Assessments During a Research Project	
Author(s): White, Marilyn Domas and Peiling Wang	
Corporate Source: College of Library and Information Services University of Maryland, College Park, MD 20742	Publication Date: January 1997

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

☑

Check here
For Level 1 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all **Level 1** documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all **Level 2** documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

☐

Check here
For Level 2 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at **Level 1**.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please Signature: Marilyn D. White	Printed Name/Position/Title: Marilyn White, Associate Professor	
Organization/Address: College of Library and Info. Services University of Maryland, College Park, MD	Telephone: (301) 405-2047	FAX: 314-9145 (301)
	E-Mail Address: whitemd@wam.umd.edu	Date: 3/5/1997

ERIC Full Text Provided by ERIC

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC / IT
Center For Science & Technology
Room 4-194
Syracuse University
Syracuse, NY 13244-4100

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to: