MODELING CHANGES IN END-USER RELEVANCE CRITERIA:  
AN INFORMATION SEEKING STUDY

DISSERTATION

Presented to the Graduate Council of the  
University of North Texas in Partial  
Fulfillment of the Requirements  
For the Degree of  

DOCTOR OF PHILOSOPHY

By

Judith Ann Bateman, B.S., M.A.  
Denton, Texas  
May, 1998
MODELING CHANGES IN END-USER RELEVANCE CRITERIA:
AN INFORMATION SEEKING STUDY

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Judith Ann Bateman, B.S., M.A.
Denton, Texas
May, 1998
Bateman, Judith Ann, **Modeling Changes in End-user Relevance Criteria: An Information Seeking Study.** Doctor of Philosophy (Information Science), May, 1998, 206 pp., 20 tables, 11 figures, references, 77 titles.

This study examines the importance of relevance criteria in end-user evaluation of valuable or high relevant information. Changes in the importance of criteria as users moved through the six stages of information seeking proposed by Kuhlthau (1994) and as users selected, obtained and read information sources were studied. A model of high relevance was proposed and confirmed using criteria that were rated as most important.

The research was conducted in two phases. In the first phase, 210 graduate students at the University of North Texas were surveyed by mail. These students were sent an instrument that instructed them to rate 40 relevance criteria by importance in their selection of the most valuable information source for a recent or current paper or project. Respondents were also asked to report their information seeking stage and whether they had obtained or read the information.

The second phase of the study involved seven graduate level classes. Respondents evaluated the importance of 40 relevance criteria as they searched for, obtained, and read highly relevant information for a research paper. They were also asked to judge their stage of information seeking. Search diaries were also used to study the relevance judgment process, the information seeking process and the use of IR systems. The importance of these criteria did not change as the respondents moved
through the information seeking stages or as they searched for, obtained and read
information.

Eleven criteria rated as highly important by respondents in the survey phase were
used to propose a model of high relevance. This model was confirmed using factor
analysis and second-order factor analysis. Three constructs were used in the model:
Information quality, Information credibility, and Information completeness can be used to
explain 48% of high relevance for these respondents. Three additional constructs:
Information availability, Information topicality, and Information currency are suggested
as important constructs of high relevance for graduate students. These constructs will
require additional criterion terms on the criteria evaluation instrument.
ACKNOWLEDGMENTS

I would like to express my thanks to Linda Schamber for her help as my committee chair and to the members of my committee: Amanda Spink, Randy Schumacker, and Kathy Swigger for their help and support. I also would like to thank Carol Kuhlthau, Rutgers University and Carol Barry, Louisiana State University for their helpful critique and encouragement. I would like to thank my husband Douglas Bateman and my many friends and colleagues for their help and support throughout the dissertation process.
TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ viii

LIST OF FIGURES ...................................................................................................... x

Chapter

1. PROBLEM STATEMENT ......................................................................................... 1

   Introduction
   Problem Statement
   Research Questions
   Importance of the Problem
   Relevance Related to Information-Seeking and Information Retrieval Models
      Information Retrieval Models
      Information Seeking Models
   Relevance Within IR and Information Seeking Models
   Structure of the Proposed Research
   Summary

2. REVIEW OF THE LITERATURE ........................................................................... 27

   Introduction
   Relevance as Situated in IR Studies
      The Archetypal or Cranfield-like Approach
      Probabilistic Approach
      Information Retrieval-Expert System Approach
   Relevance Definitions
   Relevance Studies
      Relevance Judgments after the Search Interaction
      Studies of End-user Relevance Criteria
   Information Seeking Theory
   Information Seeking Studies
      General Studies
      Information Seeking Behavior Independent of a Specific Information Problem
      Longitudinal studies of Information Seeking for a Specific Problem
3. METHODOLOGY

Introduction
Assumptions
   Assumptions Regarding Relevance
   Assumptions Regarding Information Seeking Behavior
End-user Relevance Criteria Evaluation Form
Characteristics of the Data
   Survey Data
   Class Data
Data Analysis
   Survey Data
   Class Data
Reliability Issues
Validity Issues
Limitations
Summary

4. DATA ANALYSIS AND RESULTS

Introduction
Data Analysis of Specific Research Issues
   Question One: What are the Relationships between a Group of 40 End-user Relevance Criteria in Regard to their Degree of Importance?
   Question Two: At What Stages of Information Seeking do Respondents Employ IR Systems or Information Resources?
   Question Three: At What Stages of Information Seeking do Respondents Search for, Obtain and Read Information?
   Question Four: Does the Importance of End-user Relevance Criteria Change as the Respondent's Information Seeking Stages Change?
   Question Five: Does the Importance of End-user Relevance Criteria Change as Respondents Search for, Obtain and Read Information?
Factor Analysis
Description of Changes in Criteria Importance using the Motion Index
   Class Data Described by Search Stage
   Survey Data Described by Search Stage
   Class Data Motion Indexes Between Searching, Obtaining and Reading
   Comparison of Class and Survey Means
Analysis of Class Qualitative and Information Seeking Data
Criteria Reported on the Criteria Evaluation Form
Evidence of Multiple Search Sessions
Changes in Topics
Changes in Information Sources
Trends in Search Diary Reports
Affective Issues in Information Seeking
Information Resources
Changes in Topic and Focus
Summary

5. CONCLUSIONS

Introduction
Question One: What are the Relationships between a Group of 40 End-user Relevance Criteria in Regard to their Degree of Importance?
Question Two: At What Stages of Information Seeking do Respondents Employ IR Systems or Information Resources?
Question Three: At What Stages of Information Seeking do Respondents Search for, Obtain and Read Information?
Question Four: Does the Importance of End-user Relevance Criteria Change as the Respondent's Information Seeking Stages Change?
Question Five: Does the Importance of End-user Relevance Criteria Change as Respondents Search for, Obtain and Read Information?
Implications for Future Research
Summary

APPENDIX A: End-User Relevance Criteria Evaluation Form

APPENDIX B: Search Diary

APPENDIX C: Activities that are Characteristic of the Six Search Stages

APPENDIX D: General Instructions for Class Respondents

APPENDIX E: Verbal Instructions to Class Respondents

APPENDIX F: Ideal Information Source Form

APPENDIX G: Demographic Form

APPENDIX H: Survey Letter, First Mailing

APPENDIX I: Information Criteria Evaluation Survey, First Mailing
LIST OF TABLES

Table 1  Relevance Factors Suggested in the Literature........................................... 58
Table 2  User-Defined Relevance Criteria................................................................. 61
Table 3  Discipline of Survey Respondents............................................................. 79
Table 4  Undergraduate Discipline of Class Respondents......................................... 80
Table 5  Information Seeking Stages Reported by Respondents............................... 82
Table 6  Criteria Evaluation Forms Returned by Class Respondents......................... 83
Table 7  Criteria Rated as Important by Survey Respondents................................. 104
Table 8  Information Resource Use Reported at Each Stage of Information
Seeking by Class Respondents.................................................................................. 107
Table 9  Information Seeking Stage of Class Respondents when Searching,
Obtaining or Reading................................................................................................. 108
Table 10 Inter-item Correlations and Reliability for Constructs for High
Relevance with Missing Data...................................................................................... 115
Table 11 Inter-item Correlations and Reliability for Constructs for High
Relevance with Mean Substitution for Missing Data............................................... 116
Table 12 First-Order Factors....................................................................................... 117
Table 13 Factor Inter-correlations.............................................................................. 118
Table 14 Second-Order Factors.................................................................................. 118
Table 15 Information Seeking Stages with Largest and Next Largest Motion
Indexes, Class Data ........................................................................................................... 120

Table 16 Information Seeking Stages with Largest and Next Largest Motion
Indexes, Survey Data ........................................................................................................... 122

Table 17 Criteria Mentioned by Class Respondents on Criteria Evaluation
Forms ................................................................................................................................... 126

Table 18 Number of Search Sessions Conducted on Each Information Resource
Class Respondents .............................................................................................................. 130

Table 19 Information Seeking Stages of Class Respondents Reporting a Single
Change in Topic Focus ........................................................................................................ 134

Table 20 Information Seeking Stages of Class Respondents Reporting Two
Changes in Topic Focus ...................................................................................................... 134
LIST OF FIGURES

Figure 1 Early Model of Relevance in Information Retrieval .......................... 12
Figure 2 "User-Driven" Model of the Information Retrieval Process .................. 14
Figure 3 Linguistic Level of Information Generation .................................. 15
Figure 4 Cognitive Level of Information Generation ................................ 16
Figure 5 Modification of Major Elements of Information Retrieval .................. 16
Figure 6 Relevance Within the User's Contexts ...................................... 17
Figure 7 Stratified Model of IR Interaction ........................................ 18
Figure 8 Broad Model of the Information Seeking Process .......................... 20
Figure 9 Model of the Information Search Process .................................. 22
Figure 10 Elements of the Interactive Search Process ............................... 23
Figure 11 High Relevance Second-Order Factor Model ............................... 113
CHAPTER 1

PROBLEM STATEMENT

Introduction

The goal of this exploratory research was to increase understanding of end-user relevance judgments made at different stages in the process of seeking information to resolve a particular information problem. This study examined the importance of the reasons or criteria (relevance criteria) employed by graduate students as they judged information to be highly valuable or relevant in resolving a complex information problem. In addition, the study examined the meaningful relationships between relevance criteria rated as important by graduate students, the constructs measured by these criteria, and high relevance.

One widely accepted operational definition of relevance as it relates to information science is "the user's decision to accept or reject information retrieved from an information system" (Schamber, 1994, pg. 3). In this study highly relevant information sources were operationalized to the respondents as the most valuable source or sources of information for a current or recent research project or paper. Because many graduate students who participated in the study were enrolled in library and information science classes, the term "relevant" was not used. Relevance is a central and controversial concept in the study of IR systems and it was likely that these students would be familiar with the term as it is used within library and information
science, which could have been a potential source of bias. Although the term valuable may not be entirely synonymous with high relevance, this term has been frequently used in relevance research to operationalize the concept of relevance to respondents. Highly relevant information may not always be valuable to respondents, particularly if topicality is the main criterion for determining high relevance. However, information respondents judge as the most valuable in resolving an information problem should also be judged by the same respondent as highly relevant.

In this thesis only high relevance judgments were examined. However, the definition of relevance was extended to include the broader context of information seeking. Within this context respondents rated relevance criteria importance as they searched for, obtained, and read highly valuable or relevant information. Although some respondents used IR systems to obtain most or all of their information, other methods of searching for information including browsing, the Internet, and recommendations of others were also employed by respondents. For the purpose of this study the term information source will mean a document, document representation, a personal communication, World Wide Web (WWW) site or page etc. as opposed to the broader definition of this term (i.e. an online search, a library, the Internet etc.). Information resource will refer to IR systems, people and links that respondents used in their process of locating information sources. In the context of this research an end-user is defined as an individual who has an information need or problem and who is employing an information resource to attempt to resolve this problem.

This study examined the importance of relevance criteria and the information seeking process for end-users who were searching for a real information problem using
IR systems and information resources that were available to them. Relevance is currently perceived as a complex phenomena which can not be well understood without examining the judgments of individuals who are using IR systems or resources to resolve their own information problem (Ellis, 1996; Schamber, 1994). Schamber et. al. (1990) in a discussion of relevance research and literature conclude that relevance is dynamic, multidimensional, situational, and dependent on the experience and cognition of each individual within the context of his or her information problem and situation.

This study investigated high relevance judgments by examining changes in the importance criteria or reasons that end-users employed when making these judgments as they moved through an information seeking process. These criteria or reasons will be referred to as relevance criteria. Changes in the use and importance of relevance criteria were related to the respondent's progression through the six stages of information seeking proposed by Kuhlthau (1994). Kuhlthau's model was selected because it proposes patterns of user information searching or seeking for a particular information problem. These patterns are presented as stages of information seeking characterized by sets of actions, thoughts, and feelings experienced by the end-user. The six stages are identified as task initiation (initiation), general topic selection (selection), early exploration of the general topic (exploration), focused formulation of the topic (formulation), information gathering that supports the specific topic focus (collection), and search closure (closure). Other models of information seeking have identified strategies that individuals use whenever they purposefully seek information to resolve an information problem or need (Ellis, 1992; Leckie, Pettigrew et. al., 1996). These models do not examine the specific
information seeking process for a particular information problem at the level of detail that is represented in Kuhlthau’s model.

This study was based on the assumption that when respondents’ thoughts, feelings, and actions shift as they experience different stages of information seeking it is very likely that the reasons or criteria they employ to select information will also show changes in use and importance. The thoughts, feelings, and actions that characterize each stage are an important part of the respondent’s situation and consequently are likely to affect relevance judgments. It was also assumed that respondents might not actively search for and evaluate information at all stages of information seeking.

Problem Statement

The definition and measurement of the concept of relevance is a central issue in the evaluation of information retrieval (IR) systems and in the study of information seeking behavior. The methodologies, assumptions, and results of studies evaluating IR systems are determined by how relevance is defined and operationalized. Much has been theorized about the complexity and dynamic character of relevance judgments (Schamber et. al., 1990; Schamber, 1994) but most research has continued to measure relevance as a single variable (relevant or not relevant) and at a single point in time (after the user receives results from an IR system). This study posits that relevance judgments can be better understood by measuring the importance of end-user defined relevance criteria in regard to these judgments throughout the process of seeking information to resolve an information problem. Previous relevance research has not studied importance ratings of relevance criteria by end-users or the relationships between these criteria and high
relevance. The identification of the criteria that are rated as most important by end-users contributes to our understanding of the relevance judgment process and how IR system design can best support that process.

The use of some relevance criteria may depend on the amount of information that is available to the user in regard to the item being judged. A user who is judging a full-text document has potentially more information to use in making a relevance judgment than a user who is judging only the citation for a document. This variation in available information about the information source was partially accounted for by asking respondents whether they had obtained or read the information sources that they were referring to when they reported the importance of the relevance criteria. The use of the Internet and full-text IR systems allowed some respondents to combine searching for, obtaining and reading an information source, and these respondents did not experience the variation in available information about information sources that is inherent in the structure of traditional IR systems.

Relevance judgments have been primarily studied in the context of the user's interaction with an IR system and have rarely been examined within the context of the information seeking process. Spink (1997) has examined how end-users and intermediaries employ citation and index term relevance judgments to make decisions to reformulate searches and print and accept citations for a complete online search process. Kuhlthau (1994) has studied the entire information seeking process for a particular information problem that is resolved over a period of time and may involve several information search iterations. However, it is not known how user relevance criteria change throughout the information seeking process or how important each criterion is to
the end-user at any point in this process. Further research into how individuals evaluate information at various stages in the information seeking process for an information problem is needed to help bridge the research gap between snapshot studies of information seeking at one point in time and general information seeking studies.

Only a handful of studies have attempted to examine relevance judgments for multiple information seeking sessions in regard to a particular information problem. Wang and White (in press) extended an earlier study of academic researchers' decision criteria (Wang and Soergel, 1998). The initial study examined faculty and graduate students in Agricultural Economics who were asked to think aloud while selecting citations from online database searches. Document information elements (DIEs) and criteria were identified using content analysis. The criteria were related to the DIEs that provided clues to these criteria. The second study was conducted three years later with the same respondents. Respondents were interviewed about decisions to read and cite documents from the original search and about the documents that were cited in written works but had not been retrieved by the original searches. The second study found that with one exception all criteria identified by the original study were used in decisions to read and cite information. The study also identified 14 new criteria. However, the study did not ask the respondents to rate the criteria by importance and it did not relate the criteria use to discrete information seeking stages, but only to the activities of selecting, reading and citing information.

Smithson (1994) compared relevance judgments from an online search with the citations on the final bibliography. Respondents were graduate students who were writing a research paper as a course assignment and who completed the online search to obtain
information sources for their paper. He found that most documents that were judged as relevant at the time of the online search were not cited in the final bibliography. However, he did not examine relevance judgments at any other point in the information seeking process. He measured relevance as usefulness using a six-point scale and did not investigate the reasons why these citations were judged as useful. This study did not attempt to determine the information seeking stage of the users at the time of the online search.

Kuhlthau (1994) used search diaries and interviews to study the entire information seeking process of a number of respondents in regard to a particular information problem. These respondents included high school and college students and public library patrons. The information seeking stages identified from this research were determined by the affective states, cognition, and search activities of the users. However, these studies did not examine relevance criteria used by the respondents. Specific search behaviors and interactions with IR systems were not studied; the study focused on affective changes and changes in broad user search strategy formulations. Kuhlthau's model was selected for this study because it is the only model of the information seeking process that has clearly delimited information seeking stages that have been developed from the experiences of individuals who were seeking information to resolve a specific information problem.

Studies by Schamber (1991) and Barry (1993) identified a number of end-user relevance criteria. A comparison of the two studies and the criteria groups that were common to both studies indicates that end-users employ a finite number of criteria when making relevance judgments (Barry & Schamber, in press). However, these criteria have not been rated by importance by any end-user group. Relationships between criteria have
been suggested in the literature (Schamber, 1991; Barry, 1993; Schamber & Barry, 1997), but research to support these suggested relationships has not yet been done. Respondents in this study rated the importance of 40 relevance criteria in their selection and use of valuable or highly relevant information sources. Factor analysis was used to examine the meaningful relationships between the criteria that were rated as important.

This study identified the most important constructs or dimensions of high relevance for graduate student end-users. It also identified the criterion terms that reliably and validly measured these constructs and explained the portion of high relevance that can be measured across users. This is the first study to examine in detail what high relevance means to graduate students.

A study of the complete information seeking process for an information problem and identification of changes in the importance of the end-user relevance criteria employed during this process is needed to better understand how users search for and select information. This study also identified relevance criteria that were rated as most important in selecting, obtaining, and using information sources that were the most valuable in the resolution of an information problem. Results can be used to design IR system interfaces that will present or represent information in a way that assists the end-user in evaluating the potential usefulness or value of the information that is retrieved.

Research Questions

The primary research question for this study was: What is the relationship between the criteria employed by respondents in selecting sources of information that they find to be most valuable and the respondent's stage of information seeking with
respect to a specific information problem? The following research questions address the specific issues for this study.

When graduate students make high relevance judgments of information sources for a paper or research project:

- What are the relationships between a group of 40 end-user relevance criteria in regard to their degree of importance?
- At what stages of information seeking do respondents employ IR systems or information resources?
- At what stages of information seeking do respondents search for, obtain, and read information?
- Does the importance of end-user relevance criteria change as the respondent’s information seeking stages change?
- Does the importance of end-user relevance criteria change as respondents search for, obtain, and read information?

Importance of the Problem

Almost all relevance research has studied user relevance judgments only at the point of user interaction with an IR system. Researchers have called for a more longitudinal approach to the study of information seeking and user evaluation of IR system interaction and output (Hert, 1996; Robertson & Hancock-Beaulieu, 1992; Spink, 1996). Hert found that during a single interaction with a library online public access catalog (OPAC) user information needs or goals did not significantly change. However, when the whole process of information seeking was examined, Kuhlthau (1994) found
that the information focus changed considerably during the entire information seeking process. These findings indicate that further research is needed to examine the complete information seeking process for a specific information problem.

A better understanding of changes in the importance of end-user determined relevance criteria throughout the process of solving an information problem can lead to IR system interfaces, document representations, and information retrieval results that will better meet the user's needs at their particular stage of information seeking. An examination of the relative importance of a number of different end-user relevance criteria and the meaningful relationships between the most important criteria adds to the body of research that has examined how relevance is conceptualized. This further understanding of end-user high relevance can be used to design IR systems that will assist users in the identification of relevant information.

Relevance has traditionally been measured as relevance, usefulness, value, satisfaction, or some combination of these variables. These variables have been defined by the researcher without investigating their meaning to the user who is making the relevance judgment. Recent studies have identified criteria that end-users employ in making relevance judgments (Schamber, 1991, Barry, 1993). These end-user defined relevance criteria can be used to study relevance judgments in a manner that captures what relevance means to users, rather than relying on the researcher's definition of relevance as the dependent or measured variable.

Relevance research, although it often examines user behavior, is still almost exclusively conducted within the context of IR system research. This study posits that to better understand relevance and better help users make the best relevance judgments for
their particular information problem, the study of end-user relevance must be situated within the context of the entire information seeking process rather than strictly within the IR system evaluation context. Users' relevance judgments do not cease after they leave their interaction with the IR system but continue whenever they select or reject an information source. The information that is selected, obtained, and read may affect the information problem and its focus, information seeking strategies, information source evaluation, and the criteria used in this evaluation.

Users who are able to clearly define their information seeking stage and how their interaction with an IR system or information source affects their relevance judgments and information problem may have a better chance of identifying and finding the information they need to resolve their problem. They may also be better able to find a clearer topic focus that can aid in the resolution of their information problem. An understanding of how information seeking stages interact with relevance criteria and an understanding of the stages when users are likely to be searching for, obtaining, and reading information can also help information professionals guide users. This study is a beginning step toward the study of relevance within the information seeking process and context.

User interaction with the IR system must still be considered as part of the whole information seeking and relevance judgment process. Relevance is now viewed as something that is defined by the individual user rather than an innate characteristic of a document, its representation, or of an IR system (Schamber et. al., 1990; Schamber, 1994). However, the end-user's relevance definition and judgments interact with how the IR system presents information to the user and of the characteristics of the information itself. When a user does or does not select a particular source of information he or she has
made a decision that influences the rest of the information seeking process and takes the user down a particular path to (or possibly away from) other information. It is possible to backtrack and traverse the same path again, and it appears that some users may do just that (Spink, 1996). However, when the path is repeated the user has often made changes in his or her relevance definition and information problem. Users who can judge where they are in an information seeking process and how IR systems and information sources or representation affect their relevance judgments and information problems are likely to be more effective and efficient in their information seeking process.

Relevance Related to Information Seeking and Information Retrieval Models

![Information Retrieval Models Diagram](image)

**Figure 1. Early Model of Relevance in Information Retrieval**

The earliest models of relevance within information retrieval are fairly simple (Figure 1). A user makes a query of an IR system, the system matches the terms used in the query with the terms assigned to the document and returns information or a representation of information to the user based on this match. In this model relevance is defined as an innate characteristic of the information. Relevance can be judged by anyone with some knowledge of the topic of the query and is independent of the user and his or her information problem. Relevance is synonymous with the subject or topic of the
document and is a static quality of the document or query itself. The effectiveness of the system is measured by the ratio of the number of citations judged to be relevant to the total number of citations retrieved (precision), and the ratio of the number of citations judged to be relevant to some estimate of the total number of relevant citations in the database (recall). This model attempts to position IR research within the traditional scientific experimental paradigm. Research based on this model is often conducted in an experimental setting using queries determined a priori by the researcher and databases of limited size and subject breadth.

In an important variation on this model relevance is conceptualized as something that must be determined by the user in regard to his or her information problem. Within this model relevance is sometimes called pertinence to distinguish it from the more traditional concept of relevance described above. Relevance is still defined as topicality or "aboutness" and is static and stable. However, relevance is determined by the user's perception of his or her situation and information problem and is not simply a characteristic of the document or query. IR system research based on this definition of relevance is more likely to be conducted in a library or naturalistic setting with users and their own information problems and employs production (as opposed to experimental) databases. One serious limitation of both variations on this model is the assumption that users can easily state their information needs and that these needs can be well represented in a query. These models do not accommodate either multiple search iterations or the role of feedback in search query formulation and reformulation by search intermediaries or end-users (Spink, 1996; Spink, 1997a). This model also assumes that document or information source topics can be adequately represented by indexing, keyword retrieval
from the title or abstract, or keyword and word proximity retrieval from the full-text of the document.

Figure 2. "User-Driven" Model of the Information Retrieval Process.

The system-oriented model discussed above has been largely replaced by more user-oriented models of information retrieval. A good example of this model is Taylor's (1986) (p. 25) "user-driven" model of the information retrieval process (Figure 2). Taylor posits that each step of designing and adding value to information systems should be driven by user needs. Most components of this model are system components and although user needs are a primary consideration, he makes no attempt to model how the user interacts with the system, the system interface, or a human intermediary.

An important element of the user-oriented or user-driven models has been recognition of the importance of the user's cognitive process within the IR interaction.
This recognition of the importance of the cognition of the user, and the authors, indexers, and creators of IR systems has contributed to the paradigm shift in information science theory and research which has been proposed by Dervin and Nilan (1986). Dervin and Nilan call for a shift in IR research that moves away from a system orientation to a predominantly user orientation. Saracevic (1995) argues this shift should not be absolute and that IR research must continue to include both IR system and user elements. However, Saracevic emphasizes that an understanding of user behavior and cognition is important in IR research.

Belkin (1980) (p. 191) has been an important proponent of the role of cognition in IR theory and has proposed that an information need can be represented as an individual user's Anomalous State of Knowledge (ASK). An individual with an information need often will go through this ASK state where he or she is aware of this need but may not be able to articulate it. Belkin models the linguistic level of information generation as flowing from a generator or author to a text. The recipient or user then interacts with the textual representation of the information (Figure 3).

![Figure 3. Linguistic Level of Information Generation](image)

Belkin models the cognitive level of the same information interaction as the generator's state of knowledge which creates information and the user's or recipient's ASK which interacts with that information (Figure 4).
 Ingwersen (1982) has proposed a model which incorporates the user's ASK state and adds the cognitive processes of the generator or author of a document or information, the librarian or other information intermediary, and the user. This model suggests how the major elements of the information retrieval process are each modified by human cognition and the conceptual knowledge of the creator of the information, the intermediary or librarian and the end-user (Figure 5).

Within these cognitive models of IR relevance is seen as situational, dynamic, and multidimensional and centered within the entire information seeking process, rather than
a phenomenon that only occurs at the IR system interface. It is assumed that the user may not be able to verbalize his or her information problem.

A limitation of these models is that the user’s relevance judgment process is not represented. Situational or environmental factors which have been identified in several major relevance studies as affecting relevance judgments are not included or represented (Schamber, 1994). These models do not examine any changes that might take place as the user moves through the information seeking process. The cognitive models indicate that knowledge representations influence the user and his or her cognition, but the effect of user cognition on relevance judgments and further information seeking is not modeled.

Figure 6. Relevance Within the User’s Contexts

Park (1993) (Figure 6) has modeled relevance within the user’s internal context, the user’s external context, and the problem context. Within this model relevance is
influenced by the contexts and characteristics of the user and of the document or citation. User characteristics include the perceived quality of documents. Citation characteristics include the subject, style, and keywords. This model does not situate relevance judgments within the longitudinal process of seeking information, but rather represents relevance judgments made after the time of interaction with an IR system.

Saracevic (1996) has proposed a stratified model of IR interaction that encompasses environmental and situational factors and user adaptations during interactions with an IR system and evaluation of information (Figure 7).

![Figure 7. Stratified Model of IR Interaction](image-url)
The system and the user both interact with each other through several strata or levels. He suggests, “On the surface level interaction involves dialogue through an interface. [The] user side involves adaptation through different levels - cognitive, situational and affective. [The] computer side also has levels -- content, processing and engineering” (p. 13). Saracevic concludes that relevance is actually a system of relevances that are a result of interplay or interaction between the different user and system levels. He does not suggest how these different relevances might be measured or what the importance of each relevance might be to the system’s or user’s document selection. While this model includes the situational and environmental factors that affect relevance judgments it does not move the relevance judgment process away from the IR system interface and does not consider user interaction with information that is obtained and read.

**Information Seeking Models**

Information seeking behavior has not been modeled as extensively as relevance or user interactions with IR systems. Most information seeking models can be categorized as one of two general types. The first type of model represents the usual information seeking behavior of individuals without regard to the information seeking process for specific information problems or the potential for stages or movement within this process. The second type of model examines the process for a specific information problem over the course of information seeking for this problem and may include movement through stages, multiple search iterations, feedback, and search reformulation.
Many studies of information seeking behavior describe information seeking strategies or the use of various sources of information, often with a focus on libraries and information professionals (Ellis, 1992; 1993; Chatman, 1996). These studies have not attempted to model information seeking behavior or look at it in relation to a specific information problem or as a longitudinal process which could involve multiple iterations, feedback, and interactions with IR systems. Also, the effects of IR systems and information sources on the information problem, subsequent information seeking behavior, and changes in information evaluation over time are not modeled.

**Figure 8.** Broad Model of the Information Seeking Process

Leckie, Pettigrew et. al. (1996) examined the literature on information seeking and developed a model based on numerous studies of the user information seeking
process (Figure 8). Most information seeking studies have examined how users seek
information independent of any particular information problem and their model is a
reflection of this. The information seeking process is modeled in a very broad manner.
The outcomes of information seeking are affected by work roles that influence work
tasks. These work tasks then determine the characteristics of the information need. The
information need, sources of information, and awareness of information interact with the
information being sought and the information seeking process.

A major problem with this model is the lack of representation of the process of
the individual user and his or her cognitive state. The whole information seeking process
is summed up as “information is sought” and the model does not indicate either the
process or strategies of information seeking. It does not model the evaluation of
information and does not account for changes in information needs or information
seeking strategies that might occur as the result of user interaction with IR systems and
information sources. Although feedback is included in this model, it is not clear what the
effect of this feedback is on the information seeking process or on information search
strategies.

Kuhlthau (1994) (Figure 9) models information seeking as a process where the
user experiences six stages. Each stage has a characteristic set of thoughts, feelings and
actions. Her model is longitudinal, but it does not include user interactions with IR
systems or feedback.
<table>
<thead>
<tr>
<th>Stages</th>
<th>Task</th>
<th>Topic</th>
<th>Prefocus</th>
<th>Focus</th>
<th>Information</th>
<th>Search</th>
<th>Starting</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>Uncertainty</td>
<td>Optimism</td>
<td>Confusion,</td>
<td>Clarity</td>
<td>Sense of</td>
<td>Relief</td>
<td>Satisfaction or dissatisfaction</td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td></td>
<td></td>
<td>frustration,</td>
<td></td>
<td>direction/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and doubt</td>
<td></td>
<td>confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 9. Model of the Information Search Process**

In this model the user moves from an unfocused, ambiguous, and sometimes frustrated state to a more focused, specific conclusion. Information that is initially sought is more general and is categorized as relevant to the topic. Relevance in these early stages can be determined by anyone with some understanding of the topic since the user’s information problem focus is not yet clearly defined. As the user moves through the latter stages the information problem becomes more focused and pertinent information is obtained. Pertinence refers to information that the user judges in relationship to his or her specific situation and environment. Although Kuhlthau presents her model in a linear fashion these stages of information seeking may not always be a linear process, users may go through some of the stages several times and there is some evidence from Kuhlthau’s studies that some users do not complete all stages. In particular users may not complete the topic focus formulation stage but may still complete the project. However, unlike most information seeking models, her model is based on a longitudinal process of information seeking for a specific information problem.

**Relevance within IR and Information Seeking Models**

Spink (1997) studied interactive feedback from users who were present while intermediaries conducted searches of online databases for these users. She studied
feedback as users evaluated system output and their subsequent search query reformulation. She found five types of interactive feedback; content relevance feedback, term relevance feedback, magnitude feedback, tactical review feedback, and term review feedback. Her model (Figure 10) of the elements of the search process includes cycles of search strategies which are influenced by interactive feedback.

Figure 10. Elements of the Interactive Search Process

This feedback can include search moves and tactics and user evaluations or interpretations of system output. This model clearly shows interactions between the user, the IR system, and the influence of these interactions on information evaluation. It also models searching as a cyclical process that can involve several iterations or cycles.
This research extends parts of Kuhlthau’s model in regard to interactions with IR systems and information sources and the resulting relevance judgments. The focus of this research was a more specific study of user actions within the model. The actions of searching for, obtaining, and reading information sources are of particular interest since these are points where a user is making relevance judgments. This study determined the stages when users actively seek and evaluate information; the stages when decisions to search for, obtain, and read information are made; and changes in the importance of the criteria that users employ in making relevance judgments from stage to stage.

The elements of Spink’s feedback model are also an important consideration in this research. User relevance judgments can be viewed as a part of a feedback loop. User judgments may be based on results from an IR system or as a result of browsing library shelves, obtaining or not being able to obtain information, and reading information. The user may then evaluate this information using one or more relevance criteria and may then reformulate his or her search tactics or topic and move on to another searching, browsing, obtaining, or reading cycle. The user judgments are influenced by the user’s situational and environmental factors including the user’s stage of information seeking. At some points in the overall information seeking process the user may not be involved in any search cycles, but may still be moving within the information seeking stages as a result of cognitive change. This research did not examine each cycle within this process or all elements of each cycle but focused primarily on relevance judgments and search activities reported by respondents.
Structure of the Research

Research has indicated that a number of factors influence relevance judgments. These include subject knowledge of the user, the user’s type of information problem, and research stage (Schamber, 1994). The criteria that users employ to select information sources and the characteristics of the information that they find to be valuable or highly relevant may vary considerably depending on the type of user and information problem. An undergraduate working on a term paper will likely select different information than a graduate student or a faculty member working in the same area. Individuals who are seeking information to solve a personal problem or for entertainment are likely to use different criteria than students or researchers. This study provided limited control for some variables that could affect relevance judgments by using a group of respondents who were at the same level academically and who had a similar information problem, a class paper or project.

This study was conducted in two phases. The first phase (survey phase) examined the relationships between relevance criteria rated as important by graduate students, the constructs measured by these criteria and high relevance. The selected criteria were mentioned by other end-users as their reasons for selecting an information source as something they would pursue (or not pursue) to help resolve their information problem (Schamber, 1991; Barry, 1993). A random sample of UNT graduate students were surveyed by mail and asked to rate the importance of each criterion in their selection of the most valuable information source for a recent or current paper or project. They were also asked to report their stage of information seeking and to indicate whether they had obtained or read the information source.
In the second phase (class phase) of the research, respondents were graduate students who had been assigned a research paper as a course requirement. They were asked to keep search diaries, to record the two to four most valuable information sources as they searched for, obtained, and read information sources. They rated their relevance criteria using a relevance criteria evaluation form, and estimated their stage of information seeking.

Summary

The goal of this study was to further an understanding of changes in the use and importance of relevance criteria throughout the end-user’s information seeking process for an information problem and to relate these changes to Kuhlthau’s six information seeking stages and the activities of searching for, obtaining, and reading information. This study also identifies the criteria that are important to graduate students in their selection of valuable or highly relevant information sources and the relationships between these criteria, the constructs they measure, and high relevance. An understanding of changes in relevance criteria use and importance during the course of a project will provide a better understanding of the information seeking process and the relevance judgment process from the user’s point of view. Identification of the criteria that are used to select highly relevant information sources may ultimately help design IR systems that better meet user information needs at all stages of their information seeking process.
CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

The concept of relevance has been the subject of considerable debate, theorization and research in the field of information science. In a recent essay, Saracevic (1996) stated that “150+ research articles and reports reflecting theoretical, experimental, or observational studies” regarding relevance have been published in the information science literature to date (p. 5). According to Saracevic, this does not take into account the many reports of IR research that have used relevance as the primary measure of system effectiveness.

Historically relevance was a simple measure of how well an IR system had matched queries for information with the information (or representation of information) contained in the system’s database or databases. Initially relevance appeared to be an intuitive concept that we all understand on some basic level (Saracevic, 1975). However, the concept of relevance has proven to be very difficult to conceptualize and measure. Researchers and theorists in information science have not been able to come to an agreement on the use, measurement and definition of relevance. Relevance has continued to be used as a measure of system effectiveness in the study of IR systems, often with little or no change in how it is defined and measured as compared to very early studies. Harter (1996) argues that this stream of research has virtually ignored the volume of user-
oriented relevance research that indicates relevance judgments are affected by numerous factors that should be considered in experimental design.

Relevance research began to emerge as a research area in its own right in the 1960's. The earliest studies to use relevance as a measure of IR effectiveness were the Cranfield studies conducted by Cleverdon in the late 1950's and early 1960's (Cleverdon, 1962). These studies will be described in detail later in the chapter.

A few years later, several large studies funded by the National Science Foundation (NSF) and conducted by the System Development Corporation (SDC) (Caudra & Katter, 1967) and Case Reserve University (Rees & Schultz, 1967) established a foundation for future relevance research. At this time, SDC was a major online database vendor. These studies examined relevance judgments within the context of the user's information problem or query. However, relevance was still studied only at the point of user evaluation of a citation or results list that an IR system had retrieved in response to the query. Some of the factors which affect relevance judgments that were suggested by Caudra & Katter and Rees & Schultz were a function of the user's whole information seeking process (i.e. stage of information seeking) or a function of how the system and the search results affected the user and his or her information seeking problem. Still most relevance research has continued to only examine the user and relevance judgments only at the point of receiving results from an IR system.

Relevance research has slowly become more concerned with the interaction of relevance with the user's information need and cognitive processes (Schamber, 1994), but few studies have examined relevance judgments within the broader context of
information seeking, rather than as a component of IR system evaluation. To fully understand user information seeking behavior and user relevance judgments, relevance research must be situated within the user's information seeking context. User relevance needs to be studied within the user's general information seeking context and within the context of information seeking for specific information problems. Relevance judgments are not limited to a citation list from an IR system, but occur at a number of points in any information seeking process. Some of these points are not connected with an IR system at all, but may still be influenced by what the user was or was not able to obtain from the IR system, both in respect to the available information on a subject and how the information was represented. Two studies that have collected relevance judgments at other times in the information seeking process indicate that relevance judgments made at the point of contact with the information system do not match the judgments made at other points in the process. Positive relevance judgments at the point of contact with the IR system do not appear to be good predictors of whether a user will actually obtain, read, or cite an information source (Smithson, 1994; Sandore, 1990).

This research study proposes that relevance must be studied within the user’s information seeking context for a specific information problem, not just at the point of contact with the IR system. However, how IR system interaction affects users, their information problem, and their relevance judgments is a very important consideration when studying information seeking and must continue to be a part of relevance research.

This study also examines the effect of the stage of information seeking on the use of relevance criteria. A better understanding of how relevance criteria change as
This chapter will trace the changes in the conceptualization and use of relevance within IR research. It will also examine theorists' views of relevance and relevance research within the last two decades. Research in information seeking will be described and the studies that have examined relevance at more than one point in the user's information seeking process will be discussed. Finally, the importance of the criteria studies that provide some insight into the user's conceptualization of relevance will be examined.

Relevance as Situated in IR Studies

Ellis (1996) discusses the three different approaches to the issue of measurement in IR research. A critical factor in each approach is how relevance is defined and operationalized. The three approaches are the archetypal approach (sometimes referred to as "Cranfield-like" studies), the probabilistic approach, and the Information Retrieval-Expert System approach.

The Archetypal or Cranfield-like Approach

This approach was initially developed for the Cranfield studies. These early IR system studies were the first to study indexing devices as opposed to evaluating only indexing languages. These indexing devices included indexing vocabulary such as controlled terms, single terms and simple concepts. These studies were conducted using
experimental scientific methodology employing quantitative measures. These measures were all based on a new measurement criterion, that of relevance. As Ellis states:

"The principle methodological assumption, and the key to the enterprise, was that the study of the interaction of the entities was to be based on the application of the scientific method, and that the interaction between the devices was to be understood entirely in quantitative terms in relation to the newly conceived measures of recall, precision, and fallout. These measures were themselves derived from the basic quantity, or criterion of measurement of relevance." (p. 24)

The Cranfield experiments were not the first studies to employ the concept or criterion of relevance. In 1953 The Armed Services Technical Information Agency (ASTIA) attempted an evaluation of an information retrieval system with the measurement of effectiveness based on the relevance of retrieved documents to real questions. This study, known as the ASTIA-uniterm study was never completed because the two groups of relevance judges could not agree on which documents were relevant to each question. Ellis notes that an additional problem was that there was no provision for mediation of disagreements regarding relevance. Thus, the first attempt to use relevance as a basis for measuring IR system retrieval effectiveness was aborted.

The Cranfield experiments controlled for disagreements among relevance judges by having researchers create questions from source documents and then employing the retrieval of these documents as the criterion of IR effectiveness. This approach was
widely criticized and the subsequent Cranfield studies used relevance judgments to measure IR effectiveness. Cleverdon and Keen (1966) did differentiate between "stated" relevance or relevance which could be determined by anyone with subject knowledge in the area and "user" relevance or relevance which could only be determined by the user in relationship to his or her own information need. Many other researchers have continued to differentiate between these two types of relevance resulting in considerable confusion and several terms for the same concept. User relevance has been called pertinence usefulness, satisfaction, and relevance and stated relevance has been called relevance and topicality (Schamber, 1994).

Probabilistic approach

The probabilistic approach was based on calculating the probability that a user with a given query will judge a particular document to be relevant (Maron & Kuhns, 1960 approach) or that a set of documents will be judged relevant with regard to a particular property by a given user. (Roberston & Sparck Jones (1976) approach). Recall and precision are treated as probabilities and document retrieval is ranked in decreasing order of their probability of relevance.

This research, based on mathematical modeling primarily in test collections, was seen as limited in its applicability to non-laboratory situations. Additionally the same relevance judgments were being used to refine the retrieval set and to evaluate the same set, which would not be true in a non-experimental environment.
Information Retrieval-Expert System Approach

The information retrieval-expert system (IR-ES) approach is influenced by the cognitive viewpoint which finds its expression in Belkin's theory of an Anomalous State of Knowledge (ASK). The ASK represents the user's cognitive state in regard to his or her information problem. Measurement should be based on the changes in the user knowledge state which are brought about by the information from the system. The texts within the system represent the cognition of the author and the system intelligence is a function of its emulation of the cognition of experts, incorporating expert system and artificial intelligence techniques. Relevance is viewed as more cognitive, complex and dynamic but no suggestions are made as to how to measure it.

Ellis concludes that the goal of IR system development, that is retrieving relevant documents and only the relevant document for a user or query is simply not an attainable goal. A large body of research into the nature of relevance judgment supports this conclusion (Schamber, 1994). The complexity and dynamic character of user relevance judgments which makes perfect IR system retrieval impossible provides an opportunity to study user information seeking behavior through the examination of the elements of relevance judgments and the changes in these judgments.

Relevance Definitions

A brief discussion of some of the seminal definitions of relevance and two taxonomies of relevance definitions will be presented in this section. These taxonomies were selected for discussion here because each is based on a comprehensive review of the
literature of relevance (Schamber, 1994, Saracevic, 1996). For a comprehensive
discussion of how relevance has been defined and studied within information science
literature the reader is referred to four comprehensive reviews of this literature
(Schamber, et. al 1990; Saracevic, 1975; Schamber, 1994; Mizzaro, 1997).

Vickery (1959b) separated relevance into two categories; relevance to a subject or
topicity and user relevance. He based his definition of user relevance on how far the
user would pursue the search for information and stated that this type of relevance is a
multi-staged process. This very early definition of user relevance hints at the need to
study the user’s search for information rather beyond the IR system interface.

Cooper (1973) proposed a “logical” relevance, which was based on the principles
of logical deduction. He proposed that logical relevance was the “aboutness” of a
document or query and that utility might better describe the characteristics of the
documents that the user planned to obtain and use. His definition of utility included
elements such as quality, novelty, importance and “many other things.”

Wilson (1973) proposed a situational relevance based on the individual’s
concerns, preferences, and personal knowledge. Foskett (1972) and Kemp (1974)
categorized relevance as public knowledge and pertinence as private knowledge.
Swanson (1986) has proposed a division of relevance into subjective and objective.
Subjective relevance is the individual’s cognitive definition and objective is the logical
relationship between a document and a query stated in an objective format. Situational
relevance as proposed by Schamber, et. al. (1990) and psychological relevance as
proposed by Harter (1992) will be discussed in detail later in this section.
Schamber (1994) categorizes the different views or definitions of relevance as the system view, the information view, and the situational view. The system view and the information view have often been combined into one category referred to as “system relevance” within other discussions of relevance (Saracevic, 1996). Schamber’s system view defines relevance as an algorithmic matching by the system of query terms with documents or document representations. Human judgment of relevance plays little or no role in this view. Proponents of this view maintain that systems can adequately algorithmically index and match query terms to documents with little or no human input. In reality, many of these algorithmic IR strategies (clustering, matching, term weighting) have been initially evaluated using human relevance judgments.

The information view of relevance is similar to “stated” relevance as defined in the initial Cranfield studies. This view assumes that relevance is static, is represented by the topic of the query or document, and that effective retrieval can be measured by determining how well query terms match index terms or key terms from the title, abstract, or document text. Relevance is inherent in the document or query and is independent of the user or information problem. Information problems can be easily and clearly expressed in query terms by either a user or an intermediary and any person with a reasonable knowledge of the subject can make adequate relevance judgments.

The situational view assumes relevance is dynamic and multi-dimensional and is dependent on the situation, user, and information problem. This view has evolved as a result of the cognitive approach to IR and user information seeking theory. It is supported by research that indicates that many factors affect relevance judgments and that user
relevance judgments are based on more than the topic of a document or its representation. This approach assumes that relevance is situation and user dependent, can only be judged by the user with the information problem, and resides within the user’s cognitive process. A number of researchers have posited this approach, but exactly how to measure the changes in relevance and how to control the variables affecting relevance has not been suggested within the literature (Schamber, et. al., 1990).

A large group of studies cannot be classified within this scheme. These studies employ the concept of “user” relevance as defined within the Cranfield studies. They differ from the system view in that relevance is determined by the individual user and his or her problem or situation. Thus, relevance judgments must be made by the individual with the information problem. Relevance is determined by the user and does not reside within the document or document representation. However, it is still assumed that relevance is static and the user is able to clearly state his or her information problem and translate that problem into a well-defined query. These assumptions differ from those made within the situational view of relevance. A considerable body of IR system research continues to be conducted based on these assumptions despite relevance studies that support a view of relevance as a dynamic, multi-dimensional concept which can be influenced by a number of factors and studies which indicate that user information problems are often poorly defined and difficult for users to articulate (Harter, 1996).

Saracevic (1996) proposes five different frameworks that define the nature of relevance; systems, communication, situational, psychological, and interactive. The interactive framework models Saracevic’s view of relevance and incorporates parts of the
other four frameworks. Where Schamber's views are more descriptive of existing relevance conceptualizations Saracevic’s framework forms a broad theoretical base for how relevance is defined and used.

The systems framework encompasses Schamber’s system and information views. Relevance is defined and evaluated by matching. Saracevic states that the primary weakness of this framework is that the user, his or her context and the context of information use are not considered nor is the dynamic and interactive nature of IR system use.

The communications framework was proposed by Saracevic (1975) as based on a code model of messages between source and destination, similar to Shannon’s model of communication but based on relevance as the criterion for the evaluation of communication effectiveness. He posited relevance as a relationship between the source’s subject knowledge, subject literature, and files and the destination’s cognitive structure and representation, use, context and values.

The situational framework was proposed by Schamber et. al. (1990) in response to the body of research that suggested that relevance was dynamic, situational, context dependent and multi-dimensional. This research was situated in Dervin and Nilan’s (1986) user paradigm that posited a shift in research in information science from system-based to user-based often within a cognitive context. Saracevic critiques this framework for not incorporating IR system interactions and dynamics. This framework hints at the need to study relevance beyond a “snapshot” approach, suggesting a need to “asses the full effect of the process of information retrieval on users” (Schamber, et. al., 1990, p.
They suggest examining users’ accounts of information seeking, evaluation and other user behaviors but do not suggest how this might be accomplished or exactly how relevance judgments would be studied or compared.

The psychological framework was proposed by Harter (1992) where relevance was based on changes in the users cognitive state rather than on a subject or topicality. Harter views relevance as a dynamic, cognitive response to an information need which is constantly changing. Relevance is seen as changing as the user encounters each new bit of information. This view does not incorporate the user’s interaction with the IR system or other aspects of the information seeking process.

The interactive framework as proposed by Saracevic (1996) combines the user’s interactions with the system and the user’s cognition into a stratified model where users interact with a system and information resources on a surface level, a cognitive level, a situational level and an affective level. The system is seen as consisting of a surface level, a content level, a processing level and an engineering level. His model is based on Ingwersen’s cognitive model of IR interaction and Belkin’s episode model. Ingwersen’s model includes the cognition of the user, and other cognitive elements of the system. The system’s cognitive elements are artifacts of the cognition that was employed during development of the system and user interface and in the creation of texts and representations of texts that are stored in the system. Belkin’s model views IR interactions as a series of episodes that are effected by user’s tasks, goals, and intentions. Within Saracevic’s model relevance is implied on a number of the user and system levels, but it is not explicitly defined, nor is any suggestion made regarding how to measure
interactions or relevance. He proposes an interdependent interacting system of relevances; system or algorithmic relevance, topical or subject relevance, cognitive relevance or pertinence, situational relevance or utility, and motivational or affective relevance. He does not suggest how these relevances might be clearly separated or measured nor does he take the study of relevance beyond the context of the IR system.

Mizzaro (1997) proposes a model of relevance and categorizes more than 160 relevance papers. His theoretical relevance model has four dimensions. Each dimension has several values. Dimensions and subcategories or values interact with each other through time. The first dimension is composed of the values surrogate, document and information. The second dimension is composed of the values topic, task, and context. The third component is composed of the values of query, request, information need and problem. The final dimension is the passage of time from the conception of the problem until its solution. These dimensions can interact with each other and the values within each dimension can interact with each other and with other dimensions and their values. He suggests a multi-dimensional model of the relevance judgement process based on kind of judgment, the judges, and what is being judged (surrogate, document, or information).

Mizzaro categorizes relevance papers as experimental, survey, and theoretical and orders the papers according to seven subcategories based on the approach of the paper to the study of relevance: Foundations, kinds, surrogates, criteria, dynamics, expression and subjectiveness. He presents a very complete overview of the relevance literature but has no suggestions on how to measure relevance and relevance judgments or how such measurement might fit into his model.
Spink, Greisdorf and Bateman (in press) propose a model that is a three-dimensional extension of Saracevic’s levels of relevance model. The first dimension is the degree of relationship or non-relationship of a relevance judgment to each of the Saracevic levels: systemic/algorithmic, topical/subject, cognitive inference/pertinence, utility/situational and motivational/affective. The second dimension is the degree of relevance from not relevant through partially not relevant, partially relevant to highly relevant. The third dimension is the dimension of time, which can be represented by Kuhlthau’s search stages and by the number of successive searches conducted by the user. The authors propose that user relevance judgments can be plotted on their three-dimensional model. By measuring and plotting user relevance judgments on the model, the underlying complexity of user relevance judgments can be partially captured and studied throughout a search process.

With the exception of the Spink, Greisdorf and Bateman model, none of these frameworks or views situates relevance fully within an information seeking context. The situational, psychological and interactive framework clearly recognize that relevance is ultimately defined by the user, even though a system or indexer definition of topical relevance may influence the user’s concept, definition and use of relevance. Relevance is seen as dynamic, situational and affected by the IR system and information, but how relevance might interact with an information seeking process or how changes in relevance might be measure are not discussed.
Relevance Studies

Relevance Judgments After the Search Interaction

Relevance research has identified numerous factors that can influence relevance judgments. In a review of relevance literature Schamber (1994) identifies 80 factors that influence relevance judgments. She indicates that this is not a complete list of the known factors. These factors can be divided into factors that are external to the document being judged such as search stage, choice of scale, and judgment conditions and factors or criteria that the individual employs to make the relevance judgment. These factors are related to the information itself and are in part presented to the user by the system as a representation of the document. These factors can include aboutness or topicality, currency, reputation of the author or journal and several others. Taken as a whole these criteria represent what makes a particular representation or document something that a user will or will not pursue -- i.e. relevant or not relevant, useful or not useful. This research will study relevance criteria using the external factor of search stage as an independent variable, and will attempt to control for the other external factors whenever possible, although many of the factors can be very difficult to control for in a naturalistic setting.

The vast majority of studies of end-user relevance judgments of information sources or representations of those sources have been conducted using citation lists from online database searches or OPACs. Users are asked to rate document surrogates as relevant or not relevant to the topic of their research. Often simple bipolar rating scales
are used although some studies have used a Likert-type or magnitude estimation scale (Janes, 1991). Many studies have examined relevance as a single variable, although it may have been operationalized as aboutness, usefulness, value, or satisfaction or some combination of these variables.

This method of determining relevance judgments has been used as a basis for evaluation of information retrieval systems, search strategies, and searchers and provides a basis for the traditional topical search evaluation measures of precision and recall. Precision is calculated as a ratio of the number of citations judged to be relevant to the total number of citations retrieved, and recall is calculated as the ratio of the number of citations judged to be relevant to some estimate of the total number of relevant citations in the database.

These measures are based on the assumption that the relevance of an item remains constant for the user, that relevance can be adequately measured as a single variable, and that once an item has been judged to be relevant the end-user will actually use it and subsequently find it to be of some value in resolving their information need or problem. However, research indicates that relevance is a dynamic concept for the end-user which can change from day to day and even changes with the order of presentation of a citation list and the order of presentation of the elements of each citation (Eisenberg & Barry, 1988) (Janes, 1991). The limited research that has been done regarding the evaluation and use of documents that have been judged as relevant from a search results list indicates that relevance and relevance judgments often change throughout the course of a paper or project. (Smithson, 1994).
Su (1992) correlated 20 measures of system and search effectiveness with the user’s overall evaluation of system success. These measures can be classified into four general categories; relevance, efficiency, utility and user satisfaction. She found precision as measured traditionally did not correlate well with system success and that recall was more important to users in her study. Other factors that correlated well with system success included value of search results as a whole and factors relating to interaction and effectiveness. She concluded her research with an interview where she asked users to explain their overall success ratings. These interviews resulted in 26 “success dimensions” or criteria. Some of Su’s success dimensions are similar to criteria found in studies by Schamber (1991) and Barry (1993).

Park (1993) used a naturalistic enquiry method to interview respondents concerning their information seeking context, problem area, stage of research, and reasons for selection of citations after they had examined online database search results. Respondents were also asked to verbalize their thoughts as they selected or rejected citations from the online search. The respondents were deliberately chosen from different user groups and at different research stages. She developed a model of user-based and document-based factors which affected relevance assessments. The user-based factors were categorized as internal context, external context and problem context. Document-based factors were often very similar to the relevance criteria enumerated by Barry (1993). These included accessibility, perceived quality, methodology, subject matter, and style. Park clearly states that her study is still within the IR context; “I view the determination of relevance from the perspective of how users select ‘relevant’
bibliographic citations within IR” (p. 319). She examines how user relevance assessments are derived from individuals’ information problems, experiences, perceptions, and context, and but never moves this examination away from the IR system into a more complete information seeking context.

Wang (1994) examined the document selection of users in the context of a model combining components from relevance literature and consumer choice and decision making literature. She proposes that users arrive at and employ criteria or values to select information based on the information elements presented to the user. Final decisions regarding the selection of information are made based on a number of criteria (multi criteria). Wang was able to support this model with data gathered from library users.

Gluck (1996) examined relevance judgments and end-users’ process satisfaction, product satisfaction, overall user satisfaction, response satisfaction, and cost-benefit satisfaction. His respondents completed a sense-making time-line questionnaire that was based on the user’s report of a recent need for geographical information. He found that relevance had a significant relationship with all satisfaction measures except cost-benefit satisfaction. His research indicates that relevance and user satisfaction are related to each other but that one does not measure or subsume the other.

Park (1997) examined the dimensions of relevance through a qualitative study. Twenty-four respondents were asked to examine full-text documents retrieved by an online search. They were asked to mark portions of the documents and to describe why the marked portion was relevant, helpful, useful, valuable or irrelevant. Content analysis was used to identify the dimensions of relevance and these dimensions were used to
develop a model that relates relevance and its dimensions to value, use, problem, relevance and the document itself. The use of dimensions was measured by 1) the number of uses, 2) the number of users mentioning the dimension and 3) the percentage of the total number of uses of all dimensions. Park's study did not suggest any relevance criteria or dimensions that differ from those identified by Schamber (1991), Barry (1993) and Su (1992). The number of uses of a criterion or dimension cannot be assumed to be a measure of the importance of the criterion. The study suggests some interesting relationships between what Park identifies as primary criteria or dimensions and secondary criteria or dimensions. However, these dimensions were determined at the IR interface and a way to measure these relationships and dimensions across end-users and throughout the information seeking process is not suggested.

Studies of End-User Relevance Criteria

Research which has measured relevance as a single variable has assumed that relevance is judged by the user on the basis of apparent topicality or aboutness of a source, the usefulness of the source, or the value of the source. Criteria that could be easily classified as quality of the source have been recognized for many years. For instance, credibility and accuracy were suggested as relevance factors by Cooper (1973) and scientific hardness as a factor by Cuadra & Katter (1967) and Rees &Schultz (1967). However, quality of information has not been widely used as a variable to measure or define relevance or IR system effectiveness.
Preliminary research into the concept of end-user relevance has identified a number of criteria that users employ in making relevance judgments. This research has found that although topic or aboutness is often the primary criteria that is used to evaluate information, users also employ non-topical relevance criteria in their judgment of a document's potential usefulness in the context of their information problem. These criteria can include the author or journal's reputation; the document's perceived accuracy and availability; the cost to the user in time, effort, and expense; and the number of references (Barry, 1993).

Two studies have focused on identifying end-user relevance criteria. Schamber (1991) studied the relevance criteria used in the selection of weather information sources by individuals whose work made this information important to them. This study examined how these individuals usually seek information based on Dervin's time line interview and was not tied to one specific information problem at a given time. Barry (1993) studied the criteria mentioned by respondents as the reasons they felt they would obtain or not obtain documents that were cited in an online search results list. Barry found that after nine respondents had been interviewed no new criteria were mentioned regardless of how the interviews were ordered, indicating that most users probably employ a finite number of criteria when making relevance judgments. These researchers found a considerable degree of overlap in the criteria mentioned in each study. The only criteria that were not mentioned by users in both studies were aspects of information that were characteristic of one situation or type of information source. For example, respondents in the weather study mentioned interactivity and geographical criteria that did
not apply to the textual sources in the study of online search citation lists. This indicates that the concept of end-user relevance may be sufficiently measured with a limited number of criteria.

These studies did not examine how or if information rated as relevant was actually obtained and used and if the relevance judgments of the information and the criteria used to judge the information changed at the point of usage. Also, the relative importance of each relevance criterion to the user was not examined in these exploratory studies.

Information Seeking Theory

Information seeking theory has focused on how the user’s information need is defined and perceived and provides an important foundation for information seeking and relevance studies. More broadly stated, information seeking theory is encompassed by information use and need theory which includes the relationship of cognition to the user’s use of and need for information and subsequent behavior. Dervin & Nilan (1986) define an information need as a gap in a user's knowledge that the user attempts to bridge by seeking additional information. Generally, it is hypothesized that users experience changes in their information need during the information seeking process. These changes have been modeled by several theorists as stages of an information need. It is also theorized that users will employ different information seeking strategies and require different types of information and intermediary support depending on the particular information problem stage (Kuhlthau, 1994).
Belkin, Brooks & Oddy (1982) hypothesize that users experience a critical stage of information need, often early in the information seeking process, which they describe as an anomalous state of knowledge (ASK). An individual experiencing an ASK stage is aware of an information need but cannot clearly verbalize this need. Taylor (1968) proposed four stages of information need. In the first two stages, the visceral need stage and the conscious need stage, the user is aware of an information need but unable to verbalize this need even to him or herself. As the user becomes more aware of the need and the information that is potentially available to resolve it, the user expresses this need as a compromised need. The compromised need is the information need stated in a form that the user feels can be met by available information. Kuhlthau (1994) found that users experienced anxiety during information seeking until they found a topic focus, indicating that users do appear to experience an ASK state during the early stages of information seeking. She hypothesized six different stages of information seeking based on the user’s thoughts, feelings and actions during the information seeking process.

Since by definition users in an ASK state cannot verbalize their information need, it is impossible to study the ASK state directly. By studying user search diaries and how users rank various relevance criteria at various points in seeking information for a research project it may be possible to discover indirectly how users resolve their ASK state and if certain types of system interfaces, training, or intermediary support will help users move through the early stages of information seeking more easily. It may also be possible to identify patterns in the use of relevance criteria and information seeking
strategies that can help validate the proposed end-user information seeking and
information need stages.

Information Seeking Studies

General Studies

Most information seeking studies have focused on where the user seeks
information with an emphasis on the use of or failure to use library resources. (Leckie, et.
al, 1996) Wilson (1977) has proposed two major types of information gathering behavior,
purposeful and incidental. Incidental information is obtained through conversations,
newspapers, radio, television etc. Purposeful information is obtained through a deliberate
search process by the individual. Purposeful information can be obtained through any of
the sources mentioned above but it is most likely to be obtained from libraries, journals,
professionals and books (Williamson, 1996). Studies of more general information seeking
behavior are likely to include a mixture of purposeful and incidental information seeking
and retrieval. Since the focus of this study is primarily purposeful information seeking for
a research paper this review will cover primarily the literature that examines this
particular type of information problem and information seeking behavior.

Wilson (1997) presents an interdisciplinary review of information seeking
behavior studies. The information seeking literature examines various aspects of the
information seeking process: information need; intervening variables; personal
characteristics; and economic, environmental and situational variables. His extensive
review does not identify any studies that examine the entire end-user information seeking process throughout the resolution of a complex information problem.

Few studies have examined how the user employs various resources and processes to seek and obtain information, the characteristics of the information or information sources that are actually used, and what made these characteristics important to the user and the resolution of the information problem. A limited number of studies have examined users' information seeking processes in detail. These studies have primarily taken place at the IR system interface and may be considered a user-oriented subset of IR research. However, in this dissertation, this group of studies will be discussed in the broader context of information seeking. Another group of studies has attempted to observe and describe the broad information seeking patterns of groups of users. These studies consider information seeking behavior in the most general sense and do not examine or explain the use of specific information sources or IR systems. They place information needs or problems in a very general context and do not examine information seeking for a specific information problem.

Leckie et al. (1996) did a meta-analysis of information seeking studies of professionals with specific emphasis on engineers, health professionals and lawyers. The model of information seeking which they developed from this analysis focused on the factors that they concluded represent major influences in information seeking behavior of professionals. This model does not describe the user's information seeking processes in detail. Factors in their model include work roles, tasks, characteristics of information needs, sources of information, awareness of information, and outcomes. While this model
is helpful in understanding some of the broad reasons that professionals use certain sources of information and follow certain patterns when seeking information, it does not provide much insight into the actual information seeking process of these individuals. The writers do mention factors that influence the individual's information seeking path or strategy. These include familiarity and prior success with system, strategy, or information source; trustworthiness, packaging (convenience, usefulness etc.), timeliness (found when needed), cost, quality, and accessibility. Most of these factors were also mentioned as relevance criteria or the reasons that a particular document or information source was selected in studies by Barry (1993), Schamber (1991) and Su (1992).

The focus of research in information retrieval evaluation and information seeking has shifted away from the assumption that relevance is a static innate element of information and that evaluation of IR system output is simply a matter of matching the terms in a query to system terms. Currently users and their information needs are the focus of information system evaluation studies and relevance is defined as situational and user determined. However, user centered studies are still primarily conducted at only one point in the user's information seeking process. Information seeking strategies, information needs, and the criteria used to evaluate information are usually studied at the end-user's point of contact with the information retrieval system or with an intermediary. This area of research has focused heavily on the use and evaluation of OPACs, online searches done either by an intermediary or by the end-user, and face-to-face interactions with intermediaries. The initial purpose and focus of this research was to evaluate system or searcher effectiveness. These studies have been clearly situated in IR research.
However, as IR systems have become widely available for users who may not have access to an intermediary, the study of the behavior of the user has become an important focus for this type of study and can offer some valuable insights into user information seeking behavior. Because of the historical IR context of these studies, few of them move away from the user’s interaction with the IR system and do not examine how this interaction and the IR system affect the entire information seeking process.

Methodologies used to study end-user interactions with IR systems have included examination of transaction or keystroke logs as recorded by the system, user surveys, user interviews, focus groups, and talk-aloud protocols (Hert, 1996; Matthews et. al., 1983). One study has compared end-users who initiated their search at an OPAC with those who started their search at the shelves. This study did determine if the information sources identified were subsequently checked out of the library. However, the researchers did not examine how or why the information was ultimately used or what criteria were used to select this information (Hancock-Beaulieu, 1990).

Information seeking research outside of the context of IR system evaluation has primarily examined user information seeking behavior in a general sense without regard to any particular information problem on the part of the user. Researchers have studied the information seeking behaviors of academic researchers (Ellis, 1992), how individuals organize textual information in their offices (Kwasnik, 1991), and even the information seeking and needs of janitors, elderly women, and the homeless (Chatman, 1996). These studies have not attempted to tie information seeking to a particular information problem.
Only Kuhlthau (1994) has studied the information seeking process and how it changes during the resolution of a particular information problem.

**Information Seeking Behavior Independent of a Specific Information Problem**

The primary goal of these studies has been to examine how users normally seek, organize, and select information without regard to a particular information problem or need. These studies have not examined information selection criteria in detail and have not attempted to delineate stages of information seeking. Their focus has been on the strategies employed by individuals to gather information as opposed to examining why users select and use particular information sources and information seeking strategies to resolve a particular information need or problem.

Ellis (1989) interviewed social science faculty regarding their usual information seeking behaviors. He delineated the strategies employed by these individuals in the process of seeking information for research projects or maintaining current information in their area of expertise. These interviews were not tied to specific information problems (although the users often mentioned specific examples) and did not examine why information was selected or how it was eventually used.

Ellis asked social scientists at the University of Sheffield to "describe their work and the sorts of activities they engaged in which might be understood as having an information component" (p. 174) The respondents included individuals from a university sponsored research unit and teaching faculty. Individuals who had obtained an online computer literature search from the library made up a third group of respondents. Ellis
found six major information seeking patterns; starting, chaining, browsing, differentiating, monitoring, and extracting. Differentiating, defined as “using differences between sources as filters on the nature and quality of the material examined” (p. 178), would include relevance judgments as defined in this dissertation. For Ellis’ users the most significant differentiating (or relevance) criteria were the topic of the study, the approach or perspective of the study, and the quality, level, or type of treatment of the study. As these behaviors and criteria were described without the context of a specific information problem or source Ellis did not make an attempt to determine the order or search stage when these behaviors might occur.

A clear picture of the reasons why individuals select and use information is important in understanding the information seeking process. Research into how individuals evaluate information at various stages in the information seeking process for a discrete information need or problem is needed to help bridge the research gap between snapshot studies of information seeking at one point in time and general information seeking studies.

**Longitudinal Studies of Information Seeking for a Specific Information Problem**

Although users have been studied primarily at one point in their information seeking process for a particular information need, one study has determined that many users report conducting multiple search iterations and modifying search terms when seeking information on a particular topic (Spink, 1996).
A few studies have examined users at different points in the information seeking process and some studies have found that some information judged to be relevant or useful at the point of contact with the IR system was never obtained or used. This indicates that relevance judgments can change rather dramatically during the course of information seeking. Sandore (1990) contacted users two weeks after receiving a citation list from an online database search. A number of the users in this group did not even obtain all of the documents that they evaluated as highly relevant to their need at the time of the online search. She did not determine if their relevance judgments had changed, their information problem was resolved or redefined, if additional information had been obtained, or if the user had simply lost interest.

Smithson (1994) studied the relevance judgments and use of citations by 22 students who were required to write a project report for a class in Analysis, Design and Management of Information systems at the London School of Economics. Each student was provided with a free online search and evaluated the results in terms of usefulness on a six-point scale. When the project was completed respondents were given the initial search again and asked whether they had looked for the documents, whether they had found them, and to give a revised relevance judgment for the documents they had found. This group of users did not look for or find most (81%) of the documents retrieved by the search, only three users judged more than half the initially relevant documents as finally relevant, and some initially non-relevant documents were obtained and reclassified as relevant.
Kuhlthau (1994) studied high school students throughout the process of information seeking for a research paper by using student research diaries and personal interviews. Excerpts from the diaries clearly indicate that the students engaged in more than one information seeking iteration. Kuhlthau proposed that the students went through six different stages in their information seeking for this project based on the students' thoughts, actions and feelings at various points in the project. She related these changes and stages to topic focus and search activities but did not relate the proposed information seeking stages to the criteria used to select this information. She did find in some cases that citations judged not relevant early on appeared on the student's final bibliographies. It was not clear whether this represented bibliography padding or if the student's relevance judgments regarding this information had changed.

Measuring Relevance Judgments

Historically relevance was first measured as a match between the topic of a document citation and the topic of a search of an IR system. The citation represented the document and the search query represented the user's information need or problem. This measured what was assumed to be the most important aspect of a document or information source for most users, the topic of the document. Early information searches were submitted as batch requests to a mainframe computer system. It could take several days or even weeks for users to obtain search results. Due to cost and time factors database searches were rarely repeated making it critical to retrieve the most topically precise (precision) and complete set (recall) of documents possible and then allow the
user to employ any additional criteria to select the most useful or valuable documents from the retrieved set.

Unfortunately, it soon became apparent that even experts in a subject area were unable to agree on the relevance of a document to a topic (Cleverdon, 1962; Gull, 1956). Two types of relevance judgments were then proposed; a judgment which depended on the user and his or her information problem (sometimes referred to as pertinence) and a strictly topical judgment that anyone familiar with the topic could make (sometimes referred to as relevance).

In an effort to improve information retrieval, several major studies were conducted in an effort to examine the situational factors that affect user relevance judgments. Two landmark large-scale studies of variations in relevance judgments identified a number of situational factors or variables that affect relevance judgments. (Cuadra and Katter, 1967; Rees and Schultz, 1967) These studies employed experimental, quantitative methods and consequently the relevance factors examined in these studies were determined by the researchers before the studies were conducted. The questions or topics which represented the query were also predominately predetermined by the researchers. Schamber (1994) has listed a number of the relevance factors from these and other similar studies in tabular form. (Table 1).
Table 1. Relevance Factors Suggested in the Literature
Based on (1) Casdor & Kater (1967), (2) Rees & Schulze (1967), (3) Cooper (1971, 1973) and (4) Taylor (1986)

<table>
<thead>
<tr>
<th>JUDGES</th>
<th>DOCUMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biases (1)</td>
<td>Aboutness (3)</td>
</tr>
<tr>
<td>Cognitive style (1)</td>
<td>Accuracy (truth) (3)</td>
</tr>
<tr>
<td>Concept of relevance (1,2)</td>
<td>Aesthetic value (3)</td>
</tr>
<tr>
<td>Error preference (1)</td>
<td>Authorship (3)</td>
</tr>
<tr>
<td>Expectations regarding distribution (1)</td>
<td>Credibility (3)</td>
</tr>
<tr>
<td>Formal education (2)</td>
<td>Difficulty level (1)</td>
</tr>
<tr>
<td>Intelligence (1)</td>
<td>Diversity of content (1)</td>
</tr>
<tr>
<td>Judging experience (1)</td>
<td>Importance (3)</td>
</tr>
<tr>
<td>Judgment attitude (1)</td>
<td>Informativeness (3)</td>
</tr>
<tr>
<td>Knowledge/experience (1,2)</td>
<td>Interesting content (3)</td>
</tr>
<tr>
<td>Professional involvement (2)</td>
<td>Level of condensation (1,2)</td>
</tr>
<tr>
<td>Research stage (2)</td>
<td>Logical relevance (3)</td>
</tr>
<tr>
<td>Use orientation (1,2)</td>
<td>Novelty (3)</td>
</tr>
<tr>
<td>Vigilance level (1)</td>
<td>Pertinence (3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INFORMATION SYSTEM</th>
<th>REQUESTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access (item identification) (4)</td>
<td>Availability of anchors (1)</td>
</tr>
<tr>
<td>Access (subject description) (4)</td>
<td>Difficulty level (1)</td>
</tr>
<tr>
<td>Access (subject summary) (4)</td>
<td>Functional ambiguity (1)</td>
</tr>
<tr>
<td>Accuracy (data transfer) (4)</td>
<td>Specificity/amount of information (1,2)</td>
</tr>
<tr>
<td>Browsable (4)</td>
<td>Subject matter (1)</td>
</tr>
<tr>
<td>Comprehensiveness (coverage) (4)</td>
<td>Textual attributes (1)</td>
</tr>
<tr>
<td>Convenience of location (3)</td>
<td>Weighting of components (3)</td>
</tr>
<tr>
<td>Convenience of hours (3)</td>
<td>breadth of document set (1)</td>
</tr>
<tr>
<td>Cost saving (4)</td>
<td>Definition of relevance (1,2)</td>
</tr>
<tr>
<td>Currency (updating) (4)</td>
<td>Order of presentation (1)</td>
</tr>
<tr>
<td>Ease of detection of relevance (3)</td>
<td>Size of document set (1)</td>
</tr>
<tr>
<td>Effort expended (3)</td>
<td>Social pressure toward convergence (1)</td>
</tr>
<tr>
<td>Flexibility (dynamic interaction) (4)</td>
<td>Specification of the task (1,2)</td>
</tr>
<tr>
<td>Formatting (scannability) (4)</td>
<td>Time for judging (1)</td>
</tr>
<tr>
<td>Interfacing (help, orientation) (4)</td>
<td>Use of control judgments (1)</td>
</tr>
<tr>
<td>Links to external sources (4)</td>
<td></td>
</tr>
<tr>
<td>Ordering (subject matter) (4)</td>
<td></td>
</tr>
<tr>
<td>Physical accessibility (4)</td>
<td></td>
</tr>
<tr>
<td>Precision of subject output (4)</td>
<td></td>
</tr>
<tr>
<td>Reliability (consistency) (4)</td>
<td></td>
</tr>
<tr>
<td>Selectivity (input choices) (4)</td>
<td></td>
</tr>
<tr>
<td>Simplicity (clarity) (4)</td>
<td></td>
</tr>
<tr>
<td>Time spent (3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHOICE OF SCALE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of anchors (1)</td>
<td></td>
</tr>
<tr>
<td>Ease of use (1,2)</td>
<td></td>
</tr>
<tr>
<td>Kind of response required (1,2)</td>
<td></td>
</tr>
<tr>
<td>Number of rating categories (1)</td>
<td></td>
</tr>
<tr>
<td>Type of scale</td>
<td></td>
</tr>
</tbody>
</table>

In most subsequent relevance research, relevance has been measured as a single variable, i.e. relevant or not relevant. Sometimes relevance was defined for respondents as utility, value, or topicality, but relevance was still measured as a single variable. The
characteristics or clues that users employed in judging citations as relevant or not relevant were not studied. In fact, the a priori definition of relevance to respondents is listed as one of the factors that may cause user relevance judgments to vary (Table 1).

Clearly, these situational factors tell us a great deal about what makes relevance difficult to define and measure. The research represented by these factors can be used to substantiate the nature of relevance; that it is dynamic, user-centered, multi-dimensional and situational (Schamber et. al. 1990). Even a cursory examination of the factors in Table 1 clearly indicates that the factors are related to the situation of the user, the experiment and even the document and the system. Unfortunately, knowing that these factors can affect relevance judgments does not clearly indicate which factors or criteria are actually employed by users in judging information. It can be indirectly assumed that the document factors listed in Table 1 (aboutness, accuracy, recency, usefulness) etc. represent some aspects of what makes some documents relevant to some users. However, since the factors were determined a priori by researchers who wished to see how they affected relevance judgments, rather than why users made these judgments, within the context this earlier body of research they are simply situational factors that affect relevance judgments.

The original goal of relevance research was to determine how relevance could be measured in order to evaluate and compare IR systems with each other. Some subsequent information science research has examined improving IR system interface design as an important component of IR system effectiveness (Dervin and Nilan, 1986). In order to accomplish this, it is important to understand not only why relevance judgments vary, but
the reasons the user employs when making these judgments; how the user defines relevance as he or she judges information, and the clues regarding information sources that IR systems provide to the user.

Much has changed since most of the research was conducted that determined the relevance factors in Table 1. The IR systems available to users and subsequently the situational context of users have change dramatically. Users now do their own searches, often at little or no cost on a variety of systems that provide instantaneous feedback and store millions of citations as well as full-text documents. Often the user retrieves far more topical or peripherally topical information than he or she may ever consider evaluating. Current end-users will benefit from IR systems that present information or information representations in a manner that makes judging relevance easier.

To further the understanding needed to assist user relevance judgments at the IR system interface, it is necessary to understand the criteria users employ in making these judgments. However, these criteria are not independent of the IR system. The criteria and clues that the IR system presents are the major factors that are available for the user to employ when making relevance judgments. We cannot ignore how the system and document representation affect user relevance criteria. Perhaps there are criteria or clues that users would find useful in judging documents that are not represented by the system.

Two recent qualitative studies of end-user relevance judgments (Schamber, 1991; Barry, 1993) have identified a number of criteria that users have enumerated as their reasons for selecting information. A selection of these end-user relevance criteria are listed in Table 2. These criteria show some overlap with the relevance factors listed in
Table 1. The document factors listed in Table 1 can either be considered situational factors that affect relevance (if relevance is being judged as a single variable and the factors have been determined a priori by the researcher) or end-user relevance criteria (if relevance is being studied and measured as a concept made up of several variables which is situationally defined by each end-user).

If relevance is being studied as a user-defined, multi-variable concept then situational factors listed in Table 1 can be placed into two categories based on how they

### Table 2. User-Defined Relevance Criteria

<table>
<thead>
<tr>
<th>Topicality</th>
<th>Information Source Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>About my topic</td>
<td>Prominent</td>
</tr>
<tr>
<td>Availability</td>
<td>I know the author personally</td>
</tr>
<tr>
<td>Easy to obtain</td>
<td>I know the source (author, journal, etc.)</td>
</tr>
<tr>
<td>Free or inexpensive</td>
<td>Reputable</td>
</tr>
<tr>
<td>Novelty</td>
<td>Format of source (Book, journal article, microform, etc.)</td>
</tr>
<tr>
<td>Unique or the only source</td>
<td>Interactive</td>
</tr>
<tr>
<td>Original</td>
<td>Information Characteristics</td>
</tr>
<tr>
<td>New to me</td>
<td>Describes methods or techniques</td>
</tr>
<tr>
<td>Familiar</td>
<td>Provides examples</td>
</tr>
<tr>
<td>Quality of information</td>
<td>Provides graphics</td>
</tr>
<tr>
<td>Well-written</td>
<td>Statistical or methodological approach</td>
</tr>
<tr>
<td>Credible</td>
<td>(Descriptive or Interpretive)</td>
</tr>
<tr>
<td>Accurate</td>
<td>Theoretical or applied approach</td>
</tr>
<tr>
<td>Understandable</td>
<td>Provides proof</td>
</tr>
<tr>
<td>Consistent</td>
<td>Controversial</td>
</tr>
<tr>
<td>Focused</td>
<td>Provides bibliography or links</td>
</tr>
<tr>
<td>Presentation Characteristics</td>
<td>Provides background or history</td>
</tr>
<tr>
<td>Presentation of information</td>
<td>Appeal of Information</td>
</tr>
<tr>
<td>(Good/poor physical layout/design)</td>
<td>I like it</td>
</tr>
<tr>
<td>Suitable length (long or short)</td>
<td>Validates my viewpoint</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>Interesting</td>
</tr>
<tr>
<td>Suitably general or specific</td>
<td>Enjoyable</td>
</tr>
<tr>
<td>Detailed</td>
<td></td>
</tr>
<tr>
<td>Introductory</td>
<td></td>
</tr>
<tr>
<td>Overview</td>
<td></td>
</tr>
</tbody>
</table>
will impact relevance research design and measurement. The first category consists of the factors or variables that affect how a user makes relevance judgments. This category consists of the factors which are listed in Table 1 under the headings of judges, requests, information system, judgment conditions, and choice of scale. These factors are external to the information being judged and are not part of the user’s conceptualization of relevance when making relevance judgments, although they may influence this conceptualization. Some examples of these factors are the subject knowledge of the user, research stage, and the user’s information need or problem. Some of these factors are simply artifacts of experimental design, for example choice of measurement scale and the instructions to the person making the relevance judgments (specification of task).

The second category consists of criteria or variables that are employed by the user to conceptualize relevance and make relevance judgments. These criteria include the characteristics of the information being judged and the presentation or representation of that information and will be referred to as user-defined relevance criteria. These criteria in combination represent how actual users conceptualize relevance. User-defined relevance criteria are employed by users in judging information and represent their definition of relevance in regard to their information problem at a particular point in time.

If relevance is measured using a combination of the user-defined relevance criteria then the situational variables that influence relevance judgments must be separated from variables or criteria that the user is employing to make these judgments in order to assure meaningful, valid, and reliable measurement. The factors that affect relevance and relevance criteria will show interaction. A user with subject expertise may employ
different criteria and ultimately make a different relevance judgment than a subject novice. In any study calling for the measurement of relevance judgments, good experimental design requires that the factors in category one be controlled for as much as possible. These factors will bias relevance judgments regardless of whether the study measures relevance based on a single variable definition or employs user-based relevance criteria to measure relevance. The levels of any factor or variable in this category should be controlled by the researcher. Differences in these variables should be built into the experimental design as much as possible. Whatever factors are built into the experimental design will be independent variables. The dependent variable is the relevance judgment itself since it is what is being measured. How the researcher defines relevance will become the basis for how the dependent variable or variables are measured.

This study proposes that to adequately measure changes in relevance or relevance judgments, some combination of these variables is necessary. Utility, topicality, quality, and value represent various researchers' definitions of relevance. The user-defined relevance criteria have captured what relevance means to actual users and a study of changes in the use of these criteria and the interactions and correlations among these criteria will allow for the development of an instrument to measure how a user conceptualizes relevance and changes in this conceptualization.

Measuring the importance of criteria that define relevance for actual users and seeing how these criteria correlate with each other will help further our understanding of how users make and change relevance judgments. Changes in these criteria and their importance will indicate why the relevance judgment was made and why it changed, not
just what external factors might have caused it to vary from one person to another or one moment to another.

Determining Information Seeking Stages

Kuhlthau (1994) is the only researcher who has studied information seeking over the entire course of information problem resolution. Kuhlthau’s respondents and information seeking problems were similar to those who will be studied in this research; she studied a class of high-achieving high school students who were seeking information for a research paper assignment. Kuhlthau’s information seeking stages will act as a framework for this research. However, relevance judgments may not occur at each of Kuhlthau information seeking stages. Users may pass through some stages by thinking about information they have already obtained and may not actively seek or evaluate information at all stages. Also it is not clear how users pass through these information seeking stages, some may follow a linear process, whereas others may go through several stages more than once in an iterative process. All users may not go through all stages and some users may feel they are experiencing several stages at once. In spite of these limitations, Kuhlthau’s model is the only empirically developed information seeking model that has delineated the characteristics of the various stages of information seeking within the context of a single, complex information problem. This model is fairly easy for users to understand and Kuhlthau found that users were easily able to identify their information seeking stage using this model. Respondents in a pilot study for this research
had no difficulty in using the Kuhlthau model to identify their information seeking stage or stages.

End-users are likely to make relevance judgments that result in the selection or rejection of information sources at four points in any search process. The information that the user will be able to ascertain regarding the document he or she is judging will be dependent on the user’s point in the relevance judgment process. A user who has only a citation has much less information about the document than a user who has read the document. The points where users are likely to make relevance judgments are:

1. When they decide to write down or print a citation from an IR system or from another source that points to full-text information.
2. When they actually decide to seek and obtain the full-text information. At this point availability of information can be a confounding variable, that is information that appears to be highly useful, valuable, and relevant may never be used because it cannot be obtained.
3. When the user decides to read the information.
4. A final judgment is made when the information is incorporated into the user’s solution of the information problem or cited by the user within his or her own communications to inform others.

Most of these judgment stages may take place in any or all of Kuhlthau’s six information seeking stages. Some materials may be judged several additional times; for example, the user may change a judgment of a particular document from not relevant to relevant.
Summary

Relevance research has been conducted predominantly at the users point of contact with the IR system interface where relevance is measured as a single variable, even though many researchers and theorists agree that relevance is a dynamic and multidimensional concept. Relevance is also situational and this has been reflected by the tendency for researchers to examine users who are using IR systems in naturalistic setting to resolve their own information problems. Several studies have elicited the criteria that users employ in selecting and rejecting information sources (Barry, 1993; Schamber, 1991; Su, 1992). These criteria show considerable overlap between studies and among users and can be seen as comprising a broadly drawn user-defined concept of relevance.

Relevance is rarely studied within the context of information seeking or the complete process of resolving an information problem. Information seeking research has predominantly examined very broad information seeking behaviors of specific groups of individuals with an emphasis on the use or non-use of library resources. Kuhlthau (1994) appears to the only researcher who has examined the complete information seeking process to resolve an information problem. Studies that have collected relevance judgments at more than one point in the information seeking process indicate that relevance judgments change considerably and that a judgment of relevance at the IR system is not a good predictor of whether a document will be obtained or cited (Smithson, 1994; Sandore, 1990). To better understand user relevance judgments these judgments need to be studies within the context of the entire information seeking process.

If relevance is measured as changes in the importance in criteria that have been
determined by other end-users, this will provide some hints as to why relevance judgment change and what support the IR system can provide for end-users at different stages in their information seeking process.
CHAPTER 3

METHODOLOGY

Introduction

The purpose of this study was to examine the importance of relevance criteria for a group of graduate students in the selection and use of highly relevant or valuable information sources for an academic paper or project. The study examined the meaningful relationships between criteria that were rated as important by respondents. The study also related changes in the importance of these criteria to changes in respondents' information seeking stages. Respondents were asked to estimate the importance of 40 different user-defined relevance criteria in their selection and use of information sources. These information sources had been judged as highly valuable in the resolution of their information problem. The study was conducted in two phases. The first phase will be referred to as the survey phase and the second phase will be referred to as the class phase. In both phases of the study respondents were graduate students enrolled in course work. Respondents in the class phase evaluated the importance of the criteria two to four times during their information seeking process for an assigned class paper or project. Respondents in the survey phase were sent a mail survey and evaluated the importance of the criteria at only one point in their information seeking process.
In the survey phase, a group of 500 randomly selected graduate students enrolled at the University of North Texas (UNT) during the first summer session of 1997 were sent a mail survey (Appendix H, Appendix I, Appendix J, Appendix K). Two hundred-ten usable survey forms were returned after two mailings. These respondents were asked to focus on the most valuable source of information for a current or recent research paper or project. They were then asked to rate the importance of each of 40 criteria using a relevance criteria evaluation form (Appendix I, Appendix K). The relevance criteria evaluation form employed magnitude estimation methodology. Respondents were asked to estimate the importance of each criterion on a 100-mm line. This form was developed from a series of studies (Schamber & Bateman, 1996) and was tested for usability during a pilot study conducted in June 1996. The analysis of the data from the survey phase of the study represents a step in the process of developing a reliable and valid instrument to measure different constructs that contribute to high relevance.

Survey respondents were also asked to determine their stage in the information seeking process. Each of the six information seeking stages proposed by Kuhlthau (1994) was represented by a short phrase on the survey form. Respondents were also asked to indicate their degree program (Masters, Ph.D.), their discipline, and whether they had obtained and read the information source.

Eleven criteria whose mean importance rating was high were used to propose a three-construct model of high relevance. Confirmatory factor analysis was used to confirm the reliability of the criteria or sub-scale scores in measuring these constructs. Second-order factor analysis was then used to determine how much variance of the
concept of high relevance could be explained by each construct. Gorsuch (1983)
recommends a minimum sample size for factor analysis of five respondents for each
variable. Marcoulides and Schumacker (1996) recommend a sample size of ten
respondents for each variable. The sample size of 210 respondents was adequate for the
number of variables used in this factor analysis.

The software package Amos (version 3.61, 1997) was used in the data analysis of
the survey data. Because survey data were obtained at a single point in the information
seeking process of these respondents, a second phase of the study was conducted to
gather data throughout a complete search process for a paper that was assigned as a class
project.

The class phase of this study examined trends in the importance of relevance
criteria employed in evaluating information sources that were selected by respondents as
the most valuable at the points of searching for, obtaining, and reading information.
Thirty-five graduate students from seven classes at Texas Woman's University (TWU)
and the University of North Texas (UNT) volunteered to participate in the study
throughout their process of seeking information for a research paper for a class.
Respondents were asked to complete the same relevance criteria form (Appendix A).
They were asked to focus on the two to four information sources that they found to be the
most valuable at three different points in the information seeking process: when they
evaluated search results from an IR system, when they obtained full-text information and
when they read full-text information. Respondents were asked to determine their stage in
the information seeking process using a description of the six information seeking stages proposed by Kuhlthau (1994) (Appendix C).

Class respondents were also asked to complete a search diary form (Appendix B) where they described their current topic and their information seeking experience each time they sought information. The search diary was used to study changes in the respondent’s information seeking process and to relate their use of various information resources to the six information seeking stages. For the purposes of this study information resource will refer to IR systems, people and links that respondents used in their process of locating information sources. Class respondents represent a small sample of the graduate student population at each school and they were not selected at random. The data from the criteria evaluation form was described using a motion index calculated from changes in mean criteria importance. This index was developed by Janes (1991) to examine changes in user relevance judgments. The sample size was not large enough, given the number of variables being examined to support any further analysis. The analysis suggests some trends in the use of relevance criteria at different stages of information seeking. Further research will be needed to confirm these trends.

This chapter will discuss how this study examined criteria and constructs that contributed to graduate students’ conceptualization of relevance. Assumptions, the relevance criteria evaluation form, the motion index and validity and reliability issues will also be discussed.
Assumptions

This research is based on several assumptions regarding the measurement of relevance and the meaning and use of relevance criteria by end-users. Several assumptions regarding end-user information seeking behavior are also made.

Assumptions regarding relevance

Relevance, as conceptualized by the end-user at a particular point in time, can be measured using multiple variables that are represented by end-user relevance criteria. The importance and use of many relevance criteria are dependent on the user's situation. However, some criteria appear to be common to different user groups and may be consistently employed by end-users in making high relevance judgments (Schamber and Barry, 1998). Criteria that are rated as most important are likely to be the criteria that are less situation-dependent. These criteria may be used to develop a reliable and valid measure of high relevance across users. The study of the importance ratings of these criteria throughout the information seeking process provides insight into the different dimensions of high relevance and their relationship to changes in the user's situation. The combination of criteria that ultimately cause a user to judge a document as relevant or not relevant will depend on the user, his or her cognitive state, and the elements of the situation. One important part of the situation within this study was the focus on only highly relevant or valuable information sources.

Users employ a finite number of relevance criteria when they evaluate information. The criteria used in this study represent a selection of criteria that were
mentioned frequently by users in two studies (Schamber, 1991; Barry, 1993; Barry & Schamber, 1997). Many criteria were mentioned by several respondents in each study and a number of the criteria were common to both studies even though the respondent groups for each study were seeking information in very different contexts. The criteria selected for this study are representative of the criteria that users employ when making relevance judgments. Barry interviewed 18 respondents and found that no matter how the respondents were ordered, after the ninth respondent no new criteria were mentioned.

The terms used to represent relevance criteria will have a similar meaning to users within the context of an information problem. The criterion terms were selected from a group of terms that respondents in several studies were best able to clearly understand and use consistently (Schamber & Bateman, 1996). Users employed criterion terms most consistently when these terms were related to their own information seeking situations and problems. The criterion terms on the criteria evaluation form were grouped into broad categories (Appendix A) to provide respondents with a clear context for these terms. Some terms were presented as phrases to help clarify the meaning of the criterion to respondents.

Users are able to recognize and rate the importance of the criteria they use to evaluate highly relevant information. Users in studies by Barry (1993), Schamber (1991), Su (1994) and Wang (1994) had no problems reporting the criteria they used to evaluate information. Many of the criteria were mentioned by users in all studies. Respondents in studies by Eisenberg (1988) and Janes (1991, 1994) were able to use magnitude
estimation scales consistently and without difficulty to report and rate their overall relevance judgments of documents.

Assumptions regarding information seeking behavior

Users experience recognizable stages of information seeking. If provided with a list of characteristics of the different stages of information seeking respondents can determine their current stage of information seeking. Kuhlthau’s (1994) research determined that the six stages of information seeking are characterized by distinct patterns of thoughts, actions, and feelings. Kuhlthau found that end-users had little difficulty in determining their stage of information seeking. However, she found a skew in respondent self-evaluation of information seeking stage towards later stages (4,5,6) and away from early stages (1,2,3). Kuhlthau’s respondents reported stage 5 most frequently. In this study data from both the survey and the classes showed the same skew in the reported information seeking stages and stage 5 was the most frequently reported stage. The search diaries were used to triangulate class respondents’ evaluation of their information seeking stages and respondents appeared to be reporting their information seeking stages accurately.

Users are able to self-report their information seeking behavior accurately. Kuhlthau (1994) found consistent trends in thoughts, feelings and actions across her respondent population, which she identified as representative of six stages of information seeking. Dervin’s timeline method has been found to be reliable and valid in identifying user information seeking behaviors after the information problem has been resolved.
(Schamber, 1991). Ellis (1992) found that academic end-users reported a finite set of clearly identifiable information seeking strategies. Most class respondent search diaries were not detailed, but did report thoughts, feelings and actions that were consistent with respondents' progress through their information seeking process and with Kuhlthau's information seeking stages.

Users engage in multiple search sessions throughout the process of resolving a complex information problem such as writing a research paper. These searches are likely to occur at different stages of the information seeking process. Spink (1996) found that some users reported multiple search sessions as they sought information to resolve a complex information problem. The search diaries kept by Kuhlthau's (1994) respondents indicate that respondents searched for information at different information seeking stages. All but one class respondent reported multiple information search sessions. Class respondents were less likely to report early information seeking stages perhaps because they were not able to judge information sources as most valuable or highly relevant in these early stages. These respondents also may not have been actively searching for information in early information seeking stages. Some respondents reported experiencing several information seeking stages concurrently.

**End-User Criteria Evaluation Form**

The form used to obtain respondents' ratings of relevance criteria importance has been under development over a period of several years through a series of exploratory studies (Schamber & Bateman, 1996). The first phase of this study represents a step in the
validation of this instrument. Further studies are needed to complete this validation and to examine the use of this instrument in specific disciplines and user populations.

The goal of these exploratory studies was to identify a finite group of terms that represent the majority of relevance criteria employed by end-users. The criteria used in these studies were drawn from those mentioned by respondents in the Schamber (1991), Barry (1993) and Su (1992) studies as reasons why they selected or did not select information. Criteria, which were similar in meaning (plurals, close synonyms) were eliminated. Initially an attempt was made to use one word to represent a criterion. However, in some cases it was necessary to use a phrase to convey a consistent meaning of a criterion to all respondents.

In the first stage of the Schamber and Bateman (1996) study, respondents were asked to group criteria by meaning. Respondents had considerable difficulty with this task, due in part to the number of criteria (119). Respondents were not asked to focus on a specific information source or problem. As a result, some respondents did not relate these criteria to information or an information seeking process. In later stages of the study, users were instructed to focus on their specific information problem and the information source that was most valuable to them in the resolution of this problem. These respondents were better able to select the criteria they used to evaluate information and to group these criteria by meaning. Respondents were not asked to rate the importance of the criteria.

Over the course of these studies, it was possible to eliminate 79 criteria terms or phrases. Some of the eliminated terms were confusing or ambiguous to respondents and
others appeared to be very close in meaning to other terms. Terms that appeared to have
the clearest and most consistent meaning were selected for the final criterion term list.

A 100-mm line was chosen as a magnitude estimation scale based on research in
relevance judgments done by Eisenberg (1988). He found that users were able to easily
understand and use this scale and that this type of scale contributed to more consistent
relevance judgments. Data from this scale are ratio level and continuous. This type of
scale accommodates a wider range of variation than the traditional five to nine point
Likert scales. All class respondents found the magnitude estimation scale easy to use and
understand and used the scale as intended. However, a group of survey respondents
misunderstood the instructions and used the line as a two-part scale. These respondents
were able to fill out the scale as intended once they were provided with a set of
instructions that included a middle range example and emphasized that respondents could
mark anywhere on the line.

Characteristics of the Data

Survey Data

The registrar's office provided a mailing list of 500 randomly selected University
of North Texas (UNT) graduate students. This list included only students who were
enrolled in the first summer session 1997 at UNT to exclude inactive students from the
sample.

One hundred sixty-eight respondents returned the survey after the first mailing.
Thirty-nine (23%) respondents marked the 100mm line at only two points: not important
(approx. 10 mm) and important (approx. 90 mm). One of these respondents commented,
"it would be nice to have a category of so-so or medium importance". This comment made it apparent that some respondents had misunderstood the instructions and determined that the correct way to complete the criteria evaluation form was to mark the line under the label "not important" or under the label "important" resulting in a binary selection. These respondents were sent a second mailing with additional instructions (Appendix K) indicating they could mark anywhere on the line. They were also sent a cover letter (Appendix J) which asked them to complete the form again if they had misunderstood the instructions or to return the original form if it was marked the way they had intended. Twenty-two respondents returned the survey a second time, 15 respondents completed the criteria evaluation form again and nine respondents returned the original evaluation form. All respondents in the second mailing received the updated instructions.

Two hundred-thirteen forms were returned for a 44% response rate. For 17 respondents the survey was returned with no forwarding address. One respondent's name did not have a mailing address when the mailing list was obtained. The response rate was calculated as (213 (returned))/(500-1-17(no forwarding address)). One respondent returned the survey without completing it, stating she had not yet written a paper during her graduate program. Two survey forms were returned with only the first page completed. Two hundred-ten criteria evaluation forms were included in the data analysis.

Two hundred-thirty respondents returned the form at least once for a response rate of 48%. Seventeen respondents who initially made a binary selection on the criteria evaluation forms did not return the form when it was mailed to them again. The data on
these forms could not be included in the data analysis because it was not possible to
determine whether or not these respondents misunderstood the instructions for marking
the 100mm line.

One hundred forty-six survey respondents were masters students and 48 were
doctoral students. Sixteen respondents did not report their degree program. Respondents
reported a variety of disciplines. Education was the most frequently reported discipline.
The disciplines of these respondents are reported in Table 3.

<table>
<thead>
<tr>
<th>Education</th>
<th>Business</th>
<th>Humanities</th>
<th>Social Science</th>
<th>Library Science</th>
<th>Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>88 (42%)</td>
<td>31 (15%)</td>
<td>30 (14%)</td>
<td>11 (5%)</td>
<td>13 (6%)</td>
<td>32 (15%)</td>
</tr>
</tbody>
</table>

Note. Missing data are not reported in this table.

Table 3. Discipline of Survey Respondents

Respondents were asked to focus on the most valuable information source for
their project or paper when they filled out the criteria evaluation form. Journal articles
were the most frequently selected information source. One hundred-fifteen (55%)
respondents focused on journal articles, 53 (25%) respondents focused on books, 27
(13%) respondents focused on World Wide Web pages, and only three (1%) respondents
focused on citations. Seven (3%) respondents selected more than one type of information
source to focus on (even though they had been asked to select only one source) and five
(2%) respondents did not indicate the type of information source they focused on when
completing the criteria evaluation form. The number of respondents who selected World
Wide Web pages is interesting, particularly since respondents were instructed to pick the most valuable information source to focus on as they completed the criteria form.

Class Data

Thirty-five respondents from seven different classes completed the study. Three classes were offered through the College of Health Studies at TWU. These classes included Foundations of Health Science (summer, 1996), Foundations of Health Science (summer 1997), and Epidemiology (spring 1997). Four classes were offered through the Department of Library and Information Science classes at UNT. These classes included Management of Information Agencies (SLIS 5300) spring 1997 and summer 1997; Foundations of Library and Information Services (SLIS 5000) spring 1997; and Fundamentals of Information Organization (SLIS 5200) summer 1997, Houston campus. All 22 UNT respondents were master’s students; the 13 TWU respondents included four master’s students and nine doctoral students.

Although all respondents from the classes were Library and Information Science graduate students or Health Studies graduate students they reported a variety of disciplines for their undergraduate degrees. The disciplines reported for class respondents are listed in Table 4.

<table>
<thead>
<tr>
<th>Education</th>
<th>Business</th>
<th>Humanities</th>
<th>Social Science</th>
<th>Science</th>
<th>Health Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>(20%)</td>
<td>(9%)</td>
<td>(23%)</td>
<td>(11%)</td>
<td>(14%)</td>
<td>(17%)</td>
</tr>
</tbody>
</table>

Note. Missing data are not reported in this table.

Table 4. Undergraduate Discipline of Class Respondents
Not all students completed all four criteria evaluation forms. Two respondents completed two criteria evaluation forms, seven respondents completed three criteria evaluation forms, and 26 respondents completed four criteria evaluation forms. Twenty-seven respondents listed and evaluated at least one information source on two or more criteria evaluation forms. Two respondents evaluated the same information sources four times and five respondents evaluated different information sources on each of their criteria evaluation forms. Most respondents returned between two and four search diary forms. Three respondents did not report their information sources on their criteria evaluation forms. A total of 129 criteria evaluation forms were obtained from the classes.

The number of search diary forms for each respondent ranged from none to seven, with an average of 3.3 forms per respondent. Ten respondents returned two search diary forms, nine respondents returned three forms and seven respondents returned four forms. Only one respondent did not return any search diary forms. Respondents completed a total of 116 search diary forms.

The number of criteria evaluation forms completed by information seeking stage was skewed toward later search stages. Stage 5 was the most frequently reported stage for both class and survey respondents. Information seeking stage 6 was the second most frequently reported stage for both class and survey respondents. Eighty-six (41%) survey respondents reported their information seeking stage as stage 5 and 47 (22%) reported their information seeking stage as stage 6. Thirty-four (26%) criteria evaluation forms were completed when class respondents were in information seeking stage 5 and 28 (22%) were completed when class respondents were in information seeking stage 6.
Forms completed in stages 5 and 6 accounted for 63% of the survey criteria evaluation forms and 50% of the class criteria evaluation forms.

The number of survey respondents in each information seeking stage and the number of criteria evaluation forms reported for each information seeking stage by class and survey respondents are reported in Table 5.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Survey (n=210)</th>
<th>Classes (n=129)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>% single stage</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>86</td>
<td>47</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>26</td>
</tr>
<tr>
<td>Mixed</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Information Seeking Stages Reported by Respondents

Class respondents did not report stage 1 as a single stage on their criteria evaluation forms. However, four class respondents reported stage 1 in combination with other information seeking stages. Nineteen (9%) survey respondents reported multiple information seeking stages and 23 (18%) criteria evaluation forms were completed when class respondents reported more than one information seeking stage.

One hundred ninety-eight (94%) survey respondents completed the criteria evaluation form using information sources that they had obtained and read. Eight (4%) respondents reported that they had obtained but had not read the information source. Two
(1%) respondents reported that they had obtained the source and did not indicate whether they had read the source or not.

Twenty-three (18%) class criteria evaluation forms reported information sources that had not yet been obtained or read. Twenty-four (19%) class criteria evaluation forms reported sources that had been obtained. Thirty-five (27%) class criteria evaluation forms reported sources that had been read. Eighty-two criteria evaluation forms reported a single activity, i.e. searching or obtaining or reading. Forty-three (36%) criteria evaluation forms reported some combination of searching, obtaining, reading, and writing. Thirteen (10%) criteria evaluation forms reported all three activities concurrently; searching, obtaining, and reading. Four (3%) criteria evaluation forms reported writing as a single activity or in combination with reading. Table 6 reports the number of criteria evaluation forms returned by class respondents by activity.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Number</th>
<th>Percent</th>
<th>Percent single forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Searching</td>
<td>23</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>Obtaining</td>
<td>24</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>Reading</td>
<td>35</td>
<td>27</td>
<td>43</td>
</tr>
<tr>
<td>Searching, Obtaining</td>
<td>11</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Obtaining, Reading</td>
<td>19</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Searching, Obtaining, Reading</td>
<td>13</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Number of Criteria Evaluation Forms Returned by Class Respondents (n=129)
Nineteen respondents combined searching with obtaining or obtaining with reading. One respondent combined all three activities. Fifteen respondents did not report combined activities. It appears that when respondents search an online catalog, CD-ROM index or the Internet their searching may be immediately followed by obtaining the information and sometimes by reading or scanning the information. If a user’s information problem has not been resolved they may go back and repeat the cycle of searching, selecting, obtaining, and reading information again.

No respondents were in information seeking stage 6 when they completed forms that reported searching and obtaining as single activities. No respondents were in information seeking stages 2 or 3 when they completed criteria evaluation forms that reported reading as a single activity.

Even though respondents were asked to complete twice as many criteria evaluation forms when they were searching for information, only 28% of the single forms were completed when respondents were searching for information. Twenty-nine percent of the single criteria evaluation forms were completed when respondents were obtaining information and 43% of the single activity criteria evaluations forms were completed when the respondents were reading information. Seventy-five percent of the mixed activity criteria evaluation forms reported reading as one of the activities. It appears that the frequency of information seeking stages 5 and 6 can be explained in part by respondents’ preference for completing the criteria evaluation form after they had read the information sources that they were evaluating.
Respondents were asked to focus on the most valuable information sources for their paper or project when they judged the importance of the criteria. Respondent’s tendency to complete the criteria evaluation forms in later information seeking stages may indicate that respondents may have felt unable or were unwilling to judge information as the most valuable in resolving their problem before they read the information source. This may help explain why users do not always obtain information sources that they judged as highly relevant at the IR system interface. It appears that users may make the strongest and most definitive relevance judgments during or after they have read information. Reading information sources may change the user’s conceptualization of the information that they need to complete the resolution of an information problem and can influence any further information seeking and relevance judgments.

Data Analysis

Survey Data

This study proposed and validated three important dimensions or constructs of high relevance using data from the survey. Eleven variables that had high importance ratings were used to propose a three-construct model of high relevance. Survey data were analyzed using Amos (version 3.61, 1997) software to conduct a confirmatory and secondary factor analysis.

Factor analysis is used to identify or confirm that measured variables (criteria) share variance. Measured variables that show high variance and covariance (correlation) are grouped into smaller subsets of hypothetical constructs. Correlation between
measured variables is seen as the result of a shared common construct or dimension. The inter-item correlation matrix for the criteria was used to confirm their reliability in measuring the proposed constructs.

Ten criteria had a mean importance rating of 75 or more. Respondents from the survey and respondents from the classes rated the same ten criteria with a mean importance of 75 or more. The remaining 30 criteria in both the survey and the class evaluations had a mean rating that was less than 75. Nine of these ten criteria and two other criteria that appeared to share some meaning with one of the ten highly rated criteria were used to propose the model of high relevance. The three constructs or dimensions of high relevance and the criteria that loaded on each construct are as follows. Criteria that load on the construct of Information quality include V8 (Current), V9 (Well-written), V12 (Understandable), V13 (Consistent), and V14 (Focused). Criteria that load on the construct of Information credibility include V1 (About my topic), V10 (Credible), and V11 (Accurate). Criteria that load on the construct of Information completeness include V17 (Comprehensive), V18 (Suitably general or specific) and V19 (Detailed).

V1 (about my topic) was weakly correlated with the construct of Information credibility and probably is a separate dimension of high relevance. If other criteria that measured the construct of Topicality (aboutness) were added to the relevance criteria evaluation form, it is likely that the form would measure more of the variance in the concept of high relevance for graduate students. The components of topicality or aboutness are often very situational and may be difficult to measure across users.
V8 (Current) was weakly correlated with the construct of Information quality and probably is a separate dimension of high relevance. If other criteria that measured the construct of Currency were added to the relevance criteria evaluation form, it is likely that the form would measure more of the variance in the concept of high relevance for graduate students.

Two additional criteria, V2 (Easy to obtain) and V3 (Free or inexpensive) could have been included in this model, but each variable represented a unique concept that no other variables measured. Easy to obtain and Free or inexpensive did show some positive correlation. For some respondents free or inexpensive information may also have been interpreted as easy to obtain. However, some users will want information that is easy to obtain no matter what the cost and other users will want information that is free or inexpensive independent of how easy it is to obtain the information. Easy to obtain had a mean importance rating of 76.4 and Free or inexpensive had a mean importance rating of 66.2. Since Easy to obtain had a high importance rating, it is suggested that other criteria should be added to the criteria evaluation form that measures the separate construct of Availability reliably and validly. With the addition of this construct and the criteria that measure it, the form would measure more of the variance in the concept of high relevance for graduate students.

The three dimensions of high relevance for this user population were confirmed. These three dimensions were Information quality, Information credibility and Information completeness. Three additional dimensions or constructs appear to be important to this user population but will require additional criterion terms in order to measure them.
reliably and validly. These three dimensions are Information currency, Information topicality (aboutness), and Information availability.

**Class Data**

Data from the classes were analyzed using a combination of content analysis and a motion index first developed by Janes (1991). The motion index was used to describe the changes in the mean criteria importance in relation to the six stages of information seeking. Respondent search diaries and responses to the criteria evaluation form were used to identify the information seeking stages where respondents were most likely to be making relevance judgments, using IR systems, and obtaining and reading information. The search diaries and relevance criteria evaluation forms were also examined for mentions of relevance criteria, and trends or themes in the use of relevance criteria and respondents' progression through the information seeking process. The changes in respondents' topics throughout the information seeking process were examined and discussed.

This study was an exploratory study to see if changes in information seeking stages influence the importance of relevance criteria employed by end-users when judging information as most valuable or highly relevant. The number of respondents in this study is not sufficient to permit tests of statistical inference in regard to changes in the importance of the relevance criteria between information seeking stages. Instead the data will be described using the motion index developed by Janes (1991). This motion index describes the total change in the mean of respondents' ratings of the importance of each criterion; independent of whether this change was positive or negative. It is also possible
to transform the motion index to determine what percentage of this change can be attributed to positive change or movement.

Janes (1991) developed the motion index as a descriptive method for a study that examined changes in user relevance judgments. Document representations from an online search conducted for end-users were presented to these users in increments of the complete document representation. Users were randomly assigned to one of four groups. Each group was shown increments of the document representation in the one of the following orders: TAB (Title/Abstract/Bibliographic); TAI (Title/Abstract/Indexing); TBA (Title, Bibliographic/Abstract); and TIA (Title/Indexing/Abstract). As each of the three different representations were presented to respondents they were instructed to “make a mark on the line [100-mm line] corresponding to your impression of the degree of relevance of that document to your query, from none (N) to total” (p. 634). The resulting data were similar to the data for this study except each category of representation (for example title to abstract) had equal numbers of user relevance judgments.

Janes developed the motion index to describe the total movement or change in degree of relevance between each form of document representation. This motion index is used in this study to describe the change in importance for each relevance criterion as respondents moved between information seeking stages and as they sought, obtained, and read information. Respondents information seeking stages were self-determined by respondents who evaluated their own thoughts, feelings, and actions at different points in the project. As a result all information seeking stages were not equally represented in the
data. However, all but one respondent moved through the information seeking stages from earlier to later stages. This one respondent reported mixed stages in each of the two evaluations. On the earlier evaluation form this respondent reported stages 5 and 6 and later reported stages 4, 5, and 6. All other respondents who reported single stages went from earlier stages to later stages.

Class respondents returned 105 criteria evaluation forms, which reported a single information seeking stage and 82 criteria evaluation forms, which reported a single activity (searching, obtaining, reading). Survey respondents returned 184 criteria evaluation forms reporting a single search stage. These single stage and single activity forms were used to calculate means and motion indexes for each criterion. Survey respondents returned only eight criteria evaluation forms that reported either obtaining or searching so it was not possible to calculate a motion index for the activities of searching, obtaining and reading for these respondents.

The motion indexes for the search stages were: MI stages 2 to 3, MI stages 3 to 4, MI stages 4 to 5, MI stages 5 to 6, and total MI. The indexes for the activities were MI searching to obtaining, MI obtaining to reading, and total MI. The total motion index was calculated by summing the absolute value of linear change between each stage or activity. The mean of the earlier stage or activity was subtracted from the mean of the next stage or activity. Then the sum of the absolute (non-negative) values of each difference between the stages or activities was calculated.

Increases and decreases in criterion importance between stages canceled much of the movement between stages for some criteria, i.e. the “movement is not sensitive to the
direction of change” (Janes, 1991, p. 636). The total motion index for the information seeking stages reflects the total amount of change in the importance of each criterion between each pair of information seeking stages from stage 2 to stage 6. For example if the mean importance of a criterion between each pair of information seeking stages was calculated as 30, 40, -30, and 45 the total MI would be 10 + 10 + 15 or 35.

Janes (1991) states; “Linear change (simple subtraction was used rather than proportion change (ratios), in the belief that a change in judgement from 50 to 100 is more important than that from 1 to 2, and that the MI should reflect that.” (p. 636). This rationale is applicable to this research as well.

To determine the percentage of the movement that occurs between each information seeking stage the difference between each pair of stages was divided by the total motion index. In this case the absolute value of the difference was not used. The percentage figure for each pair of stages indicated how much of the total change in each criterion mean importance could be attributed to the change in criterion mean importance between these stages and whether this change was positive or negative.

Janes (1991) also calculated the percentage of positive change in the total motion index. He called this measure delta. The delta for each variable is calculated by subtracting the mean of stage 2 from the mean of stage 6. This allows negative and positive changes to cancel each other out. This figure is then divided by the motion index. The result will be a number between -1 and 1 where -1 represents all negative change (the mean criterion importance of each stage is less than the previous stage) and 1
represents all positive change (the mean criterion importance of each stage is greater than
the previous stage).

A data transformation was done by multiplying the previous calculation by 50 and
then adding 50 to it. The transformation figure indicates what percentage of the total
movement or change is positive or negative. For example a criterion may have a total
motion index of 20 and the difference in its mean importance from stage 2 to stage 6
could be -10. The overall change in mean importance from stage 2 to stage 6 is negative,
but since a motion index of 20 is greater than the absolute value of -10 some of the
change between stages must have been positive change. The percentage of positive
change in mean importance is calculated as follows: (50 * (-10/20)) + 50, making the
percentage of positive change in this variable 25%.

Criteria where positive and negative changes in importance between stages are
equal will have a delta of 50%. For example if the MI of a criterion is 40 but the change
in its mean importance from stage 2 to stage 6 is -20 half of the change in mean
importance is negative and half is positive. The delta will be (50 * (-20/40)) + 50,
making the percentage of positive change in this variable 50%.

The motion index was also calculated with the class data using the 82 criteria
evaluation forms that reported a single activity (searching, obtaining or reading). All
respondents who reported single activities moved sequentially from searching to
obtaining to reading and did not loop back. This allowed the motion index to be
calculated using the means of the importance of each criterion for each activity. These
means were used to calculate a motion index between searching for and obtaining
information sources, a motion index between obtaining and reading information sources, and a total motion index. The delta or percent positive change for each criterion was also calculated.

Survey respondents returned 184 evaluations that reported a single search stage. It was possible to calculate motion indexes for each change in search stage. The class data and the survey data were not directly comparable since the survey measured changes between respondents who were in different search stages whereas class respondent data included changes throughout each respondent's information seeking process.

Reliability Issues

Reliability refers to the consistency and replicability of data measurement under similar conditions. The use of a pre-tested form with criterion terms that were developed through several iterations with close to 100 respondents in previous studies (Schamber & Bateman, 1996) helped assure reliability. The magnitude estimate scale methodology used in this form has been found to be highly reliable in measuring relevance judgments (Eisenberg, 1988; Janes, 1991, 1994). The terms and phrases used on the form were employed consistently by respondents within the context of evaluating information sources related to their own information problems. The reliability of the sub-scale scores or criterion terms in measuring three constructs of high relevance was confirmed using factor analysis. Cronbach's Alpha coefficients were computed for each construct and found to be acceptable for an exploratory study such as this one.
Class respondents used the relevance criteria evaluation form consistently and were able to complete it easily and quickly. Each class of respondents was given a trial form and asked to imagine their ideal information source and to evaluate the importance of the criteria based on the ideal source. Even respondents who did not take part in the rest of the study had no difficulties using this form.

Validity Issues

Validity is the degree to which the concept or phenomena purported to be measured is what is actually being measured. Factor analysis was used to begin validation of the criteria evaluation form. Survey respondents came from a random sample of UNT graduate students who were taking course work in the first summer session of 1997. The number of respondents who returned the survey was adequate to insure a reliable and valid factor analysis for the eleven variables that were analyzed. Random sampling is one the strongest defenses against threats to validity (Campbell & Stanley, 1968).

Class respondents were at a similar academic level (beginning masters and doctoral students) and had the same general type of information problem (an academic research paper). Respondents were also asked to self-report their academic level, their familiarity with their subject area and their familiarity with information seeking using various library and computer resources. Most respondents indicated a moderate level of experience and confidence in seeking information for a research paper or project. Survey respondents came from a variety of disciplines and class respondents reported a variety of undergraduate degree areas.
Class and survey respondents were representative of upper-level and graduate student users of academic IR systems who are doing research for class papers and projects. These users represent an important segment of academic library and information system patrons. Most importantly, these respondents were rating criteria suggested by other users and reporting their information seeking process in the context of their own information problems in a naturalistic setting. This is in contrast to experimental research where information problems, IR systems, and citation lists are determined a priori by the researcher.

Relevance research indicates that relevance judgments change as additional information about a document is presented to the user (Janes, 1991). To control for this threat to validity, users were surveyed regarding whether they had obtained or read the information source. This threat proved to be very difficult to control for in this non-experimental setting. Respondents followed and reported their own preferred information seeking strategies. Almost all survey respondents (94%) evaluated information that they had already obtained and read. Class respondents were instructed to complete the criteria evaluation form before they obtained information, after they obtained information and after they read information. More respondents completed the forms after they read the information and group of respondents combined these tasks in various combinations. Either respondents were not willing to evaluate information as most valuable (high relevance) before they obtained and read it or the immediate availability of some full-text information on CD-ROM or the Internet made it difficult or undesirable to separate these tasks. The tendency of respondents to combine searching for, obtaining and reading
information and the apparent reluctance of respondents to evaluate information they have not obtained or read should be considered when conducting further relevance research.

Criterion terms were chosen for the form on the basis of clarity to other respondents. Terms that were confusing and had multiple meanings to respondents in earlier studies (Schamber & Bateman, 1996) were dropped. When necessary, phrases were used to better clarify the intended context and meaning of the term. Terms were grouped into broad general categories to provide a consistent context for these criterion terms to the respondents.

Qualitative data from class search diaries was used to triangulate the data from the criteria evaluation forms. This qualitative data was examined for evidence of criteria usage and for thoughts, feeling or actions that signaled search stage changes. This data was limited but it was consistent with the data reported on the relevance criteria evaluation forms and respondents' assessments of their information seeking stages. This triangulation helped assure both reliability and validity of criteria usage and information seeking stage selection by class respondents.

Limitations

Relevance judgments are dependent on the user's situation, which includes the type of problem and the user's subject and system expertise. This study examined only one group of users and cannot be broadly generalized to other user groups. The importance of some of the relevance criteria was probably influenced by the respondent's discipline. This study was cross-disciplinary. Even though class data was gathered only from Health Studies and Library and Information Science graduate students, the
undergraduate degrees of these respondents represented a broad range of disciplines. However, general patterns of criteria use and the relative importance of the various criteria were fairly stable in both groups. The same group of 10 criteria had a mean importance of 75 or more for both survey and class respondents. Criteria importance ratings appear to be representative of graduate students who are seeking information for a research paper or project.

This study examined importance ratings of relevance criteria used to judge information sources that respondents considered most valuable. Criteria that respondents used to determine that information was not relevant or was partially relevant to their information problem were not investigated. Reasons used in making highly relevant and not relevant judgments are often antonyms, i.e. about my topic, not about my topic. However, some criteria used in judging partial relevance and non-relevance may differ from those used to judge high relevance. Recent studies (Spink, Greisdorf, & Bateman, 1997; Spink & Greisdorf, 1996) indicate that the number of items judged partially relevant from mediated online database search results is significantly correlated with changes in information problems as judged by both the end-user and the intermediary. The study of relevance criteria and judgments of partially relevant information appears to be an important emerging area of relevance and information seeking research.

This study did not attempt to examine in detail how a user's topic or subject changed, although broad changes in topics were determined from the search diaries. In this study topicality was simply defined as the subject or aboutness of the information source and was measured as a single variable (about my topic). Both respondent groups
rated this criterion as very important and reliability and validity of the relevance criteria evaluation form could be improved if other criterion terms that measure topicality were added. A study of how users employ topicality in making relevance judgments is another important area for further research. Previous research indicates that topicality is a very complex issue for many users (Barry, 1993) and that many document representations do not provide the user with adequate information about the topic of the document (Weinberg, 1987).

Summary

A relevance criteria evaluation form was mailed to a random sample of UNT graduate students in the summer of 1997. These students were asked to focus on their most valuable information source for a current or recent paper or project. They were asked to rate the importance of 40 relevance criteria in regard to their selection and use of this information source.

Survey data from this study were then used to propose a model of high relevance. Eleven criteria whose mean importance rating was high contributed to this three-construct model of high relevance. Confirmatory factor analysis was used to confirm the reliability of the criteria or sub-scale scores in measuring these constructs. Second-order factor analysis was then used to determine how much variance of the concept of high relevance could be explained by each construct.

Respondents who were enrolled in graduate classes at TWU and UNT were asked to complete the same relevance criteria evaluation form as they searched for, obtained, and read information sources for a class paper or project. These respondents were also
asked to keep search diary forms to report their search process and the information resources they used in this search process. The ratings of criteria importance and search diary data obtained from the classes were used to determine changes in the importance of the relevance criteria as related to Kuhlthau's six stages of information seeking.

Data were analyzed using Amos (version 3.6, 1997) for factor analysis of the survey data. Jane's (1991) motion index and content analysis were used to analyze the class data. Assumptions, reliability issues and validity issues, the development of the relevance criteria evaluation form, and the motion index were also discussed in this chapter.
CHAPTER 4

DATA ANALYSIS AND RESULTS

Introduction

The study was conducted in two phases. In phase one (survey) a mail survey that included the criteria evaluation form (Appendix H, Appendix I, Appendix J, Appendix K) was sent to a random sample of 500 University of North Texas (UNT) graduate students during the 1997 summer sessions. In phase two (classes) 35 graduate students from seven different classes at Texas Woman’s University (TWU) and UNT volunteered to complete search diary forms (Appendix B) and criteria evaluation forms (Appendix A). Class respondents completed these forms during their information seeking process for a paper or project assigned as a class requirement.

The survey data was used to propose a three-construct model. These constructs contributed to survey respondents’ conceptualization of high relevance. Eleven criteria that survey respondents rated as important contributed to these constructs. The reliability of these criteria or sub-scale scores for the measurement of their corresponding construct was determined using confirmatory factor analysis. A model of meaningful relationships between these criteria, the three constructs, and high relevance was proposed. This model was confirmed using second-order factor analysis.

Analysis of the class data was highly exploratory and inferential statistics were not used in this part of the study. This phase of the study examined respondents who
evaluated and reported their own information seeking stages during the process of seeking information to resolve their information problem. Most respondents reported information seeking stages that were in sequential order but respondents did not go through the information seeking stages in a uniform manner. The result was an uneven distribution of criteria evaluation forms across Kuhlthau's six information seeking stages.

Class respondents were instructed to focus on the most valuable information source(s) for their paper or project and to rate the importance of each relevance criterion on the evaluation form. They were instructed to complete two criteria evaluation forms before obtaining the information source(s), one form after obtaining the information source(s) and one form after reading the information source(s). Many respondents had information seeking patterns that deviated from these instructions and returned criteria evaluation forms that reported combinations of searching for, obtaining, and reading information sources. Respondents also chose to complete more criteria evaluation forms when they were obtaining and reading information. This more natural and apparently more accurate reporting resulted in an uneven distribution of criteria evaluation forms which were completed while students were searching for, obtaining, and reading information sources.

Descriptive techniques and content analysis were used to analyze the class data. These descriptive techniques suggest some trends in the data that contribute to a better understanding of how and when these students made high relevance judgments and how their information seeking stage affected these relevance judgments. The data also
suggest some factors that influenced these respondents' progress through the information seeking stages.

Data Analysis of Specific Research Issues

The primary research question for this study was: What is the relationship between the criteria employed by respondents in selecting sources of information that they find to be most valuable and the respondents' stage of information seeking with respect to a specific information problem? This study examined graduate students who were seeking or had sought information for a research paper or project. The study surveyed graduate students at only one point in their information seeking process. This resulted in a snapshot of the importance of each relevance criterion in the selection of the most valuable source of information for a paper or project. Data were also gathered from a separate group of graduate students who were enrolled in classes. These respondents were asked to complete search diaries and multiple evaluations of the importance of relevance criteria as they sought, obtained, and read information sources throughout the course of a project. These respondents were instructed to focus on the most valuable information sources for their paper or project at four different times during their information seeking process.

In the following sections five research questions that were specifically addressed by this study will be discussed. The phase or phases of the study and the methodology used to analyze data for each question will be discussed for each question.

**Question One:** What are the Relationships between a Group of 40 End-user Relevance Criteria in Regard to Their Degree of Importance?
A survey of 210 UNT graduate students was conducted in the summer of 1997 to gather data to answer this question. These students were asked to focus on the most valuable source of information for a research paper or project and to complete a relevance criteria evaluation form with this information source in mind.

The criteria that respondents rated as most important in their evaluation of the information source were used to propose a three-construct model of high relevance. Confirmatory factor analysis was used to confirm the reliability of the criteria or sub-scale scores in measuring the constructs. Second-order factor analysis was then used to determine how much variance of the concept of high relevance could be explained by each construct.

The criteria used in this model included nine of the ten criteria that had a mean of 75 or more (on a 0 to 100 scale). Two additional criteria (V17 Comprehensive, and V19 Detailed) that had means of 73.3 and 71.3 respectively were also used in this model. Criteria rated as important and their role in the model are reported in table 7.

V17 and V19 were included because they were rated as only slightly less important than the top ten criteria. Also these criteria share some meaning with V18 (Suitably general or specific) and should contribute to the reliable measurement of the shared construct of Information completeness.

V2 (Easy to obtain) was dropped from the analysis, even though it had a mean of 76.0, because it did not share meaning with any other variable. V2 had some positive correlation with V3 (Free or inexpensive) but these two variables represent different
### Table 7. Criteria Rated as Important by Survey Respondents

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Mean Importance</th>
<th>Included in Model</th>
<th>Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1 About my topic</td>
<td>87.8</td>
<td>yes</td>
<td>Information credibility</td>
</tr>
<tr>
<td>V2 Easy to obtain</td>
<td>76.0</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>V3 Free or inexpensive</td>
<td>66.2</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>V8 Current</td>
<td>82.8</td>
<td>yes</td>
<td>Information quality</td>
</tr>
<tr>
<td>V9 Well-written</td>
<td>81.5</td>
<td>yes</td>
<td>Information credibility</td>
</tr>
<tr>
<td>V10 Credible</td>
<td>89.4</td>
<td>yes</td>
<td>Information credibility</td>
</tr>
<tr>
<td>V11 Accurate</td>
<td>89.5</td>
<td>yes</td>
<td>Information quality</td>
</tr>
<tr>
<td>V12 Understandable</td>
<td>86.0</td>
<td>yes</td>
<td>Information quality</td>
</tr>
<tr>
<td>V13 Consistent</td>
<td>83.4</td>
<td>yes</td>
<td>Information quality</td>
</tr>
<tr>
<td>V14 Focused</td>
<td>82.7</td>
<td>yes</td>
<td>Information quality</td>
</tr>
<tr>
<td>V17 Comprehensive</td>
<td>71.3</td>
<td>yes</td>
<td>Information completeness</td>
</tr>
<tr>
<td>V18 Suitably general or specific</td>
<td>75.5</td>
<td>yes</td>
<td>Information completeness</td>
</tr>
<tr>
<td>V19 Detailed</td>
<td>73.3</td>
<td>yes</td>
<td>Information completeness</td>
</tr>
</tbody>
</table>

Concepts since information that is easy to obtain may or may not be free or inexpensive and free or inexpensive information may or may not be easy to obtain. If other terms or criteria that share meaning with V2 and that relate to the construct of availability were added to the relevance criteria evaluation form, the reliability and validity of the form in measuring the concept of high relevance would likely increase.

V1 (About my topic) was grouped with V10 (Credible) and V11 (Accurate). V1 is weakly related to this construct (Information credibility) and was rated as very important by respondents (mean 87.8). No other items on the relevance criteria evaluation form were related to topicality, primarily because relevance criteria research has not yet identified components of topicality that are not highly dependent on the individual user's information need and situation.
V8 (Current) was grouped with V9 (Well-written), V12 (Understandable), V13 (Consistent) and V14 (Focused). V8 is weakly related to this construct (Information quality), but also represents a different, separate concept. To improve the reliability and validity of the relevance criteria evaluation form additional terms that describe the concepts of currency, topicality (aboutness), and availability should be added.

Since the concept Free or inexpensive (V3) was not rated as highly (66.2) this concept had a more limited contribution to the concept of high relevance judgment for this respondent group. This variable is probably a member of the larger group of variables that were rated as important within the context of specific information problems or situations and therefore cannot be used to measure the concept of high relevance judgment consistently or reliably across respondents.

Question Two: At What Stages of Information Seeking do Respondents Employ IR Systems or Information Resources?

Respondents were asked to indicate how many times they had used information resources on the search diary form (Appendix B). For the purposes of this study information resource refers to IR systems, indexes, people or any other resource that functions as a pointer to information. The information resources that were listed on the search diary form included library online catalog, print index to journals, CD-ROM index to journal articles, World Wide Web, librarian, reference or link from another document, other Internet resources, other people, and other. Respondents were asked to indicate the information resource by name or type for other Internet resources, other people, and other. The search diary forms were matched with the relevance criteria evaluation forms
to determine the information seeking stage of respondents at the time of completing each search diary. Table 8 lists information resource use by information seeking stage.

The overall use of information resources was highest in information seeking stage 5, but respondents used information resources in all information seeking stages. Use of each type of information resource was fairly even across search stages with the exception of CD-ROM indexes, which had a proportionally higher rate of usage in stages 2 and 3. The search diaries also indicate that some respondents did not continue to search for information after they had begun to read and write their papers. However, another group of respondents continued to search for information throughout the reading and writing process. This may be an artifact of the amount of time available for students to complete the paper, especially for summer session students who had a maximum time of five weeks to complete the paper. Two patterns of information seeking were observed in students who completed the study during a long semester. One group started their information seeking early in the semester and continued to seek information over a two to three month period. The other group waited until two or three weeks before the paper was due and did intensive information seeking, reading, and writing during that period.

Only two respondents completed less than two search diary forms. This indicates that 33 of the 35 respondents conducted multiple search sessions on their topic. Additionally, respondents were asked to complete only one search form each time they went to the library, used the Internet etc. and most respondents reported more than one search session on each search diary form.
<table>
<thead>
<tr>
<th>Stages</th>
<th>Two</th>
<th>Three</th>
<th>Four</th>
<th>Five</th>
<th>Six</th>
<th>Mixed Stage more than 4</th>
<th>Mixed Stage less than 3</th>
<th>Other mixed stages</th>
<th>Missing</th>
<th>Total</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPAC</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>13</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>49</td>
<td>23</td>
</tr>
<tr>
<td>Print index</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>CD-ROM index</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>40</td>
<td>19</td>
</tr>
<tr>
<td>WWW</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>51</td>
<td>24</td>
</tr>
<tr>
<td>Librarian</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Link</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Other Internet</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Other people</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>29</td>
<td>22</td>
<td>42</td>
<td>38</td>
<td>18</td>
<td>26</td>
<td>14</td>
<td>10</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>Percent</td>
<td>7</td>
<td>14</td>
<td>10</td>
<td>20</td>
<td>18</td>
<td>8</td>
<td>12</td>
<td>7</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8 Information Resource Use Reported at Each Stage of Information Seeking by Class Respondents
Question Three: At What Stages of Information Seeking do Respondents Search for, Obtain, and Read Information?

Respondents were asked to complete the criteria evaluation form twice as they searched for information but before they had obtained information, once when they obtained information but before they had read it and once after they had read information. Respondents were also asked to indicate their stage of information seeking on the same form. Of the 35 respondents 20 combined at least two of these activities (searching, obtaining, reading) on at least one form. Four forms were filled out when the respondent was writing or reading and writing.

Respondents in the classes returned 23 criteria evaluation forms that reported more than one information seeking stage. Table 9 reports the information seeking stage of the respondents at the time of seeking, obtaining, and reading.

<table>
<thead>
<tr>
<th>Information seeking stage</th>
<th>Searching</th>
<th>Obtaining</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 2</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Stage 3</td>
<td>9</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Stage 4</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Stage 5</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Stage 6</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>24</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 9. Information Seeking Stage of Class Respondents when Searching, Obtaining, or Reading

Early stages (2 and 3) of information seeking were reported only when respondents were searching for or obtaining information. Although later stages of information seeking were most frequently reported when respondents were reading or
writing, six forms reported stage 5 when the respondent was searching for information and eight forms reported stage 5 when the respondent was obtaining information. Stage 6 was not reported until respondents were reading information.

It appears that respondents may have started searching for information with a clear idea of their topic. Only two respondents changed to a completely different topic as they moved through their information seeking process. Several factors are probably responsible for this. All students were graduate students and may have already had some familiarity with their topic, even though classes selected for this study were primarily core classes for students just beginning the graduate programs. However, some students in the classes were further along in their program and they may have had a clearer idea of the topics they wanted to study.

The skew towards later information seeking stages was also probably due to several factors. Four classes were held during five-week summer sessions where students may have been forced into the later stages of information seeking by the course time constraints. These students frequently complained about their frustration with the short time period for completing the paper. Several students indicated that their information seeking behavior would have been more deliberate and comprehensive if they had been allowed more time to complete the paper. One student ordered a book on interlibrary loan only to find out that the time frame for ILL was too long.

Respondents may have wanted to be in a more “comfortable” information seeking stage and reported later stages than they were actually in. Kuhlthau (1994) found that respondents in first three stages of information seeking experienced anxiety and frustration until they identified a focused topic. Respondents experienced some relief in
stage 2 when a general topic was identified. No respondent from the classes reported being in stage 1 as a single stage. This may be because stage 1 was uncomfortable for them or because they felt they could not identify a valuable source of information in regard to their information problem at this time. However, 17 respondents from the survey reported being in stage 1 as a single stage and only four respondents reported being in stage 2. In both the class and the survey early search stages were underrepresented.

Ninety-four percent of the respondents from the survey reported that they had obtained and read the source that they focused on when filling out the criteria evaluation form. In this group of respondents all information seeking stages were reported when respondents were reading information. This may be an artifact of asking respondents to focus on highly relevant information. Respondents may not be willing to judge information as highly relevant until they have read it.

**Question Four: Does the Importance of End-user Relevance Criteria Change as the Respondent’s Information Seeking Stages Change?**

Respondents from the classes were asked to rate the importance each of 40 end-user relevance criteria by marking a 100-mm line. On this line, 0 was not important and 100 was most important. Respondents were asked to list the most valuable sources for their paper or project and to focus on these sources while completing the criteria evaluation form. They were asked to complete this form at four different times during their information seeking process and to indicate their information seeking stage on the form. Two respondents filled out only two forms and seven respondents filled out three
evaluation forms. Respondents who turned in only one evaluation form were dropped from the study. Only the forms that were returned when the respondents reported being in a single information seeking stage were used in this descriptive analysis. One hundred-five criteria evaluation forms reported a single information seeking stage and were used in this analysis. No class respondent reported being in stage one when completing this form.

A descriptive motion index developed by Janes (1991) to describe the movement of user relevance judgements was used in this part of the analysis. This index was used to describe the total motion of each criterion from stages 2 to 6 and to describe the motion of each criterion from stages 2 to 3, stages 3 to 4, stages 4 to 5, and stages 5 to 6. The motion index expressed the absolute change in the importance of each criterion.

The survey respondents’ evaluation forms were also described using the motion index. Only forms that reported single information seeking stages were used in this descriptive analysis. Only four respondents reported information seeking stage 2 so information seeking stages 1 and 2 were combined for a total of 21 respondents. One hundred eighty-four survey respondents reported single information seeking stages.

**Question Five: Does the Importance of End-user Relevance Criteria Change as Respondents Search for, Obtain, and Read Information?**

Respondents in the classes returned 82 criteria evaluations that reported a single activity (searching, obtaining, and reading). The means for each criterion were calculated for each of these activities. These means were used to calculate the total motion index.
and motion indexes between searching for and obtaining information sources and obtaining and reading information sources.

Almost all survey respondents focused on an information source that they had already obtained and read. The low incidence of sources that respondents had not obtained and/or read precluded the calculation of a motion index for survey respondent judgments using the activities of searching, obtaining, and reading.

The motion index is described and discussed in a later section of this chapter. Since this is an exploratory study with a large number of variables no further analysis was done to answer research questions four and five. The results of the analysis using the motion index indicate some trends and patterns in the data. To confirm these trends further research with a greater number of subjects will be necessary.

Factor Analysis

A model of high relevance and three constructs that contribute to the concept of high relevance for graduate students was proposed using eleven criteria that survey respondents rated as important. The development of this model was guided by the similarities in criteria and criteria groupings that Barry and Schamber proposed from their research (Barry & Schamber, in press). They elicited relevance criteria from two very different respondent groups: respondents who were searching for weather information using a variety of media and respondents who were evaluating online searches which resulted in text-based citations. Confirmatory factor analysis was used to confirm the model and verify that the constructs would produce reliable sub-scale scores.

The model that was proposed is shown in Figure 11. Within this model three constructs or dimensions of high relevance are proposed. The model proposes that V8
Figure 11
High Relevance Second-Order Factor Model
($\chi^2$/df=2.02, p<.05, GFI=.94, RMSEA=.07)
(Current), V9 (Well-written), V12 (Understandable), V13 (Consistent), and V14 (Focused) will be highly related to a construct called Information quality. Each of these criteria is related to some aspect of the quality of the information content. Other criteria or clues may be used to help judge these quality criteria (a peer-reviewed journal, a particular author), but the criteria themselves deal with the information content. V8 (Current) can be viewed as a quality criterion for information content in some situations. Some users will interpret the currency of information as a criterion that measures some part of information quality. However, for other users currency may simply be a characteristic of the information source. This helps explain the low reliability score of this variable for this construct. Since this variable was rated as highly important (82.8) it was not dropped from the model. If other criteria were added to the form that shared meaning and importance with the construct of Current another construct could be added and a model that explained more of the concept of high relevance for graduate students could probably be proposed and confirmed.

The second construct in the model is called Information credibility and is measured by three criteria: V1 (About my topic), V10 (Credible), and V11 (Accurate). About my topic is weakly related to the construct and shows poor reliability. Since it was a very important criterion (87.8) it was included in the model of high relevance, but this research did not attempt to measure topicality in any other way. If other variables that measured the construct of Topicality (aboutness) were added to this form, then it would be possible to have a separate construct that measured topicality and a model that explained more of the concept of high relevance for graduate students could probably be proposed and confirmed.
The last construct, Information completeness was measured by three criteria: V17 (Comprehensive), V18 (Suitably general or specific), and V19 (Detailed). This model also proposes that the constructs of Information quality, Information credibility, and Information completeness will be correlated with each other and that they will explain some of the concept of high relevance judgments for graduate students.

To verify that the three constructs would produce reliable sub-scale scores, an inter-item correlation matrix measurement of Crohnbach's Alpha was computed for each construct. Table 10 indicates the construct name, inter-item correlation matrix and Cronbach Alpha internal consistency reliability coefficient for each construct with missing data on some items.

<table>
<thead>
<tr>
<th>Construct Name</th>
<th>Inter-item correlation matrix</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Quality</td>
<td>1.00</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>.27</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.24</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>.20</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>.22</td>
<td>.33</td>
</tr>
<tr>
<td>Information Credibility</td>
<td>1.00</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>.28</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.31</td>
<td>.66</td>
</tr>
<tr>
<td>Information Completeness</td>
<td>1.00</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td>.39</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.30</td>
</tr>
</tbody>
</table>

Table 10. **Inter-item Correlations and Reliability for Constructs for High Relevance with Missing Data**

Since only a few variables had missing data, this reliability analysis was repeated using mean substitution for missing data. Table 11 shows the inter-item correlations and reliability coefficients for the constructs with mean substitution for missing data.
<table>
<thead>
<tr>
<th>Construct Name</th>
<th>Inter-item correlation matrix</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Quality</td>
<td></td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.27</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.24</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>.20</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>.22</td>
<td>.33</td>
</tr>
<tr>
<td>Information Credibility</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.28</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.31</td>
<td>.66</td>
</tr>
<tr>
<td>Information Completeness</td>
<td></td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.36</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.30</td>
</tr>
</tbody>
</table>

Table 11. Inter-item Correlations and Reliability for Constructs for High Relevance with Mean Substitution for Missing Data

The differences between the Cronbach’s Alpha reliability coefficients for the analysis conducted with missing data and the analysis conducted with mean substitution for missing data were less than .001 for the constructs of Information quality and Information credibility. The difference between the Cronbach’s Alpha reliability coefficients for the analysis conducted with missing data and the analysis conducted with mean substitution for missing data was less than .01 for the construct of Information completeness. The criteria that were used to measure Information completeness had a slightly greater number of missing values. These reliability coefficients indicate that mean substitution for missing data is appropriate for this analysis.

Since this is an exploratory study, which examines the relative importance of high relevance criteria for the first time, the Cronbach’s Alpha scores of .71, .67, and .56 were acceptable. Dropping criteria V1 (about my topic) and V8 (current) from the analysis could raise these scores. However, relevance research and the high mean
importance ratings of these criteria suggest that they are important components of high relevance for graduate students and that other criteria that scale reliably and validly with the constructs topicality and current should be added to the form.

The results of the first order factor analysis for each construct are shown in Table 12.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Information Completeness</th>
<th>Information Quality</th>
<th>Information Credibility</th>
<th>( h^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>V19 Detailed</td>
<td>.60</td>
<td></td>
<td>.36</td>
<td></td>
</tr>
<tr>
<td>V18 Suitably general or Specific</td>
<td>.55</td>
<td></td>
<td>.30</td>
<td></td>
</tr>
<tr>
<td>V17 Comprehensive</td>
<td>.50</td>
<td></td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>V14 Focused</td>
<td></td>
<td>.62</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>V13 Consistent</td>
<td></td>
<td>.63</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>V12 Understandable</td>
<td></td>
<td>.72</td>
<td>.52</td>
<td></td>
</tr>
<tr>
<td>V9 Well-written</td>
<td></td>
<td>.64</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>V8 Current</td>
<td></td>
<td>.36</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>V11 Accurate</td>
<td></td>
<td></td>
<td>.83</td>
<td>.69</td>
</tr>
<tr>
<td>V10 Credibility</td>
<td></td>
<td></td>
<td>.80</td>
<td>.64</td>
</tr>
<tr>
<td>V1 About my Topic</td>
<td></td>
<td></td>
<td>.35</td>
<td>.12</td>
</tr>
<tr>
<td>Factor Variance Explained</td>
<td>30%</td>
<td>37%</td>
<td>48%</td>
<td></td>
</tr>
</tbody>
</table>

Table 12. First-Order Factors

The factor loading of each variable with the construct is shown. These factor loadings are then squared to determine the amount of variance in the construct that is explained by that variable \( (h^2) \). Variables 1 (About my topic) and 8 (Current) both show weak factor loadings and therefore explain very little of the factor variance (12 and 13 % respectively).

After determining that the factor scales were sufficiently valid and reliable a second order factor analysis was performed because the factors significantly inter-
correlated. Table 13 shows the correlation of each factor with each other factor or the factor inter-correlations.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Information Quality</th>
<th>Information Credibility</th>
<th>Information Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Quality</td>
<td>1.00</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Information Credibility</td>
<td>.54</td>
<td>1.00</td>
<td>-----</td>
</tr>
<tr>
<td>Information Completeness</td>
<td>.46</td>
<td>.33</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 13. Factor Inter-correlations

Table 14 shows the second order factor analysis factor loadings and communality ($h^2$). The variance explained by each first order factor indicates how much of the variance in the second order concept high relevance is explained by the three factors: Information quality, Information credibility, and Information completeness. These three constructs together explain 48% of the concept of high relevance judgments for graduate students.

<table>
<thead>
<tr>
<th>Factor</th>
<th>High Relevance Factor Loading</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Quality</td>
<td>.87</td>
<td>.76</td>
</tr>
<tr>
<td>Information Credibility</td>
<td>.63</td>
<td>.40</td>
</tr>
<tr>
<td>Information Completeness</td>
<td>.53</td>
<td>.28</td>
</tr>
<tr>
<td>High Relevance Variance Explained</td>
<td></td>
<td>48%</td>
</tr>
</tbody>
</table>

Table 14. Second-Order Factors

The three constructs can be measured reliably and accurately by criteria: V9 (Well-written), V10 (Credible), V11 (Accurate), V12 (Understandable), V13 (Consistent), V14 (Focused), V17 (Comprehensive), V18 (Suitably general or specific) and V19 (Detailed). Criteria V1 (About my topic) and V8 (Current) were included in the model. Each of these two criteria probably measures part of a separate construct, and if other criteria were added to the form that measured these constructs, a model that
explained more of the concept of high relevance judgements for graduate students could probably be proposed and confirmed.

This second order factor analysis fits well with relevance theory. A limited number of variables can be used to measure a portion of the concept of high relevance, but a larger number of criteria are very situational and only explain a part of high relevance for some individuals in some situations. Except for variables 1 (About my topic) and 18 (Suitably general or specific), all criteria in both phases of the study were rated from a low of 10 or less to a high of 95 or more. The standard deviations of the criteria means ranged from 10.6 to 33.8 in the survey data and 9.1 to 30.7 in the class data. This variability and range of values can be explained as the effect of situational relevance. The importance of each criterion varied depending on the respondent’s situation. This situation included the respondent’s information problem, the respondent’s knowledge of the topic, and the respondent’s topic and discipline. The high standard deviations and range of the 30 criteria that had mean ratings of 75 or less indicates that these criteria are dynamic and situational. They can be very important in the context of some information seeking problems. The criteria used in the model and factor analysis were not as dependent on the individual respondent’s information situation and thus can be used to measure a portion of high relevance for graduate students.

**Description of Changes in Criteria Importance using the Motion Index**

**Class Data Described by Search Stage**

After the total motion index and the motion indexes between search stages were calculated the stages were ordered from the greatest change in mean criteria importance
to the least change in mean criteria importance for each criterion. Because the greatest amount of actual change between any two stages was only 20.5 it was necessary to examine the amount of actual change to decide if the change was meaningful. An actual change of 10 represents a movement of 10% on the 100-mm line. Since respondents did not have any guides on the lines, they were visually estimating their rating of the importance of the criteria. This visual estimation could show some variation from evaluation to evaluation even though a respondent intended to mark the line at the same point on each evaluation form. Since the motion indexes were calculated from the means for each criterion, a movement of 10 to 20 percent indicates a slight trend in the data. However, this is still a relatively low amount of change.

In the class data the greatest change between any two stages for 22 criteria was between 10 and 20%. The next greatest change between any two stages was greater than 10 for 5 criteria. Only one criterion (V36 Provides background or history) changed more than 10% in the third greatest change between stages.

Table 15 reports the number of times a stage had the highest motion index and the number of times a stage had the second highest motion index for the class data.

<table>
<thead>
<tr>
<th>Stages with largest MI</th>
<th>Number of criteria</th>
<th>Stages with next largest MI</th>
<th>Number of criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI23</td>
<td>17</td>
<td>MI23</td>
<td>9</td>
</tr>
<tr>
<td>MI34</td>
<td>6</td>
<td>MI34</td>
<td>9</td>
</tr>
<tr>
<td>MI45</td>
<td>13</td>
<td>MI45</td>
<td>11</td>
</tr>
<tr>
<td>MI56</td>
<td>4</td>
<td>MI56</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 15. Information Seeking Stages with Largest and Next Largest Motion Indexes, Class Data
In the class data seventeen of the 40 variables had the largest motion indexes between stages 2 and 3 and 13 of the 40 variables had their largest motion index between stage 4 and 5. Of the 17 variables where stage 2 to 3 had the largest motion index, ten of these variables showed an actual change of 10 or more. Of the thirteen variables that had the largest motion index between stages 4 and 5, six showed an actual change of 10 or more.

The deltas for each variable indicate that the percent of positive movement in the variable. Variables with small or large deltas have either a larger positive or negative change associated with them and changes between stages do not tend to cancel each other out. Deltas of between 40 and 60 indicate that the variable was showing almost as much positive change as negative change, with a delta of 50 indicating an equal amount of positive and negative change. Nineteen variables had a delta of 70 or more and seven variables had a delta ranging between 60 and 70. Five variables had a delta of less than 40. Only nine variables had deltas between 40 and 60.

This description indicates that some change was occurring in the mean importance of the criteria with the strongest changes between stages 2 and 3, followed by changes between stages 4 and 5. However, actual changes in the criteria importance between search stages and overall changes in criteria importance are quite small. This indicates that criteria for high relevance did not change much as respondents moved through the different stages of information seeking.

In a recent study Spink and Greisdorf (1997) found that the number of items judged partially relevant during an online search correlates significantly with indicators of change in the end-user’s information problem. This may partially explain why
relevance criteria showed little change in importance since respondents in this study were focusing on valuable or highly relevant information. Further study is needed to examine if or how criteria importance changes for information that is judged partially relevant.

**Survey Data Described by Search Stage**

In the survey data there was less evidence of meaningful changes between search stages. The greatest motion index was 10 or more for only ten criteria from the survey data. The next largest motion index was 10 or more for only four criteria. For three of the four criteria, the largest change was a negative change between stages 3 and 4. For these three criteria the next largest change was a positive change between stages 4 and 5.

Table 16 reports the number of times a stage had the highest motion index and the number of times a stage had the second highest motion index for the class data.

<table>
<thead>
<tr>
<th>Stages with largest MI</th>
<th>Number of criteria</th>
<th>Stages with next largest MI</th>
<th>Number of criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI23</td>
<td>10</td>
<td>MI23</td>
<td>11</td>
</tr>
<tr>
<td>MI34</td>
<td>20</td>
<td>MI34</td>
<td>6</td>
</tr>
<tr>
<td>MI45</td>
<td>7</td>
<td>MI45</td>
<td>13</td>
</tr>
<tr>
<td>MI56</td>
<td>3</td>
<td>MI56</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 16. Information Seeking Stages with Largest and Next Largest Motion Indexes, Survey Data

The greatest number of changes in criteria importance in the survey data occurred between stages 3 and 4 where 20 criteria had the largest motion index. Between stages 2 and 3 ten criteria had their largest motion index. Of the ten criteria whose largest change was ten or more seven changes were between stages 3 and 4, two changes were between stages 2 and 3, and only one change was between 4 and 5.
More criteria in the survey data had changes that cancelled each other compared to the class data.

21 criteria had a delta of between 40 and 60. More criteria had an overall negative change. Thirteen criteria had a delta of 40 or less. Fewer criteria had an overall positive change and this change was less pronounced than in the class data. Only one criterion (V1 About my topic) had a delta of 75 or more and its total MI was very low (13.6). Six criteria had a delta of 60 or more.

Since some class evaluations were completed before information was read and 94% of the survey criteria evaluations were completed after the information was read it is not surprising that the survey data showed less overall change. The survey data had a slight tendency towards greater changes in mean criteria importance between stages 4 and 5.

Both class and survey respondents had the lowest number of changes in mean criteria importance between stages 5 and 6. More respondents reported being in stage 5 or stage 6 and more evaluations were filled out when respondents were in stage 5 or 6. Seventy-three percent of the survey respondents who reported a single information seeking stage were in stages 5 or 6 and 64% of the class criteria evaluation forms that reported a single information seeking stage reported stage 5 or 6. It is possible that there is less variability between these two stages simply due to greater numbers. However, Kulthau’s model predicts that stages 5 and 6 should show less variation in the information problem. It is logical that if the information problem were less variable then the mean importance of the relevance criteria would vary less too. However, the data from this study showed weak trends in the tendency of relevance criteria to vary in
importance with information seeking stage. Further study is needed to fully examine and answer this question.

**Class Data Motion Indexes between Searching, Obtaining, and Reading**

Actual changes in mean criteria importance were very slight. The largest absolute change between any two activities was 23.4. Only 13 criteria showed an actual change of 10 or more between any two activities. No criterion showed an actual change in mean criteria importance of 10 or more between both searching and obtaining and obtaining and reading. The largest motion index for 17 criteria was between searching and obtaining and the largest motion index for the other 23 criteria was between obtaining and reading. Of the criteria that had an actual change in mean importance of 10 or more five criteria changed between seeking and obtaining and eight criteria changed between obtaining and reading.

Twenty-two criteria had a delta of more than 60, six criteria had a delta of 40 to 60, and 12 criteria had a delta of less than 40. Thus the changes in importance of criteria as respondents sought, obtained and read information were likely to be mostly positive or mostly negative.

**Comparison of Class and Survey Means**

The difference between the class mean importance for each criterion and the survey mean importance for each criterion was slight. The range of difference in mean criteria importance between the class and the survey was .6 to 13.6. Only six criteria had a difference of more than eight between the two means. The same ten criteria had a mean criteria importance of 75 or greater for both respondent groups. Since respondents from the classes used the criteria evaluation form several times and respondents from the
The stability of the means does seem to indicate that the same criteria contributed to high relevance decisions in both respondent groups. This similarity in mean criteria importance and the very low changes in mean criteria importance across information seeking stages and across the information problem activities of searching, obtaining, and reading indicate that some aspects of high relevance remained stable for this group of respondents across information seeking stages.

Analysis of Class Qualitative and Information Seeking Data

Criteria Reported on the Criteria Evaluation Form

Respondents in the classes were asked to "describe in a few sentences why you think the information sources you have listed will be the best for your project". They were to complete this part of the criteria evaluation after listing the two to four most valuable information sources for their paper or project and before marking the importance of each criterion. Reasons or criteria were selected from this description and broadly categorized.

Most responses were direct and often mentioned criteria that were being evaluated on the criteria evaluation form using the 100-mm lines. Some respondents mentioned the same criteria on all evaluation forms. No respondent mentioned more than four criteria in this portion of the form. Each mention of the criterion was counted, so if a respondent mentioned topic on all four criteria evaluations, topic was counted four times. The criteria mentioned on this part of the form and the number of times each criterion was mentioned are listed in table 17.
<table>
<thead>
<tr>
<th>Criterion</th>
<th>Number of times mentioned</th>
<th>Criterion</th>
<th>Number of times mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>60</td>
<td>Appeal (to audience)</td>
<td>1</td>
</tr>
<tr>
<td>Current</td>
<td>19</td>
<td>Appropriate</td>
<td>1</td>
</tr>
<tr>
<td>Reliable author</td>
<td>13</td>
<td>Best</td>
<td>1</td>
</tr>
<tr>
<td>Relevant</td>
<td>9</td>
<td>Case study</td>
<td>1</td>
</tr>
<tr>
<td>Pertinent</td>
<td>5</td>
<td>Clear</td>
<td>1</td>
</tr>
<tr>
<td>Easy to obtain</td>
<td>4</td>
<td>Complete</td>
<td>1</td>
</tr>
<tr>
<td>Examples</td>
<td>4</td>
<td>Comprehensive</td>
<td>1</td>
</tr>
<tr>
<td>Supports my argument</td>
<td>4</td>
<td>Detailed</td>
<td>1</td>
</tr>
<tr>
<td>Focused</td>
<td>3</td>
<td>Direct</td>
<td>1</td>
</tr>
<tr>
<td>Helpful</td>
<td>3</td>
<td>Easy to read</td>
<td>1</td>
</tr>
<tr>
<td>Personal interest</td>
<td>3</td>
<td>Format</td>
<td>1</td>
</tr>
<tr>
<td>Practical approach</td>
<td>3</td>
<td>Foundational</td>
<td>1</td>
</tr>
<tr>
<td>Specific</td>
<td>3</td>
<td>General</td>
<td>1</td>
</tr>
<tr>
<td>Study</td>
<td>3</td>
<td>New to me</td>
<td>1</td>
</tr>
<tr>
<td>Useful</td>
<td>3</td>
<td>Practical</td>
<td>1</td>
</tr>
<tr>
<td>Authoritative</td>
<td>2</td>
<td>Reliable journals</td>
<td>1</td>
</tr>
<tr>
<td>Consistent</td>
<td>2</td>
<td>Suggests solutions</td>
<td>1</td>
</tr>
<tr>
<td>Defines terms</td>
<td>2</td>
<td>Technical</td>
<td>1</td>
</tr>
<tr>
<td>Easy to read</td>
<td>2</td>
<td>Thorough</td>
<td>1</td>
</tr>
<tr>
<td>Explanation</td>
<td>2</td>
<td>Unbiased</td>
<td>1</td>
</tr>
<tr>
<td>Graphic, diagrams</td>
<td>2</td>
<td>Understandable</td>
<td>1</td>
</tr>
<tr>
<td>Informative</td>
<td>2</td>
<td>Unusual perspective</td>
<td>1</td>
</tr>
<tr>
<td>Overview</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17. Criteria Mentioned by Class Respondents on Criteria Evaluation Forms

Several Library and Information Science respondents mentioned pertinence or relevance and then proceeded to qualify these criteria as “pertinent to my topic” or “relevant to my topic”. The topic of the information source was also mentioned in regard to several other criteria for example: “overview of my topic” and “specific topic focus”.

Several class respondents who were judging information sources before they had obtained or read them qualified their judgments based on title or citation information.
One respondent stated “titles sound appropriate”. Another stated, “at this point very superficial criteria – haven’t read sources, only going on the title and journal”. This supports the idea that respondents were uncomfortable judging information sources as most valuable before they had read or at least skimmed those information sources.

Sometimes respondents were looking for information sources that matched the class guidelines. Others were looking to information sources for help in structuring and organizing their paper. One respondent was concerned with the “appeal of information to the audience” he or she was writing for. Several respondents were looking for new or unusual information, even though this criterion (V6 New to me) only had a mean of 42 for the class respondents. Respondents mentioned “new to me”, “unusual perspectives”, and “tells me something I don’t know”.

The reliability of the author was mentioned 13 times and the reliability of the journal was mentioned once. Only mentions of topic (60 times) and currency (19 times) were more frequent. Criteria listed on the criteria evaluation form that seem to share some meaning with this concept include V22 (Prominent, mean 46.3), V24 (I know the source, mean 43.3) and V25 (Reputable, mean 70.8), and V11 (Credible, mean 83.3). Respondents stated: “frequently referenced author”, “name came up repeatedly when searching for citations”, “reputable source” and “reliable source”.

The approach of the information was also mentioned frequently. “Technical approach”, “practical approach”, “case study”, “research study and not a review of literature”, “practical”, and “good selection of different approaches” are some examples of this category. One respondent mentioned “restructures a report, found previously in another source”. It appears that some respondents were looking for information sources
that presented information in a particular manner, whereas other respondents wanted to find something that they did not already know.

Several respondents mentioned information that supported a particular viewpoint, even though V34 (Controversial, mean 29.6) and V38 (Validates my viewpoint, mean 49.8) were not rated as important. This concept seemed to be highly situational and dependent on the respondent and their information problem.

The mention of criteria cannot be assumed to be an indicator of criterion importance. However, it is interesting to note that no new criteria were mentioned and that many respondents focused on the topic of the information or some aspect of that topic.

**Evidence of Multiple Search Sessions**

On the search diary forms respondents were asked to report the number of times they used different information resources including people, links, and IR systems that they used or consulted to lead them to information sources. Information resources that were specifically mentioned on the search diary forms included: online public access catalogs (OPACs), print indexes, CD-ROM indexes, World Wide Web, librarian, link or citation, other Internet resources, other people, and other. Respondents were asked to indicate who or what the other resources were.

During the first semesters that the study was conducted neither school’s OPAC had a gateway on the World Wide Web. Also neither school’s web site offered a gateway to CD-ROM indexes. It was possible to telnet to both schools’ OPAC. By the end of the study, UNT’s OPAC and CD-ROM indexes could be searched on the World Wide Web from the UNT web site and CD-ROM indexes were available from the TWU
web site. The TWU OPAC still did not have a gateway on the World Wide Web. Respondents seemed to understand the difference between a CD-ROM index that was available on the World Wide Web and searching the World Wide Web itself. However, this aspect of information resource use was not specifically addressed on the search diary forms. The rapid development of World Wide Web gateways to OPACs and CD-ROM indexes could not have been easily predicted at the beginning of this study. This development suggests a need for research into how end-users perceive and use these World Wide Web gateways.

Respondents conducted frequent multiple search sessions on their topics. Thirty respondents conducted multiple search sessions on the same system within one search session or used multiple systems. The other five respondents did not report the number of times that they searched using a particular system and it was not possible to tell whether these respondents were conducting multiple search sessions or not. Ten respondents did not report the number of times that they used a system on all search diary forms and four of the other respondents failed to report the number of times they used a system on some of their search diary forms.

Table 18 lists the number of multiple search sessions class respondents reported on one search diary by type of information resource. Respondents were instructed to complete one search diary for each library trip or search session. Respondents who reported the number of times they used a system reported a total of 73 single searches (57%) on a given system and 55 multiple search sessions (43%). Respondents who used the World Wide Web reported more multiple search sessions (17) than single searches
<table>
<thead>
<tr>
<th>Information resource</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
<th>Four</th>
<th>Five</th>
<th>Six</th>
<th>Seven</th>
<th>Eight</th>
<th>Total multiple searches</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPAC</td>
<td>25</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Print Indexes</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CD-ROM Indexes</td>
<td>18</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>World Wide Web</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Librarian</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Link or Bibliography</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Other Internet</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other People</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>73</td>
<td>24</td>
<td>13</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td>Grand Total</td>
<td>128</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18 Number of Search Sessions Conducted on Each Information Resource, Class Respondents
Respondents who used links or citations also reported more multiple search sessions than single searches. In all other search system categories respondents did a greater or equal number of single searches as compared to multiple search sessions. The World Wide Web facilitates and encourages multiple search sessions, but it is also possible that some respondents did not easily find the information they needed on the World Wide Web. This may be an alternative explanation of the reported frequency of World Wide Web multiple search sessions.

Some of the comments made by respondents on the search diary forms seem to support the notion that respondents were not easily finding the information they needed. Although respondents were asked to report World Wide Web use separately from other Internet resources, respondents reported World Wide Web use on the form but referred to it as “the Internet” in their written comments. Respondents’ comments included: “no luck on Internet”, “I wish I knew where to find stuff on the Internet”, “One problem I am having with the Internet is that some sources do not identify the author and it is hard to pin down dates: i.e. how current some of the information is”, and “frustration with no exact hits off the Internet”.

Information seeking and information resource use was more frequently reported in the last three search stages. These stages accounted for 60% of the reported use of information resources. Search stages 5 and 6 were the most active information seeking stages with 20% of the searches reported in stage 5 and 18% reported in stage 6. Some respondents continued to search for information well into their writing process. This may be explained in part by the number of respondents who were in short 5 week summer
sessions and who may have felt they had to combine searching, obtaining, reading, and writing.

The most popular information resources were OPACs (23%), CD-ROM indexes (19%) and World Wide Web (24%). CD-ROM indexes received a high percentage of their use in the early stages compared to the percent of the searches that were conducted in the early stages. Twelve percent of the CD-ROM index use occurred in stage 2 where only 7% of the searching occurred and 20% of the CD-ROM use occurred in stage 3 where only 14% of the searching occurred. CD-ROM indexes were the most popular resource in stage 2 and tied with OPACs for the most popular resource in stage 3. The distribution of the use of the other major information resources (OPAC, print indexes, World Wide Web, links, and people) was similar to the distribution of the number of searches across information seeking stages. It appears that some respondents were using CD-ROM indexes to begin their information seeking process.

Librarians and other people accounted for 15% of the reported use of information resources. The respondents used people infrequently, but evenly throughout the information seeking process. Instructors and library personnel were used either to begin searching or to obtain help after some initial searching had been attempted.

Changes in Topics

Eighteen respondents did not change topic statements during the research. It is possible that some respondents either had clearly formed a topic focus before they began their paper. It is also possible that some respondents did not feel ready to complete the evaluation forms until they had a clear topic focus and were in later information seeking
stages. Eighteen respondents changed topics or topic focus and one respondent did not clearly indicate their topic on the evaluation or search diary forms.

Fifteen (57%) of the 35 respondents were in long semester classes, the remaining 20 respondents were in five-week summer sessions. Eleven (61%) of the 18 respondents who changed topic statements during the research were from the summer session classes. The shorter semester did not seem to influence the likelihood that a respondent would keep the same topic throughout the research since respondents in the summer sessions were actually slightly more likely to change topics. The search diary forms from these respondents indicate that some of these respondents changed topics early in their information seeking because they could not find adequate information on their initial topic.

Twelve respondents reported one change in their topic focus and six respondents reported two changes in their topic focus. The changes in topic were primarily changes to a more focused aspect of the original topic. For example one respondent changed from "health belief model" to "HIV/AIDS prevention programs for adolescents using the health belief model". One respondent changed topics entirely between the first two evaluations. This respondent reported being in stages 1 and 2 on the first evaluation and stages 1, 2, and 3 on the second evaluation. Table 19 reports the information seeking stages where topic changes occurred.

Most topic changes occurred between stage 3 and a later stage. Two respondents completely changed their topic between their first two information seeking sessions but did not report their information seeking stages. Respondents who had a second topic
Information seeking stages | Number of respondents
---|---
2 to 3 | 2
3 to 4 | 2
3 to 5 | 1
3,4 to 3,4 | 1
3 to 4,5 | 1
2,3 to 5,6 | 1
4 to 5 | 1
5 to 6 | 1
stages missing* | 2

*These two respondents had complete topic changes early in their information seeking but did not report their information seeking stage.

Table 19. **Information Seeking Stages of Class Respondents Reporting a Single Change in Topic Focus**

focus change usually had this change between stage 4 and a later stage. Table 20 reports information seeking stages of respondents who had two topic focus changes.

<table>
<thead>
<tr>
<th>First change</th>
<th>Second change</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 to 4</td>
<td>4 to 6</td>
</tr>
<tr>
<td>3 to 4</td>
<td>4 to 5</td>
</tr>
<tr>
<td>3 to 4</td>
<td>4 to 4</td>
</tr>
<tr>
<td>3,4 to 4,5</td>
<td>4,5 to 5,6</td>
</tr>
<tr>
<td>1,2 to 1,2,3**</td>
<td>1,2,3 to 3,4,5</td>
</tr>
</tbody>
</table>

**Complete change in topic**

Table 20. **Information Seeking Stage of Class Respondents Reporting Two Changes in Topic Focus**

These topic changes occurred most frequently around stage 3. Kuhlthau's model indicates that changes in topic focus should occur between stage 3 and a later stage.

Some respondents who reported a second change in topic focus appeared to be honing their topic before finishing the paper.
Changes in Information Sources

Twenty-seven respondents choose at least one of the information sources that they had identified as the most valuable in their first evaluation (when they were searching for information or searching for and obtaining this information) as the most valuable on all their relevance criteria evaluations. Only five respondents evaluated entirely different information sources on all their relevance criteria evaluation forms. Three respondents did not include citations for the information sources they were evaluating.

Only two respondents evaluated the same information source every time. One of these respondents conducted one search, found what he or she wanted and did no further information seeking. The changes in information sources were more frequent and pronounced when respondents were searching for and obtaining information. Since these activities occurred in respondents' earliest information seeking stages, changes in information sources often occurred in early information seeking stages. It appears that at some point in the later information seeking stages respondents decided on at least a portion of their most valuable or relevant information sources and these sources remained valuable throughout subsequent information seeking. These sources probably influenced respondents' relevance judgments and criteria when they evaluated additional information sources. The characteristics of the information sources that are judged as valuable in early stages and that remain valuable throughout information seeking may set a benchmark that additional information sources are compared to. This helps explain why the mean importance of the relevance criteria in this study showed very little change as respondents' information seeking stages changed.

Most respondents reported their first information seeking stage as stage 2 or 3.
Seven respondents reported stage 2 as their first information seeking stage, seven respondents reported stage 3, three respondents reported stage 4, six respondents reported stage 5, and no respondent reported stage 6. Information seeking stage 1 was not reported as a single stage but was reported in combination with other stages for seven respondents. Three of these respondents reported stage 1 in some combination with stage 5. Six respondents reported stage 5 on their first criteria evaluation form. The respondents who reported later information seeking stages on their first evaluation form may have been pressured by time to move into later information seeking stages as quickly as they could. Also they may have been unwilling or unable to evaluate the most valuable information sources for their projects until they reached later information seeking stages. They also may have been forcing themselves to move through the early, more anxiety-ridden stages to avoid the discomfort associated with these stages. Some of Kulthau’s respondents also avoided reporting early stages of information seeking and she attributed this behavior to the discomfort associated with early stages. In this study some combination of the three previously mentioned factors is probably responsible for respondents who only reported later information seeking stages.

Twenty-five respondents reported information seeking stage 4 or less on their first evaluation and 19 respondents reported stage 3 or less on their first evaluation.

Respondents went through the information seeking stages sequentially with only one respondent moving back to a combination of stages 4, 5, and 6 from a combination of stages 5 and 6. One respondent did not report his or her information seeking stages.
Trends in Search Diary Forms

The search diary forms were examined for trends in respondent thoughts, actions, and feelings. The criteria evaluation forms provide some indication of each respondent’s thoughts about a particular set of valuable or highly relevant information. The reported use of different IR systems and people as links or guide to information provide some evidence of respondent actions. On the search diary form respondents were asked to “Describe your information seeking experience today. Include anything you think might be appropriate: feelings, actions, strategies, problems etc. Discuss any changes in your topic development or plan for the paper and what influence these changes” (Appendix B).

Most respondents discussed their information seeking process in a few sentences or phrases in their search diaries and did not report many details. A content analysis of these diaries was conducted and three different trends were identified. Each of these issues will be identified and then discussed in more detail. Respondents discussed affective issues freely: how frustrated, anxious, or relieved they were. Other search diary comments discussed the information seeking resources used by respondents and their changes in topic or topic focus. Earlier discussion indicates that across the respondents all traditional information resources were used, with many respondents reporting frequent use of OPACS and CD-ROM indexes. The use or non-use of Internet resources was often discussed and will be examined in greater detail.

Affective Issues in Information Seeking

Many respondents expressed anxiety in the early stages of their project. This was particularly true of the students who were enrolled in the five-week summer sessions.
Some respondents were frustrated with the very narrow time constraints: "I was frustrated somewhat today in that the time frame ... is too short". The respondents in the full semester classes also tended to report anxiety and frustration in early stages of the information seeking process and were frustrated because information sources were not always readily available. "The library did not carry most of the journals I needed". Some of the Library and Information Science students did not express much anxiety or frustration, even when they were starting the project only two or three weeks before it was due. These students may have been thinking about their topics, may have been embarrassed to mention any frustration or anxiety about searching for information, or they may have been very confident in their information seeking skills. Fewer Library and Information Science students discussed their feelings. These students tended to discuss how they developed a topic focus or what information seeking resources they used. The full semester students were often satisfied at the end of their information seeking.

Proportionally more Library and Information Science students were in full semester sessions (12 of the 22 Library and Information Science students were in enrolled in a full semester, whereas 10 of the 13 Health Studies students were enrolled in the five-week summer sessions). Some students in the full semester completed their project in less than a month. Two students returned all search diary and criteria evaluation forms and appeared to have completed their information seeking in the first six weeks of the full semester. Other students reported information seeking throughout the semester or choose not to start their project until about three weeks before it was due. This group of students, who had imposed their time constraints on themselves, did not report being anxious and frustrated. This is in contrast to the students who had time
limits imposed on them by the short summer sessions and reported anxiety and frustration, particularly with the time constraints.

Some respondents were frustrated because they thought they did not have adequate time to obtain information. They felt, that given more time, they could have found better information sources and produced a better paper. "Time is too short to obtain ILL materials, decided against the use of one article because it could not be obtained in time" "If I had more time I would go to more libraries". Sometimes respondents felt that the short time frame influenced them to simply take whatever information was available. Some respondents also felt that this information might not have been what they would have chosen, without the narrow time constraints. "[I] might have made a different selection if I'd had time to read several [articles] and compare contents".

Some respondents were very anxious about having adequate information seeking skills. This was particularly evident in the Health Studies respondents. Some of these respondents were very comfortable using the library and print sources and not very comfortable with computers and the Internet. "I'm more comfortable here [in the library], it is familiar but not always convenient". However, one respondent was very comfortable with the Internet but not with library and print sources. This respondent had considerable initial difficulty because she could not find anything on the Internet. "no luck on Internet using key words. I wish I were more adept at library resources". However, another respondent reported, "My best and strongest sources came from the Internet".
Other respondents were frustrated with the mechanics of using libraries, particularly within such a short time frame. One did not like the way one library was laid out and another couldn't stay in one library very long because it made her claustrophobic (interestingly this was a Library and Information Science student).

Respondents felt relieved part way through the process – after they had passed the earliest search stages, as predicted by Kuhlthau's model. However, in the summer sessions some respondents were frustrated during most of their information seeking. Time pressure and problems obtaining information sources usually caused this frustration. These respondents often went through a second wave of anxiety very close to when the paper was due. This anxiety was always attributed to the short time frame and the non-availability of critical information sources. As one respondent reported, "I needed some further resources in full text. I felt desperate, helpless and frustrated when I couldn't get what I needed. Paper due tomorrow. What will happen?"

Except for the last burst of anxiety in some respondents in the summer session these respondents' comments indicated that they experienced a pattern of anxiety followed by satisfaction. Respondents seemed to skip over the very early information seeking stages or combine these stages with later stages. This may have been due in part to the short summer sessions; however, the pattern of movement through the information seeking stages was similar in the both summer and full semester respondents. These respondents, like Kuhlthau's respondents, may have felt anxious and uncomfortable in early search stages and pushed themselves to later search stages. How this affects information seeking and relevance criteria and judgments is an area for further study. It
appears that the summer respondents felt pushed through the information seeking stages and experienced anxiety due to the imposed time constraints.

**Information Resources**

The following section reports on respondents' use of the Internet and people as information resources during their information seeking process.

**Use of the Internet**

Respondents were either very successful with the Internet or had very little success. Most complaints were about the difficulty of searching; "I feel like my searching is totally unstructured" and the quality of the information on the Internet; "Internet has a lot of junk", "Internet has a lot of articles that are not scientific enough for research – more appropriate for general information". One respondent dealt with the quality issue by deciding to do a search on the Internet for "sensational" information about her topic.

Internet use and the value of the information that is available on the Internet appears to be very situational, much like relevance itself. Thirteen percent of the respondents from the survey rated an Internet site or page as their most valuable source of information. The Internet is clearly affecting how students search and may be affecting how students evaluate information and their relevance criteria. The number of Internet resources reported in this study was too small to develop a set of "Internet relevance criteria". Therefore "Internet criteria" were not compared to "print criteria" in this study.
Use of People

Respondents often started their searches and relieved some of their anxiety by consulting people. Some respondents consulted librarians or library teaching assistants for help in narrowing their topic focus and for instruction in searching for information or using computer equipment. They also requested searches from library personnel. Some respondents used the instructor to either help them find a topic focus or to provide an information source that they could use as a starting point. Respondents were satisfied and comfortable with their interactions with librarians and instructors.

Changes in Topic and Focus

Respondents seemed to follow the pattern of topic change predicted by Kulhthau’s model, even when they did not report being in early stages of information seeking. When topic focus was reported respondents started with a topic that was usually “too broad” and had to find ways to narrow, limit, or focus their topic. “I started out with a vague sense of my topic, after working with the TA I narrowed my focus for my topic.”

Sometimes the availability and amount of information affected students’ topics. Some respondents found too little information and had to change their topic focus. “Changed topic from infectious disease/Colonias health belief model to HIV adolescent population [health belief model]. Easier to find information on.” “I might need to change some of my search words to more accurately define the information I am looking for”. Other students picked broad topics and soon found they had too much diverse
information. "The first topic had several hundred citations ... I think the topic in general is too hard to focus".

Some respondents in the long semesters quit searching when the information became redundant to them. "I think I have located enough references that say the same exact thing to end my search now and finish my article".

Summary

This chapter described characteristics of the data obtained from a mail survey of 210 UNT graduate students and from 35 graduate students in seven classes at UNT and TWU. Eleven criteria that survey respondents rated as important were used to propose a model of high relevance. This model was tested for reliability and validity using confirmatory factor analysis and second order factor analysis. Two criteria V1 (About my topic) and V8 (Current) were weakly correlated with their respective constructs and it is suggested that additional items be added to the criteria evaluation form in order to measure the constructs of Topicality (aboutness) and Currency separately. Availability is also a construct that appears to contribute to high relevance judgements and items that measure this construct should also be added to the criteria evaluation form.

The data from the classes were not analyzed using inferential statistics, but were described using a motion index first developed by Janes (1991). This motion index was used to examine the changes in mean importance of each criterion between information seeking stages and between the activities of searching for, obtaining, and reading information. Survey data was also described using the motion index to examine the changes in mean importance of each criterion between information seeking stages.
Students did not go through the information seeking stages or the activities of searching for, obtaining, and reading information in a uniform manner. The importance of the criteria employed by respondents in judging highly valuable or relevant information sources did not appear to change much as respondents moved through an information seeking process.

The criteria that were judged as most important by respondents showed less variability and less change in mean importance. Respondents employed information resources at most stages of information seeking. CD-ROM index use was more frequent in earlier information seeking stages. Respondents searched for and obtained information in the earlier information seeking stages and read information in the later information seeking stages. Some respondents combined information seeking stages and activities. Both respondent groups preferred completing criteria evaluation forms after they had read their information source or sources. This seems to indicate that this group of users preferred to wait until they had read information before they judged the information as most valuable in their resolution of their information problem. Respondents went through Kuhlthau’s stages sequentially and experienced anxiety and changes in topic focus as predicted by the Kuhlthau model. Some respondents who were enrolled in five-week summer classes experienced a final stage of anxiety just before they turned in their paper. This anxiety was associated with their difficulty in locating adequate information sources in the short time frame.
CHAPTER FIVE

CONCLUSIONS

Introduction

This chapter will discuss the results of this research for each specific research question, discuss other findings from this study and suggest areas and questions for further research. The primary research question was: What is the relationship between the criteria employed by respondents in selecting sources of information that they find to be most valuable and the respondent's stage of information seeking with respect to a specific information problem? The results indicate that the criteria used in the selection of valuable or highly relevant information do not show much change in importance as respondents moved through the stages of information seeking. Criteria did not change in importance as respondents obtained and read the information. The results also identify the criteria that graduate students rated as most important. A model of the relationships between these criteria, the constructs of Information quality, Information credibility, Information completeness, and high relevance was proposed and confirmed.

Question One: What are the Relationships between a Group of 40 End-user Relevance Criteria in Regard to their Degree of Importance?

A three-construct model was proposed. This model incorporated the criteria that were rated as important by graduate students. The three constructs were Information quality, Information credibility, and Information completeness. This model was
confirmed using factor analysis. Second order factor analysis indicated that 48% of high relevance as judged by these graduate students could be explained using these three constructs (Figure 11). Ten criteria used in this model describe an aspect of the information content: About my topic, well-written, Credible, Accurate, Understandable, Consistent, Focused, Comprehensive, Suitably general or specific, Detailed. The criterion Current is an aspect of both the information content and the information source.

Users employed other relevance criteria that described the information source. However, this group of respondents did not rate these descriptive criteria as highly as information content criteria. Respondents used some of the attributes of the information source as clues for judging information content. For example, when respondents knew the journal or the author of an information source they used this as a clue regarding the credibility of the information. Some of these attributes were also listed as criteria on the criteria evaluation form.

These respondents were graduate students who were not naive users but had not yet become experts in their respective disciplines. Results indicate that quality and credibility criteria are very important for this group of users. They wanted information that was accurate, credible, well-written, understandable, consistent, and focused. Users who are experts in their discipline may not rate these criteria as highly because expert users may have sufficient subject knowledge to be comfortable with their own judgments of accuracy, credibility and other information quality criteria. Graduate students are still learning about their discipline and are likely to want to obtain and read information that has been accepted as accurate, reliable, and credible within their field.
The criterion of Current (V8) was only weakly correlated with the construct Information quality. Currency may be viewed as an aspect of information quality in some situations. However, in other situations this criterion may be viewed as an attribute of the information source. This criterion is a separate, but important construct: if more items that measured this construct were added to the criteria evaluation form it would probably be possible to measure a larger percentage of high relevance for these graduate students.

The criterion of About my topic (V1) was weakly correlated with the construct of Information credibility and also is a separate and very important construct. Topicality (aboutness) is usually highly situational and dependent on the user, his or her situation, and the information problem. This may make it difficult to identify criteria to measure this construct that are valid and reliable across users and information situations and problems. Topicality is thus a very difficult construct to separate into its most important criteria.

Topicality may also have components that are dependent on the discipline of the user and it is likely that topicality interacts with many of the other criteria on the form. In this study, About my topic was only weakly correlated with the criteria Credible and Accurate and probably represents a separate construct. Development of an instrument that can reliably and validly explain and measure the components of topicality to end-users cannot be developed until research has identified these components. This study did not attempt to use more than one criterion (About my topic) to measure topicality.
Study respondents also wanted information that was Comprehensive, Suitably general or specific, and Detailed. These criteria were components of the construct Information completeness. Most respondents needed to locate and read sufficient information about their topic in a fairly short time. The search diary forms returned by class respondents indicated that they could not always obtain the information sources they wanted within the time constraints of the class. This may explain why graduate students wanted comprehensive, focused information in the information sources they could easily obtain.

Easy to obtain was an important criterion for both survey and class respondents but it did not correlate well with any other criteria. If other criteria that measured the construct of Information availability reliably and validly could be added to the criteria evaluation form this construct could be used to explain a percentage of high relevance. The constructs of Information quality, Information credibility, Information completeness, Information currency, Information topicality, and Information availability are suggested as the constructs that would measure a proportion of high relevance for graduate students if adequate criteria to measure each construct reliably and validly were added to the criteria evaluation form.

In this study almost all criterion means had a range from 10 or less to 95 or more on a scale of 0 to 100. In other words each criterion was judged as very important by at least one respondent and not at all important by at least one respondent. This phenomenon supports the situational view of relevance, which posits that the importance
of a criterion to a user will be determined by the user's situation. The user's information
problem and stage in the resolution of that problem are important aspects of the situation.

The criteria that had means below 75 also had higher standard deviations. This
indicates that the importance of these criteria varied more from user to user and were
probably the criteria whose importance was most influenced by the user's situation. An
exception to this was the criterion that was rated as least important; I know the author
personally (V23). This criterion had a lower standard deviation and this group of
respondents appeared to be rating this criterion as not important with some consistency.
This is not surprising since graduate students are less likely to know researchers and
authors personally.

Many criteria that were rated as most important by these respondents were also
criteria that this respondent group probably did not feel comfortable judging until they
had read the information. Not surprisingly, respondents strongly preferred to judge
information as the most valuable after they had read it. This pattern of preferring to wait
to judge value or high relevance until after reading may not be as strong in other user
groups. Overall, graduate students should be motivated to learn. They also may not yet
be experts in their disciplines and may not know how their discipline regards certain
authors and journals. Users who are experts in their fields may be more willing to make
judgments of high relevance and the value of information before they read it.

Some criteria may be important to one discipline and not important to another.

This study was conducted across disciplines; graduate students from all academic units at
UNT participated in the survey. Although students who participated in the research with
the classes were Library and Information Science and Heath Studies students, their
undergraduate majors represent a variety of disciplines. Many of these students were also
just beginning their graduate studies and were probably not yet completely familiar with
the information seeking criteria and methods commonly employed by their respective
disciplines. Further study with respondents who are well versed in their particular
discipline will be necessary to identify criteria that have importance within the different
disciplines.

Question Two: At what Stages of Information Seeking do Respondents Employ IR
Systems or Information Resources?

Respondents used most IR systems or information resources evenly across
search stages 2 through 6. Respondents were somewhat more likely to use CD-ROM
indexes in the earlier stages of information seeking. This seems to indicate that some
respondents may have used CD-ROM searches to initiate their information seeking
process. CD-ROM indexes were also used somewhat less frequently in later information
seeking stages. The frequency of use of other information resources did not decrease in
the later search stages.

The Internet affected the information seeking of these respondents, but it was
not clear exactly how. Some respondents found the Internet a very good source of
information, others stated that the Internet has “junk” and that it was difficult for them to
find credible information on the Internet. A few respondents decided to learn to use the
Internet as part of their information seeking process. One respondent was very comfortable with the Internet and had some difficulty with other information resources; other respondents stated they were more comfortable with traditional information resources. The respondents seemed to be aware of the difference between the Internet as an information retrieval resource (i.e., Internet search engines) and the Internet as a gateway to other information retrieval resources (OPACs and CD-ROM indexes). As more institutions add gateways to OPACs and CD-ROM indexes on their Web sites the distinction between Internet search engines versus OPACs and CD-ROM indexes that are accessed via Internet gateways may not be clear to many end-users. This use of Web sites as gateways to other IR systems will have an impact on end-user information seeking strategies and use of information resources. This is an important and interesting area for further research.

Question Three: At what Stages of Information Seeking do Respondents Search for, Obtain, and Read Information?

Respondents reported searching for information primarily during the early stages of information seeking, which is not surprising. However, evidence in the search diaries and from respondents who reported mixed information seeking stages indicates that some respondents searched for information throughout all stages of their information seeking process. Many respondents combined searching for, obtaining, reading, and writing. This may be due in part to the respondent’s personal information seeking style. However, the immediate availability of full-text documents from some CD-ROM indexes
and from the Internet encourages this style. Many respondents did successive searching or used multiple information resources.

This successive searching behavior and respondents' tendency to combine the activities of searching, obtaining, reading and writing have implications for IR system design and relevance research. Results from this research further support the criticism of the use of relevance judgments based on a citation list from a single search session to calculate precision and recall as measures of search effectiveness. The overall effectiveness of any search is influenced by the entire information seeking process for the stated information problem. When users engage in successive searching relevance judgments and search effectiveness cannot be reliably or validly measured without examining all searches and the use of the information identified by these searches. However, this type of longitudinal research is difficult to conduct and requires considerable cooperation from respondents. Further research that examines the various aspects of relevance judgments, and information use is necessary.

Question Four: Does the Importance of End-user Relevance Change as the Respondent's Information Seeking Stages Change?

The importance of the relevance criteria did not change very much as respondents moved through their information seeking stages. The largest change in mean importance ratings for any criterion between any two stages was 20.5 on a scale from 0 to 100. Respondents were asked to focus on the most valuable sources of information as they completed the criteria evaluation form. They rated criteria importance based on
information that they had already judged as valuable or highly relevant. The criteria that had the highest importance ratings did not tend to change very much across information seeking stages. Criteria that had lower mean importance ratings tended to change more across information seeking stages. However, the criterion rated as least important did not change much across information seeking stages.

Criteria importance means for class respondents and survey respondents were very similar, with a maximum difference of 13 or less on a scale of 0 to 100. The greatest number of criteria had their largest change in importance between stages 2 to 3 for the class respondents and stages 3 to 4 for the survey respondents. For both groups, the least number of criteria had their largest change in importance between stages 5 and 6.

Recent research by Spink, Greisdorf & Bateman (in press) has indicated that the number of items judged partially relevant by end-users is positively correlated with changes in the user's information problem, search intermediaries' perception that the user's topic has changed, and the user's personal knowledge due to the search interaction. The study notes that most relevance research has focused on measuring relevance in a binary fashion; relevant or not relevant. In studies that have measured degrees of relevance partially relevant judgments have been collapsed into high relevance. The research that has identified end-user relevance criteria has also focused on either highly relevant or not relevant information and has not examined the criteria used in partial relevance judgments. The correlation of the frequency of partial relevance judgements with changes in topic and problem focus suggests that the importance of criteria for
partially relevant information might change between information seeking stages. Spink, Greisdorf & Bateman (in press) identify a few criteria that were employed by users, who made partial relevance judgments, but these were obtained from a limited number of users and these users were not interviewed or probed to ensure that all criteria were mentioned. The study of partial relevance judgments and the identification of partial relevance criteria is an important area for further research.

This part of the study was exploratory and descriptive and only suggests some trends in changes in the importance of high relevance criteria throughout an information seeking process. Further research with additional respondents will be necessary to confirm these suggested trends.

Question Five: Does the Importance of End-user Relevance Criteria Change as Respondents Search for, Obtain, and Read Information?

Overall, the importance of criteria did not change very much as respondents judged information before they obtained it, after they obtained it and after they read it. The largest change in the importance of any criterion between any two activities was 23.36 on a scale from 0 to 100. More criteria had a greater change in importance between the activities of searching for and obtaining information. Conversely, fewer criteria had a greater change in importance between the activities of obtaining and reading information. The criteria did not change much in importance and these are only slight trends.
Respondents indicated a strong preference for judging information as most valuable or highly relevant after they read the information. Ninety-four percent of the survey respondents had read the information source that they focused on when completing the criteria evaluation form. Class respondents were instructed to fill out two forms before they obtained information, one after they obtained and before they read the information and one after they read the information. However, only 18% of the class forms reported seeking as a single activity, 19% of the class forms reported obtaining information as a single activity and 27% of the class forms reported reading as a single activity. Overall 46% of the forms reported seeking or obtaining or a combination of these activities and 54% of the forms reported reading or writing as one of the activities. The most important criteria for these graduate students included quality and reliability criteria that may have been very difficult for them to judge until they had read the information. Of the criteria that were rated 75 or more on a scale of 0 to 100, only About my topic, and Current could be easily determined by a non-expert user from the IR system information representation. The other important criteria: Well-written, Credible, Accurate, Understandable, Consistent, Focused, Comprehensive, Suitably general or Specific, and Easy to obtain may be determined from some clues in the information representation such as the author or the journal. To be able to use these clues users would need to have a certain degree of knowledge about the discipline, which these graduate students may or may not have had. It appears that these graduate students were most comfortable judging information as the most valuable after they had read it.
More changes in the information sources cited on the forms occurred between searching for and obtaining information. However, the changes in the importance of criteria used in judging these sources were slight. This suggests that the primary criterion that may have been causing these sources to change was topicality. This study did not attempt to correlate changes in topicality with changes in information sources or criteria.

Topic changes appeared to be influenced in part by the amount or quality of information that the respondents could obtain within their time limits. This was especially true for students in the five-week summer sessions. Some respondent’s topics did not change at any point in the study, indicating that these respondents may have already had some knowledge of their topic and how they would present it. Topics that did change became more focused as respondents moved through information seeking stages. The Kuhlthau (1994) model predicts a more focused topic developing in later information seeking stages.

Implications for Future Research

The importance of criteria used to judge valuable or highly relevant information appear to remain fairly stable throughout the information seeking process, although the overall relevance of individual information sources can change. This implies that changes in the user’s topic focus may be responsible for a large proportion of the changes in relevance judgments that have been observed in previous studies. In this study, the criteria that were rated as most important tended to remain stable. Criteria that were rated as less important had a slight tendency to change more. The results of this study also
indicate that high relevance is ultimately decided by the end-user after the information has been obtained and read. However, topics and topic focus did change for many respondents. The components of the topic or subject that cause relevance judgments to change and that plays a role in shaping the user's ultimate topic focus have not yet been identified. These components may be so situation-dependent that it could be very difficult to model and measure them reliably across users, situations, and disciplines.

A question for further research is what is the role of partially relevant information in regard to the user's topic focus and movement through information seeking stages? Are many partially relevant documents used to create a topic focus or is one document central to this focus? The area of partial relevance is an intriguing aspect of relevance research that has not been well examined. It appears that partially relevant information can play an important role in information seeking and this is an area of the relevance judgment process that deserves further research.

The Internet is influencing end-user information seeking patterns. Respondents in this study had a wide range of experience and success with the Internet as a search tool. At least one respondent was very familiar with Internet searching and not at all familiar with other information tools. This respondent was unable to find the information she wanted on the Internet and experienced a high level of anxiety in early information seeking stages. How end-users are searching the Internet, their criteria for judging Internet information sources and the success or failure of the Internet to provide end-users with information that helps them resolve an information need are all important areas for future research.
Research to identify and validate additional sub-scale items or criterion terms on the criteria evaluation form, is also needed. The constructs of Information currency, Information availability and Information topicality all should be represented on the form with additional scale items or criteria that reliably and validly measure these constructs. Relevance criteria importance may vary with the subject and discipline knowledge of the end-user and with the end-user's discipline. Respondents for this study were graduate students from a variety of disciplines. Future research should examine the importance of relevance criteria to both subject novices and experts. Future research should also examine the importance of relevance criteria to experts in different disciplines. This could identify criteria that were important to users in one discipline but not as important to users from different disciplines and suggest features for IR systems that were discipline specific.

This study indicates that high relevance and high relevance criteria (exclusive of changes in topic) do not change much as graduate students move through the stages of information seeking for a paper or project. Further study is needed to see if partial relevance and partially relevant criteria change as end-users move through the stages of information seeking. Before this type of study can be done, research that identifies end-users' partial relevance criteria should be conducted. Most changes in high relevance appear to be related to topic change and not to other relevance criteria. Further study into the complex area of how users develop topics and topic focus is needed.

Ultimately any study of relevance and information seeking will have implications for IR system design and development. The respondents from this study wanted
accurate, credible, well-written, focused, understandable, consistent information that was
easy to obtain, current and on their topic. Current IR systems provide little or no
direction to help novice users make judgments about the credibility of information. Even
though some respondents defaulted to information that was easy to obtain, their
comments and their high ratings of quality and credibility criteria indicate that ultimately
the quality information was more important than its availability. This user population is
not being well served by information that is easy to obtain but does not have clues about
its reliability and credibility in either the information representation or it full-text format.
This describes much of the information that is available through the World Wide Web.

As Web gateways to CD-ROM indexes and OPACs become more prevalent and
more transparent to the end-user the net effect is that non-expert users have even fewer
clues to the source, accuracy, and credibility of the information that they locate on the
Internet. This creates a challenge for intermediaries, librarians, system designers,
researchers, and the producers of information. There is a critical need for additional
research regarding how users search for, evaluate, and use information on the Internet.
This research can lead to the development of Internet interfaces and other IR systems that
provide end-users with better clues about the effectiveness of their information seeking
and the value of the information that they find.

Most respondents for this study had some experience with information seeking
and some experience in using traditional information resources such as OPACs, online
searches and CD-ROM indexes. These information resources used human indexing and
cataloging and controlled subject vocabularies. The information sources that are included
in OPACs and traditional indexes are selected and edited. The selection chain often
includes publishers, editors, peer-reviews, and librarians. Perhaps these respondents have
learned to expect information that is high in quality as a result of using traditional
information resources. In the past, the cost of printing, cataloging or indexing, and the
physical space to store information led to an information selection process that was often
transparent to the user.

An automated relevance-ranking algorithm simply cannot determine or rate some
of the relevance criteria that were most important to these respondents. Even for experts,
rating information as accurate, credible, reliable, and well-written is problematic. The
amount of information that is easily available continues to increase exponentially,
particularly as the Internet expands. This expansion increases the difficulty for the end-
user searching for quality information. Even if users are able to recognize clues to
quality (such as the credibility of the author or party responsible for the information),
they may not have the time or patience to look through hundreds or thousands of Web
sites or links.

Individuals who have not been exposed to information resources that have been
subjected to rigorous selection standards may not rate the importance of criteria that
contribute to the constructs of quality and credibility as highly. Unfortunately, they also
may not be aware of clues to quality that some respondents in this study used. A serious
question for this group of younger information seekers is whether they are able to or care
to distinguish information that is available on the Internet through gateways to traditional
information resources (OPACs, CD-ROM indexes). These users may simply choose
Internet information primarily because it is easy to obtain. Respondents in this study rated the criterion Easy to obtain as important. However, this criterion was one of 10 criteria that were also rated as important, if 75 (on a scale of 0 to 100) is used as a cutoff point. Respondents in the five-week summer session were also under considerable time pressure. If the study were repeated with respondents in long semester, Easy to obtain may not be rated as high in importance.

If the respondents were given a choice between availability and quality it would be interesting to see which construct would be rated as most important. A study using comparative methodology such as multi-dimensional scaling would force a choice between constructs and this could increase our understanding of the relative importance of the relevance constructs to end-users.

Criteria studies, including this one, have all focused on highly relevant or not relevant information. As a result, the criteria for partial relevance judgments by users have not been identified. An assumption that partial relevance judgments can be collapsed into high relevance judgments has often been made in relevance research. This type of data condensation reduces the variability in the data and has limited our understanding of what partial relevance judgments are and how they are made. Perhaps these judgments and criteria are similar to high relevance judgments and criteria and differ only in degree. However, without additional research in this area, it is impossible to know for certain.

The most difficult dimension of relevance appears to be the degrees and shades of topicality. The mean importance rating for About my topic was highest for class
respondents and second highest for survey respondents (Accurate had the highest mean importance rating). Topicality was clearly central to most relevance judgments for these respondents. Changes in shades of meaning due to changes in the user cognition may be the primary reason that relevance judgments are so variable. Any comparison of relevance judgments of the same information will be separated by some passage of time. Changes in a user’s understanding of his or her topic may take place in a very short time frame and are likely influenced by interactions with IR systems, other citations or information sources, and the user’s own previous relevance judgments. The study of the factors, which influence users’ understanding of their information problem and topic, is a difficult, but important area for further research.

Researchers have operationalized relevance a priori as value, usefulness, topicality or pertinence. Interestingly, it has rarely been operationalized as quality. Value, usefulness, topicality and pertinence also share meaning with the concept of relevance. Information that is valuable or useful should also be relevant to users. Quality is clearly not a synonym for relevance. Information that is judged to be high in quality may not be relevant. However, it is clear that this respondent group considered quality criteria to be highly important. Would the results of relevance research have been different if respondents were asked to judge information by its quality and its relevance, topicality, usefulness or value?

Some theorists and researchers have suggested that a criterion other than relevance might have been used to evaluate IR system effectiveness. One criterion that has been suggested is a user rating of reduction in uncertainty (Saracevic, 1996). How
would end-user relevance judgments and criteria be influenced if they were asked to evaluate information based on how well it decreases their uncertainty?

Summary

The results of this research indicate that the importance of end-user relevance criteria employed in judging valuable or highly relevant information did not change as the user's information seeking stage changes. Additionally, the importance of the criteria did not change as users searched for, obtained and read information. Most respondents searched for and obtained information in the early stages of information seeking and read and obtained information in the later stages. Some respondents reported being in several information seeking stages concurrently and some respondents reported engaging in the activities of searching for, obtaining and reading information concurrently. Information resources were used evenly across information seeking stages, except CD-ROM indexes, which were used more frequently in early stages. Three constructs: Information quality, Information credibility, and Information completeness can be used to explain 48% of high relevance for these respondents. Three additional constructs: Information availability, Information topicality, and Information currency were suggested as important constructs for high relevance which should be measured by additional criterion terms on the criteria evaluation form. Further research is needed to develop the criteria evaluation form, examine partial relevance criteria and determine if the importance of relevance criteria for partially relevant information changes with users' information seeking stages. The Internet is now an information resource for most academic
information seekers. The presentation and selection of information on the Internet is not
serving these graduate students well. These respondents wanted high-quality, credible,
and complete information. Although this type of information can be found on the
Internet, it is often hard to identify and may offer few clues to aid the user in determining
how well these criteria are met. On the other hand, the Internet can be a source of topical,
current, and available information. Further research is clearly needed to see how the
Internet is influencing information seekers, relevance judgements and relevance criteria.
APPENDIX A

END-USER CRITERIA EVALUATION FORM
(Form 1)

INFORMATION EVALUATION

Your code:
Date: __________

Current statement of topic or subject for your project:

Estimate your progress toward completion of the following using a scale of 1 to 10 where
1=have not begun and 10=finished.

_____ Information seeking for your project.
_____ Reading for your project.
_____ Writing your project.

What project activity are you currently doing?

_____ Evaluating references from a list or a library catalog
_____ Obtaining information in full-text
_____ Reading information for project

Which of the following search stages best describes your current overall information seeking activity for the project.
For information about the search stages see the page titled "Activities that are characteristic of the 6 search stages" (form 3 -- yellow).

_____ Task initiation
_____ General topic selection
_____ Early exploration of general topic
_____ Focused formulation of topic
_____ Information gathering that supports specific topic focus
_____ Search closure

List 2 to 4 sources of information that you feel will be the best for your project at this point:

1. Author: ________________________________________ Pub. Date: __________
   Title: ____________________________________________

2. Author: ________________________________________ Pub. Date: __________
   Title: ____________________________________________

3. Author: ________________________________________ Pub. Date: __________
   Title: ____________________________________________

4. Author: ________________________________________ Pub. Date: __________
   Title: ____________________________________________

Describe in a few sentences why you think the information sources you have listed will be the best for your project:

Note. Original form was on blue paper. Original lines were 100-mm in length.
1. Topicality
   a. About my topic

2. Availability
   a. Easy to obtain
   b. Free or inexpensive

3. Novelty
   a. Unique or the only source
   b. Original
   c. New to me
   d. Familiar

4. Currency
   a. Current

5. Quality of information
   a. Well-written
   b. Credible
   c. Accurate
   d. Understandable
   e. Consistent
   f. Focused

6. Presentation Characteristics
   a. Presentation of information (Physical layout/design)
   b. Suitable length (long or short)
   c. Comprehensive
   d. Suitably general or specific
   e. Detailed
   f. Introductory
   g. Overview
### 7. Source Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Prominent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. I know the author personally</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. I know the source (author, journal, etc.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Reputable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Format of source (Journal article, microform, etc.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Interactive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 8. Information Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Describes methods/techniques</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Provides examples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Provides graphics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Statistical approach (Descriptive or Interpretive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Research approach (Theoretical or applied)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Provides proof</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Controversial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Provides bibliography or links</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. Provides background or history</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 9. Appeal of Information

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. I like it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Validates my viewpoint</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Interesting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Enjoyable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 10. Other

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comments**

Thank you for your participation in this important study.
APPENDIX B

SEARCH DIARY FORM
(Form 2)

SEARCH DIARY

Your code: __________________________

Date: ____________________

Current statement of topic or subject for your project:

Estimate your progress toward completion of the following using a scale of 1 to 10 where
1=have not begun and 10=finished.

___ Information seeking for your project.
___ Reading for your project.
___ Writing your project.

Where did you seek information today? Please indicate the number of times you used each of the
following:

___ Library online catalog
___ Print index to journals
___ CD-ROM index to journal articles
___ The World Wide Web (Internet)
___ Librarian
___ Reference or link from another document

___ Other Internet resources
Please indicate the resource__________________________________________________________

___ Other people
Please indicate who -- for example "Professor, classmate"________________________________

___ Other
Please indicate the resource: ________________________________________________________

Describe your information seeking experience today. Include anything you think might be appropriate:
feelings, actions, strategies, problems, etc. Discuss any changes in your topic development or plan for the
paper and what influenced these changes. (Please remember you can contact me for help in resolving
problems finding the information you need). Use the back of this sheet or additional pages if needed.

Note. Original form was on pink paper.
APPENDIX C

ACTIVITIES THAT ARE CHARACTERISTIC OF THE SIX SEARCH STAGES
ACTIVITIES THAT ARE CHARACTERISTIC OF THE SIX SEARCH STAGES

Please remember that although these stages are presented in a linear fashion, research has indicated that some people may go through some stages more than once. Often, people will "loop back" through some earlier stages. Also, people may do all kinds of things between stages, including simply thinking about the information that they have obtained and/or read.

1. Task Initiation
   - Considering possible topics
   - Contemplating assignment
   - Brainstorming
   - Talking with others
   - Browsing library shelves or journals

2. General topic selection
   - Evaluating possible topics in light of assignment, personal interest, information available, time allotted etc.
   - Preliminary searching in library or other information sources

3. Early exploration of general topic
   - Learning about topic in general
   - Identifying possible focuses
   - Obtaining information
   - Reading information

4. Focused formulation of topic
   - Listing possible focuses
   - Choosing possible focuses
   - Evaluation of topic focus in light of assignment, personal interest, information available, time allotted etc.
   - Identifying ideas in information

5. Information gathering that supports topic focus
   - Seeking supporting information
   - Extensive searching for information
   - Extensive obtaining full-text information
   - Organizing reading and information obtained

6. Search closure
   - Considering time limit
   - Finding little additional good information
   - Exhausting available information resources

Note. Original form was on yellow paper.
APPENDIX D

GENERAL INSTRUCTIONS FOR CLASS RESPONDENTS
General instructions:

The purpose of this study is to examine the criteria that people use when they decide whether to accept or reject information they find when they are looking for information. I am particularly interested in how these criteria are used over the course of a project like yours and what activities individuals participate in when looking for information.

Please fill out the search diary form (form 2 -- pink) each time you search for information. A search consists of one trip to the library, one Internet session, or any other activity that results in pointers to full-text information or full-text information. If you do several activities in the same search session, please fill out only one form. For example, you may use the library online catalog, then browse the shelves, then use a CD-ROM index all in the same trip to the library.

Just be sure to indicate each activity on the search diary form. I will periodically collect these forms in class. I will be happy to answer any questions or concerns at that time. You are welcome to call or email me at any time during this project.

You will need to fill out the information evaluation form (form 1 -- blue) at 4 different times. 1) and 2) after you have done a search session that results at least 3 references to full-text information (books, journal articles etc.) that you think you would like to obtain. 3) after you have actually obtained (but not thoroughly read) the full-text of at least 3 references, and 4) after you have read at least 3 of these full-text materials. You will complete the information evaluation form for two separate search sessions (time 1 and 2). Everyone should fill out four information evaluation forms. Evaluate your 2 to 4 best information sources at the time you are filling out each form.

If, at any time, you have any difficulty with filling out these forms please let me know immediately. Also let me know if you need any help with information seeking for this project (learning about the library, learning how to use CD-ROM or other indexes, planning a search strategy, using the World Wide Web etc.). I have 14 years of experience as a professional reference librarian and I hope I can help. You may call me at home (817-566-6335) or send me email (I usually check several times each day). jbateman@unt.edu

I appreciate your participation in this project. THANK YOU again.
INSTRUCTIONS FOR EVALUATION INSTRUMENT

Please use the criteria listed on the following pages to help describe how you have evaluated your information sources. Indicate with a mark on the line how important the criterion was to your evaluation and selection of this document or documents.

For example if you were selecting a kitten and intelligent was an important criterion for you, you might mark the line as follows:

<table>
<thead>
<tr>
<th>a. Intelligent</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If intelligence was a criterion you used in selecting the kitten, but it was not as important, you might mark the line as follows:

<table>
<thead>
<tr>
<th>a. Intelligent</th>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you do not find a criterion that you used please indicate it on the lines for "other" at the end of the form.
APPENDIX E

VERBAL INSTRUCTIONS TO CLASS RESPONDENTS
INTRODUCTORY SCRIPT

Introduce self and library background

- UNT Doctoral Candidate in Information Science interdisciplinary program
- Professional reference librarian for 14 years, experience with database searching, CD-ROM use, Internet resources. Experience in both academic and public libraries.
- Undergraduate degree in Dietetics with emphasis on clinical nutrition

Explain research

- General information about dissertation
  The purpose of my research is to study how users look for information over the course of a project such as the paper you have been assigned for this source. I am particularly interested in the criteria users employ to evaluate information or the criteria they use to accept or reject sources of information during the project.

- General information about the research project
  During this project I will be asking you to fill out a search diary form on a regular basis. I will also ask you to use an information evaluation instrument to indicate your reasons for choosing particular information. You will use this form several times. This may look like a lot of work initially, but the forms have been designed to take as little of your time as possible. I welcome and encourage any comments at any time.

Advantages to them

As a trained, experience professional librarian I hope that I can provide some instruction and help if you have any problems with using library or other information resources during the course of this project.

You will also get to observe the doctoral dissertation process firsthand. I would also be happy (if you would like) to individually give you some of my perceptions of the dissertation process.

What I need from them

- Go over consent forms

This is a standard consent form. I would like to emphasize that participate is voluntary and you may quit at any time. I also want to emphasize that Dr. Ward will not see any results of this research until grades have been recorded for the
course and that participation or non-participation will not affect your grade in any manner. Also I am studying what people do when they are looking for and evaluating information, I am not studying people's information seeking skill. I will not be judging your search behavior or strategies at any time. Please sign two forms, I will collect one for my records and you may keep the other form for your own information.

- Go over demographic form. (Appendix G)

The purpose of this form is so I can understand how familiar you are with using library resources and what you feel your degree of subject expertise is in this type of project and in your subject area.

- Go over data collection packet

Before I go any further I want to make sure that you understand some terms I am going to use. The library online catalog is the computer terminals that you use to find books in the TWU (or other) library. The catalog may also have the titles and locations of journals, reports etc. But it will not index individual journal articles. CD-ROM indexes usually specialize in some general subject area such as education, business or medicine and allow you to find journal articles in the subject area using subjects, author names etc. You may be familiar with ERIC, the education CD-ROM.

Please fill out the search diary form (pink form) (Appendix B) each time you do a search for information. If you do several searches in the same session i.e. you use the library online catalog, then browse the shelves, then use a CD-ROM index in the same session please fill out only one form. Just indicate all the sources you have used. I will collect these diary forms and any evaluation forms that you have done at the end of the Thursday class each week. I will also be happy to answer any questions or concerns at that time, although you are welcome to call or email me at any time during this project.

You will need to fill out the information evaluation form (blue form) (Appendix A) at 4 different times. The first 2 times will be after you have done your own searching and have a group of references that you think you would like to obtain. The third time will be after you have actually obtained (but not thoroughly read) the full-text of a group of references (i.e. Books, journal articles, printouts from the Internet etc.) and the fourth time will be after you have read a group of full text materials. For this research a "group" is a minimum of 3 references and a reference is a pointer to a full-text document (i.e. Author, title, Journal name etc.)

I would also like you to evaluate and mark your stage in information seeking. Research has identified six information seeking stages and the characteristics of
these stages are listed on the yellow form (Appendix C). You may not go through all these stages, you may not go through them sequentially and you feel you are in several stages at once. Please indicate what you feel your information seeking stage on the information evaluation (blue) form.

Provide an example of evaluation of information using selecting a kitten (Appendix D)

Questions

Have them fill out an evaluation form for an "ideal" information source. (Appendix F)

So that you can fill out the information evaluation form once while I am available to answer any questions and to see if you have any problems using it I would like you to fill out one now. In order to have something to evaluate you will have to use your imagination. I would like you to imagine the ideal information source for your project. If you don't have a topic yet, think in terms of your general interest or in terms of similar projects you have done. A source can be one or more books, journal articles, people, Internet resources etc. First describe this ideal source in paragraph using your own words. Then use the information source evaluation form to indicate the criteria that would describe your ideal source and how important each criterion is to you. Feel free to ask questions at any time.

Collect ideal evaluation form, demographic form, and consent form.

Let them know they can call or email to ask questions at any time.
APPENDIX F

IDEAL INFORMATION SOURCE FORM
IDEAL INFORMATION SOURCE

Date______________

Code______________

So that you can fill out the information evaluation form once while I am available to answer any questions and to see if you have any problems using it I would like you to fill out this form now. In order to have something to evaluate you will have to use your imagination. I would like you to imagine the ideal information source for your project. A source can be one or more books, journal articles, people, Internet resources etc. If you don’t have a topic yet think in terms of your general interest or in terms of similar projects you have done. First describe this ideal source in a paragraph using your own words. Then use the information source evaluation form to indicate the criteria that would describe your ideal source and how important each criterion is to you.

Describe your ideal information source below:

Note. This form was originally on green paper.
1. Topicality
   a. About my topic

2. Availability
   a. Easy to obtain
   b. Free or inexpensive

3. Novelty
   a. Unique or the only source
   b. Original
   c. New to me
   d. Familiar

4. Currency
   a. Current

5. Quality of information
   a. Well-written
   b. Credible
   c. Accurate
   d. Understandable
   e. Consistent
   f. Focused

6. Presentation Characteristics
   a. Presentation of information
      (Physical layout/design)
   b. Suitable length (long or short)
   c. Comprehensive
   d. Suitably general or specific
   e. Detailed
   f. Introductory
   g. Overview

7. Source Characteristics
   a. Prominent
   b. I know the author personally
   c. I know the source (author, journal, etc.)
   d. Reputable
   e. Format of source (Journal article, microform, etc.)
   f. Interactive

8. Information Characteristics
   a. Describes methods/techniques
   b. Provides examples
   c. Provides graphics
   d. Statistical approach (Descriptive or Interpretive)
   e. Research approach (Theoretical or applied)
   f. Provides proof
   g. Controversial
   h. Provides bibliography or links
   i. Provides background or history

9. Appeal of information
   a. I like it
   b. Validates my viewpoint
   c. Interesting
   d. Enjoyable

10. Other
   a. Other
   b. Other
   c. Other
   d. Other
   e. Other
   f. Other

Comments

Thank you for your participation in this important study
DEMOGRAPHIC INFORMATION

Degree sought: __________________________

Degree(s) currently held/subject area of degree(s): __________________________

__________________________

Gender: F    M

Please circle your level of experience or expertise with the following library or information-finding resources, where 1 = no experience and 5 = expert:

<table>
<thead>
<tr>
<th>Resource</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWU library</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any library (public, academic)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Library electronic catalogs</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Print journal citation indexes</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Journal citation indexes on CD-ROM</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Internet or World Wide Web</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Please circle your level of experience or expertise with the following tasks, where 1 = no experience and 5 = expert:

<table>
<thead>
<tr>
<th>Task</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conducting literature searches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conducting research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing research papers or reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing papers or reports for publication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other types of writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Respondent code: ______
Hello,

I am a doctoral candidate in Information Science at the University of North Texas. I am conducting the enclosed survey as part of my dissertation research. I would appreciate your taking a few minutes out of your busy schedule to complete and return this survey. Your survey is important for my research. I know your time is valuable and I thank you for your help. If you have any questions about this survey or research you may call me at 940-566-6335.

Judy Bateman
Doctoral Candidate
Information Science
University of North Texas
APPENDIX I

INFORMATION CRITERIA EVALUATION SURVEY, FIRST MAILING
INFORMATION EVALUATION SURVEY

I am conducting research to find out how people evaluate information. Please focus on the most valuable source of information for a recent or current research paper. This source could be a book, journal article, Internet page, or citation for a book or an article. Please indicate whether you have obtained and read this information. Check the information seeking stage that applies to you when evaluating this information.

GENERAL INFORMATION
1. Information source: ___Book ___Journal article ___Internet page ___Citation
2. Information seeking stage
   - Starting my search for information
   - Selecting my general topic
   - Investigating my general topic
   - Selecting my specific topic
   - Gathering information about my specific topic
   - Ending my search for information
3. I have obtained this source in full text ___Y ___N
4. I have read this source ___Y ___N
5. I have seen an abstract of this source ___Y ___N
6. Degree sought: ______________________________
7. Area of study: ______________________________

DIRECTIONS

Please use the reasons listed below on the following pages to describe how you evaluated the information source. Indicate with a mark (I) on the line how important the reason was to your evaluation.

For example, if you were selecting a kitten and intelligence was important to you, you would mark the line as follows:

a. Intelligence

<table>
<thead>
<tr>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If intelligence was a reason you used to select the kitten but it was not important you would mark the line as follows:

a. Intelligence

<table>
<thead>
<tr>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now go to the next pages and complete the information evaluation instrument. Please mail your completed forms to me in the enclosed self addressed stamped envelope.

This project has been reviewed and approved by the UNT Committee for the Protection of Human Subjects (940) 565-3940

Note. Lines on original form were 100-mm in length.
# INFORMATION EVALUATION INSTRUMENT

As you complete this form, focus only on the information source (e.g. book, citation) you identified on the previous page; if possible, have this source in front of you.

## 1. Topicality
   a. About my topic

## 2. Availability
   a. Easy to obtain
   b. Free or inexpensive

## 3. Novelty
   a. Unique or the only source
   b. Original
   c. New to me
   d. Familiar

## 4. Currency
   a. Current

## 5. Quality of Information
   a. Well-written
   b. Credible
   c. Accurate
   d. Understandable
   e. Consistent
   f. Focused

## 6. Presentation Characteristics
   a. Presentation of information
      (Physical layout/design)
   b. Suitable length (long or short)
   c. Comprehensive
   d. Suitably general or specific
   e. Detailed
   f. Introductory
   g. Overview

<table>
<thead>
<tr>
<th>Not important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
INFORMATION EVALUATION INSTRUMENT

7. Source Characteristics
   a. Prominent
   b. I know the author personally
   c. I know the source
      (author, journal, etc.)
   d. Reputable
   e. Format of source
      (Journal article, microform, etc.)
   f. Interactive

8. Information Characteristics
   a. Describes methods/techniques
   b. Provides examples
   c. Provides graphics
   d. Statistical approach
      (Descriptive or Interpretive)
   e. Research approach
      (Theoretical or applied)
   f. Provides proof
   g. Controversial
   h. Provides bibliography or links
   i. Provides background or history

9. Appeal of Information
   a. I like it
   b. Validates my viewpoint
   c. Interesting
   d. Enjoyable

10. Other
    a. Other
    b. Other
    c. Other

Comments

THANK YOU FOR YOUR PARTICIPATION IN THIS IMPORTANT STUDY
Thank you for returning my survey. Because the instructions on the original survey were confusing to many people I am sending out the survey again. I apologize for having to ask you to look at your survey again.

The marks on the survey can be placed anywhere on the line. This mark will indicate the degree of importance of each reason when you were evaluating information. You are not limited to the two labels on the ends of the scale on the line (not important/important). I have enclosed your original survey form. If you understood the instructions and the marks are where you want them to be please return the original survey to me. Otherwise please complete the survey again on the enclosed form and mail it to me. The new form has more extensive instructions on the first page which I hope will be easier to understand. I appreciate whatever help you can give me. Thank you again.

Judy Bateman
Doctoral Candidate
Information Science
University of North Texas
APPENDIX K

INFORMATION CRITERIA EVALUATION SURVEY, SECOND MAILING
INFORMATION EVALUATION SURVEY

I am conducting research to find out how people evaluate information. Please focus on the most valuable source of information for a recent or current research paper. This source could be a book, journal article, Internet page, or citation for a book or an article. Please indicate whether you have obtained and read this information. Check the information seeking stage that applies to you when evaluating this information.

GENERAL INFORMATION

1. Information source: ___ Book ___ Journal article ___ Internet page ___ Citation

2. Information seeking stage
   ___ Starting my search for information
   ___ Selecting my general topic
   ___ Investigating my general topic
   ___ Selecting my specific topic
   ___ Gathering information about my specific topic
   ___ Ending my search for information

3. I have obtained this source in full text ___ Y ___ N

4. I have read this source ___ Y ___ N

5. I have seen an abstract of this source ___ Y ___ N

6. Degree sought: ____________________________

7. Area of study: ____________________________

DIRECTIONS

Please use the reasons listed below on the following pages to describe how you evaluated the information source. Indicate with a mark (I) on the line how important the reason was to your evaluation. You may mark anywhere on the line. Mark the place that indicates how you rate the reason.

For example intelligence was important in selecting a kitten, you might mark the line this way:
   a. Intelligence

   Not Important
   __________________________________________
   Important

If intelligence was a reason for selecting the kitten but it was not very important, you might mark the line this way:
   a. Intelligence

   Not Important
   __________________________________________
   Important

If intelligence was a reason for selecting the kitten but it was not important you might mark the line this way:
   a. Intelligence

   Not Important
   Not Important

Now go to the next pages and complete the information evaluation instrument. Please mail your completed forms to me in the enclosed self-addressed stamped envelopes.

This project has been reviewed and approved by the UNT Committee for the Protection of Human Subjects (940) 565-3940

Note: Lines on original form were 100-mm in length
**INFORMATION EVALUATION INSTRUMENT**

As you complete this form, focus only on the information source (e.g. book, citation) you identified on the previous page; if possible, have this source in front of you.

<table>
<thead>
<tr>
<th>1. Topicality</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. About my topic</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Availability</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Easy to obtain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Free or inexpensive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Novelty</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Unique or the only source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Original</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. New to me</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Familiar</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Currency</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Current</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Quality of Information</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Well-written</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Credible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Accurate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Understandable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Consistent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Focused</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Presentation Characteristics</th>
<th>Not Important</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Presentation of information (Physical layout/ design)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Suitable length (long or short)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Comprehensive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Suitably general or specific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Detailed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Introductory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Overview</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### INFORMATION EVALUATION INSTRUMENT

#### 7. Source Characteristics
- **Prominent**
- **I know the author personally**
- **I know the source**
  - *(author, journal, etc.)*
- **Reputable**
- **Format of source**
  - *(Journal article, microform, etc.)*
- **Interactive**

#### 8. Information Characteristics
- **Describes methods/techniques**
- **Provides examples**
- **Provides graphics**
- **Statistical approach**
  - *(Descriptive or Interpretive)*
- **Research approach**
  - *(Theoretical or applied)*
- **Provides proof**
- **Controversial**
- **Provides bibliography or links**
- **Provides background or history**

#### 9. Appeal of Information
- **I like it**
- **Validates my viewpoint**
- **Interesting**
- **Enjoyable**

#### 10. Other
- **Other**

**Comments**

---

**THANK YOU FOR YOUR PARTICIPATION IN THIS IMPORTANT STUDY**
REFERENCES


Foskett, D. J. (1972). A note on the concept of "relevance". *Information Storage and Retrieval, 8*(2), 77-78.


