REFERENCES
Linked references are available on JSTOR for this article:
http://www.jstor.org/stable/4308837?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

THE NATURE OF RELEVANCE IN INFORMATION RETRIEVAL: AN EMPIRICAL STUDY[1]

Taemin Kim Park[2]

Experimental research in information retrieval (IR) depends on the idea of relevance. Because of its key role in IR, recent questions about relevance have raised issues of methodological concern and have shaken the philosophical foundations of IR theory development. Despite an existing set of theoretical definitions of this concept, our understanding of relevance from users' perspectives is still limited. Using naturalistic inquiry methodology, this article reports an empirical study of user-based relevance interpretations. A model is presented that reflects the nature of the thought processes of users who are evaluating bibliographic citations produced by a document retrieval system. Three major categories of variables affecting relevance assessments—internal context, external context, and problem context—are identified and described. Users' relevance assessments involve multiple layers of interpretations that are derived from individuals' experiences, perceptions, and private knowledge related to the particular information problems at hand.

Introduction and Purpose

The concept of relevance has played a major role in information retrieval (IR) research since the 1950s. The goal of an IR system is to provide users with access to "relevant" documents. Typically, performance is measured by how well the system succeeds in retrieving relevant documents. Recall (the proportion of relevant material actually retrieved) and precision (the proportion of retrieved material actually relevant) are often employed as standard measurement criteria in experimental studies.

Despite a rich literature on "relevance," there is no commonly accepted definition. Although much experimental research in IR is based on the idea of relevance, various interpretations of the concept have been made. Because of its key role in information retrieval research, the lack of consensus on the meaning of relevance and "relevance assessment" raises issues of methodological concern. These have been most thoroughly explored by David Ellis [1] in a thought-provoking article concerning the use of relevance in information retrieval research. Ellis questioned the validity and reliability of relevance as a measuring instrument in IR experimental studies. Furthermore, our understanding of "relevance" is limited due to our lack of understanding of what real users mean by relevance.

This article attempts to develop an empirical understanding of the concept of relevance as viewed by users of real information systems. I view the determination of relevance from the perspective of how users select "relevant" bibliographic citations in IR. That is, the purpose of this research was to learn how users make relevance assessments when evaluating citations that are produced by a document retrieval system: What are their decision contexts? What are the characteristics of those contexts? and How are they related to one another? I also investigated the contextual changes that occur as a result of an examination of citations: whether a citation stimulates a reformulation of the user's information problem or provides a connection between the problem and other ideas. The exploratory nature of the study is reflected in the design of the research.

Background

Two comprehensive literature reviews on the subject of relevance already exist, by Tefko Saracevic [2] and Linda Schamber, Michael B. Eisenberg, and Michael S. Nilan [3]. Only a selection of literature on the various relevance theories and variables affecting relevance judgments is summarized here. The role of relevance in relation to traditional and cognitive models of IR research is discussed in the following sections.

Historical definitions of relevance in information science are well documented by Saracevic [2]. One of the obvious dimensions of relevance is "about," "on the topic," or "on the subject." Eisenberg and Schamber [4] called this notion *topical relevance:* whether the topic of the information retrieved matches the topic of the request. In a recent study, Schamber and her colleagues [3] provided an excellent literature review on this topic, including users' views of relevance. They viewed *user-based*

relevance as follows: "The locus of relevance is within individuals' perceptions of information and information environment—not in information as represented in a document or some other concrete form" [3, p. 771].

In distinguishing between relevance and pertinence, D. A. Kemp [5] suggested that pertinence requires that a document must have a bearing on a user's particular situation to be useful. Therefore, pertinence must be decided by the person with the information need. In other words, relevance (or topical relevance) is assessed in relation to a request—a statement of information need—while pertinence is assessed subjectively by the user. According to D. J. Foskett, "the true distinction is that 'relevant' should be taken to mean 'belonging to the field/subject/ universe of discourse delimited by the terms of the request, as established by the consensus of workers in that field'; while 'pertinence' should be taken to mean 'adding new information to the store already in the mind of the user, which is useful to him in the work that prompted the request' " [6, p. 77].

W. S. Cooper proposed the concept of utility and distinguished the term *logical relevance* from the utility that measures the ultimate usefulness of a piece of information to the user. Utility was conceived as beyond logical relevance, as "a catch-all concept involving not only topic-relatedness but also quality, novelty, importance, credibility, and many other things" [7, p. 92]. The term *novelty* was originally introduced by F. W. Lancaster and E. G. Fayen [8]. Novelty is represented as measuring the proportion of relevant retrieved documents that are new to the requester.

Patrick Wilson [9] suggested that relevance is not a single notion but multiple concepts and proposed the idea of *situational relevance*. Situational relevance is based on the relation between an item of information and a particular individual's personal view of the world and his or her situation in it. The piece of information is seen as situationally relevant if it brings a significant change in one's view of a situation.

Making a distinction between subjective and objective relevance, Don R. Swanson [10] defined *subjective relevance* as the mental experience of an individual person who has an information need. *Objective relevance* was defined as a logical relationship between a document and an objectified written statement of a request. Relevance, in this sense, is a connection between a written request and a document and belongs to the world of objective knowledge.

Finally, based on an analysis and interpretation of a theory of relevance proposed by Dan Sperber and Deirdre Wilson [11], Stephen P. Harter [12] introduced *psychological relevance* in IR. A retrieved citation, as a psychological stimulus, is found to be relevant to an individual if it causes cognitive change in that individual. Some of the essential con-

cepts and principles of psychological relevance are: context is a psychological construct that represents one's assumptions about the world at any given moment; it is dynamic, changing as contextual effects occur; and a necessary condition for an assumption to be relevant in a context is having contextual effects in that context. Three possible ways of how contextual effects can take place were explained. First, the assumptions held in the present context can combine with the existing (old) assumptions to create new assumptions. Second, a new assumption can cause a modification in one or more assumptions from the existing context. Third, a new assumption can strengthen or weaken one or more assumptions in the existing context. Involved with changes in an individual's cognitive state, psychological relevance is represented as dynamic and goes well beyond topical relevance.

Our empirical understanding of relevance has been slow to develop. Two large-scale projects on "relevance judgments" were undertaken experimentally by Carlos A. Cuadra and Robert V. Katter [13, 14] and Alan M. Rees and Douglas G. Schultz [15]. The purpose of their research was to identify the possible variables affecting relevance judgments and to determine the effects of these variables. In a series of more than fifteen experiments, Cuadra and Katter listed the major categories of variables studied: (1) people (judges), (2) documents, (3) statements of information requirements, (4) judgment conditions, (5) form of response, and (6) judgmental attitudes. The definition of relevance used by their judges was "how good an answer the journal article would be to the requirement statement" [14, p. 296]. Rees and Schultz added the utility concept to their definition of relevance: "the degree to which the document bears on or has application to the research you have heard described"; and "by usefulness we mean the degree to which the document would be useful to you as an individual. In other words, usefulness should take into account your interest, knowledge, experience, etc., in doing this research" [15, p. 16]. However, neither study examined real users and their information problems. Summarizing the conclusions of relevance experimentation, Saracevic [16] highlighted several major factors affecting the relevance judgment process: (1) judges' subjective expertise at the various stages of research, (2) judges' subjective knowledge, (3) judges' academic and professional training, (4) a document's intended use, (5) stylistic characteristics of documents, and (6) aspects and definitions of relevance. More recently, Eisenberg [17] investigated an application of magnitude estimation, an open-ended scale used in psychophysics, for the measurement of relevance judgments. His findings suggest that the magnitude estimation method is less influenced by potential biases than category-rating scales. However, Eisenberg and Carol Barry [18] found in a later investigation that

both measuring types (magnitude and category) were affected by the order in which documents are presented. Lorraine M. Purgailis Parker and Robert E. Johnson [19] found that the order of presentation does not influence relevance judgments when fewer than fifteen documents are evaluated. In a comparison of the relevance judgments between actual users and secondary judges, Joseph W. Janes and Renée McKinney [20] reported that the secondary judges tended to rate documents more highly than actual users. Their decisions were based on "perceived topicality": for example, the topical information such as indexing and abstracts was used more often than the nontopical information— bibliographic data.

In IR experimental research, the results of relevance judgments are usually treated as binary ("relevant" and "not relevant") or on a scale measuring degrees of relevance. Relevance judgments for the evaluation of search results have been made either by search intermediaries or subject experts. The flaw of relevance assessments made by nonusers in online searching was well stated by Charles T. Meadow [21]; he raised essential issues concerning how relevance judgments are made as well as by whom, and claimed that experimental data concerning relevance are ambiguous.

Relevance and Information Retrieval Testing
The Cranfield tests [22] have, with some variations, been the primary model of experimental IR research during the past twenty-five years. The research design of the Cranfield project provides a conceptual framework for reexamining the underlying assumptions of the traditional IR model and the implications of using "relevance judgments" in IR evaluation. Cranfield's experimental design involved four major steps: (1) building a test collection, (2) gathering users' questions, (3) obtaining relevance judgments, and (4) conducting tests of retrieval. The third component—obtaining relevance judgments—deserves critical attention, in particular in examining the impact of relevance, since the test results are dependent on these judgments.

There have been some cases in which the relevance assessments used in experimental studies have been critically evaluated. Swanson [23] criticized design flaws in the experimental methodology of the Cranfield project, particularly that the relevance assessments made during the screening process by a group of graduate students seemed to have missed many relevant documents. Harter [24] empirically confirmed Swanson's hypotheses and questioned the reliability of relevance assessments.

Design problems and problems with relevance judgments in particular are still present, as more recent research in online searching con-

firms. For example, Carol H. Fenichel [25] and Trudi Bellardo [26] used the same search questions, which were drawn from the ERIC ONTAP file on DIALOG, for their dissertation research. However, the two researchers disagreed radically on which citations should be considered "relevant" to the two search questions. In a recent work, Swanson [27] recalled the enormous disagreement in relevance judgments between the two research teams—Armed Services Technical Information Agency (ASTIA) and Documentation Inc.—in a large-scale IR test done in 1953. "Both groups agreed that 1,390 documents were relevant to one or more of the 98 questions, but there are another 1,577 documents that one team or the other, but not both, considered to be relevant" [27, p. 92].

Examining the underlying dimensions of relevance in IR experimentation, the Cranfield model views relevance as "on the topic," the relationship between the topic of a question and that of a document. The nature of the relationship between a document and a user's question is very precise and fixed. There are no concerns with the individuality of users. The model also assumes that users' information needs are conceptually well defined and that they know how to express them.

The concept of relevance as topical relevance, assumed in most IR tests, may not be either adequate or sufficient if we consider users as a main focus of an information system. Lauren B. Doyle's early statements illuminate the limitation of topical relevance in IR evaluation. "The flaw is that there may be a great difference between relevance to a given request statement and relevance to a person's real information need" [28, p. 199]. "Relevance will serve its purpose, but will decline as the realization slowly comes that an individual's information need is so complex that it cannot be accurately stated in a simple request. The fact that people do request information in simple terms is a reflection of the inadequacy of both people and systems, and not a reflection of the true structure of the need" [28, p. 200].

The difficulty of specifying an information need is well stated by Robert S. Taylor: the user has "a certain incompleteness in his picture of the world—an inadequacy in what we might call his 'state of readiness' to interact purposefully with the world around him, in terms of a particular area of interest" [29, pp. 180–81]. If we accept Taylor's model of information need, there is a significant difference between an information need and a question statement. Taylor classified an information need into four stages: Q1, the visceral need; Q2, the conscious need; Q3, the formalized need; and Q4, the compromised need. Taylor defined the visceral need as an actual, but inexpressible, need for information. The conscious need is a conscious mental description of an ill-defined problem area. At the formalized need, a user creates a formal

statement of his or her information question. Finally, the question is transformed from the formalized level into the compromised level for approaching an information system. As Taylor's model indicates, the original need for information tends to be distorted or changed in this process. However, despite Doyle's and Taylor's observations, the nature of relevance to real users and their information needs has not been seriously treated in IR research (except in recent studies by Schamber and her colleagues).

Relevance in Cognitive Models of Information Retrieval
Cognitive studies of IR are concerned with users' problems, knowledge structures, tasks, and goals. Nicholas J. Belkin's [30] cognitive communication model considers the IR situation as a communication system in which the roles of the participants are primary. In Belkin's model, the user evaluates the information based on his or her anomalous state of knowledge (ASK) and the underlying problems and image of the world. Hence, the user's information need has an important role in this model of IR. P. Ingwersen [31] describes the cognitive view of IR as: "Any processing of information—whether perceptual or symbolic—is mediated by a system of categories or concepts, which—for the processing device—are a model of its world whether the device is a human being or a machine. . . . It is the individual knowledge structures of the information processing mechanism . . . which provide the basis for decisions on which ambiguities (or problems) should be eliminated. Information is seen as supplementary to the conceptual system representing the mechanism's world model" [31, p. 87].

As the ASK model suggests, the cognitive approach views the information need as a way to solve an anomaly in which something is wrong in one's state of knowledge. In this view, people recognize the need for information but the user cannot specify it exactly because of lack of knowledge. Susan E. MacMullin and Taylor [32] examined information need from the perspective of its underlying problem context and speculated on eleven problem dimensions beyond subject matter as new criteria for evaluating information provision.

Some recent research efforts have taken a cognitive approach. Saracevic et al.'s study [33] reported that users' information-seeking contexts such as individual's perceptions of the problem definition and the perceived estimation of public knowledge (the available literature on a subject) had a significant effect on relevance judgments of users. In a study on users' information seeking and retrieving, Belkin and others [34] investigated users' tasks, goals, and their impact on information-seeking behavior in the context of an online public access catalog. The study combined methods of transaction logging, questionnaire, and observing

and interviewing users about their behavior. In a report on modeling
and measuring user-intermediary-computer interaction in online
searching, Louise T. Su and Saracevic [35] presented a user's informa-
tion-seeking context as including six attributes: problem, work stage,
intent, personal or internal knowledge, public or existing knowledge,
and information activities.

Methodology

The ultimate goal of an IR system is to provide relevant documents to
users. Users are the prime concern of a system. This view is especially
appropriate as more IR systems are designed to be used by end users.
If we place users and their information problems as a main focus of
IR evaluation and research, relevance should be something that a user
perceives in his or her information-seeking process. Relevance, in this
context, totally belongs to an individual's situational and psychological
state. From a user's perspective, relevance is intrinsically related to a
mental process, to the criteria used in the evaluation of citations re-
trieved in the IR process. In this sense, relevance may be equated with
users' selection behavior. Relevance is something that users perceive as
they select citations in an IR context. This view is consistent with a view
of psychological relevance that emphasizes an individual's cognitive state
and context in the judgment process [12].

   The choice of methodology for a study should fit the nature of the
research problem. The research philosophy and accompanying method-
ology chosen in this study is naturalistic inquiry. Constance A. Mellon
stated the rationale for the use of naturalistic inquiry: "Naturalistic re-
searchers are interested in knowing all about each characteristic, or ele-
ment, of the social phenomenon and how the elements work together
to create the situation under study" [36, p. 5]. Naturalistic inquiry is an
appropriate methodology to understand how end users make the selec-
tion decisions of choosing whether or not to look at a document pro-
duced by a document retrieval system; these are potentially complex
phenomena in a real-life situation over which a researcher has no con-
trol. This approach is in agreement with Brenda Dervin and Nilan's
comprehensive literature review of information needs and uses [37], in
which the authors pointed to a lack of definition or understanding of
basic concepts such as "information," "information need," and underly-
ing assumptions about IR systems, and suggested a paradigm shift in
library and information science research, away from the traditional ex-
perimental model.

   The data collection method chosen for this research was based on the

belief that the naturalistic inquiry method is appropriate to reflect the underlying meaning of relevance in users' contexts because of its potentially complex and subjective nature. I did not make an attempt to provide an a priori definition of relevance; people arrived at their own definitions using a naturalistic approach. The subjects for the study were users in real academic search situations who received an online search during the summer session or fall semester, 1990.[3] A letter concerning my possible contact of subjects was sent to all users at the time they received the results of their searches. The subjects were purposefully selected. They participated in this project voluntarily. Sharan B. Merriam stated a rationale for choosing purposive sampling: "Purposive sampling is based on the assumption that one wants to discover, understand, gain insight; therefore one needs to select a sample from which one can learn the most" [38, p. 48]. I decided to use purposive sampling rather than random sampling of subjects because I hypothesized that relevance from a user's perspective would vary by different user groups and an individual's information-seeking problem and context.

The sampling selection strategy followed Judith Preissle Goetz and Margaret Diane LeCompte's [39] criterion-based method. A quota selection was used in selecting the initial group. I chose a small number of cases (one faculty member and one doctoral student) as an initial purposeful sample, a subset of the whole. The subsequent cases were selected based on the "typical-case selection" method. Using this method, subjects are selected who display attributes and meet criteria to maximize information gained from previous cases. Yvonna S. Lincoln and Egon G. Guba [40] emphasize that naturalists are concerned that context is critical. The decision to stop sampling was based on information redundancy, not facilitation of generalization. Different user groups and problem-oriented searches at different stages of research were chosen deliberately.

The ten subjects included three faculty members, a visiting scholar, five doctoral students, and a master's student. Their disciplines were education (three), library and information science (three), sociology (two), anthropology (one), and criminal justice (one). Brief information concerning each subject is included in the Appendix. The databases searched were ERIC, Sociological Abstracts, Child Abuse and Neglect, MEDLINE, PsycINFO, Social Scisearch, Dissertation Abstracts Online, Philosopher's Index, MLA Bibliography, U.S. Political Science Documents, and National Newspaper Index.

---

3. Online searches were conducted by students in L644, Information Storage and Retrieval, a course offered by the School of Library and Information Science, Indiana University.

The interview itself started with a brief introduction and my explanation about the purpose of the interview and the research. Each subject was assured that the result of data collection would be kept confidential. Reasons for using a tape recorder to record the interview were explained and permission was requested. The interview was conducted in a semistructured way. It was guided by a list of questions to be covered, but the sequence or the exact wording was not followed. Open-ended questions concerning each subject's information-seeking context, including the problem area, the stage of research, the expected product of research results, and behavior related to selection of citations, were asked. Each interview took between forty-five and sixty minutes. The interview tape was transcribed.

The printed output of the search results served as secondary sources of needed data for this study. This source helped me as well as the subjects trace the thinking process in response to each citation on the list. The verbalization of thoughts along with each citation helped individuals to be motivated and to organize their own reasoning behind the assessment of a citation. The design process was flexible enough to maximize the free flow of thinking and to view the natural approach of each subject to each citation. I allowed the subject to choose his or her own method of going through the list of retrieved citations.

The data collected through the interviews were a collection of users' verbal descriptions about their own reasoning and thoughts behind their perceptions and evaluation judgments for each citation. The full description of each case was read and examined in depth for the purpose of unitization of data and eventual categorization. Following Lincoln and Guba's [40] strategy, unitization is a process of identifying the units of information that serve as the basic idea of formulating categories. The unit is the smallest piece of information that can stand by itself; a unit can be a phrase, sentence, or paragraph. The analysis of data focused on representing the recurring themes of each subject's relevance interpretation for each citation retrieved. The development of categories was guided and influenced by the research's purpose and the researcher's knowledge in addition to her educated and intuitive hunches.

As with other approaches to research, the naturalistic researcher should be sensitive to the biases inherent in this type of research. The researcher must admit that some of her own biases may have been filtered through the data collection and interpretation stages. In particular, Ross Atkinson's [41] typology of librarians' decision contexts in book selection inspired me prior to and during the data analysis of this research. Atkinson proposed a theoretical model that illustrates the decision contexts of librarians as they practice collection development of library materials. Atkinson's model describes the decision contexts of

the selection process influenced by a citation and shows how and what decision contexts are generated by and reflected in this process. He defined three contexts: the "syntagmatic context" concerns variables within the citation itself; "the context of supplementation" is sensitive to the selection sources (for example, book reviews); and "the context of resolution" derives from the librarian's experience with the collection, the patron, and the subject literature. The main thrust of Atkinson's model is defining a citation as intertext, which is interwoven to a large network, including personal, contextual, and institutional factors.

Atkinson's model provided an initial structure and source for classifying my data for the purpose of the pilot project. The model provided useful insights for looking at a citation in relation to a text and its intertextual function. However, as Atkinson's model is hypothetical without supporting empirical data, it was not precise enough to capture the richness of my data. The decision to use Atkinson's model for the purpose of data analysis was changed after the pilot project because of the incongruence between the data in the present study and Atkinson's typology. These differences will be discussed later in this article.

Since a naturalistic inquiry starts with different premises from the conventional quantitative method, the conventional measures of reliability and validity are not directly applicable. The primary rationale for using qualitative research lies in understanding and interpretation rather than predicting cause and effect relationships and generalization of findings. Lincoln and Guba [40] proposed substitute criteria for assessing the quality of naturalistic inquiry: credibility, transferability, dependability, and confirmability. Credibility is used to refer to internal validity and is concerned with credible findings and interpretations. Merriam [38] stated that validity must be assessed in terms of interpreting the investigator's experience. Transferability is the naturalist's version of external validity. The definition of reliability is similar to dependability.

For naturalistic inquiry some alternative techniques, described by Lincoln and Guba [40], were applied. For the purpose of increasing credibility, the technique of member checking was used to test the validity of data. Members were sent raw data for their reactions. Nine of the subjects were involved in testing the analytic categories and interpretations that emerged from the research. An attachment of the interview transcription was sent to the subjects who evaluated it for overall accuracy and credibility. They were asked to check whether my representations of their perceptions and realities were adequately reflected, giving them an opportunity to react and confirm or change them. Five members reported and corrected some typographical errors in interview transcriptions and the rest of them agreed with my data analysis and interpretation.

The peer debriefing technique is a process of exposing one's research to colleagues for comments on the findings to increase the credibility of findings. Peer debriefing was used to test and clarify my bias for data analysis and interpretation. The goal of peer debriefing was achieved by preparing and distributing a manuscript, based on my dissertation and prepared for publication, to three of my colleagues at the Indiana University Library. Also, some members of my dissertation committee provided feedback on the clarification of categories and interpretation regarding the data analysis. Inconsistencies between models of earlier and later versions were pointed out by colleagues and committee members and were adjusted accordingly. The adequacy of the labels of the variables were also noted and modified.

Data Analysis and Findings

The goal of the data analysis was to capture the recurring themes that emerged as users evaluate bibliographic citations. The variables in the model summarized here are not based on the previous theory but, rather, were developed from data themselves. Specific findings will be discussed later.

On the surface, relevance can be looked at as a relationship between a user and a citation. However, one's interpretation of a citation is rooted in various sources. A user's evaluation of a citation can be analyzed in two ways, from the attributes of the citation itself and from the contexts in which the evaluation takes place. The data analysis represents those two dimensions, which I call citation-based and user-based. The first part of the analysis is concerned with users' interpretation of the aspects or components of a citation. An interpretation of a citation is presented to illustrate how a user views and filters an element or interconnections among elements of a citation. The second part of the analysis focuses on an examination of underlying factors influencing the choice of a citation. Those factors seem to fall into three major categories, based on an individual's experience, perception, and knowledge of the information problem. The second part of the analysis characterizes those sources of relevance assessments as based on internal factors, external factors, or factors related to the problem context.

The internal context reflects a user's interpretation of a citation based on his or her own prior experiences or perceptions in the information problem area. The external context refers to factors that are particular to this search and research. The external context contrasts with the internal context in the sense that the origination of the external context stems from an individual's perceptions and situations in relation to the search process, the stage of research, and the priority of information

needs related to the present search. There are other perspectives on the interpretation of a citation. A user is continuously assessing the value of information and interpreting a citation in order to learn and shape the research problem. The problem context is a content-oriented context and indicates the various implied uses of information in relation to expanding one's thoughts and constructing one's ideas in the problem area. The three categories are not discrete but, rather, interconnected, and they influence one another.

The following sections describe the major categories that emerged from the data and shows a sample of the raw data with minimal interpretation in order to allow readers to make their own interpretations. In this article I have listed only some examples of my raw data. A full description of instances for each category and subcategory is found in my dissertation [42].

Interpretation of a Citation
This section of data analysis illustrates a user's initial approach to citation selection—how the user assesses, interprets, and understands a citation. The data analysis shows how a user perceives the value of each element of a citation in relation to his or her interest and need for information. The analysis is further subcategorized to reflect how the individual combines and balances elements of a citation. Citation variables seem to serve as stimuli in relation to the major contextual variables.

A. *Title.*—The title plays a leading role in a user's understanding of the subject matter of a citation. Key terms in the title alert the user to the area of interest. When the words in the title represent the subject matter of the information need well, the title independently becomes a sole decision source. The following statements by subjects show the importance of the title: "I know exactly what I'm looking for . . . the social construction of meaning, pragmatism, video. There are these certain words. I worked on it for two years before I did this search. I have a much better idea of what I'm looking for" (subject 8).[4] "I'm looking for phrases or terms that alert me to something relevant to me. The best analogy I can give is when I teach a field work course, we talk about using sensitizing concepts and these are terms that point you in the direction of something like . . . for an anthropologist a sensitizing concept is kinship or society or family. Well for me, these concepts aren't very precisely defined but they orient you to looking for that kind of thing. Well, I have those in my head too for this study. Those concepts

4. See the Appendix for a brief description of each subject.

are things like liberalism or social criticism or moral theory. I look for those kinds of terms in the title" (subject 10).

B. *Style of the title.*—The readability of the title and the use of language can be reasons for rejection: "It has to be readable; it can't be too technical. If it sounds too technical I'm not going to read—and all of these sound very technical; my searcher said these might be possibilities. They're all very technical-sounding to me" (subject 1).

C. *Author name.*—The author's status plays a distinctive role in the evaluation process. A prominent scholar in the field tends to become an independent decision source regardless of subject matter. An individual's perception of the author influences the decision. Previous connection with an author seems to evoke the interest of the user. Conversely, a citation on the subject might be rejected if an individual's perception of the author is poor: "This one, 'Race and Crime: What Evidence Is There That Race Influences Results in the Criminal Justice System?' I got this because he (Norval Morris) is a very distinguished scholar and a very thoughtful man. This is not an empirical study at all, but I was sure it would be interesting, because everything he writes is interesting" (subject 2).

D. *Journal name and document type.*—The perceived quality of the publication influences the selection process and may become a sole source for negative or positive decisions. The user's knowledge and perception about the journal and its status in the field help formulate the decision. The accessibility of the journal also seems to influence the selection decision: "There are some I can rule out. In other words . . . we chose journals in philosophy, management, planning, and political science. We didn't choose journals in medicine or in economics or in history, religion. The reason [I didn't look at journals in religion] is because they would take a particular slant on this or a particular approach to it that I'm not interested in. I'm interested more in that in relation to political science or evaluation" (subject 10). "The next article is with the *Chronicle;* let's see what it's all about. I may do this too because we have here a congressperson involved with it so I'd like to see. This article is fairly interesting. If it wasn't in the *Chronicle,* I might not take this one at all. . . . For my field the *Chronicle of Higher Education* is the newspaper that describes the day-to-day goings on and so it's interesting for me to read, it's easily accessible, and there will be a lot of very good things in the *Chronicle*" (subject 9). As I will show later, the name of the journal also plays a supportive role in the decision process, through interconnections with other elements in a citation.

The user's discrimination as to types of documents indicates his or her perception about the value of the information. For example, availability, up-to-dateness, and the scope and the quantity as well as quality of information, as indicated by document type, influence the selection consideration: "A dated conference paper, if it contains research of interest to me, should eventually be published, and if it's not published, then it's not of interest to me. This is another conference paper, 'Meaningful Research and Program Evaluation in Correction,' that's even older (1977) and it's a very weird conference. It's a very small, backwater association. I'm not interested in that" (subject 2). Comments on ERIC documents also illustrate this point: "One thing that depresses me is that these are all 'ED's,' which are usually microfiche. And if it's long, like this one ninety-eight-page 'ED,' if this is not available, if I can't order this from the house wherever it's from, I don't think I'm going to use the microfiche unless it is fascinating or brand new" (subject 9). " 'The Time for Assertive Action: School Strategies for Promoting the Educational Success of At-Risk Children,' sounds very good, but it's a report of a commissioner's task force in New York on an ERIC document, and I doubt that it would be good. Getting the document might be difficult and also you know when you get something that specific, like someone's put out a document of what they've done, it might not be that useful in terms of being written for a wider audience. It may be very specific. I don't know. It would be a lower value of priority" (subject 6).

E. Abstract.—Individual users showed different perceptions and perspectives about abstracts. Abstracts play a supporting role in the interpretation of a citation if a citation otherwise has enough bearing on an individual's state of knowledge. However, if the elements of a citation do not serve to alert the researcher, the abstract plays a major role in the interpretation of a citation's content: "Some of the abstracts are more descriptive than others. From some abstracts, I can get a pretty good idea of what the study involves. Others are less clear, more vague. Here we have the title, the journal, the person, and the year. I don't think I even looked at the abstract here, because I knew immediately I wanted this one. Again, because I have information about this that helps me decide whether to use it or not. If it is an uninformative title, and an unusual journal and a person I don't know, then I'll read the abstract more carefully" (subject 2).

Users have different approaches in using an abstract in the selection of a citation. A doctoral student who is at the beginning stage of her research mentioned that "Key words in the abstract is the number one criterion. Then I look at the title of article and journal, and the author would be the fourth criterion" (subject 1). A graduate student who is

working on her master's thesis said, "I normally just go to the abstract. Not until after I see if I'm interested in the study will I look at the journal and the author. It leaves you more open to less-known authors, lesser-known journals for your information" (subject 3). Also, abstracts are used with other elements, discussed below.

F. Interconnections among elements in a citation.—The value of a citation tends to be enhanced by interconnections among several elements within a citation: author/title; author/title/journal name; title/journal name/publication date; title/journal name/author/publication date; abstract with other elements. Each element supports other elements to become decision sources. Different combinations of the elements of a citation seem to reflect the user's filtering process in terms of assessing the quality and value of a citation. The process also seems to stem from the user's balancing his or her perception about the literature in a field and the priority of the need for information. Hence, key words in the title tend to be reinforced or reevaluated in relation to other elements of a citation. The following illustration indicates various interconnections among elements in a citation: "I'm drawing on ideas from a lot of different perspectives so when I look through this I look at the title and I look at the author and I look at the kind of publication, and from my knowledge of those three things, already I can pretty much say that's relevant to what I want or that's not. 'English Social Criticism in the Spirit of Reformation': I just know first of all that the journal *Clio* is a history journal and I know that this is going to be a historic focus on the Reformation. It's going to talk about social criticism but it's being talked about in a time and in a way that I'm not interested in. I can just tell from the title and the journal" (subject 10).

A researcher's comments on the publication date may reflect the individual's priorities and awareness of the literature in the field. The importance of currency was also mentioned: "I'm familiar with some of the earlier literature. Also, since my study focuses on drug prosecution, I'm less interested in earlier studies. When you have a recent article, and it's a good article, it's likely to have a better bibliography and more references that I can again consult" (subject 2). In contrast, a researcher in anthropology mentioned that "I always check the topic and journal. The year, I'm not that interested in" (subject 7).

Internal (Experience) Context
The internal context indicates various sources deeply rooted in an individual's previous experience with literature in the field and perceptions (or beliefs) about the problem area. These include level of expertise in a subject literature, awareness of published literature, previous research

experience, and educational background. The user's knowledge of the research area and of appropriate search terms influences the decision process and justifies selection or rejection decisions.

A. *User's previous experience and perceptions.*—The following instances show that an individual researcher's perceptions and knowledge about journals, as well as about authors and their previous works and affiliated academic programs or institutions, can influence a decision. When the user's perception about the journal is marginal, previous association with the author or the author's affiliated academic program tends to support a decision: "This one is interesting, the article titled 'Impact of a Restricted Natural Language Interface on Ease of Learning and Productivity.' And this one was, I think, it's published in *Communications of the ACM,* which it's hard to tell whether it's research or not—because sometimes they publish some research studies, but not all the time. But I guess there are four authors and almost all of them are, yeah, they are from Houston. They are faculty members at the Rice University in Houston, Texas. And I got some citations, some of the articles, or some dissertations that were done at Houston about user interface, so they might be highly relevant" (subject 4).

The user's assumptions or beliefs about the information problem area also tend to influence his or her interpretation of the citation and selection attitude: " 'The Library Assistant: High Level Paraprofessional.' I think that an awful lot of our clerical routines, repetitive things, are no longer even done in academic libraries by staff people. I think that we've turned much of that over to student assistants just as we have turned over much of what we used to do in the past; much of what librarians used to do in the past has been turned over to what were formerly clerical people, but now they're expected to operate at a paraprofessional level and make some decisions that we would have made historically. That's why I want to look at this article, because it's high-level paraprofessional in reference to the library assistant and I think that is true, especially as we add more and more technology to libraries. I want to see if that study refers to a change in the level of our expectations of staff in libraries in general. Underneath all of this, you see, I have a belief that people who are paraprofessional in libraries probably need different reward systems than people who are doing clerical jobs" (subject 5).

B. *User's level of expertise in the problem area.*—An individual researcher's interpretation of a citation in terms of its subject matter seems to be rooted in her knowledge of the problem area: "There are a lot of different things going on here in evaluating this [citation]. So I'm sort of

drawing on previous knowledge of the field and what I know about those things. In some cases I know them really well, in the cases of the regular education initiative. Some areas, like dealing with model programs for the developmentally disabled, I have some knowledge in that area, but it's not as extensive. So I'm not aware of as many of the authors and articles in that area" (subject 6).

The researcher's awareness of the published literature in the problem area also affects the decision process: "Corporate source from Sidney, Australia. Conference title: 'Information Online '88.' Interestingly enough, the people in Australia and England seem to be much more current with the identification of information skills than we are here in the United States. So that's a tip-off to me. There have been more dissertations done of this nature in England. I've read a couple of things out of Australia that talk more about this than the literature here in America" (subject 1).

C. User's previous research experience.—The researcher's previous experience in research and research methodology in the problem area contributes to his or her interpretations about the literature and impacts on the decision: " 'Correlates of Institutional Misconduct among State Prisoners: A Research Note.' 'Institutional misconduct,' that is, these are people who are in prison and they get into trouble while they are in prison. I'm not interested in that at all. Again, this word 'Correlates' is a key word for me. Studies that have correlates in the title are often atheoretical, that is, exploratory studies which are often very limited. This is not helpful to me. Whenever I see 'correlates' I become suspicious" (subject 2).

D. User's education (or training).—The individual's paradigm preferences in the published literature differ as a result of educational background. For example, my data from subjects in political science, sociology, and anthropology indicate that there is a tendency by researchers in those fields to avoid psychological interpretations: "I'm not going to consult this one. The reason is that this is a psychology journal and psychologists have a different kind of approach to doing research that is not helpful to me. . . . I do not think it is likely that I could learn anything from this. They are more interested in individuals. I'm trained as a political scientist. I'm more interested in patterns, and institutions, and the way that individuals relate to institutions. My interpretation might be different from others, based on disciplines, and also the institutions where they were trained" (subject 2). "This would be interesting: 'Visitational Dreams among Moroccan Jews in Israel,' and also for contrast. It's going to be a study of dreams in another culture. The problem is that it's

in the journal *Psychiatry* and anthropologists tend to avoid psychiatric interpretations. Because we do different kinds of interpretation. So I may read this, but I wouldn't read it first; I would read it later" (subject 7).

External (Search) Context
This category indicates factors that stem from an individual's search and current research. Those variables are not directly related to the user's problem area but tend to originate from the individual's view about the search goal, search process, research stage, or research product.

A. Perception of the search quality.—An individual's perception of the search quality seems to affect his or her attitude toward selecting a citation. When the user perceives a search to be incomplete, she tends to select key citations that are very recent and to rely on them to lead her in other directions to solve information problems: "This search is incomplete, because I'm aware of some studies that should have come up and that did not. Therefore, I don't view this as an exhaustive search; and I'll view this as something that will lead me in directions I don't know about. For that reason, it's most important for me to get recent citations, because they will again offer me other directions" (subject 2).

B. Purpose of search (or search goal).—The search goal and the user's perception of the search process affect the individual's behavior in making decisions. Some decisions are selective; others are comprehensive: "Just the first time through this and I probably have about fifty to sixty references that I think are probably useful. On a second pass through maybe I'd find another ten or fifteen. But the first time through I'm more selective" (subject 10).

C. Perception about the availability of information.—The perception about the information available in the problem area seems to influence the decisions: "I know anthropologists have done very little work on dreams in the Middle Ages in Europe. As an anthropologist the first thing I have to do is go through the literature in my geographic area; but there really isn't any on this topic. So that's why I'm interested in other areas of the world" (subject 7).

D. Priority of information needs.—An individual's priorities in terms of information needs and uses seem to influence the decision process. The degree of importance of information and the scarcity of information in relation to the current need tend to play roles in the situation. The need situation takes priority over the quality of information: "I was looking

for the ones that I was having difficulty finding things on; so that 'resilient individual,' and the 'life history approach'—I was just real happy to get little tidbits on that, because I had almost nothing; and they're good leads. This life history is really where I need more information right now so I was eager to get into these first. She [the searcher] was concerned that she didn't get a lot of apt citations, but I was real pleased with what she had gotten" (subject 3).

E. Stage of research.—The research stage and the degree of focus of the individual's problem area seem to result in differences in approaching the selection and interpretation of a citation: "My area of interest is not really specified at this point. So I try to get as much as I think I can, and try to read, and try to come up with any arguments whether I should add or I should delete this from my literature search. So, I'm sort of, like, fishing whether this has some potential, or high potential, to be part of the things that I need to read in order to make decisions whether I should include, or I should believe" (subject 4).

F. End product of the research.—An individual may show a different approach in selecting a citation because of her anticipated end product of research. For example, there are differences in time allocation or interest in the timeliness of the material based on whether an individual's research product is a dissertation or a report for a program sponsor: "I'm doing two things. One, I'm writing a report for the program sponsor, and that's to be fairly specific. He is not interested in the literature. But I'm also preparing an article and I'm sort of in the midpoint here. I don't want to take too much of my time reviewing the literature if it's not going to help me to prepare the report; but I also want to learn the literature. Multiple goals. So after I get into this more deeply, I may go back and try to get more of these studies" (subject 2).

Problem (Content) Context
The decision process for selection of a citation depends on the degree of understanding of the information problem the user faces. This context involves learning about one's problem, constructing various facets of problem formulation, expanding one's thinking, obtaining new ideas, becoming familiar with key concepts, defining major concepts, examining research methods, and gaining a broader picture of a problem area. The researchers' comments reflect the existence of problem contexts: "I might go back to the marginal ones. Depends on whether I feel I have a good understanding not only of the literature but of the problem. If I'm reading the literature, if I'm having trouble solving something, or putting things into place, . . . then I may go back and try to find

more things to resolve the anomalous findings" (subject 2). "I'm going to make connections between things like issues in moral philosophy or moral discourse, issues in policy analysis and program evaluation, and issues in what might also be called the narrative tradition that has been covered in both ethnography and social sciences in general. Those three things I'm trying to fit together in terms of constructing a new view of evaluation in policy analysis" (subject 10).

The following illustrations show various motivations underlying the intended uses of a citation. Instances are categorized into "same (similar) problem context" and "different problem context" to reflect how users construct or shape their information problems. The attributes of information itself, for example, "old, new or insufficient information," are also observed in the problem context.

A. Same (similar) problem, for definitions.—The user learns his or her information problem by interpreting and understanding a citation in relation to the problem area. If a citation implies discussion of the major concepts or topics in the problem area, it shows high potential of use. A user searches for information in order to get help in defining and crystallizing the key concepts in the information problem: "This one is not an empirical study, but it's interesting because they talk about the issues of functionality and usability. I think, in terms of libraries, we have talked a lot about how to design, or how to build a user-friendly system. We have talked about 'user-friendliness'; but if you want to pin down what constitutes the concept of user-friendliness, we were quite vague. Nancy Goodwin, the lady who wrote this article, tried to capture the concept of 'user-friendly' and tried to actually throw away that term. Because it doesn't really identify clearly what constitutes user-friendliness. So what she came up with, she came up with the terms 'functionality,' and 'usability.' . . . So I really think that's very helpful for me, at least, in terms of crystallizing my concept toward user-friendliness" (subject 4).

B. Same (similar) problem, as background.—Researchers are constantly seeking major issues or approaches that are addressed by or related to their direct problem area in order to have a complete picture of their problem area: "I am interested in this one because they're talking about adults from child abuse and then some of the effects here. And it would just give me more background information. And then they do say that as adults they were more inclined to be a rape victim or in a battering relationship. So this would give me more information about what would be behind the subjects' personality here, or possible problems" (subject 3).

C. Same (similar) problem, for the methodology.—" 'Library Technical Assis-
tants: A Survey of Training Programs and Employment in Selected
Libraries.' Again, I think that 'technical assistants' is another term for
'paraprofessional,' or it can be in some libraries. The survey methodol-
ogy which I'm giving very serious consideration to anyway, so I wanted
to look at that article to see if it's related" (subject 5).

D. Similar problem, off the target.—The problem's focus or boundary is
also an important criterion for acceptance or rejection: " 'Causality and
Verbal Forms—Reflections on the Linguistic Structure of Explanation.'
Oh well, it sounds very interesting but I'm not really interested in lin-
guistic structure of explanation. And you know semiotics is a little differ-
ent than sociology—they study symbols and that means meaning—but
it's different—it's more like literary analysis" (subject 8).

E. Different problem, for the methodology.—Researchers show high interest
in using information from different problem areas or fields to learn
more about their problems and to cast a different light on their own
problem area. The following instances show the user's interest in other
problem areas in terms of their methodology, framework, process, and
the problem boundaries: "What I found myself most interested in when
looking through here is the sort of things that went a little bit far afield.
I would have liked to have a few more of those. These kinds of things
that got a little bit far afield were providing some input from outside
that might help us to cast a different light on some of this stuff" (sub-
ject 6).

F. Different problem, for the framework.—The user may also be seeking and
relying on a technique or a framework in a different problem area to
illustrate or reinforce his or her own ability to explain: " 'Reading and
Writing with Computers: A Framework for Explaining Differences in
Performance.' Reading and writing with computers are probably not
[what] I'm interested in. But I'm interested in a framework for ex-
plaining differences in performance. Because I think my study will re-
ally deal with more than one group of users; at least novice searchers
and non-novice (I call them at this point 'non-novice'), or those who
have more experience than first-time users. So I need some work of,
like, framework, or some sort of techniques that help me identify how
I can explain differences in performances" (subject 4).

G. Different problem, as an analogy.—This category involves the interpre-
tation and use of a citation derived from the user's looking at a citation
as a parallel application for solving the problem at hand. Researchers

try to find a relationship or an analogy between their own problem and other problem areas to shed light on ideas in their own field: " 'Synergistic Action Research in Large Systems Change: The Ideal, the Plan, and the Reality.' It's from another area, organizational development, that I don't know anything about, but they're talking about large systems change and they say 'the ideal, the plan, the reality.' This synergistic action research is what fascinates me because I have no idea what they're talking about, but I hear some other things like large system change that sound exactly like what I'm looking for. So, basically what this is telling me is that here's a possible model that I could draw upon from another discipline to shed some light on some of the things we've pulled out of our own field here" (subject 6).

H. Different problem, as background.—The user may rely on a different problem area to get ideas and give a direction to his or her own problem or to put the background information in the problem area in a broader context: " 'Non-Academic Employees in Higher Education: A Historical Overview' is not a library thing; it doesn't have 'library' in the title, and it was again a presented paper, but it would give me the historical background for job classification on university campuses. And that's an umbrella under which the library always has to fit. So there are some overriding principles there that I need to be aware of. And the rules that are established within the classification of university employees will necessarily apply to classification of library employees and job assessment and reward systems and performance appraisal and so forth" (subject 5).

I. Different problem, not of interest.—The following instances show citations rejected because they are not of interest to the user's information need and context: "I don't know what 'informatics' means, and this is a medical paper and I don't know much about medicine. I'm reading the abstract right here. It's all about medicine . . . I don't think I would use this one at all. They're all—nothing in it looks like information skills" (subject 1).

J. New information, in the problem context.—The newly received information can stimulate the user's internal states of information need by facilitating new ideas or new directions: "This is for a population like mine, it could be, because they're talking about the adults, and this astounded me: 'the subjects entered therapy on an average of seventeen years after the abuse terminated,' almost two decades after the abuse stopped. And just that astounded me; that fact alone; I mean, that's a long time to go without treatment, and I'm like, why? why? why?" (subject 3).

K. Old (that is, repetitive) information, in the problem context.—Repetitive information tends to decrease the value and usefulness of a citation: "A lot of the articles that came up are stuff that we already have. That was, if anything, maybe the main problem, that maybe what we should have done was told her to go farther afield, so that we wouldn't end up with so many articles we already had. In a sense, that was good in terms of the validity of the search because of the fact that she found articles that we were already using and she was on the right track" (subject 6).

L. Insufficient information, in the problem context.—Some of the citations were too vague in providing the information necessary to make a con-crete decision. Certain words in the title are not meaningful enough to inform one about the contents of a citation. In addition, the less well defined problem context tends to contribute to vagueness in interpret-ing and understanding the citation: "It's pretty vague from the abstract. I'm thinking that probably it won't be any help. But if I got really des-perate for information, I probably would go and see what it said, simply because it uses the phrase 'skill set' and I'm looking for a set of skills as well. But the whole abstract is just pretty abstract to me and it doesn't make a whole lot of sense" (subject 1).

Discussion

Figure 1 presents a model that summarizes the factors and thought processes found in this study contributing to selection of a citation pro-duced by a document retrieval system. The model presents the catego-ries and contexts that are present in this process and that describe how users make relevance interpretations. The three major categories of user-based characteristics are internal, external, and problem contexts. Several variables that are citation-based are also illustrated.

   The categories of internal, external, and problem context support a user-based relevance and a cognitive approach to IR research. The model demonstrates what the attributes of the information-seeking con-text are and how relevance interpretations are affected. The internal context category presents those perceptions that are linked at the cita-tion level. For example, an individual's experience with and perceptions of authors, academic programs, journals, and his or her own beliefs about the information problem reflect a user's conceptual state of knowl-edge and image of the world. An individual's level of expertise, previous research experience, and educational background can also be regarded as a part of a private knowledge structure.

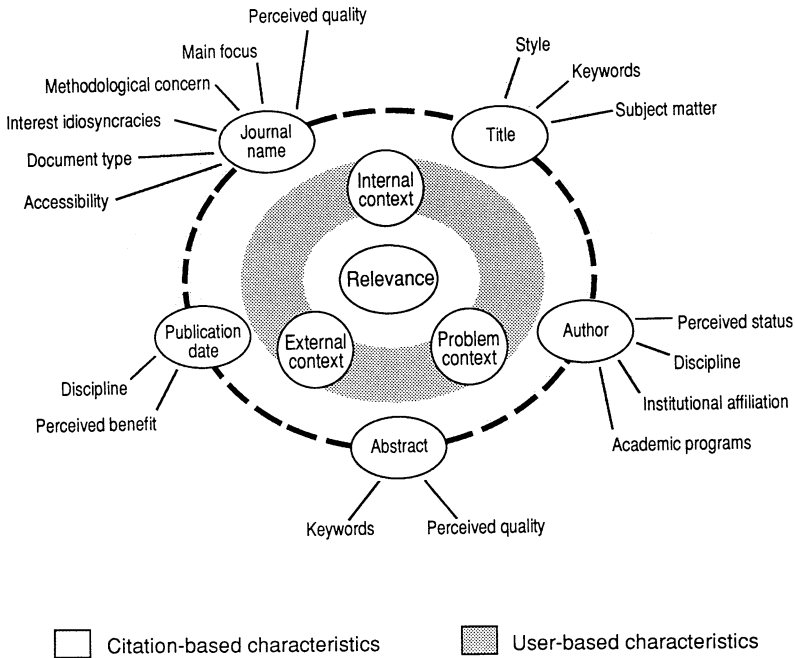   The external context category describes the contexts in which an indi-

FIG. 1.—Factors contributing to user-based relevance

vidual's research and search for information are situated. They provide an explanation of how context can be influenced by an individual's particular situation. Perceptions about the search in terms of the search quality, search goal, and availability of information can affect an individual's selection behavior.

The desired outcome of the research influences the search goal. Also, the possible end product of the research will affect aspects of information-seeking behavior including the individual's time schedule and the timeliness of the information. The priority of information needs in terms of one's preference for certain kinds of information sources depends on the importance of the information, scarcity, availability of information, and the focus of the problem. Those aspects support the idea of situational relevance proposed by Patrick Wilson [9]. The external context is not visible, but its conditions implicitly affect the decisions made at the citation level.

The problem context category illustrates why and how the user employs information to construct and solve the information problem. Users constantly seek leads, ideas, and issues in order to understand the prob-

lem at hand. They rely on citations not only in the same problem area but also in related areas or even different disciplines to expand their ideas, ways of thinking, and what they do. The original problem can be reformulated and reshaped by bringing new insights and changing thoughts on research approaches and methods during this process. The category of problem context describes an information problem as evolving. An instance of the problem context reflects this aspect: "My problem has definitely changed, because when I started I was going to look only at paraprofessionals. But I'm going to have to go back and do another search now, because of the change in my thought on what my methodology may be. If I survey, you know, if I throw out this paraprofessional in terms of me defining it and then try to make people fit within that and instead turn that over to Delphi technique and let somebody else define it, then I'm going to have to survey both paraprofessional and clerical staff. So, this has affected how I'm going to have to even define my topic. Just looking over the titles has affected that" (subject 1). This aspect seems to be consonant with the views of William B. Rouse and Sandra H. Rouse [43]. These authors viewed "information seeking [as] part of the process of decision making, problem solving, or resource allocation" [43, p. 129].

Compared to Atkinson's model of the book selection process of librarians, the two models share some common characteristics. The selection decisions belong to an individual's cognitive, private activities that do not allow precise observation or delineation. The main difference is derived from the nature of contexts between the two models. Contingencies of contexts vary; the librarians are situated differently or susceptible to other environmental factors than are end users. For example, they have to consider such factors as the type of an institution, collection development policies, needs of multiple users, and budget. However, the contexts of relevance from an end user's perspective mainly originate from an individual's psychological, situational nature.

Citation-based characteristics reflect users' thoughts and interpretations at the citation level. The model provides a possible explanation of why certain elements of a citation appear to receive more attention than others. For example, a prominent author's name in a field becomes an essential indicator: an individual's perceived status and previous experience with an author can play a decisive role in the selection of a citation regardless of the subject of a document. In addition, an individual's perceptions of a journal are important and include the following aspects: the status of the journal, subjects treated by the journal, main focus of the journal, and the journal's methodological concerns, accessibility, and idiosyncracies, such as biases toward a certain subject area. Users also have certain perceptions toward types of serial publications

(for example, journals, conference papers, newspaper articles, micro-fiche). These interconnections among various elements of a citation seem to be supported by the reported characteristics of academic social scientists' information-seeking behavior. In Ellis's [44] study, the stage of "differentiating" referred to users' filtering activities on information sources or types of sources.

A citation is evaluated in relation to the user's interest and predetermined sets of evaluation criteria. Each attribute of a citation is interpreted within an individual's particular context and motives. Data from this research provides evidence that the capability of recognizing the attributes of a document originates from the user's ability to capture the "sensitizing concepts" (a subject's words) in a problem area. However, conceptual terms may be very weak and not well defined even if an individual has a high degree of familiarity with a field. Relevance assessments at the citation level are seldom based on a single piece of information. A user's perspective on a citation is supported by layers of contextual factors, and decisions are made based on an individual's own perspective, interwoven into various contexts. Relevance is not a simple relationship between a document retrieved and a user's question but, rather, is psychological and contextual, involving an individual's cognitive states, perceptions, experiences, and knowledge about the problem at hand. It goes much deeper than simple topical relevance.

The empirical model of user-based relevance reported here supports a subjective, personal meaning of relevance: Kemp's pertinence [5], Foskett's private knowledge [6], Cooper's utility [7], Lancaster's novelty [8], Patrick Wilson's situational relevance [9], and Harter's psychological relevance [12]. My empirical findings are compatible with and supported by an interpretation of psychological relevance in particular. Reading a new citation acts as a stimulus, possibly causing an individual to create a new context and a set of cognitive changes to occur in that context. A citation is found relevant if the citation and the context interact together to produce new ideas. Psychological relevance can serve as a theoretical framework that explains the empirical findings of the present study.

Other aspects of user-based relevance demonstrate its nature as multidimensional and dynamic, as rooted in an individual's particular situation that involves one's state of knowledge, task, goals, and environmental conditions. As the problem context of this model indicated, a user's information need state may be changed as he or she encounters relevant citations. As a result of experiencing new relevant citations, an individual's approach to the problem area, including the ideas of defining the problem itself, definitions of the major concepts, and methodologies, is changed or modified. These findings are consistent with other recent

work in this area. For example, Carol C. Kuhlthau's [45] study also found that cognitive and psychological changes occur in individuals during the library search process. Kuhlthau's model describes affective as well as cognitive changes in six stages: task initiation, topic selection, prefocus exploration, focus formulation, information collection, and search closure. Jeffrey Katzer and Herbert Snyder [46] pointed out that IR evaluation should take into account the aspects of a user's increased understanding about her information problem and the effects of those changes on search processes.

Conclusions and Future Research

The empirical model of user-based relevance developed in this article demonstrates that "relevance" is interpretative; it is based on an individual's interpretations, which stem from his or her own experiences, perceptions, and knowledge in the information problem area. Relevance assessments are complex phenomena and cannot be represented as a static and precise relationship between documents and a user's question.

Although previous experimental research on relevance judgments was involved with subjects other than real users, my model supports and confirms some of the experimental findings by Rees and Schultz [15] and Cuadra and Katter [13, 14] in particular. For example, an individual's subject knowledge, professional training, and educational background were found to be influential factors in user-based relevance as well. Other variables such as an individual's anticipated usage of documents, research stage, and document styles were influential also.

However, some of the variables named in the internal context of my model show the individualistic nature of user-based relevance. These include individuals' previous perceptions of authors, institutional affiliations, journals, and academic programs, as well as individuals' previous research experience. New factors were also discovered, in the external context: perceptions about search quality, search goal, and the anticipated end product of the research. Users' search goals have been considered as potential variables in online searching; however, they have not been previously connected to the concept of relevance in IR. Variables related to the information itself, such as scarcity, availability, timeliness, and scope, are found to be affecting factors. Variables in the problem context demonstrate that the information in other fields is used for an analogy, framework, or methodology in the creation and formulation of an individual's information problem.

The model shows that relevance is not fixed but is a temporal and fluid concept that is sensed or observed at a specific moment for a partic-

ular need. Those views are also consistent with a dynamic and situational approach to understanding users' criteria of relevance: Schamber and her colleagues state that "relevance judgments are users' evaluations of information . . . in relation to their information need situations at particular points in time" [3, p. 771]. As Thomas J. Froehlich wrote, "The criteria for relevance are heterogeneous and different criteria, or sets of criteria, can be applied successively or dynamically. . . . Relevance judgements . . . can be based on vague criteria, are polythetic in character, and are based in part on the environment, task, and background of the judgement maker" [47, p. 124].

The empirical findings of this research support the views of Ellis [1] and Meadow [21] that relevance is not a variable that allows precise measurement. This issue of measurability deserves further attention. Nearly all writers have taken the position that relevance *is* measurable, even including recent authors who are concerned with relevance from users' viewpoints. Although this study is involved with a specific moment in an individual's information-seeking process, it seems that it would be possible to trace an individual's view of relevance and its changing patterns during the whole process of one's information seeking and use. For example, data concerning a user's relevance judgments at each stage of research could be collected and recorded using a qualitative research methodology. Relevance interpretations gained at the various stages of a research project could be compared to each other as a way of analyzing an individual's historical data regarding the relevance-judgment process.

The relationship between a user and a document is not a simple one but, rather, interdependent on other variable contexts. The characteristics of user-based psychological relevance presented in my model have implications for IR research and practice. The model provides a theoretical insight into understanding the misinterpretation of the treatment of relevance judgments in particular and design flaws in IR experimental studies in general. These problems derive from the nature of relevance itself. The empirical understanding of relevance gained from analysis of the users' perspectives obtained in this study implies the conceptual limitations of relying on relevance in IR testing and guides us to think critically before we accept loosely and vaguely the usage of the term "relevance" in information retrieval. The individual and interpretative nature of user-based relevance also demonstrates a serious problem in the interpretation of relevance judgments that are made by others (for example, subject experts or search intermediaries) than the actual users themselves.

However, the findings of this research should be interpreted with caution. The conclusions are based on a small number of subjects, a specific group of faculty and students from the social sciences. The

derived model may not describe all users. Further understanding of the information retrieval process may be developed fruitfully by concentrating our efforts on users of information systems. Most IR research has focused on components of information systems or on the intermediary, disregarding the complex information behavior of end users. More research in the IR process, such as user's self-evaluation, including the activities of constructing one's information problem and its impact on searches, is needed to understand how and why a user behaves in certain ways in the information-seeking process. The use of alternative research inquiry methodologies in IR research should be encouraged, to understand users' contexts and their meanings in different situations and to develop explanatory concepts in the information-seeking process. The variables identified in this article should also further efforts toward system and interface design that considers the actions and thought processes of users.

Appendix

Subjects and Their Stated Information Problems

Subject 1, Doctoral Student
My research area concerning this search is in identifying information skills and ranking of importance of those skills of students who are graduating from high schools. My research involves a Delphi study to identify information skills and to rank the importance of them using experts in the field. The results of my research will be a dissertation. I'm just starting my literature search. I have a beginning knowledge of information skills literature right now.

Subject 2, Faculty Member
My major area of research is in the criminal justice area. The research concerning this search is to examine how drug crime cases are produced or handled in the courts. I am interested in two research questions. One is whether cases that go to trial in an American court receive different sentences in cases where the defendant agrees to plead guilty. The second question is whether there are any racial differences, that is, the sentences that defendants get, whether nonwhite defendants get different sentences than white defendants get. The result of the research will be a report to the project sponsor and an article for a research journal. I'm familiar with the literature on criminal court processing and plea bargaining in general.

Subject 3, Graduate Student
I'm studying the recovery process of adults who were abused as children. I did my proposal; I have been doing a literature search and

reading and now I'm ready to start interviewing the subjects. I was having some problems because it is difficult to find information on this area of research. The literature seems to focus on the abuse, the history of the abuse, and on the negative effects. There is very little about what are the things that help a person recover. I'll be using a life history method and I have very little leads on that. I'm not an expert, but I feel I know the literature well enough. The literature I found is multidisciplinary: sociology, psychology, medical fields, psychiatry, and I've been getting literature in child development. The result of the research will be my master's thesis.

Subject 4, Doctoral Student

My research interest is on the user interface. I'm very interested in any empirical studies that try to test hypotheses about the user interface. The search is for the literature review for the proposal for my dissertation. I did a study of the user interface of ERIC on CD-ROM and published an article on that. It was a lengthy process, and this kind of thing was very difficult to identify in a lot of databases because it is very interdisciplinary. I don't know at this point how many topics I should include and how many I should exclude. I don't think it's a very straightforward kind of thing; it will involve hunting around.

Subject 5, Doctoral Student

I am looking for research done in the area of paraprofessionals in academic libraries, and the use of performance appraisal and reward systems with that group. I was doing preliminary searching for a possible dissertation topic. I'm in a very beginning stage. I'm not particularly familiar with the subject literature in this field. The studying that I have done about performance appraisal and reward systems has been through the School of Business. I don't know what has been done in terms of clerical staff and paraprofessionals in our field.

Subject 6, Faculty Member

My general area of research is school psychology, intervention and consultation in school psychology, working with students with behavior problems, within special education or in regular education. Of late, I have been working with integrating children with disabilities into the mainstream, regular classrooms. The current research problem concerning this information search is trying to identify atheoretical models or implicit theoretical models and systems change in education. The purpose of the search was to find articles that either gave us or had some clear ideas about what a model would be for systems change in education or discussed some attempt at broader system reform in an educational setting, especially dealing with children who have been served in special education, but not necessarily limited to that. The result of this specific project will be a literature review article.

Subject 7, Faculty Member
I am doing research on visions, religious visions of medieval European women, which is not an area that anthropologists have written a lot about. It's a new topic for me; I've read very little. I was surprised; there weren't many references on dreams, which is surprising because I thought anthropologists had done a lot of work on dreams, dreams in other cultures.

Subject 8, Doctoral Student
I'm working on my dissertation and the specialty area is social psychology. I'm developing a theoretical model of meaning in social interaction—the organization of meaning in social interaction. The purpose of searching was to have a literature review that covers as much ground as possible on the concept of meaning in the social sciences. My problem area involves sociology, psychology, and linguistics.

Subject 9, Doctoral Student
I'm working on the higher education Reauthorization Act and I'm interested in finding out information about the Act and other information related to financial aid and higher education. I am at the beginning stage of my dissertation, just beginning some preliminary research. I know a fair amount but not very much in-depth knowledge.

Subject 10, Faculty Member
I have a broad area that I am working with: the influence of moral inquiry in social sciences and particularly on policy analysis and program evaluation. The idea was to cast a pretty broad net because I was not ready to narrow in yet. I already have some familiarity with the various literatures—in both the philosophy and political science literature, such as moral philosophy and ethics. It is a very interdisciplinary kind of project, and is not well defined. I am going to write some papers for journals but I'm also aiming at a book. It is in progress. I have already written several papers that deal with this topic.

REFERENCES

1. Ellis, David. "Theory and Explanation in Information Retrieval Research." *Journal of Information Science* 8 (1984): 25–38.
2. Saracevic, Tefko. "Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science." *Journal of the American Society for Information Science* 26 (September/October 1975): 321–43.
3. Schamber, Linda; Eisenberg, Michael B.; and Nilan, Michael S. "A Re-examination of Relevance: Toward a Dynamic, Situational Definition." *Information Processing and Management* 26 (1990): 755–76.
4. Eisenberg, Michael, and Schamber, Linda. "Relevance: The Search for a Definition." *ASIS Mid-year Proceedings* (1988): 164–68.

5. Kemp, D. A. "Relevance, Pertinence and Information System Development." *Information Storage and Retrieval* 10 (February 1974): 37–47.

6. Foskett, D. J. "A Note on the Concept of 'Relevance.' " *Information Storage and Retrieval* 8 (April 1972): 77–78.

7. Cooper, W. S. "A Definition of Relevance for Information Retrieval." *Information Storage and Retrieval* 7 (June 1971): 19–37.

8. Lancaster, F. W., and Fayen, E. G. *Information Retrieval On-line.* Los Angeles: Melville, 1973.

9. Wilson, Patrick. "Situational Relevance." *Information Storage and Retrieval* 9 (August 1973): 457–71.

10. Swanson, Don R. "Subjective versus Objective Relevance in Bibliographic Retrieval Systems." *Library Quarterly* 56 (October 1986): 389–98.

11. Sperber, Dan, and Wilson, Deirdre. *Relevance: Communication and Cognition.* Cambridge, Mass.: Harvard University Press, 1986.

12. Harter, Stephen P. "Psychological Relevance and Information Science." *Journal of the American Society for Information Science* 43 (October 1992): 602–15.

13. Cuadra, Carlos A., and Katter, Robert V. *Experimental Studies of Relevance Judgments: Final Report.* Vol. 2, *Description of Individual Studies.* Santa Monica, Calif.: System Development Corp., 1967.

14. Cuadra, Carlos A., and Katter, Robert V. "Opening the Black Box of 'Relevance.' " *Journal of Documentation* 23 (December 1967): 291–303.

15. Rees, Alan M., and Schultz, Douglas G. *A Field Experimental Approach to the Study of Relevance Assessments in Relation to Document Searching.* Vol. 1, *Final Report.* Cleveland: Case Western Reserve University, 1966.

16. Saracevic, Tefko. "Ten Years of Relevance Experimentation: A Summary and Synthesis of Conclusions." *Proceedings of the American Society for Information Science* 7 (1970): 33–36.

17. Eisenberg, Michael B. "Measuring Relevance Judgments." *Information Processing and Management* 24 (1988): 373–89.

18. Eisenberg, Michael, and Barry, Carol. "Order Effects: A Study of the Possible Influence of Presentation Order on User Judgments of Document Relevance." *Journal of the American Society for Information Science* 39 (September 1988): 293–300.

19. Purgailis Parker, Lorraine M., and Johnson, Robert E. "Does Order of Presentation Affect Users' Judgment of Documents?" *Journal of the American Society for Information Science* 41 (October 1990): 493–94.

20. Janes, Joseph W., and McKinney, Renée. "Relevance Judgments of Actual Users and Secondary Judges: A Comparative Study." *Library Quarterly* 62 (April 1992): 150–68.

21. Meadow, Charles T. "Relevance?" *Journal of the American Society for Information Science* 36 (September 1985): 354–55.

22. Cleverdon, Cyril; Mills, J.; and Keen, M. *ASLIB Cranfield Research Project: Factors Determining the Performance of Indexing Systems.* 2 vols. Cranfield, Bedfordshire: College of Aeronautics, 1966.

23. Swanson, Don R. "Some Unexplained Aspects of the Cranfield Tests of Indexing Performance Factors." *Library Quarterly* 41 (July 1971): 223–28.

24. Harter, Stephen P. "The Cranfield II Relevance Assessments: A Critical Evaluation." *Library Quarterly* 41 (July 1971): 229–43.

25. Fenichel, Carol H. "Online Information Retrieval: Identification of Measures That Discriminate among Users with Different Levels of Types of Experience." Ph.D. dissertation, Drexel University, 1979.

26. Bellardo, Trudi. "Some Attributes of Online Search Intermediaries That Relate to Search Outcome." Ph.D. dissertation, Drexel University, 1984.

27. Swanson, Don R. "Historical Note: Information Retrieval and the Future of an Illu-
    sion." *Journal of the American Society for Information Science* 39 (March 1988): 92–98.
28. Doyle, Lauren B. "Is Relevance an Adequate Criterion in Retrieval System Evalua-
    tion?" In *Automation and Scientific Communication: Short Papers Contributed to the Theme
    Sessions of the 26th Annual Meeting of the American Documentation Institute*, edited by
    H. P. Luhn. Washington, D.C.: American Documentation Institute, 1963.
29. Taylor, Robert S. "Question-Negotiation and Information Seeking in Libraries." *Col-
    lege and Research Libraries* 29 (May 1968): 178–94.
30. Belkin, Nicholas J. "Anomalous States of Knowledge as a Basis for Information Re-
    trieval." *Canadian Journal of Information Science* 5 (1980): 133–43.
31. Ingwersen, P. "Psychological Aspects of Information Retrieval." *Social Science Informa-
    tion Studies* 4 (1984): 83–95.
32. MacMullin, Susan E., and Taylor, Robert S. "Problem Dimensions and Information
    Traits." *Information Society* 3 (1984): 91–111.
33. Saracevic, Tefko; Kantor, P.; Chamis, A. Y.; and Trivison, D. "A Study of Information
    Seeking and Retrieving. 1. Background and Methodology." *Journal of the American
    Society for Information Science* 39 (May 1988): 161–76.
34. Belkin, N. J.; Chang, Shan-Ju; Downs, Trudy; Saracevic, Tefko; and Zhao, Shuyuan.
    "Taking Account of User Tasks, Goals and Behavior for the Design of Online Public
    Access Catalogs." *Proceedings of the ASIS Annual Meeting* 27 (1990): 69–79.
35. Su, Louise T., and Saracevic, Tefko. "Modelling and Measuring User-Intermediary-
    Computer Interaction in Online Searching: Design of a Study." *Proceedings of the ASIS
    Annual Meeting* 26 (1989): 75–80.
36. Mellon, Constance A. *Naturalistic Inquiry for Library Science: Methods and Applications for
    Research, Evaluation and Teaching.* New York: Greenwood, 1990.
37. Dervin, Brenda, and Nilan, Michael. "Information Needs and Uses." *Annual Review
    of Information Science and Technology* 21 (1986): 3–33.
38. Merriam, Sharan B. *Case Study Research in Education.* San Francisco: Jossey-Bass, 1988.
39. Goetz, Judith Preissle, and LeCompte, Margaret Diane. *Ethnography and Qualitative
    Design in Educational Research.* Orlando, Fla.: Academic Press, 1984.
40. Lincoln, Yvonna S., and Guba, Egon G. *Naturalistic Inquiry.* Beverly Hills, Calif.: Sage,
    1984.
41. Atkinson, Ross. "The Citation as Intertext: Toward a Theory of the Selection Pro-
    cess." *Library Resources and Technical Services* (April/June 1984): 109–19.
42. Park, Taemin Kim. "The Nature of Relevance in Information Retrieval: An Empirical
    Study." Ph.D. dissertation, Indiana University, Bloomington, 1992.
43. Rouse, William B., and Rouse, Sandra H. "Human Information Seeking and Design
    of Information Systems." *Information Processing and Management* 20 (1984): 129–38.
44. Ellis, David. "A Behavioral Approach to Information Retrieval System Design." *Jour-
    nal of Documentation* 45 (September 1989): 171–212.
45. Kuhlthau, Carol C. "Developing a Model of the Library Search Process: Cognitive
    and Affective Aspects." *RQ* 28 (Winter 1988): 232–42.
46. Katzer, Jeffrey, and Snyder, Herbert. "Toward a More Realistic Assessment of Infor-
    mation Retrieval Performance." *Proceedings of the ASIS Annual Meeting* 27 (1990):
    80–85.
47. Froehlich, Thomas J. "Towards a Better Conceptual Framework for Understanding
    Relevance for Information Science Research." *Proceedings of the ASIS Annual Meeting*
    28 (1991): 118–25.