

## References

- Cuadra, C. A., & Katter, R. V. (1967). The relevance of relevance assessment. In *Proceedings of the 30<sup>th</sup> Annual Meeting of the American Documentation Institute* (pp. 95-99). Washington, DC: Thompson.
- Rees, A. M., & Schultz, D. G. (1967). *A field experimental approach to the study of relevance assessments in relation to document searching* (Final report. Center for Documentation and Communication Research). Cleveland, OH: Case Western Reserve University.
- King, D. W. (1968). Design and evaluation of information systems. In C. A. Cuadra (Ed.), *Annual review of information science and technology* (vol. 3, pp. 61-103). Chicago: Encyclopaedia Britannica, Inc.
- King, D. W., & Bryant, E. C. (1971). *The evaluation of information services and products*. Washington, DC: Information Resources Press.
- Schamber, L. (1994). Relevance and information behavior. In M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (vol. 29, pp. 3-48). Medford, NJ: Learned Information, Inc.

---

---

## The User-Centered Approach: How We Got Here

◆ Raya Fidel ◆

### Introduction

I started my professional career in library and information science because of my great interest in knowledge organization. The more experience I gained in the profession, the more I realized how crucial it is to understand which organization would be best for each group of users. This in turn requires an understanding of how users seek information. And so now my focus is on studying information seeking and searching behavior. Throughout the relatively long course of changing my focus, I followed Pauline Cochrane's writings. Now I can say that she has been among the first to have a "user-centered approach" to knowledge organization, and she has used the term three years before it became a mainstream phrase. The following is a short discussion about the user-centered approach which was presented in a workshop in 1997.

### What is the User-Centered Approach?

The basic assumption of the user-centered approach to the design of an information system is, rather than teaching a user how to adapt to an information system, discover how users look for information and design systems that conform to the users' searching behavior. That means that information systems are designed according to what users need, not only according to universal rules. It also means that different groups of users may require different types of information systems.

---

---

Let me give you an example from my own experience. I worked as a librarian in a geohydrology department at a university. Most of the faculty in that department were well-established scientists. When they wanted a paper from a conference proceedings, they always remembered the scientific body that organized the conference and where the conference took place. They would have requests such as: Find me the paper that Smith presented in the U.S. Geological Survey's conference in Seattle. They rarely remembered the title of the conference or the year (I think it was easiest for them to remember conferences that way because they themselves participated in the conferences and could easily remember where they were held).

Such questions were difficult to answer because the cataloging rules we used at that time (AACR) did not instruct the cataloger to have an entry for the city in which the conference took place. So I searched through all conferences that were organized by the U.S. Geological Survey (and there were many of them) and discovered for each of them where they took place. Soon I developed my own "rule": Whenever I cataloged a volume of conference proceedings, I assigned an added entry for the city in which it took place. That is, I had to "bend" the cataloging rules because they were not the best rules for some of the users in the library.

But we also had students in our department, and they did not go to conferences and therefore looked for conference proceedings in different ways, most of which were accommodated by the standard cataloging rules. This brings up the following point: Whenever we discuss users and their information-seeking behavior, we focus on a specific group of users. My library, for example, had two groups of users: faculty and students. These groups were different because they had different tasks. The faculty needed information primarily for research projects and for teaching. The students needed information to write papers. In addition, faculty members were well established in the scientific community, whereas the students were novices.

Another assumption in the user-centered approach, then, is that each group of users may have its unique pattern of looking for information. That means that when we study users, we want to first define the user groups and then investigate the seeking and searching behavior of that group.

### Why is This Approach Necessary?

For centuries, people have used libraries that were designed according to general rules. Further, the rules themselves have evolved to accommodate the changes and developments in information technology. Why all of a sudden the emphasis on the user-centered approach?

There are many reasons. The most obvious reason: the more the system fits a user's needs and seeking and searching behavior, the more likely is the user to find useful information. This is an important reason, and because of it we might advocate the user-centered approach. The question is: why is it more central now than it was ten or twenty years ago?

User needs have always been important, but they have not been addressed as a central issue previously for two reasons: (1) in recent years, an increasing number of users search by themselves without the mediation of a librarian; and (2) the flexibility required to design systems according to individual needs became available only when computers were widely used.

When information systems are not flexible, it is hoped that there is a librarian to mediate between the user and the system. In the past, most library users preferred the help of a librarian when they were not completely sure how to find what they needed. The librarian always knew very well how to search the inflexible system and, for each request, the task was to discover the user's need. Through the reference interview and other means, librarians discovered exactly what a user needed. They then formulated it to fit the system. The inflexibility in the system was compensated for by the professional expertise of the librarian.

Today, many users in the Western world search for information without the mediation of a librarian. The systems are actually not much better than the inflexible card catalog, but there is more flexibility, and users can find *something* about 70 percent of the time. Even though reference librarians are still busy, and probably will always be, it is important to recognize that much of information retrieval is being performed directly by users. Because so much is done without the help of a librarian, it is important to design systems in which users can easily find information that is useful—that is, systems that are responsive to the way users search them.

Fortunately, we can design such systems today much more easily than we could before. At the time when academic libraries used card catalogs or any other form of a printed catalog, it was almost impossible to introduce flexibility into the design of these catalogs. It was best to follow standards because this way one at least knew what to expect from the system. For example, because I knew that each conference proceedings would have an entry under the scientific body that organized it, I could look under "U.S. Geological Survey" and know that the conference I was looking for would be found. In a large university library, it would have been impossible to tailor cataloging to different kinds of researchers and students and type all these different cards.

There is much more flexibility with the computer. One can have all the data associated with a volume of conference proceedings in one place, such as title, year, editor, scientific body, place, or the name of the keynote speaker. When individual users search for such a volume, they can use

whatever information they have. Another example is the design of the interface. Many online catalogs have several interfaces: one for the complete novice, one for the experienced searcher, another for the cataloger, and yet another for the librarian working in the acquisition department. This variety of access can be made available only in computerized catalogs. No printed catalog has this flexibility. Today, therefore, we can be more responsive to user needs because we can actually design systems that can meet these needs.

## Basic Concepts

The most basic concept in the user-centered approach is *information need*. This is, however, an ambiguous term, and not every one agrees on its definition.

Why do people look for information? Usually because they have to solve a problem or make a decision. Therefore, it is safe to say that the process of information seeking results from some sort of problem solving or decision making.

Consider, for example, a database for a library school. It has information about courses, time schedules, students, faculty, staff, and so on. Suppose a student has to make a decision whether or not to register for a course taught by a Professor Baker. To make this decision, he or she needs some information. The information that is needed we call an *information need*. Unfortunately, this is not that simple because this *information need* has many facets, and it is not always easy to determine what is the objective need. That is, what information will make it possible for the student to make a decision about the course.

Therefore, the following defines a number of basic concepts:

**Information Want:** The information the user thinks he or she needs. Before a person uses an information system or talks to a librarian, the person has some idea (not always well defined) what information will help to make a decision.

**Information Demand:** The information the user *says* he or she needs. Users do not always express what they think they need. Sometimes they think they have to present the information system with a request that the system or the librarian will “understand,” or they feel uncomfortable to express what they want.

**Information Need:** The information that is objectively needed to solve a problem. This assumes that through problem analysis and other “scientific” means, it is possible to determine objectively what information will make it possible to make a decision. The purpose of the reference interview, for example, is to do just that: find out what is the objective need and then translate it into a language that the information system can understand.

Let us consider the example of the student who has to make a decision about Professor Baker’s course. To make the decision, the student thinks he or she *wants* to know if Professor Baker is a favorable instructor. The student may feel uncomfortable to ask this question directly or may think that the database does not have this information, and he or she may present the request in an information *demand*, asking for a list of the courses that Professor Baker teaches. The student’s advisor, however, may discover that what the student really *needs* to know is the average grades that Professor Baker gave last year (because professors who give high grades are usually favorites among students).

This ambiguity in the concept of *information need* makes it difficult to base the design of information systems on this need. What should be considered—wants, demands, or needs? It seems that all should be taken into account, but it is still unclear how this can be accomplished.

Because of this complexity, there were some attempts to adopt more pragmatic approaches that define *information need* according to the use and impact of the information:

**Information Use** defines *information need* according to how the information is used. For example, suppose the student wants to use the information he or she retrieves from the database to ask other students about Professor Baker. The student’s *information need*, then, is a list of students who took Professor Baker’s courses.

**Information Impact** on task performance or decision making. The *information need* is that part of the information that actually affected decision making. If the student looks, for example, at course descriptions and average grade, but the course descriptions have no effect on his decision, then the average grade is his *information need*.

But these definitions have their own various problems. For example, how can one determine that the course descriptions had no effect? After the student reads them, they probably have some effect because the student knows more than he or she knew before, even if it does not seem to have any direct impact. Most problematic, however, is the fact that, according to these two definitions, we can determine what the *information need* was *after* a decision was made or a problem was solved. This is not helpful for designing information systems. In design, it is necessary to know information needs in order to design the system—i.e., *before* the specific information needs are addressed.

Because of the difficulties in defining the concept *information need* and the difficulties in actually understanding how it is being manifested in real life, research focuses on information seeking and searching behavior.

**Information Seeking and Searching Behavior** relates to how users look for information. Studying this behavior focuses on the *process* of looking

for information, not only on the *object* that is being looked for—that is, not only on the information need.

The concept *information seeking and searching behavior* is relatively new. It is not stable yet, and it means different things to different people. I understand *information seeking behavior* to mean what a user does from the moment she or he realizes that a decision is about to be made until the time the decision is made. *Searching behavior*, on the other hand, refers only to the interaction with the information system, which includes the interaction with a librarian.

### User Studies as a Tool for the User-Centered Approach

Researchers who study information seeking and searching behavior believe that information systems should be designed to accommodate seeking and searching behavior. This makes sense because individual information needs are difficult to define and are constantly changing. Therefore, it is almost impossible to design long-term systems based on needs. On the other hand, it is likely that we can find patterns in seeking and searching behavior that are common to all users of a certain group. Once these patterns are discovered, one can design a system that accommodates the behavior of users from that group.

To find these patterns, researchers conduct user studies. Such studies explore how individual users behave when they look for information in real-life situations. User studies have several distinct characteristics:

**User studies are usually field studies.** It is common to distinguish between studies that are carried out in the laboratory, and those that are performed in the “field.”

In a laboratory study, investigators select “subjects”—that is, people who will participate in the study. They create an artificial and highly controlled environment in the laboratory and ask the subjects to perform some tasks. Investigators then measure various things. For example, if a librarian wants to compare two online catalogs, he or she can select a group of sixty users and ask them to be subjects. A procedure the librarian might follow is: have two catalogs cover the same database and make up five questions. Ask thirty subjects to search the five questions on one catalog and the other thirty on the other catalog and then compare their searching. The librarian may measure which system provided results faster and/or which system retrieved more relevant items and fewer irrelevant. He can also ask the subjects to complete a questionnaire reporting about how well they liked the system they searched.

Because the study is performed in the laboratory, the librarian can decide which questions to use, he can even instruct subjects how to search (for example, with subject headings or with keywords), and he can com-

pare the results because all subjects search the *same* five questions. Such studies are very promising for evaluating specific features of systems. But they are not very useful if one wants to discover how people look for information because the subjects are assigned questions. They search someone else’s questions rather than satisfying their own information needs.

Therefore, user studies are often done “in the field,” that is, in the natural environment where users are actually looking for information to satisfy their own needs. There are three major procedures that are used in such studies: questionnaire, interview, and observation. These are described in the next section.

**User studies examine both users and nonusers.** When researchers study information seeking and searching behavior, they usually focus on people who are already actively looking for information. It is possible, however, that some people with information needs are not looking for information because they do not know how to even start, or they are intimidated by the information system. These are nonusers. It is desirable to design systems that serve even those who do not know how to look for information. Therefore, user studies often include investigations of people who are nonusers.

**User studies have to be carried out continuously.** User studies are definitely necessary for the design of user-centered information systems. But this creates a dilemma. Information seeking and searching behavior is determined a great deal by the information system itself. Users look for information they think can be found in the system and may avoid looking for data they need if they think it is not in the system. Or they search in a way they think would be fruitful in that particular system, which may not be the way they would prefer to search. Therefore, the patterns of seeking and searching behavior researchers uncover are partly influenced by the system the users searched.

Should a designer of a new system take into consideration patterns of behaviors that were observed under the influence of an old system? Studies of card catalogs in academic libraries, for example, found that most searches were by author and only a small percentage were subject searches. Should a designer of an online catalog assume that there will be a relatively small proportion of subject searches on the new catalog? If the answer is “no,” how can designers know which patterns are stable across systems?

The “solution” to this problem is a continuous study. In the life-cycle of an information system, a user study is carried out before the system is being designed. Once the system is designed and is fully operational, the system designer’s job is not complete. With the new system, there will be new patterns of seeking and searching. These must be studied so the new

system can be improved. And this process is continuous because behavior patterns change with the development of new systems. As a result, no design of an information system is ever complete. Further, no information system is ever perfect—every system can always be improved.

## Instruments Employed in User Studies

Because user studies are usually field studies, researchers use one or more of three major instruments:

1. **Questionnaire.** Users answer a list of written questions. An open-ended question leaves some room for interpretation and asks users to respond in their own language. A closed question asks users to answer very precisely or to select one or more answers among a list of possible responses. The question, "Please list the problems you encountered when using the system," is an open-ended question. The question, "Please mark how long you waited for a system response: very little, little, normal, long, very long," is a closed question.

When composing a questionnaire, researchers have clear and specific ideas of what they want to find out. They also understand that the responses they receive reflect what users *perceive* the situation to be and not necessarily an accurate picture of reality. For example, users may respond that a system was rather slow, but the particular system tested might be the fastest that has been developed.

2. **Interview.** Users answer a list of oral questions. Presenting questions orally is somewhat different than having them in a written form. Usually interviews handle open-ended questions better than questionnaires do. Users might be reluctant to write a lengthy response in a questionnaire, but they might find responding easier when talking to an interviewer. In addition, a skilled interviewer may interact with users during an interview and elicit information that the user initially did not volunteer to report.

Because interviews facilitate open-ended questions, they can be exploratory in that researchers do not need to have specific ideas about what they will find. But like questionnaires, interviews collect users' *perceptions* of reality rather than *facts* about reality.

3. **Observation.** The researcher is physically present with the users and observes them at the time that they perform their regular job. When studying information seeking and searching, quite often the researcher asks users to think aloud in order for the researcher to understand the users' decisions and actions. In observation studies, researchers collect data by audio taping or videotaping what is occurring in the field, by writing notes, and by collecting paper documents such as search transcripts or filled request forms.

Most observation studies are exploratory because researchers have no control over what information will come their way. On the other hand, such studies collect data about what is actually occurring in the field and are not limited to users' perceptions. The remainder of this discussion will describe various user studies as they have been carried out in England and the United States.

## Early User Studies: Correlation Among Observable Variables

Large-scale user studies were first carried out in the 1960s. The purpose of most of these studies was to discover how personal attributes of users correlated with their information-seeking behavior. The basic framework relied on the rationale that if researchers collected enough data, they would be able to predict seeking and searching behavior before systems were in place. For example, suppose we found that engineers who have just graduated from college, who work in teams, and who work in large companies prefer short articles. We could use these data when we design an information system for such users and have only short items.

To find such relationships, large-scale studies used questionnaires and interviews. The questions in these instruments were very specific and precise because they were designed to define variables. After data were collected on the variables, researchers performed statistical analyses to discover the preferences of each group of users. Even though all research at that time focused only on engineers and scientists, there was some inconsistency in the results. Nevertheless, these first studies gave a first glimpse at seeking and searching behavior.

B.C. Vickery (1973) summarized the variables that were examined in his book *Information Systems*. Here are some examples, as he arranged them in five categories.

### A. Environment

1. Nature of the institution at which the user works (academic or industrial)
2. Size of the institution in terms of employees
3. Duration of the project on which the user is working
4. Rank of user within the institution
5. Nationality of the user

### B. User

6. Age
7. Education (highest degree, first degree, technical qualifications)
8. Linguistic ability
9. Nature of work activity (management or research)

10. Subject field of work
11. Length of experience in the job
12. Volume of search activity (number of library searches per month, reading hours per week, etc.)
13. Stage of the project

*C. Message*

14. Subject field
15. Date of publication
16. Type of message (theoretical statements, results and data, method, etc.)
17. Extent of message (full text, abstract, or index record)
18. Perceived complexity of message

*D. Channel*

19. Type of channel (personal, such as oral or correspondence, or impersonal)
20. Directness of channel (primary or secondary)
21. Bibliographic form of impersonal media (book, journal, report, etc.)
22. Language of medium
23. Perceived accessibility

*E. Source*

24. Type of source (government department, research organization, etc.)

Vickery also listed the relationships that were discovered:

As far as the *user* is concerned, the volume of search is related to:

- the nature of the institution in which he works
- the size of his work team
- the duration of his work project
- his rank in the institution
- his age
- his education
- the nature of his job
- the subject field of his job
- his scientific productivity
- the bibliographic form of medium used

As for message, the age of message used is related to:

- the nature of the institution in which the user works
- the subject field of the message
- its type
- the form of the secondary medium used

## Second Generation of User Studies: Identifying Patterns of Behavior

The early seeking and searching studies aimed at detailed analyses of individual users and their personal attributes, hoping that these would give clues about behavior patterns. The next stage, which started primarily in the United States during the early 1970s, produced studies that examined the patterns themselves rather than the individual user.

These new studies were carried out in locations where users looked for information—i.e., in libraries. There, researchers investigated how users search the catalog. Investigators in these large-scale studies went into a library and interviewed users before and/or after they searched the card catalog and sometimes even looked over users' shoulders to observe how they actually performed searches.

Unlike the studies in the first stage, the aim of these catalog-use studies was to identify patterns of searching in a library of a certain type without taking into account many individual attributes of users. The main targets were academic and public libraries. Examples of such studies were published in Lipetz (1970, 1972) and Tagliacozzo (1972). The first catalog-use studies resulted in a variety of findings. The studies found, for example, that:

- most of the searches were for known-items, for which the user had some bibliographic information;
- undergraduate students searched more by subject than faculty did; and
- most users preferred to search by author name.

These were new insights at the time. It is possible that the librarians in those libraries knew all along that their users prefer to search by author name. The studies, however, used scientific and systematic methods to prove that these are indeed the patterns of searching, and that they are consistent from one library to another. The studies had a major limitation, however. It was not clear how the findings could guide the design of new catalogs—those that would be better for users in these libraries. The lack of “why” questions in the studies was the main reason for this limitation. Catalogs cannot be improved unless designers understand the reasons that caused users to follow certain patterns of behavior.

Let us examine, for example, the finding that most searches were for known items. Given this piece of data, how can a catalog or library services be improved? They cannot without knowing the reason for that pattern. If users prefer known-item searches because most of their assignments in the library require such searches, then an improvement would be to guarantee that such searches are as easy as possible to perform. For instance, an author name can be entered twice: once in a direct manner (e.g.,

Darlene Baker) for those users who do not know that a name is usually inverted, and another in an inverted form (e.g., Baker, Darlene) for experienced users.

But if most of the searches are known-item searches because users gave up on subject searches, the library should take a different approach. Librarians may want to investigate what problems users had in subject searching and act to help users to overcome such problems. Librarians may find, for example, that users avoided subject searching because they could not find the subject headings they needed. Investigating the reasons, it may be found that the terms users looked for were more specific than the subject headings used in the library. Librarians may then create new lead-in terms ("see" references) that point users with specific terms to the broader subject headings they could enter.

User studies that ask "why" questions, and thus aim at a somewhat deeper understanding of user behavior, are not very common even today. But those that do are very insightful and usually provide guidelines that are promising for system improvements.

## A Summary of General Results

By the end of the 1970s, many user studies that employed questionnaires and interviews were carried out. At that time, some general patterns of seeking and searching behavior began to emerge. In 1976, Sylvia Faibisoff and Donald Ely presented a summary of research results. The following are some of them:

- people tend to seek out information which is most accessible;
- people tend to follow habitual patterns when seeking information;
- users and potential users of information are often unaware of sources and how to use them;
- face-to-face communication is a primary source of information;
- different types of persons use different sources of information;
- the nature and content of information needed is variable and complex, varying from discipline to discipline and from group to group;
- there is a wide range of need among users in the quantity of information required;
- the quantity of information often exceeds the capability of the individual to use it; and
- the information needs of the individual change at different stages of his or her career and with changes in his or her project.

These generalizations show patterns that are common to information seeking and searching, but they also indicate that behavior is situational—i.e., it depends on the particular situation of a user when he or she searches for information.

For me, the most promising finding here is that people tend to follow habitual patterns when seeking information. I believe that once we identify the seeking patterns for a certain group of users, an information system can be designed to accommodate these patterns. Therefore, studying patterns of seeking and searching behavior is most promising for the development of information systems, and this is the type of research I do.

## The CLR OPAC Studies in Early 1980

With the 1980s, the first online catalogs started to appear in American libraries. The Council on Library Resources (CLR) immediately funded a number of studies to help design the future OPAC (online public access catalog). Several teams of researchers investigated users and nonusers, as well as librarians, in a variety of settings and libraries. These studies were different from previous user studies because they employed a variety of instruments and sometimes in combination. The main methods were:

- Focus-group interviews with users and library staff. In such interviews, a moderator leads a group of about five to twelve individuals through an open in-depth discussion (see, for example, Markey, 1983).
- Individual and group interviews with library staff at research libraries (see, for example, Ferguson et al., 1982).
- System monitoring with transaction log analysis. This analysis uses a computer printout that records all the transactions that took place during a search (see, for example, Tolle, 1983).
- Self administered survey online. This survey targeted users after they completed their searches. A screen with an electronic questionnaire appeared at the end of a search, and the user was asked to fill it in. One of the new aspects of such a questionnaire was that the results were automatically transferred to a statistical package which was then used to perform statistical analyses.

An excellent review and analysis of these methods can be found in Cochrane and Markey (1983). The discussion here about the OPAC studies is based on this article.

## Examples of Studies that Used a Questionnaire

The results collected with the online questionnaires at the CLR study can illustrate the type of information about users that can be uncovered with a questionnaire. Here are some examples:

1. Demographic characteristics of OPAC users:
  - Academic library users visit the library more frequently than do public library or federal and state library users.
  - OPAC users visit the library more frequently than do OPAC nonusers.



—A greater percentage of men than women are OPAC users.

—Older adults are likely to be OPAC nonusers.

## 2. Online searches of the OPAC:

—OPAC users search by subject more than any other search type.

—OPAC searches at academic libraries are course-related and at public libraries are for personal interest.

—At least 40 percent of OPAC users find all that they were looking for or more than they were looking for.

## 3. Problems with OPAC features:

—Finding the correct subject term.

—Increasing the results when too little is retrieved.

—Reducing the results when too much is retrieved.

Such findings can be very helpful for making decisions about managing a library. For example, a library manager may want to increase the frequency in which women and older adults search an OPAC by reaching out to them and providing more training classes. Or, after realizing that only 40 percent of users find all that they want, a manager can create a new goal for the library—e.g., raise this figure to 60 percent.

Some data, however, are only indications that improvements can be made, but they cannot by themselves show how improvement could be accomplished. For example, a manager may realize that users have difficulties with finding the correct subject term, but the results of the questionnaire cannot provide a suggestion for a solution. Other, more in-depth, methods will be needed for that.

## Examples of Studies that Used Interviews

The CLR studies used interviews with groups as well as with individuals. For example, the focus-group interviews revealed the problems users had with OPACs and their databases, what advantages they found, and suggestions for improvements. The following are examples of results relating to existing systems:

### 1. Problems:

- Terminals were not always available
- Periods of downtime and slow response time
- Commands difficult to use
- Equipment was not useful

### 2. Advantages:

- Sophisticated search capabilities
- Time saver
- Printout was available, no need to take notes
- Circulation information

### 3. Improvements:

- New services, such as community bulletin boards, information about the institution
- Enhancement of commands
- New features
- Access at home or locations other than the library

Some examples of results relating to databases are:

### 1. Problems:

- Difficulties in finding the right subject heading
- Ambiguous codes and abbreviations
- Coverage not current
- Indexing and database not accurate

### 2. Improvements:

- Useful subject headings and subject headings display
- Shelflist-related displays
- Name and subject cross references
- Holdings information

These and other results were useful in interpreting the findings from the questionnaires. For example, by finding out what users perceived the advantages of the existing systems to be, researchers understood why the attitude of OPAC users was favorable toward the OPAC. Similarly they discovered why users were not satisfied when they examined the nature of the problems users said they had with searching systems and databases.

## Examples of Studies that Used Transaction Log Analysis

The use of computer-based systems for information seeking and searching provided another advantage to researchers: users can be monitored without affecting their searching behavior. Because the computer can log every transaction, we can record various pieces of data, such as what terms a user entered, how much time lapsed before another set of terms was entered, how many items were displayed, and which were printed. The CLR studies used analyses of such transaction logs to collect a large variety of data. For example, they collected data about:

- total number of online transactions and/or online search sessions over a given period;
- frequency of commands entered by users;
- average length of an online search session;
- most frequently used access points (author, title, subject, keywords, etc.);



- number of errors by type and number of errors per online search session; and
- average five minutes of catalog activity during a week (number of active terminal, number of logons and logoffs, average response time, etc.).

By themselves, data collected from transaction logs can be essential to support administrative decisions. They can identify the busy and slow hours in the library, the most frequently used features of the systems, and the type of errors made. But in combination with findings arrived at by other methods, the data so collected can help to create a more complete picture of seeking and searching behavior.

For example, transaction log analysis revealed that many users retrieve nothing. It is estimated that currently, on average, 30 percent of the searches in a catalog retrieve zero hits. Obviously, this shows that online catalogs require improvements. But how do you help users avoid zero hits?

If transaction log analysis showed that many errors were spelling errors, a spell checker can be installed to correct such errors. Or, if users frequently entered an author name as a direct entry, rather than inverted, a program should be written to change the order of names in an author search after a zero hit.

Such improvements may eliminate some of the zero hits. It is also likely that many zero hits were the result of a problem mentioned in the focus-group interview: users did not know what subject terms to use. To help users here, we can rely on some of the questionnaire's finding. In an academic library, we can assume that most searches are course related, and that they are subject searches. For each course that is taught, then, we can add relevant vocabulary lists to the online catalog. Once a user types in the title of the course for which he or she is looking for information, the machine can display automatically the relevant lists to help the user find new terms.

## Examples of Studies Using Observation

The CLR studies did not use observation. In this method, the investigator is physically next to a user observing him or her while seeking and searching for information. In fact, not many studies to date used this method, primarily because it is labor intensive and because it requires in-depth analysis. One of the earliest examples of an observational study was carried out in England by Tom Wilson. He and his colleagues explored how employees in local social services departments seek information. The aim of the project was to discover patterns in information seeking and searching behavior and to design an information system based on these patterns. The three stages of the study are reported in Wilson and Streatfield (1977), Wilson et al. (1979), and Steatfield and Wilson (1982).

Wilson and his colleagues were interested in information-seeking behavior as part of the communication process in the organization—i.e., for the first time, information seeking was looked upon as a *process*, rather than just a collection of facts.

To explore this behavior, project members spent an entire work week with selected employees of the department. Each project member was assigned to one person and followed him or her all day long. Project members were present when participants talked on the phone, when they participated in meetings, when they talked with a colleague or with their boss, when they went to visit clients, or when they wrote memoranda. Team members were silent but, whenever a communication occurred, they recorded it as a communication event and reported some information about it.

The study included twenty-two employees in various positions in the department and a total of 6,000 communication events. Methods to analyze the data were developed as the analysis was performed, and the findings provided a rich description of the seeking and searching behavior of several user groups. Some of the general findings were:

- 70 percent of information-transfer encounters took up to five minutes. From that the investigators concluded that the working day was fragmented. Therefore, information products should be designed in such a manner that they could be skimmed in less than five minutes.
- 61 percent of the information-transfer encounters were oral communication, face-to-face or by phone. The researchers suggested, then, that the librarian or information specialist should keep direct contact with users.

The team also found patterns that were typical of users in certain positions in the organization and suggested the kind of information services they require. For example, directors need to keep information on a range of subjects which may change frequently. They require, therefore, a personal and individualized information service. Specialists are another source of information for the rest of the staff. They require, therefore, a well organized research library.

Social workers' main difficulty was discovering who are the experts. They need primarily a directory of official and unofficial experts. Tom Wilson and his colleagues actually designed the information system for the social services department. In their evaluation, they found that the new system was much better than the previous one.

Observation can be carried out in various ways. Annelise Mark Pejtersen, for example, observed reference librarians in public and school libraries when they helped users to find books of fiction. Based on the analysis of the data she collected, she designed a retrieval system for fiction with the graphic interface called BookHouse. Other studies are carried

out where the investigator is next to a user while the user looks for information and expresses vocally the thought processes. These are taped and analyzed later.

Observation is a powerful method because it records events as they occur and does not have to rely on the user's perception or memory. In addition, it provides for a comprehensive and in-depth analysis of the situation of the user and the process of seeking and searching as a whole. On the other hand, observation is time consuming, and data analysis is labor intensive.

### Examples of Studies that Used a Combination of Instruments

Today, it is accepted that the more instruments one employs in combination, the more reliable and insightful a user study is because each instrument reinforces the others. Nevertheless, most user studies are limited to one method, mostly because of practical considerations. An example of an early study that employed a variety of instruments in combination is the study of information transfer among engineers that was conducted by Tom Allen of MIT in the mid 1970s.

This was a complex study that collected rich information and generated important findings. In particular, Allen was interested in the networks that engineers created to disseminate information. His study coined the concept *information gatekeeper* which describes a person who collects information in order to disseminate it to other people. The research project and its results are reported in the book by Allen (1977).

Allen focused on a certain situation where a number of teams had to solve the same problem. A government agency gave contracts to several companies and asked them to propose a preliminary design of a certain hardware product that met certain specifications. This situation was a great advantage to Allen. Because several teams were each working to solve the same problem, he could make a comparison among the teams and find out what was common in their seeking behavior and what was different. In addition, when the teams delivered their proposals to the agency, it ranked the teams by performance. Allen then could compare the teams that performed well with those that performed poorly.

The method Allen used was the "case study" method—that is, a detailed analysis of an individual case. Here, each team was an individual case, and he investigated in great detail the process that led the team to the solution.

To investigate the case, Allen broke down the problem that the engineers were supposed to solve into subproblems. He then collected data for each subproblem. For that he used several instruments:

- **Time allocation forms.** At the end of the day, each engineer indicated

on a form how much time was spent on: (1) literature search; (2) consultation with colleagues within the organization; (3) consultation with persons outside the organization; and (4) analytic design.

- **Solution development records.** Engineers provided a weekly estimate of the probability that a certain solution would be chosen by the team for a certain subproblem. Suppose a hypothetical team considered two possible solutions, A and B. The solution development record for the first three weeks might be:

	Week 1	Week 2	Week 3
Solution A	0.8	0.75	0.4
Solution B	0.2	0.25	0.6

This shows that Solution A was leading in the first week because the engineer estimated that there was 0.8 probability that it would be selected. Toward the third week, however, Solution B seemed more probable. With these tables, Allen could follow the development of a solution.

- **Interviews.** *Before* the initiation of the project, engineers described what alternatives they considered for solutions. *During* the project, Allen interviewed them by phone if there was a marked shift in the probabilities to find out what was the cause for the shift. *After* the project was completed, he asked the engineers for the cause of each change in probability and what information source was used to induce the change.

Data analysis involved comparing data from teams and uncovering communication and information networks. The data reinforced the previous finding that most of the sources were person-to-person. Allen also showed that several sources were used in solving the same sub-problem. In addition, he found several differences between the lower and higher performing teams. For example, lower performing teams had the heaviest use of literature at the beginning of the project. Higher performing teams had the heaviest use of literature at a point about one-third into the duration of the project and a short period of heavy use at the end of the project.

### Conclusion

Researchers in library and information science have been looking for methods to study information seeking and searching behavior for the last

forty years. At the start, there was a belief that some universal patterns could describe how users seek and search for information. The way to uncover these patterns was to study as many variables as possible and to find correlations among them. Once these were established, one can create a huge table with all the relevant variables and show how they relate to one another. Such a table could then be used for the design of information retrieval systems.

Today, very few researchers believe in such detailed universal patterns. While all agree that there are some common attributes among all users, it seems that the relationships between this multitude of variables are too complicated and dynamic to be put into well-structured tables regardless of how large they can be. The current trend is to look at specific groups of users and to identify the patterns of seeking and searching behavior that are typical to each group. This means that each group of users may require its own information system.

To design user-centered information systems, developers must investigate the information-seeking and searching behavior of future users. A natural question here is: What is the best instrument to employ in a user study? The only answer is: It depends.

Each instrument provides a different type of results. Therefore, the desired type of results will determine what instruments to use. For example, a questionnaire will probably be the choice if one needs results that are specific and detailed that lend themselves easily to statistical manipulations. On the other hand, an analysis of a user group on which the developer knows very little would require observation.

I hope that this short overview illustrated, however, that a comprehensive and in-depth study always requires a combination of instruments. For example, when exploring a new group of users, an investigator can carry out observation first to discover what questions to ask in a questionnaire. Interviews can help later on in the interpretation of the questionnaire's findings. Suppose an investigator discovered, through observation, that users copied information from the screen rather than printing it. A series of questions in a questionnaire further investigated how users in general preferred to receive the system's output. Suppose the questionnaire revealed that one portion of users preferred to copy and the other to print, interviews with representatives from each camp might then uncover the reasons for these preferences.

We are just starting to adopt the user-centered approach to the design of information systems. Yet, we have a rich experience and tradition of studies and methods. Even though no one study is perfect, and no one method is the best, each contributes to our knowledge and our ability to design new studies. Once we learn how to understand information seeking and searching behavior, we will be able to design systems according to user needs.

## References

- Allen, T. J. (1977). *Managing the flow of technology: Technology transfer and dissemination of technological information within the R&D organization*. Cambridge, MA: MIT Press.
- Cochrane, P. A., & Markey, K. (1983). Catalog use studies since the introduction of online interactive catalogs: Impact on design for subject access. *Library and Information Science Research*, 5(4), 337-363.
- Faibisoff, S., & Ely, D. (1976). Information and information needs. *Information Reports and Bibliographies*, 5(5), 2-16.
- Ferguson, D.; Kaske, N. K.; Laurence, G. S.; Matthews, J. R.; & Zich, R. (1982). The CLR public online catalog study: An overview. *Information Technology and Libraries*, 1(2), 84-97.
- Lipetz, B.-A. (1970). *User requirements in identifying desired works in a large research library. Final report*. Washington, DC: Office of Education (ED 042-479).
- Lipetz, B.-A. (1972). Catalog use in a large research library. *Library Quarterly*, 24(1), 129-139.
- Markey, K. (1983). Thus spake the OPAC user. *Information Technology and Libraries*, 2(4), 381-387.
- Streatfield, D. R., & Wilson, T. D. (1982). Information innovations in social services departments: A report on Project INISS. *Journal of Documentation*, 38(4), 273-281.
- Tagliacozzo, R. (1972). Some relations between queries and search-terms generated by catalog users. *Journal of the American Society for Information Science*, 4, 278-280.
- Tolle, J. E.; Sanders, N. P.; & Kaske, N. K. (1983). Determining the required number of online catalog terminals: A research study. *Information Technology and Libraries*, 2(3), 261-265.
- Vickery, B. C. (1973). *Information systems*. Hamden, CT: Archon Books.
- Wilson, T. D., & Streatfield, D. R. (1977). Information needs in local authority social services departments: An interim report on project INISS. *Journal of Documentation*, 33(4), 277-293.
- Wilson, T. D.; Streatfield, D. R.; & Mullings, C. (1979). Information needs in local social services departments: A second report on project INISS. *Journal of Documentation*, 35(2), 120-136.