Using ethnography to investigate life scientists' information needs*†

By Diana E. Forsythe, Ph.D.
Associate Professor

Medical Anthropology Program
University of California, San Francisco
San Francisco, California 94143-0850

and

Visiting Scholar

Center for Biomedical Ethics
Stanford University
Stanford, California 94305

Designing information resources that actually meet the information needs of individuals requires detailed knowledge of these needs. This poses a challenge for developers. Because the meaning of particular terms can vary by field, professional knowledge differs to some extent in different disciplines, and the questions that people ask assume a certain amount of unarticulated background knowledge, understanding the information needs of life scientists is not a trivial undertaking. One source of help in meeting this challenge is ethnography, a set of research methods and an associated conceptual stance developed and used by anthropologists for investigating uncontrolled real-world settings. Drawing on the author's experience in using ethnographic techniques to study clinicians' information needs, this paper describes why such research is necessary, why it requires particular research methods, what an ethnographic perspective has added to the study of information needs, and what this broader approach has revealed about the types of information sought by clinicians in the course of their daily practice.

INTRODUCTION

Information resources for the life sciences are intended to help meet people's information needs. Designing resources that actually achieve this goal poses a challenge for developers. Because the meaning of particular terms can vary by field, professional knowledge differs to some extent among disciplines, and the articulated questions assume a certain amount of unarticulated background knowledge, understanding the information needs of life scientists is not a trivial undertaking.

One source of help in meeting this challenge is ethnography, a research method developed by cultural anthropologists and applied in this country for well over a century [1-3]. Despite the stereotypical assumption that anthropologists work in nonliterate societies, many are trained expressly to study social and cultural phenomena in Western industrial nations. Of these, a growing proportion apply their skills in the rapidly developing subfield known as the anthropology of science and technology [4-5]. Understanding the meaning and use of information in the life sciences is relevant to this developing area of research. Since the mid-1970s, anthropologists and other social scientists have contributed a good deal to the study of medical informatics [6-19]. A substantial body of work applying anthropological techniques to support the

* Based on a presentation at the Life Sciences Reference Research Institute, University of Pittsburgh, School of Information Sciences, Department of Library and Information Sciences, Pittsburgh, Pennsylvania, May 7, 1996.
† The final preparation for publication was done by Judith Barker, Ph.D., University of California, San Francisco.
design and evaluation of computerized technologies for medical settings now exists.

From an anthropological stance, this paper discusses some central reasons why the information needs of life scientists are not obvious but require thorough investigation in real-life contexts. Then, using examples derived from a variety of ethnographic studies of information needs in medicine, the paper demonstrates the complex variety of information needs that can exist, and the importance of context for understanding the production and use of information. Information needs can be verbal or nonverbal, general or specific, formal or informal, or literal or metaphoric. In their daily practice, physicians routinely seek and make use of this variety of types of information and information needs to translate normative generalizations and rules into particular actions that accommodate the contingencies and variability of individual settings and cases. The complexity and contextuality of the information needs of life scientists challenge information resource development about how best to incorporate into information systems the tacit, taken-for-granted, non-standardized information so essential to comprehension in particular situations.

WHY DO WE NEED TO INVESTIGATE INFORMATION NEEDS?

Investigating the information needs of life scientists is important for two main reasons. First, we know that physicians, and others, are having trouble obtaining information; meeting their information needs is a significant problem for many practitioners. To cite two pioneering studies from the extensive literature, Covell et al. have found that 70% of questions arising during patient visits remained unanswered and that in 25% of cases doctors did not know where to find the necessary information [20]. Williamson et al. have found that two-thirds of practitioners feel that “the current volume of scientific literature is unmanageable” [21].

Second, the development of automated tools is often proposed as a solution to the problem of practitioners' information needs. But computerized technology alone cannot solve this problem: people who design information resources require data on the information needs and work practices of their intended users. Whether automated or paper-based, no tool is likely to be effective and appropriate unless it has been designed with end users in mind.

Designing information resources raises many sorts of questions about end users and the nature of the knowledge to be made available to them. One important set of issues involves their conscious wishes: what do end users want in an information resource? These days, designers are increasingly sensitive to the importance of consulting those for whom a given resource is intended. Another set of important questions involves epistemological issues. Because these issues receive much less attention from designers, four will be mentioned here: selectivity, audience, form and style, and utility.

Selectivity

As a consequence of the fact that it is impossible to put all possible knowledge on a given condition in a single information resource, selectivity is a problem. Thus, one must make choices. But how to decide which knowledge should be included and which should be left out?

From an anthropological standpoint, knowledge is positioned. That is, knowledge does not exist in a vacuum: it is known by particular people who occupy particular positions in society [22-24]. Given pieces of knowledge can thus be characterized as reflecting the perspectives of particular individuals in particular social locations. In relation to the necessity of choosing which knowledge to include in a given information resource, the positioned nature of knowledge implies such questions as the following:

- Whose knowledge should go into a particular information resource? Should it be that of one expert or many? Which expert(s) should be debriefed for design purposes? If more than one expert contributes, what should be done if they disagree—whose knowledge should count as the “gold standard?”
- Styles of practice differ from place to place and experts have been trained in different traditions. Which tradition should a given resource encode? If experts somewhere else see and do things differently, should a project try to develop some kind of consensus?
- There are many different types of knowledge in the world, including formal and informal, universal and local, global and specific knowledge [25, 26]. What kind(s) of knowledge should be encoded in a given system?

Audience

Information resources are designed to be helpful—to meet people’s information needs about particular issues. One could imagine more than one potential set of users for any given resource, however, none of which is homogeneous. So whose information needs should a particular tool be designed to meet? What do these people want to know? Is the same information likely to be helpful and suitable for all potential users? If different categories of patrons want to know different things, what should go in a given information resource?

Form and style

In addition to questions of content, designing information resources also raises questions of form. How
should information be presented to be both comprehensible and consistent? Whose language and categories should be used to present the information? For example, medical domain specialists have specialized vocabularies suitable to their subject matter, but not all of it is likely to be accessible to primary care physicians, let alone other care givers, patients, and other members of the public. How much explanatory background material should accompany the technical information presented in a given resource? Specialists have a great deal of background knowledge about their field that they do not need to have spelled out in order to apply information correctly. But what about nonspecialists or users from a different field? What about members of the public if they obtain access to a resource? Different audiences need different tacit knowledge spelled out, so what should be encoded in a given resource?

Utility

Finally, having made our way through all these design issues, constructed a given information resource, and made it available to users, how will we know whether or not it is actually useful? Although people often take it for granted that any information will be helpful, it is a good idea to check whether a resource is really meeting the information needs of the users for whom it is intended. This means undertaking some sort of evaluation process.

WHY ARE SPECIAL METHODS NEEDED TO STUDY INFORMATION NEEDS?

Common sense suggests that if we wish to understand the information needs of life scientists, we should simply ask them. Why should special concepts and methods be needed for this purpose? The answer is that we should certainly ask people about their information needs. This alone will not provide a full understanding of the problem, nor will it produce sufficient data to serve as the basis for designing information resources to meet all information needs. Common to the problems of selectivity, audience, form and style, and utility, is a need for information about what specific categories of people in the real world actually want to know, treat as knowledge, take for granted, and find useful. Investigating such questions in real-world settings is the purpose of ethnography.

The problem of information needs is more complex than appears to the common sense view. Four reasons for this are described here.‡


First, information produced by people in answer to direct questions about their own activities is known as self-report data. Self-reports are not completely accurate. Conscious models of practice tend to be incomplete, at best; sometimes they are incorrect. While it is desirable to ask people about their own information needs, then, it is also desirable to collect other types of data on this topic. Without such supplemental data, the researcher has no way of assessing whether or not a given respondent's self-report is accurate and complete.

Second, while self-report data are always problematic when used alone, self-reporting is particularly suspect in the case of research on information needs. This is because the issue itself may be interpreted as reflecting upon the competence of the life scientists under investigation. In other words, respondents who feel that citing a long list of information needs may reflect badly on their competence have a motive to underreport these needs.

Third, in addition to the problem of detecting conscious information needs, Williamson et al. have shown that physicians may have significant unperceived information needs, that is, information deficits of which they are unaware. Williamson et al. note that substantial numbers of practitioners are "unaware of many newly established medical advances" [27]. Obviously, simply asking will not produce direct data about unperceived information needs: even the most cooperative respondents cannot volunteer information about something they are unaware of.

Fourth, even with respect to the subset of information needs that respondents are aware of and are willing to discuss, investigators are faced with issues of interpretation and completeness. How do we know that we have correctly understood what respondents are telling us? What about items they do not think to mention, e.g., things they take for granted?

To gather reasonably complete and reliable data on real-world information needs, we need methods that allow us to go beyond the partial and unreliable information produced by self-reports. We need to be able to check those self-reports, to gather information needs people cannot or will not tell us about, and to assess our understanding of the meaning of our respondents' taken-for-granted terms and assumptions. Ethnographic methods enable the trained investigator to do all of these things.

USING ETHNOGRAPHY TO INVESTIGATE INFORMATION NEEDS

Ethnography is grounded in a substantial body of theory, which provides a useful conceptual structure for understanding the information needs of life scientists and other end users of information resources. Among other matters, this theory distinguishes between dif-
different kinds of knowledge. For example, anthropologists pay close attention to the distinction between what people believe should occur, what they believe does occur, and what can be observed to occur in particular circumstances [28, 29]. Anthropological theory also emphasizes the contextual nature of knowledge—the fact that knowledge is held and used by specific people who live and work in specific cultural and organizational contexts [30]. Understanding the meaning of a given piece of knowledge means illuminating how, when, and by whom it is used; in other words, putting it into context.

Anthropology also offers systematic methods for investigating and documenting what people know and how they use that knowledge in their lives and work. Ethnographic methods include participant observation, formal and informal interviewing, and analysis of documentary material. These methods are flexible and designed to be adapted to real-world situations. In addition to producing detailed understanding of real-world social processes, they also provide insight into the concepts and premises that underlie what people do, but of which they are often unaware. For example, ethnographic techniques have been used successfully to investigate tacit assumptions about the meanings of “knowledge” and “work” held by system developers in artificial intelligence and medical informatics [31, 32], the different meanings of “knowledge about migraine” taken for granted by neurologists and migraine patients [33], and the cultural and disciplinary values unintentionally built into medical information systems [34, 35].

Ethnographic studies of information needs

Presented below are the information needs of life scientists as revealed by ethnographic research. Examples come from several projects over the past decade. The first study in 1988 investigated physicians’ information needs in internal medicine, and was directed by Randolph A. Miller, M.D. (then at the University of Pittsburgh) and involved Bruce Buchanan, Ph.D., Jerome Osheroff, M.D., and others [36, 37]. Findings from two later studies are included: one of neurologists and migraine sufferers, directed by Dr. Buchanan [38, 39], and a second one of emergency medicine conducted jointly with Bern Shen, M.D.

Examination of the information needs literature in the late 1980s showed that investigators approached the problem on the basis of several assumptions questionable to an anthropologist. First, “information” was generally taken to mean bibliographic information or textbook-type information, that is, formal facts and relations deemed relevant to clinical practice. It followed that practitioners’ information needs were assumed to involve a need for formal information. Both assumptions are notable because they differ from the experience of field anthropologists that a great deal of the knowledge people need and apply in their daily lives is actually informal in nature [40].

Second, investigators assumed that information needs were easily discovered and were manifested in daily work in the form of syntactic questions. Thus, most studies were directed toward gathering people’s conscious questions about particular topics. Again, this approach conflicted with what experienced fieldworkers know about the complex nature of human communication. In daily interaction, a great deal of information is communicated non-verbally. In addition, verbal communication itself is highly variable and nuanced; many messages are sent and received in forms that do not necessarily entail the explicit spelling out of meaning [41]. Rather than assuming that information needs were necessarily expressed as syntactic questions, a central question in the research became: “How exactly are information requests made?”

Accordingly, an empirical study of physicians’ information needs was conducted, incorporating a clear definition of “information need,” a broader notion of “information,” and a more comprehensive methodology than had heretofore been adopted in medical informatics. The goal was to observe and interview people systematically in order to discover patterns in information-seeking, to understand what knowledge would provide the answers sought, and to explore underlying (generally tacit) disciplinary perspectives and practice contexts in terms of which particular information needs and potential answers had meaning.

Between 1988 and 1994, ongoing medical practice and interpersonal communication in a variety of settings in internal medicine, emergency medicine, and neurology was observed.§ Interaction between physicians, medical students, and patients was recorded in written fieldnotes and in some cases audiotaped as well. Fieldnotes and transcribed audiotapes produced over 1,000 pages of text containing verbatim expressions of information needs plus interpretive comments and contextual data recorded by the observer. In this research, “information need” was defined to mean the conscious seeking of further information. No assumption was made about how information needs are communicated. If an information-seeking message was communicated in a way that an observer could detect and that interlocutors responded to, it was recorded as an information need. This approach treats information needs as distinct from the unconscious information deficits described by Williamson et al. [42].

§ Myra Brostoff, Nancy Bee, and Linda Morrison contributed to the field work and or data analysis in the study of migraine sufferers and neurologists.
Some generalizations

Fieldwork revealed that some assumptions in the information needs literature were incorrect. First, ethnographic observation revealed that information needs were both much more frequent and much more diverse than was discovered through self-reporting [43]. Self-report, what people believe is occurring, leads to underreporting and selective reporting. In other words, self-report produces incomplete and biased data. This method should not be used alone for the study of information needs.

Second, most information needs expressed by physicians in clinical practice do not primarily concern bibliographical information or textbook facts and relations (i.e., what people often feel should be or are predominant information needs). Indeed, “MEDLINE questions” are actually quite rare in ongoing discourse. This is not to suggest that MEDLINE is not valuable, but rather that there is a need for information resources to provide other sorts of information as well.

Third, information needs are not necessarily expressed as syntactic questions, nor do syntactic questions necessarily express information needs. Identifying information needs in normal communication flow is not the straightforward matter that many investigators seem to assume. On the contrary, as the material below illustrates, interpretation is often required to identify information-seeking messages and to understand what information is actually being sought.

Complexity of information-seeking messages

In ongoing discourse, information needs are expressed in many and often complex ways. They are not necessarily verbalized as syntactic questions. Here are two examples of utterances coded as information needs that are not syntactic questions:

- “I heard you’re feeling much better.”
- “I need to know her creatinine level.”

Information needs are not necessarily verbalized at all; they may be communicated tacitly without being labeled as inquiries either verbally or nonverbally. For example, it was noted one day on work rounds in internal medicine that a physician pointed in the door of a patient’s room. The resident immediately began reporting on the status of that patient. Without uttering a word, the senior physician had communicated the inquiry:

- “Is there anyone in this room whom we need to discuss?”

If information needs are not necessarily communicated as questions, the case is also that syntactic questions do not necessarily indicate information needs. Teaching questions, veiled commands, and rhetorical questions all have quite different purposes in ongoing discourse than the inquiry implied by their surface meaning. Thus, for example, the following query was directed by a senior physician to a junior colleague on work rounds:

- “What causes av nicking?”

The purpose of this question was not to find out what causes av nicking; the senior doctor already knew the answer. Instead, this teaching question was directed toward finding out if the person being tested knew this answer.

Similarly, the following question is not really a query at all:

- “Can you sit up?”

Instead, it is a politely-worded command designed to get a hospital patient to sit up in bed in order to be examined by the physician. The intended response to this utterance is not “Yes, I can sit up” but rather the action itself.

Finally, the following syntactic question recorded on work rounds is not the straightforward information request it may at first appear to be:

- “What do you do for the treatment of breast cancer?”

This message was uttered by a senior resident knowledgeable in the treatment of breast cancer. She was expressing her frustration at a situation in which all available treatment for a given patient had failed. This is a rhetorical question whose intended answer may be religious or philosophical. It is certainly not what it appears to be when taken out of context: a request for basic treatment information on breast cancer.

As these examples illustrate, many information needs are communicated in ways that are not simple. In order to interpret them correctly, contextual information is needed. This point applies especially in the case of information-seeking messages that are ambiguous, such as the following:

- “He had right facial?”
- “Where do you put the lesion in her?”
- “What are the drugs you’re supposed to use on pregnant women?”

These messages are easily decoded by both their recipients and the observer. Without contextual information, however, the questions are either meaningless or open to incorrect interpretation. In the first, one caregiver asks another on which side a patient is experiencing facial weakness. In the second, a senior doctor queries listeners where they think the lesion has occurred in a stroke patient’s brain. In the third, in the context of a discussion of treatment for tuberculosis patients, a senior doctor asks a teaching question about drugs that can safely be used for pregnant patients with tuberculosis.

Literal interpretation of an information-seeking message can miss the speaker’s real intent. To respond appropriately to such messages, one must be able to gather appropriate contextual information and interpret it correctly. Human beings do this so automatically that we often fail to notice it and, for the most
part our interpretations, are correct. In contrast, such interpretation poses a major challenge for automated communication.

Types of information sought

Fieldwork has shown that physicians express a need for different types of information in the course of their daily work. Major dimensions by which information varies are formality and specificity.

To date, information resource developers have focused on access to two types of information. The first, formal, general information is the sort found in textbooks and accessible through MEDLINE (e.g., widely-recognized diagnostic categories). The second, formal, specific information is the sort accessible through hospital information systems (e.g., the diagnosis of a specific patient). Information of both sorts is essential to medical practice and to medical record-keeping. In practice, physicians seek and make use of other types of information as well [44].

A third type is informal, general information. For example, this includes knowledge about what constitutes appropriate behavior with colleagues and with patients, as well as knowledge about how to live with the stress and the inevitable errors of medical practice [45, 46]. An example of an information need in this realm is the heartfelt cry mentioned above: “What do you do for the treatment of breast cancer?” Answers to this sort of information need are necessary to carrying out and surviving daily work in a medical setting, but they are not normally found in a textbook. Some informal, general information is what anthropologists describe as basic cultural knowledge, which varies to some extent between societies, is learned during childhood and professional socialization, and is generally transportable across settings. How, when, and where to discuss a patient with colleagues without breaching confidentiality is an example of such knowledge inculcated during medical training and professional socialization. The basic rules for successfully and ethically discussing patients are essentially the same, whether in a hospital in Boston or a clinic in Albuquerque or Seattle.

Another kind of informal, general information is experiential evidence accumulated and transmitted in the course of medical practice. For example, a physician might prefer to prescribe, and teach students to prescribe, a particular treatment regimen for a specified condition over other equally appropriate protocols because experience suggests fewer patient complaints about unwanted side effects will result. This is one mechanism by which physicians come to develop “usual and customary procedures” that vary somewhat from location to location. Some informal, general information may be in the process of being formalized. For example, preliminary results from a clinical trial remain informal until statistical validation, peer review, and replication allow them to be formalized through publication.

A fourth type of information for which physicians expressed a need is informal, specific information. This too is unlikely to be found in a textbook or list of formal rules. Also, it is unlikely to have been learned during childhood or professional training. Instead, it is picked up during socialization into particular institutions or simply through the experience of working in specific environments. For example, emergency medicine residents who rotated through different hospitals were often observed to seek information about the location of specific supplies and items of equipment in particular settings. Because different emergency departments arrange supplies and equipment differently, the residents lost track of such information between rotations. Similarly, an internist supervising the care of a number of patients divided between several patient care units at the same hospital inquired about the procedure for dealing with decubitus ulcers in one particular unit. Guidelines for patient care varied somewhat between clinical settings even in the same institution, therefore, this physician needed a reminder about specific local preferences.

CONCLUSION

Investigating information needs in real-life settings is a challenging and complex task. Applying anthropological concepts and methods can help us to understand the information needs of life scientists better and to develop better information resources to meet these needs.

A word of caution: perhaps because ethnographic methods are largely qualitative in nature and are intentionally unobtrusive, people without formal training in these methods often mistakenly assume that ethnography is something that anyone can do [47]. Doing valid and reliable ethnographic research requires considerable training and practice. While anthropological concepts and techniques are very useful in the study of information needs and in the design of technologies to meet those needs, professional expertise is required to obtain these benefits. Rather than attempting “do-it-yourself” ethnographic research, designers of information resources are urged to include anthropologists in development teams.

This discussion of life science information needs has focused on just one group, physicians. As soon as the information needs of other health care providers, and of patients and their families, are brought into view alongside that of physicians, however, the complexity—and importance—of the task confronting information resource developers becomes clear. Fundamental epistemological issues of selectivity (whose knowledge), audience (whose needs), form and style (com-
prehensibility and consistency) and utility pose enormous challenges to the successful development of computerized information systems.

Currently, educational resources for medicine tend to focus on the transmission of formal, general information. Written procedures and guidelines for individual institutions are designed to convey formal, specific information. If one looks only at textbooks and rule books, one may receive the impression that all medical knowledge is of this formal nature. But observing medical practice provides quite a different impression. In their daily work, physicians also routinely seek and make use of informal information to accomplish their work. They use such information to translate normative generalizations and rules into particular actions that accommodate the contingencies and the variability of individual settings and cases. One of the challenges for information resource development in the future will be to investigate whether and how computerized technologies can be used to facilitate access to all of these kinds of information.

REFERENCES

25. GEERTZ, Local knowledge, op. cit.
27. WILLIAMSON, op. cit.
28. GEERTZ, Local knowledge, op. cit.
29. GEERTZ, The interpretation of cultures, op. cit.
30. ROSALDO R, op. cit.
33. FORSYTHE, Using ethnography in the design of an explanation system, op. cit.
34. FORSYTHE, Blaming the user in medical informatics, op. cit.
38. Forsythe, Using ethnography in the design of an explanation system, op. cit.
40. Geertz C. Local knowledge, op. cit.
41. Tannen D. That’s not what I meant!: how conversational style makes or breaks your relations with others. New York: Morrow, 1986.
42. Williamson, op. cit.
43. Osheroff, op. cit.
44. Forsythe, Expanding the concept of medical information, op. cit.

Received July 1997; accepted September 1997