

Chapter 4. Foundations for Personal Information Infrastructures: Information-Seeking Knowledge, Skills, and Attitudes

"Knowledge is recollection." (Plato)

"The best way to stay young is to have a bad memory." (Miles Davis)

Information seeking, like learning and problem solving, demands general cognitive facility and special knowledge and skills, and is influenced by attitudes and preferences. General cognitive facility--what is commonly called intelligence--includes our abilities to remember, make inferences, and monitor our intellectual activity. Special knowledge and skills of three types also interact to determine information-seeking performance: knowledge and skills related to the problem domain, knowledge and skills specific to the search system and setting, and knowledge and skills related to information seeking itself. Attitudes such as motivation, confidence, tenacity, tolerance for ambiguity and uncertainty, curiosity, and preferences for social interaction and media influence when and how we apply information-seeking knowledge and skills. Taken together, these types of knowledge, skills and attitudes compose our personal information infrastructures. Information problems are always embedded in a context that determines which facets of our personal information infrastructure are brought to bear in that specific situation. Personal information infrastructures develop as we gain knowledge of the information-seeking factors and skills in managing the information-seeking process. Information professionals apply their general cognitive abilities to building knowledge and skills about sources and systems that contain information, techniques for mapping user needs to tasks, and strategies for seeking and representing information. Knowing what knowledge and skills are useful in manual environments and today's electronic environments will lead to better design for future information systems and to better training for professionals and end users alike.

This chapter considers these various types of expertise in turn, with particular emphasis on knowledge and skills related to information seeking. It examines how people seek information using existing information systems and the various knowledge, skills, and attitudes that constitute information-seeking expertise. We begin with consideration of the general cognitive skills that are important to information seeking and then consider domain, system, and information-seeking knowledge as types of expertise. The final section outlines four levels of cognitive and behavioral activity specific to information seeking. These discussions set the stage for the following two chapters that examine analytical and browsing strategies.

GENERAL COGNITIVE FACILITY

Information seeking can be characterized as a set of behaviors controlled by interacting levels of conscious (logical) and unconscious (intuitive) mental activity. At the most

general (pattern) level, humans demonstrate intelligence, preferences, and styles. Defining and understanding these qualities are central goals of cognitive psychology and many theories and models of cognitive faculty have been proposed, with most of the emphasis on the notion of intelligence (see Sternberg, 1982 for a taxonomy of theories of intelligence and a collection of views). A classical view of intelligence proposed by Guilford (1967) identifies three dimensions of intelligence, each with several factors that interact to give 150 discrete components of intelligence. The operation dimension has five factors: cognition, memory, divergent production, convergent production, and evaluation. The content dimension has five factors: visual, auditory, symbolic, semantic, and behavioral. The product dimension has six factors: units, classes, relations, systems, transformation, and implications. Sternberg (1985) proposed a triarchic theory of intelligence made up of three subtheories. In addition to a components subtheory similar to the components view of Guilford and others, he adds a sociocultural setting theory that considers how we adapt to and change the environmental context. The third subtheory considers experience, knowledge, and our ability to deal with novelty. Regardless of the theory adopted, any human activity is influenced by general cognitive abilities, especially complex processes such as information seeking.

At the strategic level, people use many perceptual and cognitive processes to plan, solve problems, make decisions, monitor progress, and reflect on past activity. These processes may be rooted in belief or in logic, but in practice, most intellectual activities--including information seeking--depend on interactions among various logical and intuitive actions. A fundamental form of logical reasoning is making inferences--"the process of deriving new information from old information (Davis, 1990, p.3)." Clearly, information seeking can support inferencing by delivering the old information. More significantly, however, the process of information seeking requires many types of inferences to be made as information seeking progresses. For example, we make inferences to assign meaning during problem definition, to determine which system best matches the problem, to choose query terms and map queries to database structure, to judge relevance, to monitor progress continually, and to decide when to terminate. Inferences can be deductive, inductive, or abductive.

Deduction yields conclusions that are precise and certain, since each step in the process is validated by logical combinations of known facts or generalizations. For example, all men are mortal, Socrates is a man, therefore, by the modus ponens rule of inference, Socrates is mortal. Thus, if the premises are true and one accepts the axioms of the logical system used, then truthful conclusions must follow. Unfortunately, most plans, decisions, and problems do not yield to these assumptions and pure deduction cannot be applied. Consider, for example, whether the term "profit" captures all aspects of the concept "economic gain," or whether a given document is completely relevant to a topic. Deduction is applied at the tactical level but is rarely applicable at strategic or pattern levels. For example, deduction allows us to conclude that if we have retrieved two sets

of documents for two specific terms, that joining these two sets with an AND operator must yield a set with cardinality (number of elements) less than or equal to the smallest cardinality of the two sets. Relational database management systems apply deduction and the relational algebra to assure that records matching precise specifications are retrieved. This is advantageous to support the most rote and basic types of information seeking and retrieval, but cannot be applied to the more general types of information seeking of interest in this book.

Induction yields general conclusions that are based upon specific examples or cases. Inductive reasoning is closely related to empirical methods that draw conclusions based on testing controlled cases. For example, if there are several telecommunications carriers from which to choose when going online and carrier A is selected five times and found to be busy each time, induction leads to the conclusion that carrier A is not a good service to use as a default. Of course, induction may lead to incorrect conclusions if antecedent information is not accurate or representative, and systematic validity checks should be conducted to maintain confidence. Induction leads to many of the defaults that information seekers use at all levels of activity.

Adduction is the kind of inferencing most commonly done in all but the most technical situations¹. Adduction is plausible reasoning, e.g., providing examples, arguments, or appeals to authority to increase the probability of some truth. All but the most rote forms of information seeking require adduction. Plausible reasoning has many forms and is influenced by different intuitive beliefs. Various heuristics or forms of plausible reasoning have been identified and information seekers are well advised to reflect on which heuristic is used at various phases of information seeking.

Simon (1979) described "satisficing" as a psychological form of the law of diminishing returns--we do not seek optimal solutions to problems since the costs are too high, so we settle for solutions that are satisfactory given the cost. This principle is clearly demonstrated in libraries, where users show high levels of satisfaction because they are able to find at least some relevant items with minimal investments of time and effort. This result is even more typical in electronic environments since time and effort can be minimal and the scope of sources can be broader. Satisficing is essential to most information seeking since all pertinent information for open-ended problems can seldom be assembled and assimilated optimally.

As opposed to conscious application of satisficing, people often engage in "wishful thinking" when engaged in complex reasoning. People sometimes neglect relevant information if it leads to conclusions that counter their beliefs, and sometimes give high weights to information that reinforces stereotypes and low weights (or totally forget) information that disconfirms them. People selectively recall information to explain events or make decisions and reinforce beliefs with redundant information rather than seek additional new information. People are biased toward "first impressions", giving

too much weight to initial hypotheses and too little weight to new evidence. (see Kahneman, D. Slovic, O. & Tversky, A. (eds), 1982, for a collection of papers related to biases in decision making) These biases are not themselves inferences, but influence plausible reasoning.

In addition to biases, people take actions to protect themselves from discomfort and danger. In fact, there is often wisdom in selective ignorance. For intellectual activity, this tendency leads us to protect ourselves from information overload and from topics that are beyond our knowledge or grasp. In the first case, we purposefully ignore bodies of information or systems to avoid taxing or diverting our memory, time, or financial resources. In what may be thought of as the "installed base syndrome," experienced users of computer systems resist upgrading to the latest iteration of an operating system because they understand the many ramifications of such changes. They recognize the integrated nature of their systems and that special peripherals, old but reliable software, well-established databases and archives, and well-learned procedures may be affected as a seemingly straightforward change in one part of the system propagates throughout the system. Likewise, institutions resist upgrades and changes in application packages because of the heavy investments in employee training in existing systems. In the second case, we avoid information that is too technical if we are not experts in the area, or we simplify information so that we may manipulate it more readily. We also may avoid formal information systems in preference to informal systems such as colleagues or ready-reference handbooks or services.

Davis (1990) provides a set of types of plausible reasoning and develops probabilistic models for representing plausible reasoning in a computer. Issues that affect plausible reasoning include: degree of belief, strength of argument, applying rules of general but not universal validity, avoiding all qualifications of a problem, inference from the absence of information, limiting the extent of inference, inference using vague concepts, finding expected utility, inferring an explanation, scheme-based inference, analogy, and inferring a general rule from examples. Degree of belief and strength of argument are dependent on experience and knowledge of the world. For information seeking, inferring the best selections of vocabulary at the proper level of granularity may use such knowledge. Applying general but not universal rules is the basis for default values and is a kind of inference that may be made automatically. For example, a specific database has been useful in finding information related to legal aspects of business practice in the past, and is the first source used for similar problems in the future. Avoiding all qualifications of the problem is a way of de-constraining the problem so that it is approachable. This heuristic is used in culling out what are the essential features of a problem during problem definition.

Inference from absence of information is very important in information seeking. Davis illustrates three forms of assuming a statement to be false based on not knowing it to be true. First, use world knowledge about the probability of truth, for example, assume

that a particular dermatological disease is not fatal because it is believed that most dermatology diseases are not fatal; i.e., classes inherit nonproperties from superclasses as well as properties. Second, (autoepistemic inference) assume a statement to be false based on the fact that one does not know it to be true, for example, assume a communication package does not support macros because one is an expert in the package and has not encountered such a facility. Third, assume a statement is false due to it not being present in an authoritative report, for example, since malaria is not discussed in a document describing travel procedures for a developing country, assume malaria is not among the health risks to worry about. In fact, some searches are conducted to explicitly determine that something has not been done (e.g., in developing a significance argument for original research or patent searching for a new chemical structure). Stielow and Tibbo (1988) have termed this the "negative search". At a tactical level, queries for specific terms may be executed to assure that all facets of a concept have been considered.

Limiting the extent of inference refers to recognizing when an inductive inference has gone too far, and using vague concepts allow us to assign items to multiple classes in gradations--the basis for fuzzy reasoning. Finding expected utility refers to our abilities to do "back-of-the envelope" cost-benefit analyses, for example, to determine whether to consult a second source after extracting information from one source. Inferring explanations points to the human characteristic of justifying or rationalizing events, for example, assuming someone picked up the phone extension when one is disconnected from a system while using a modem. Scheme-based inference allows us to infer more general or more detailed information based on world knowledge and relationships. For example, if a menu has an edit choice, we infer a variety of delete, insert, cut and paste operations according to the specific operating system we are using. Analogy is a powerful heuristic that allows us to transfer knowledge or skills to other settings and may be helpful in using our experiences with past problems or systems to solve immediate problems at hand. Analogy is related to the final heuristic addressed here, inferring a general rule from examples. This is a loose application of induction, leading us to jump to conclusions. For example, erroneously assuming that there are no documents related to Rocky Mountain Spotted Fever in a bibliographic medical database because the explicit term "fever" does not appear as a text word in any of the citations.

These examples illustrate how abduction and other forms of inference influence information seeking. In most cases, specific inferences and heuristics are used as information-seeking tactics either spontaneously or by design as part of a larger information-seeking strategy. There is promise in exploring ways that systems can assist users recognize biases and abductions and to select and apply them judiciously during information seeking.

DOMAIN EXPERTISE

Expertise in a domain requires knowledge and skills that allow the expert to solve problems in the area quickly and effectively. Studies of expertise have typically focused on the distinctions between novice and expert within a content area or on descriptions of what constitutes expert behavior (see Chi, Glaser, & Farr, 1988 for a collection of papers on expertise). Several important characteristics of expertise have been identified. Experts perceive large meaningful patterns in their own domains (Chase & Simon, 1973) that reflect the organization of a knowledge base rather than superior perceptual skills (Glaser, 1987). Experts see and represent problems according to fundamental principles that structure the domain knowledge, while novices tend to represent problems at a more superficial level (Chi, Feltovich, & Glaser, 1988). When they begin to solve a problem, experts typically spend a great deal of time qualitatively trying to understand a problem, effectively building a representation of the problem from which inferences can be drawn (Glaser, 1988). This mental representation allows the expert to "see things differently" and to guide behavior in purposeful ways. For example, in a study of expert/novice diagnosis of X-Ray pictures, Lesgold et al (1984) found that experts had a truly different perception of what they saw, which enabled them to discard non-pertinent information and quickly make a correct diagnosis. Studies of skill transfer among experts show little evidence that one domain is transferable to another domain. In fact, experts in one domain solve problems in much the same way as novices when confronted with problems outside their own subject domain.

In studies of computer scientists, economists, and attorneys searching full-text databases specific to their expertise, experts demonstrated how their facility with the information-seeking factors guided their searches (Marchionini, Dwiggin, Katz, & Lin, 1993). They were able to identify and critique types of information typically found in journals or databases (e.g., technical reports vs. reviews vs. news reports) and match information problems to those sources. They were able to identify key facets of the information problem and use technical terminology in formulating queries, and made rapid relevance judgments about results, in which they expressed high levels of confidence.

Through experience, workers in a field develop information-seeking expertise for information problems in that field; they are familiar with the various sources and systems devoted to the field; have extensive knowledge of the domain and its major concepts, problems, organizational schemes, and vocabulary; are comfortable in settings where work in the field takes place (e.g., laboratories, classrooms); possess skill in solving typical classes of information problems; and are able to differentiate relevant and irrelevant information. Domain experts form hypotheses about answers and focus information-seeking activity on testing those hypotheses. Thus, experts in a domain have greater facility and experience related to the information-seeking factors specific to

the domain and are able to execute the subprocesses of information seeking with speed, confidence, and accuracy.

SYSTEM EXPERTISE

To be successful, information seekers must have some basic facility with the physical interface to a search system. For example, abilities to read, use an index, physically turn pages, and knowledge of alphabetical and numeric orderings are essential to finding information in printed sources. Likewise, there are some fundamental abilities required to use electronic systems. These include: abilities to begin and end, read from a screen, to scroll and page, to use keyboards or a mouse, to make selections from various types of menus, to recognize the difference between prompts and commands, to get help, and to manage windows (e.g., move, resize, hide/show). Minimal levels of expertise necessary to use electronic search systems can be acquired quite easily with a few minutes of demonstration, depending on the system interface. Remote online systems typically require more skill to use because they require knowledge about establishing network connections. These additional requirements can be minimized by interfaces that automate the connection process, once the local parameters have been set. For example, the Grateful Med system automatically connects end-users to various medical literature databases once the phone numbers and account codes have been set up.

Expertise with a system goes beyond these minimums and includes the ability to customize the system and use most of the system's features. More significantly, online systems require users to understand how the database is organized, what level of coverage is provided, and how documents are structured. Organizational considerations include how the database is indexed (e.g., alphabetical, conceptual) and what indexing policies apply (e.g., most specific concepts assigned), and how morphological variations, abbreviations, and special codes are handled. Coverage includes sources and selection criteria, and document structure includes how records are fielded or how objects are delimited.

System expertise is less influential for information seeking success than is domain or information-seeking expertise. Results with elementary school and high school students searching full-text electronic encyclopedias demonstrate that only very brief introductions are necessary for users to be able to apply minimal search system features to find information. Although these information seekers did not apply all the power of the system and used naive information-seeking strategies, they were able to use these powerful search systems to find information relevant to their needs (Marchionini, 1989a, 1989b). In studies of students using the Perseus hypermedia corpus, previous computer experience and experience with the Perseus system were less influential in determining performance than interface effects (Marchionini & Crane, 1994; Marchionini, Neuman, & Morrell, 1994; Evans, 1993). On open-ended responses, only those students with no previous computer experience at all reported significant

problems in accomplishing assignments and dissatisfaction with their experience. These results suggest that the system interface was more important for user confidence and satisfaction than were previous computer experience or actual use of Perseus itself. These results parallel those of Kahn and Robertson (1992) who found that previous computer experience and training had little effect on job satisfaction and motivation for workers in travel agencies. In fact, the actual job tasks were by far the most potent predictors of satisfaction and motivation.

Although basic familiarity with electronic systems appears to be a requisite for successful use of various information-seeking tasks, well-developed mental models for specific systems are far less influential on information-seeking outcomes than more salient factors such as task, domain knowledge, and setting.

INFORMATION-SEEKING EXPERTISE-

All humans develop some degree of expertise in general cognitive abilities such as problem solving, learning, planning, decision making, and information seeking. In the case of information seeking, our personal information infrastructures develop as we gain expertise in various domains and operate in an information-intensive world. Note that expertise in a general cognitive ability applies across all domains of knowledge; we make specific adjustments and select particular parameters for instantiation when applying our mental models for information seeking to specific problems and domains. This is in contradistinction to domain knowledge that does not transfer to other domains. Expertise in information seeking can be characterized by a person's general knowledge of information-seeking factors and their skills and attitudes for executing the information-seeking process.

Knowledge of how various domains are organized helps an information seeker to determine how the domain of the problem at hand may be organized. The value of specific domain knowledge clearly aides information seeking (the basis for professional intermediaries specializing in particular domains), but our general experiences with how different domains are organized provide important cues to examine the organization of an unfamiliar domain when seeking information in that domain.

As with all intellectual activity, individual cognitive and perceptual characteristics influence information-seeking performance. People who have more experience with making inferences and recognizing the appropriateness and limitations of plausible reasoning should be more effective information seekers. Greene, Gomez, & Devlin (1986) found reasoning ability to be a determinant of information seeking success and Vincente, Hayes, & Williges (1987) reported that spatial visualization was a determinant of successful database searching. It seems reasonable that capabilities such as superior memory and visual scanning abilities interact to support broader and more purposive

examination of text. One goal of human-computer interaction research is to apply computing power to amplify and augment these human abilities.

In addition to knowledge and skills with a variety of search systems, expert information seekers recognize that every information-seeking activity is guided by the particular information problem at hand². Recall that the information problem is manifested as an information-seeking task that drives the information-seeking process and may change as search progresses. The task as stated by the information seeker and the "answer" or solution to the problem may each be defined in detail or anomalously. Consider the following cases that vary according to how well-defined are the tasks and solutions. To retrieve or verify a fact, an information seeker has an explicit and well-defined image of both the task and the solution--this is a production task that does not involve the information seeker in learning. This is a task typical of database management applications, where the object is to recover a specific object previously stored in a systematic fashion. The task requires minimal decision-making and reflection during search. In the case of finding information to write an assigned essay, a student may have a well-defined task statement, but few expectations about what the literature will hold. This is a learning task and requires significant decision-making and reflection during the search process. In another case, an information seeker may come across an interesting concept incidentally and look for a task to which it applies. This is a solution in search of a problem, a different type of learning that requires less decision making and reflection and more open-ended, synthetic hypothesis generation. Finally, an information seeker may initiate search without a specific task or preconceptions about what solutions may emerge, for example, a scholar examining literature to monitor progress in a research area. In this case the information seeker must use both synthetic and analytic cognitive strategies to determine the information need (topic) and find information relevant to the topic that may change and emerge as information seeking progresses. This last case, which we term "accretional information seeking", where both task and solution are ill-defined, includes much of the information seeking we do to stay abreast of fields of interest or to develop new interests.

Naturally, the ways that problems are recognized and accepted, and how individuals articulate them as information-seeking tasks are dependent on the individual's characteristics and the specific context and setting. The point is that information-seeking is highly variant in nature, more like general problem solving than it is like word processing or statistical analysis that are complex and challenging, but also circumscribed. In spite of these variations in factors, humans have developed both general-purpose information-seeking strategies and highly specialized strategies for specific classes of tasks or systems. We all learn general-purpose strategies during our basic education and learn specialized strategies through our advanced education or our professional work. Electronic information systems have begun to make some specialized strategies more generic and have led to new general and specialized information-seeking strategies.

Journalists, librarians, detectives, and other professionals devoted to information seeking develop collections of approaches and actions for their work. Journalists learn interviewing techniques, nurture relationships with sources, and monitor information-seeking with the goal of organizing results for formal dissemination. Librarians also learn interviewing techniques, although for the purposes of eliciting questions rather than answers, learn about and use a large number of reference sources and indexing techniques, and monitor information-seeking with the goal of organizing results for presentation to an end user. Results depend primarily on the librarian's ability to elicit and clarify the information problem, experience with similar questions, and knowledge of various indexes and reference sources. Katz (1987) outlines a general manual strategy that focuses on the reference interview, using reference sources to identify potential primary sources, retrieving and assessing potentially relevant information, and including additional exchanges with the end user. These general steps are influenced by the personal creativity of the librarian in matching the problem to sources and by time constraints.

Information specialists must be cognizant of a wide variety of information systems and have skill in using the most representative systems. Much of the training in reference that librarians receive, for example, is devoted to the many types of bibliographic sources and indexes that point to primary information sources (see for example the chapter headings for reference textbooks such as Katz, 1987). Likewise, much effort in the field of library and information science is aimed at creating compilations and organizations of information sources. Thus, librarians and information specialists have assumed important roles in the development of electronic networks (Buckland & Lynch, 1987). Likewise, information specialists must be able to apply emerging forms of hardware and software as electronic information systems become more pervasive.

Expert information seekers possess substantial knowledge related to the factors of information seeking, have developed distinct patterns of searching, and use a variety of strategies, tactics, and moves. Because the most extensive use of electronic environments has been made by professional intermediaries, valuable insights into the information-seeking process has been gained by studies of professionals using such systems³. Since the earliest and still the most pervasive electronic systems are bibliographic in nature, studies of intermediaries searching bibliographic online databases provide the bulk of our knowledge of expert information seeking in electronic environments. These results must be tempered by our understanding that they are based on the experience of professionals who adapted to primitive and specialized search systems. They define a collection of analytical information-seeking strategies and will be considered in detail in Chapter 5.

The development of electronic search systems led the information industry to provide new services and products and caused significant changes in how professionals seek

information and how librarians and other information professionals work and are trained. Changes in how information is accessed and managed subsequently changed the expectations of users, thus affecting subsequent system development and use. These changes in information systems and user behavior reflect the larger ongoing evolution of actions and expectations that affect the behavior of users and the design of new systems in general. For example, patent searches for new pharmaceuticals take significant time and effort in manual systems, but are necessary to make investment decisions. The ability to do patent searches online and by way of chemical structures leads to more searching because researchers come to expect routine searches to check and verify developments as work progresses.

The earliest systems provided secondary information (e.g., bibliographic data) and many of today's online systems still provide secondary information separate from primary information. Furthermore, these systems have typically been complex and difficult to use, requiring significant levels of training on the part of users. The value of specialized systems to support the practice of medicine, law, business, and engineering have changed the way professionals in those fields work and what expectations they have about information access, but the complexity of using such systems led to increased dependence on information specialists as collaborators in the information-seeking process. Although this state of affairs has begun to change toward end-user-oriented systems that contain primary information, the early experiences with developing online systems and training professional intermediaries to use them led to numerous studies of information-seeking behavior and the identification of strategies for efficient and effective use. An important transition in electronic environments that is currently underway is the augmentation of analytical information-seeking strategies by end-user-oriented browsing strategies.

PATTERNS, STRATEGIES, TACTICS, AND MOVES

Our thoughts and behaviors related to information seeking can be considered from different levels of granularity. Four levels of activity are defined to illustrate how early electronic systems have affected how experts seek information and to consider what these effects should have on emerging end-user oriented systems. Electronic environments first affect the most fine-grained activities such as the physical actions taken when executing searches (e.g., pressing keys or a pointing device instead of turning pages), but eventually affect all levels of information-seeking activity.

At the most coarse level, people exhibit information-seeking patterns. *Patterns* are sometimes conscious, but most often reflect internalized behaviors that are discernable over time and across different information problems and searches. Patterns may be due to chunked strategies or tactics that people internalize through repetition and experience. Psychologists have described automatic information processing for specific skills and tasks (e.g., "skilled memory", Ericsson & Staszewski, 1988), and automatic

information processing has also been studied for complex skills such as reading (Schneider & Shriffrin, 1977; Shriffrin & Schneider, 1977). Automaticity for groups of related actions may aggregate as default actions for general classes of problems, for example, high school students in our studies most often reported books and encyclopedias as first sources for information for a variety of search problems (Marchionini, 1989b). There is ample evidence in a variety of fields that people tend to use proximate sources (e.g., colleagues, personal collections) before libraries or other formal sources. Note that default patterns need not reflect accurate or efficient performance, but do serve to reduce immediate cognitive load. As more secondary and primary sources become physically proximate due to network access from homes or offices, new default patterns for source selection will necessarily emerge.

Patterns may also be considered to be due to personality traits, attitudes, or cognitive style. Some people are highly tolerant of ambiguity and uncertainty and others demand specificity and completeness. Likewise, some individuals enjoy social interactions and adopt information-seeking patterns that maximize interactions with colleagues or experts, whereas, others prefer the challenge of personal discovery and immerse themselves in books or electronic systems. (e.g., see Turkle, 1984 for characterizations of computer hacker personalities). Bellardo (1985) characterized professional intermediaries according to field dependence/field independence measures and compared their searching behavior, but found no statistically significant differences in their searching. Fidel (1984) conducted case studies of professional intermediaries and described operationalist and conceptualist styles. Operationalists tend to manipulate the search system to further search and aim for high precision. Conceptualists tend to manipulate concepts and vocabulary and aim for high recall. Patterns are highly individual, dependent on complex interactions, difficult to identify and classify, and are resistant to change. The introduction of electronic systems clearly affect the behavioral aspects of patterns and eventually the underlying cognitive and affective determinants as well. Patterns are mainly information-seeker specific but are also influenced by domains and search systems.

A strategy is the approach an information seeker takes to a problem. *Strategies* are sets of ordered tactics that are consciously selected, applied and monitored to solve an information problem. Thus, strategies are concerned with multiple subprocesses of the information-seeking process and are applied to specific information-seeking tasks. Strategies can be general and flexible or highly specialized and well-defined. Because electronic systems operate according to well-defined algorithms and many systems charge fees on time-based schedules, information-seeking strategies that are systematic and specialized have been adopted by expert information seekers. These formal, batch-oriented approaches to information seeking are called analytical strategies and are distinguished from browse strategies that are informal and interactive. These two classes of strategy are the extremes of continua as illustrated in Figure 4.1. Analytical strategies tend to be more goal driven whereas browsing strategies proceed according

to cues found in the data as search progresses. Analytical strategies require planning and thus are appropriate when information seeking is highly time sensitive. On the other hand, browsing strategies are opportunistic, beginning with an entry point and proceeding according to what occurs along the way. Well-defined goals and planning make analytical strategies more deterministic while browsing strategies tend to be heuristic. Analytical strategies are more precise and methodical and browsing strategies are informal and depend on interaction. Expert information seekers typically tend toward analytic strategies and novices toward informal strategies, although it will be argued that this is mainly due to primitive search systems and setting factors such as time pressures. Electronic environments make strategies used in manual environments more explicit and extensive, and have enabled new strategies for information seeking. Strategies are mainly search specific.

[Insert Figure 4.1 about here]

Tactics are discrete intellectual choices or prompts manifested as behavioral actions during an information-seeking session. Tactics most often are applied within one or two information-seeking subprocess. Tactics are thus more focused than strategies, for example, restricting search to a specific field or document type to narrow search results. Tactics often apply inferences that further information seeking. For example, asking a question that requires a binary answer that allows the search space to be limited-- consider a physician who orders a specific test that does not solve the diagnosis problem but eliminates many possible diagnoses. There have been numerous explications of tactics used by expert intermediaries; most have been influenced by Bates' (1979a, 1979b) taxonomy of search tactics and idea tactics. New tactics have been enabled by electronic systems and tactical skills with online systems most obviously distinguish expert and novice users of those systems. Improving systems at the tactical level is one of the most practical and promising areas of interface design. Tactics are mainly task-system specific.

Moves are fine-grained actions that are manifested as discrete behavioral actions such as walking to a shelf, picking up a book, pressing a key, clicking a mouse, or touching an item from a menu. Moves are manifestations of tactics and although they are conceptually uninteresting individually, taken in context, they offer observable evidence for interface assessment and in mapping the intellectual activity at higher levels of action. This level of study lends itself to electronic environments since discrete acts taken during information-seeking can be automatically and unobtrusively logged by the computer together with time and system status data. Uses of computer monitored data was first applied to information retrieval by Penniman (1975); has been described by Rice & Borgman (1983); and illustrated for comparing search strategies by Borgman (1986), Marchionini, (1989a, 1989b, 1990), Lin, Liebscher, & Marchionini (1991) and Evans (1993). Clearly, the physical moves taken by information seekers are radically affected by electronic environments, since moves are mainly system specific.

These levels of information-seeking actions are not mutually exclusive, as a set of specific moves with a given system may be chunked (e.g., as a macro) as a tactic, and sets of tactics or strategies may be chunked as patterns by searchers using the same system regularly. These levels are partially analogous to the Goals-Operators-Methods-Selection Rules (GOMS) model of human-computer interaction (Card, Moran & Newell, 1983). In this model of human-computer interaction, users set goals and work through a search space that consists of possible operators and methods for applying and combining operators. They aim to reach the goal state by judicious selection of operators and methods in the space. GOMS operators are conceptually analogous to tactics and behaviorally analogous to moves, methods are analogous to tactics, and selection rules are guided by overall strategy.

The development of electronic environments that provide primary information via high-speed networks and that offer highly interactive interfaces requires new information-seeking patterns, strategies, tactics, and moves. A major goal of research and design in the human-computer interaction involved in information seeking is to determine what patterns, strategies, and tactics are useful and how they can be supported by new systems. This is an example of a chicken-or-egg problem of user-centered design: we cannot discover how users can best work with systems until the systems are built, yet we should build systems based on knowledge of users and how they work. This is a user-centered design paradox. The solution has been to design iteratively, conducting usability studies of prototypes, and revising the system over time (e.g., Egan et al, 1989; Nielsen, 1989).

One approach is to take the various online searching patterns, strategies, and tactics identified in studies of professional online searchers and develop systems that optimize those activities. In fact, this is what has occurred with today's systems. Early systems were cumbersome but only experts used them and they developed creative workarounds to use the system efficiently. These workarounds became strategies and tactics over time and as end-user oriented systems were planned, they were designed to assist novices in using these strategies and tactics rather than supporting strategies and tactics appropriate to casual users. For example, the document set-oriented approach offered by Boolean-based systems lends itself to analytical strategies such as building sets that are then aggregated (building block strategy) or divided (successive fractions strategy). Preliminary experience with systems that provide ranked output suggests that new analytical strategies are needed when document sets are ordered (Marchionini, Barlow, & Hill, 1994).

Alternatively, systems can be designed to improve upon the naive patterns, strategies, and tactics exhibited by novice users, since those are presumably more natural and intuitive and computing power can be used to leverage and expand them. Clearly, the inertia of the marketplace and the user-centered design paradox will insure that system

design blends both approaches in iterative ways over time. We argue for starting with systems that amplify naive patterns, strategies, and tactics and then enhance performance with features that support expert patterns, strategies, and tactics. This approach addresses learnability first, and usability second, a natural progression when both requirements are met.

Chapter 4 Notes.

1. Davis uses the term abduction--lead from, rather than adduction--lead to
2. In the case of professional information specialists, additional specialized skills are needed since they are often removed from the problem and must also develop skills in helping clients articulate the problem.
3. There has also been considerable study of business databases in business settings that are designed to support well-defined fact retrieval tasks. This work has led to significant improvements in interfaces, especially query language improvements and alternative interaction styles such as query by example.