KNOWLEDGE ELICITATION

Nancy J. Cooke

New Mexico State University

WORD COUNT = 14,983

Chapter submitted to *Handbook of Applied Cognition*

Contact Information:

Nancy J. Cooke
Department of Psychology, 3452
New Mexico State university
P.O. Box 30001
Las Cruces, NM 88003

office: (505) 646-1630
fax: (505) 646-6212
email: cooke@crl.nmsu.edu

In this chapter the enterprise of *knowledge elicitation*, the process of explicating domain specific knowledge underlying human performance, and the cognitive issues that surround this practice are reviewed. Knowledge elicitation had its formal beginnings in the mid to late 1980's in the context of knowledge engineering for expert systems. *Expert systems* are computer programs that embody domain-specific knowledge and that perform (e.g., decision making, problem solving, design) at levels typical of human experts, (but not necessarily in exactly the same manner as human experts). *Knowledge engineering* is broadly defined here as the process of building knowledge-based systems or applications. These include expert systems, as well as intelligent tutoring systems, adaptive user-interfaces, and even knowledge-oriented selection and training devices.

The process of knowledge engineering involves *knowledge acquisition* which includes knowledge elicitation and other activities such as knowledge explication and conceptual modeling (Regoczei & Hirst, 1992), as well as the coding of the resulting knowledge, the design of a usable interface, and the testing and evaluation of this interface (Diaper, 1989b). Thus, knowledge elicitation is a subprocess of knowledge acquisition, which is itself a subprocess of knowledge engineering. In order to fit knowledge elicitation into the larger context of applied cognitive psychology it is necessary to understand its brief evolution.

Some Background

The push for expert systems in the 70's and 80's was motivated by (1) the technological capability, (2) the growing specialization of the workforce and cognitive complexity of jobs (Howell & Cooke, 1989), (3) the interest in creating artificial intelligence in machines, and (4) rejection of alternative general problem solving approaches (Feigenbaum, 1989). Instead of relying on search strategies, this new form of machine intelligence was "knowledge-based" or powered by facts and rules. The realization that "knowledge is power," triggered a flurry of interest in knowledge, and particularly in its elicitation and representation (Feigenbaum, 1989). Meanwhile, parallel developments in the psychology of problem solving were taking place. It was becoming clear that expert problem solving could not be attributed to strategy, as much as to domain-specific facts and rules (Glaser & Chi, 1988).

With a new focus on knowledge, questions regarding knowledge elicitation became central to both applied and basic endeavors. How can knowledge be effectively elicited from an expert? Interestingly, the typical (or thought-to-be-typical) transfer of cognitive theory and principles to application did not hold here. Although the cognitive literature had addressed the issue of knowledge, the focus was largely on the question of representation and various theoretical conceptualizations of knowledge structure such as semantic networks, scripts, prototypes, and schemata (e.g., Anderson, 1995; Best, 1995, chapters 5 & 6). The most relevant cognitive research on expert problem solving and memory organization did not directly address elicitation, but provided some hints or guidelines that would help guide the future development of the methods. Additionally, the favored cognitive measures of reaction time and error rate were inadequate as a solution for knowledge elicitation (e.g., Bailey & Kay, 1987).

Thus, researchers and practitioners began to develop knowledge elicitation methods. Many of these techniques were adapted from cognitive methods or methods in other disciplines including anthropology, ethnography, counseling, education, and business management (Boose & Gaines, 1988; 1990; Cooke, 1994; Diaper, 1989a; Hoffman, 1987). Although, initial conceptualizations of knowledge elicitation portrayed the process as one of direct "extraction" (e.g., LaFrance, 1992), it quickly became obvious that the problem was not so simple (Cullen & Bryman, 1988). Error and bias were common, and experts' verbal reports and intuitions were often flawed. Thus, more recent conceptualizations of knowledge elicitation view the process as one of constructing a model of the expert's knowledge -- the outcome of which may reflect reality to varying degrees (Compton & Jansen, 1990; Ford & Adams-Webber, 1992).

These methodological developments and applied questions fueled research in cognitive psychology and the new field of cognitive engineering[2] (Woods & Roth, 1988; Vicente, 1997). The genesis of cognitive engineering has been in response to the need for research addressing cognition in complex contexts, such as those found in knowledge engineering applications. Similarly, Hoffman, Shadbolt, Burton, and Klein (1995) point out that the study of expertise "has recently gained impetus in part because of the advent of expert systems and related technologies for preserving knowledge" (p. 129). These new developments in research and methodologies no longer neatly fall within the boundaries of the basic or applied.

Concurrently, other applications have surfaced that demand knowledge elicitation, including intelligent tutoring systems, adaptive computer interfaces, and intelligent agents. In addition, developments in human resources and the increasing cognitive complexity of many jobs have led researchers and practitioners in that area to a stronger focus on the cognitive components of job performance (Howell & Cooke, 1989). Training and selection research has looked to knowledge elicitation techniques for answers. Note that unlike performance-critical applications such as expert systems, applications like training that go beyond knowledge use to the transfer of knowledge, require more attention to the psychological validity of the elicited knowledge. Similar emphases on knowledge and cognition underlying complex task performance has surfaced in other areas, such as human-computer interaction, human factors work (e.g., Benysh, Koubek, & Calvez, 1993) and cognitive engineering in general. These areas have also made use of knowledge elicitation methods.

This wide array of applications broadened the early focus on knowledge to include other aspects of cognition such as decision making, perception, planning, and design processes. The practitioner's tool kit was once again inadequate, and work was and is being devoted to developing additional methods. Many of these methods were also adapted from cognitive psychology, and are referred to as cognitive task analysis, cognitive engineering, cognitive modeling, or naturalistic decision making methods (e.g., Hutchins, 1995; Klein, 1989; Randel, Pugh, & Wyman, 1996; Sundstrom, 1991; Woods & Roth, 1988;). Although terms are different, there is substantial overlap among the methods and tools associated with them. In this chapter the focus is on knowledge elicitation, but many of the methods that are described are also used by those who take this broader focus. Before describing these methods and some of the newer developments in knowledge elicitation, the major cognitive issues that have influenced knowledge elicitation are reviewed.

Cognitive Influences

Although mainstream cognitive research and theory offered little in the way of direct solutions to knowledge elicitation, they were nonetheless influential in the development of methods for knowledge elicitation, particularly in the areas of problem solving expertise and knowledge representation. In addition to the influence from these two content areas, was the influence of verbal report methodology. In this section, each of these three influences is briefly reviewed.

Problem Solving Expertise

Early research on problem solving in the information processing tradition was dedicated to investigating strategies that individuals used to solve problems such as Tower of Hanoi puzzles or anagrams (Greeno, 1978). This research helped to identify some general strategies of problem solving, such as means-ends-analysis and working backwards, and to highlight the importance of problem representation. Then, in the mid-seventies a new problem solving paradigm emerged that focused on expert problem solving of complex tasks such as chess, bridge, geometry, and physics. In their seminal work, deGroot (1966) and Chase and Simon (1973) found that expertise in chess was associated not so much with search strategies like looking ahead, as with skilled pattern recognition based on the storage of many specific chess configurations in memory. Additionally, it was found that how that domain-specific knowledge was organized in memory was critical for expert problem solving. For instance, Chi, Feltovich, and Glaser (1981) found that experts in physics categorized physics problems according to laws or principles of physics, whereas those with less physics experience categorized the same problems according to the surface features of the problem. From this result Chi, et al. (1981) inferred that the physics experts represented physics problems according to deep principles, whereas less experienced individuals represented physics problems at a surface level.

A flood of research on expertise followed that replicated the now famous expertise effect (i.e., the finding that experts recall domain-related information better than novices) across many domains (e.g., Engle & Bukstel, 1978; Reitman, 1976; Sloboda, 1976). Other research on expertise explored more fully the actual distinctions between expert and novice knowledge organization (e.g., Cooke & Schvaneveldt, 1988; Gillan, Breedin, & Cooke, 1992; Housner, Gomez, & Griffey, 1993a; Schvaneveldt, Durso, Goldsmith, Breen, Cooke, Tucker, & DeMaio, 1985). One side effect of experimentation in this area was the need to more clearly define expertise or at least to distinguish experts from novices. This issue also surfaces in knowledge engineering applications (Hoffman, et al., 1995) and is addressed in this volume (see Charness & Schultetus, chapter XX for a review of the expertise literature). Although some of the methodology used in the early expertise experiments to explore knowledge organization (e.g., card sorting, relatedness ratings, think aloud problem solving) has been adopted by knowledge engineers, the major impact of research on problem solving expertise was that it provided scientific justification for the knowledge engineering enterprise. That is, it provided evidence for the importance of knowledge, in terms of both its content and structure, for expert performance.

More recently, the literature on problem solving expertise has included tasks that go beyond puzzles, games, and academic domains to include complex job-related tasks such as radiology (Myles-Worsley, Johnston, & Simons, 1988) and avionics troubleshooting

(Rowe, Cooke, Hall, & Halgren, 1996). This new emphasis has been led by applied researchers, faced with understanding these more complex problems. Indeed, the expertise associated with complex real world tasks is often impossible for basic researchers to study in factorial experiments due to problems obtaining experts or studying realistic task scenarios in the laboratory. This is a case, therefore, in which a true synergy is required between the basic and applied in order to understand the complexities of expert problem solving.

Thus problem solving research, in its attempts to describe and explain problem solving expertise, revealed the importance of domain-specific knowledge and the organization of this knowledge. This emphasis dovetails nicely with research focusing on knowledge representation.

Knowledge Representation

As the importance of knowledge representation for expert problem solving was recognized, other research in artificial intelligence and the psychology of memory focused on knowledge representation or how meaningful associations are organized in memory (e.g., Minsky, 1975). One of the first network models of memory organization was proposed by Quillian (1969), an artificial intelligence researcher interested in creating a program that could understand language. Psychologists elaborated upon Quillian's model, tested it empirically (Collins & Quillian, 1969) and added processing assumptions (Collins & Loftus, 1975). Other network models of memory organization were developed, as well as feature models in which concepts were represented in terms of a feature list (Smith, Shoben, & Rips, 1974).

In order to test these models and to explore knowledge representation empirically, several existing psychometric scaling techniques were employed including cluster analysis (e.g., Johnson, 1967) and multidimensional scaling (e.g., Shepard 1962a; 1962b). Other techniques were developed specifically for this purpose (e.g., Pathfinder network scaling: Schvaneveldt, Durso, & Dearholt, 1989; Schvaneveldt, 1990). These methods were also being used to study knowledge representation underlying expert problem solving (e.g., Schvaneveldt, et al., 1985). Knowledge engineers adopted these methods, and others like them such as concept mapping (Sanderson, McNeese, & Zaff, 1994) and the repertory grid technique (e.g., Shaw & Gaines, 1987; 1989) for the purpose of knowledge elicitation. However, because the theoretical goals did not focus on elicitation, but rather representation, the methods required some additional tinkering. For instance, in regard to Pathfinder network scaling, it was necessary to develop methods to elicit an initial set of domain concepts (Cooke 1989) and to identify the meaning of links in a Pathfinder network (Cooke, 1992b). In sum, the theoretical work on memory organization, the methods developed for exploring it empirically, and the concomitant importance of knowledge representation for expert problem solving, provided impetus for new methodological developments in knowledge elicitation.

One other related issue that surfaced simultaneously in both basic and applied camps has to do with the differential access hypothesis (Hoffman, et al., 1995) or the assumption that different knowledge elicitation methods may tap different types of knowledge. Along these lines, the phenomena of dissociations in memory performance under different test conditions is a well-studied topic in memory research today (e.g., Roediger, 1990.) Further, some knowledge measures may tap knowledge that is more predictive of performance that others. For example, Broadbent, Fitzgerald, and Broadbent (1986) have found dissociations between verbal reports and performance. Similarly, Cooke and Breedin (1994) found dissociations between individuals' written explanations for physics trajectory problems and their predictions of those trajectories. Together these results suggest that all measures of knowledge are not equal and that in particular, they may differ in terms of the connection between knowledge and performance.

The connection between knowledge and performance is critical in applications that utilize knowledge to improve or aid performance (e.g., training, expert systems). It has thus become important to map out the relationship between elicitation method and type of knowledge and performance. For example, Rowe, et al., (1996) compared various knowledge elicitation methods used to elicit knowledge about an avionics system (i.e., mental models). They found that a relatedness rating method and a hierarchical concept listing interview were superior to diagramming and think-aloud methods at eliciting knowledge that corresponded to avionics troubleshooting performance. Some general assumptions (some untested) about the type of knowledge elicited by various methods is presented later in this chapter in the context of the methods.

One way to systematize the comparison of the numerous knowledge elicitation methods available is to identify one or more dimensions along which they differ. Questions about type of knowledge elicited and connection to performance can then be addressed for these unifying dimensions. One such dimension is the degree to which the method relies on verbal reports, with methods like think-aloud and interviews relying heavily on them compared to other methods such as observations and relatedness ratings. Some cognitive issues relevant to verbal reports shed some light on this dimension.

Verbal Reports

Verbal reports have been used in research ranging from decision making and text comprehension to applications ranging from accounting to user testing in computer systems (Ericsson & Simon, 1996). Although their use in some form dates back to the early 1900s in the heyday of structuralism, verbal reports have been recently revived as a legitimate form of psychological data after a hiatus during the stimulus-response era of psychology.

Throughout this history, verbal report methodologies have undergone much scrutiny. Criticisms of verbal reports, have been around as long as verbal reports themselves (e.g., Nisbett & Wilson, 1977). Although some of the earlier critiques were misguided or incorrect, most recent criticisms are based on the grounds that "the TA [think aloud] procedure changes subjects' thought processes, gives only an incomplete report of them, and mainly reports information that is independent of, hence irrelevant to, the actual mechanisms of thinking" (Ericsson & Simon, 1996, p. 61). Much of these arguments, however, lose their steam when one places verbal reports in the context of other forms of behavioral data, each of which has strengths, weaknesses, and methodological pitfalls.

Ericsson and Simon (1996) have developed a theory of verbalization processes under think aloud instructions and have been able to account for most of the data suggesting verbal interference, completeness, and relevance within this theoretical framework. Furthermore, their theory suggests conditions under which verbal report procedures should succeed or fail. For instance, verbal reports are not as effective for eliciting knowledge when the problem is novel or the reporter has low verbal ability or is inhibited in some way. Guidelines such as these are relevant to the use of verbal reports by knowledge engineers and are highlighted later in this chapter. Too often however, practitioners are unaware of, or for practical reasons fail to adhere to these recommendations and it is in these cases, that the knowledge elicited using verbal report methodology should be questioned.

<u>Summary</u>

Research in cognitive psychology has been influential in the development of knowledge elicitation methods. This research has demonstrated the centrality of knowledge in human performance and specifically the importance of the content and structure of knowledge and the context surrounding elicitation of knowledge. Further some methodologies for studying knowledge organization and utilizing verbal report data have been adopted and adapted by those interested in knowledge elicitation.

In the next section, four groupings of knowledge elicitation techniques are described. Each grouping is illustrated by way of a specific example of knowledge elicitation for the design of an expert system in the area of student advising.

Knowledge Elicitation Methods

Reviews of knowledge elicitation methods and various categorization schemes for these methods abound (Benysh, et al., 1993; Boose, 1989; Boose & Bradshaw, 1987; Cooke, 1994; Cordingley, 1989; Geiwitz, Klatsky, & McCloskey, 1988; Geiwitz, Kornell, & McCloskey, 1990; Hoffman, 1989, Hoffman, et al., 1995; Kitto & Boose, 1989; McGraw & Harbison-Briggs, 1989; Meyer & Booker, 1990; Olson & Biolsi, 1991; Olson & Rueter, 1987; Shadbolt & Burton, 1990; Shaw & Woodward, 1989; Wielinga, Schreiber, & Breuker, 1992). This preponderance of reviews is probably a reaction to the eclectic nature of the body of methods and the tendency for practitioners to develop methods specifically suited to their application, often with little documentation of their efforts.

In this section, four categories of knowledge elicitation methods are identified and briefly described. Recent methodological developments associated with a particular category are also highlighted. Within each grouping there are a number of knowledge elicitation methods and variations on individual methods. Space precludes the description of each specific method and the variations within each category (but see Cooke, 1994 or McGraw & Harbison-Briggs, 1989 for details). Instead, in this chapter, breadth is traded for depth. In particular, each knowledge elicitation category is illustrated through an enumeration of the procedural steps involved in applying a single method within that category to a hypothetical problem. The problem involves the development of an expert system that gives advice to university students regarding course registration. The system should be competent in the mundane aspects of advising such as degree requirements, course availability, and scheduling, as well as some of the more expert issues such as career considerations, course content, and course substitution. Specifically, the illustration focuses on the knowledge elicitation aspect of the development of this system in which knowledge is elicited about university advising from experts (professors, advising staff, experienced students). Although the domain of university advising is not as technologically complex as some other potential knowledge elicitation applications (e.g., avionics troubleshooting, nuclear plant operation), it is assumed that most readers would have some experience or knowledge within this domain, and would therefore be less likely to lose the knowledge elicitation message in the terminology and technical details of the example.

Throughout this section, is important to keep in mind that due to the wide ranging problems, domains, tasks, and knowledge types, multiple knowledge elicitation methods are warranted for nearly any problem. As mentioned previously, different elicitation methods may tap different types of knowledge (Hoffman, et al., 1995), not all of which may correspond to task performance (Rowe, et al., 1996). Equally important is the fact that there is no single definitive procedure for applying each of the methods. Although a method and an associated procedure is specified for the hypothetical problem, there are most assuredly other methods and procedures that would also be reasonable. Knowledge elicitation is a modeling enterprise and the methods can be thought of as tools to facilitate the modeling process. These tools may need to be modified to fit the specific situation.

<u>Observations</u>

Knowledge elicitation often begins with observations of task performance within the domain of interest. Observations can provide a global impression of the domain, can help to generate an initial conceptualization of the domain, and can identify any constraints or issues to be dealt with during later phases of knowledge elicitation. Observations can occur in the natural setting, thus providing initial glimpses of actual behavior that can be used for later development of contrived tasks and other materials for more structured knowledge elicitation methods. However, there are some tasks that cannot be observed in the natural settings (e.g., flying a one-seater aircraft) and in these cases it may be necessary to observe performance in a simulated context or through use of a contrived task (Hoffman, et al., 1995). Aside from where they occur, observational methods also vary in terms of what is observed (ranging from everything to specific predefined events), the observer's role (ranging from passive and nonintrusive to participatory), and the method of recording (writing, video, photos, audio). See Hoffman (1987), Meyer (1992), and Suen and Ary (1989) for additional information on observational methods.

Observational methods, like other knowledge elicitation methods, are associated with cost-benefit tradeoffs. For instance, on the benefit side, observations tend to interfere minimally with task performance. On the other hand, this is only the case if the observer is nonintrusive. Furthermore, observations can be a rich source of data, however, the interpretation of the data can become unwieldy.

The most recent innovations in this area come from adopting specific observational methods used by other fields such as anthropology and ethnography (Hutchins, 1995, Suchman & Trigg, 1991). Of most relevance, video analysis tools such as VANNA (Harrison & Baecker, 1991) and MacSHAPA (Sanderson, et al., 1994; Sanderson, Scott, Johnston, Mainzer, Watanabe, & James, 1994) have been developed to facilitate data analysis of observational videos. In general, these tools allow a video recorder and monitor to interface with a computer so that while the video is viewed, events can be identified and coded or categorized using the computer. Later, summaries of events, their time course, and frequencies can be generated from the software record. In some cases, particular events on the video monitor can be easily located through the software record.

How would observational methods be applied to the advising problem? The most straightforward way to approach this and many other knowledge elicitation problems is to simply nonintrusively observe experts at work in their natural setting while taking notes with pen and paper. Several types of information should emerge from this process including the scope of the advising task and the role that an expert system might play in this task. Through observation of several sessions, typical topics that are discussed in advising or specific questions that are asked of the advisor should similarly surface. These topics and questions can provide or refine objectives for the knowledge-based system in terms of areas of knowledge (i.e., facts, rules, strategies) in which a knowledge-based system should be proficient. In other words, observations should provide guidance in generating or refining the functional requirements of the knowledge-based system.

A procedure for implementing the naturalistic passive observation in the context of the advising example is presented in Table 1. In addition, some hypothetical data that may be collected in the course of applying this procedure are listed in Table 2.

[Insert Tables 1 and 2 about here]

Interviews

The most direct way to find out what someone knows is to ask them. This, in a nutshell, is the approach of unstructured interviews, the most frequently employed of all elicitation methods (Cullen & Bryman, 1988). Like observations, unstructured interviews are good for early stages of elicitation when the elicitor is trying to learn about the domain and does not yet know enough to set up indirect or highly structured tasks.

Unstructured interviews are free -flowing, whereas structured interviews have predetermined content or sequencing. The form of structured interview questions can range from open-ended (e.g., how, what, or why questions) which impose minimal constraints on the response to closed (e.g., who, where, or when questions), imposing somewhat greater constraints (Shaw & Woodward, 1990). In addition, question content can vary greatly, each type targeting a slightly different type of knowledge (e.g., Ford & Wood, 1992; LaFrance, 1987). Thus, interviews can be used to elicit a wide range of knowledge types depending on the specific interview task.

There are many varieties of structured interviews. Some are focused on a specific topic such as a case, the task goals, or a diagram. For example, forward scenario simulation interviews, make use of verbal simulation to focus on a case. The expert is walked through the problem verbally by the elicitor who presents system and environmental events to which the expert is asked to respond (Cordingley, 1989; Diederich, Ruhmann, & May, 1987; Grover, 1983). This procedure is likely to generate some conditional if-then rules, the "if" part stemming from the elicitor's problem statement and the "then" part comprising the expert's response. Other types of interviews such as goal decomposition involve having the expert work backwards from a single goal to the evidence leading to that goal. A result of this method can be a set of rules associated with each goal (Grover, 1983, Hart, 1986, Schacter & Heckerman, 1987). In other cases, the interview may focus on diagrams. These diagrams may reveal the structure of a task or system. For instance, the elicitor may have the expert draw information flow or functional diagrams (Hall, Gott, & Pokorny, 1994) or charts of task activities (Geiwitz, et al., 1988) or system state diagrams (Bainbridge, 1979). The information elicited may reveal system or task models held by the experts.

Other structured interview techniques are less focused on a specific type of interview material and instead suggest an interview procedure. The "teachback" method for example, is a technique in which the expert explains something to the elicitor, who in turn explains the same thing back to the expert for verification. This process continues until the expert is satisfied with the elicitor's explanation (Johnson & Johnson, 1987). This method serves to bring the elicitor up-to-date with the information in the knowledge-base and the way it is presented. Another structured interview technique, the "twenty questions" method, involves having the expert try to guess a domain concept targeted by the elicitor. As in the traditional parlor game, the expert can ask the elicitor yes/no questions about the concept (Breuker & Wielinga, 1987; Cordingley, 1989; Grover, 1983; Shadbolt & Burton, 1990; Welbank, 1990). The yes/no questions that are asked reveal information about distinguishing attributes within the domain.

In general, structured interviews are thought to provide more constraints on the expert's responses and consequently more systematic coverage of the domain. The additional constraints also tend to facilitate the dialog between the expert and the elicitor as compared to unstructured interviews. Although elicitor training in interview techniques is valuable regardless of interview type, it is much more critical for unstructured interviews than for structured interviews. On the down side, structured interviews require more preparation time and more knowledge of the domain than unstructured interviews.

Interviews, whether structured or unstructured, are relatively easy to administer compared to other knowledge elicitation methods. However, the tradeoff occurs at the data analysis and interpretation phases. The tasks of summarizing and drawing conclusions from open-ended interview responses are not trivial. Depending on the degree of structure inherent in the interview and the amount of preplanning regarding questions that are asked, the analysis of the responses may be relatively straightforward (i.e., frequencies of various responses, similarities of diagrams, lists of features). On the other hand, if the interview is unstructured or structured only slightly, tools and techniques used for observations and protocol analysis (described in the next section) can be helpful. If the interview is recorded on video, then video analysis tools may be used. If it is not taped, then it is still possible to develop and apply a code to audio or written transcripts. Pidgeon, Turner, and Blockley (1991) recommend the use of "grounded theory" to analyze interview data. Grounded theory is social science's version of protocol analysis, in which conceptual models are generated from qualitative data.

Recent trends in knowledge elicitation interviews include the development of highly specific interview methodologies in the context of particular domains and problems that target very specific types of knowledge. For instance, the Critical Decision Method (Klein, Calderwood, & MacGregor, 1989) requires a series of questions to be asked about an important past event such as a near accident in the case of an aviation domain. This information is used to better understand decision making, and the focus on the specific and real case is said to facilitate elicitation. Another methodology labeled PARI (Hall, et al., 1994) is associated with questions to get at each of four aspects of a problem (i.e., Precursors, Actions, Results, Interpretations) that are associated with declarative, procedural, and strategic knowledge. PARI has been used primarily for instructional design in the domain of avionics troubleshooting.

The method illustrated in the advising example is the forward scenario simulation structured interview method. As described previously, this method focuses on one or more cases which the elicitor provides to the expert in some initial form. The expert "walks through" the way in which each case would be handled. The elicitor provides information relevant to the scenario only as it is requested by the expert.

Using forward scenario simulation in the context of the advising example, one would expect to elicit from the expert the relation between relevant features of the situation such as the student's major, career plans, years in college, and grade-point-average and the advice given. These features could comprise the "if" part of some if-then rules. Sequential dependencies among these features may also surface in the order in which the expert requests particular information. This is the information that the knowledge-based system will have to request from the user. The responses of the expert to the information presented by the elicitor should also reveal some "then" parts of the if-then rules. Together this information is needed by the expert system to give advice using a production rule architecture.

A procedure for implementing the forward scenario simulation in the context of the advising example is presented in Table 3. In addition, some hypothetical data that may be collected in the course of applying this procedure are listed in Table 4.

[Insert Tables 3 and 4 about here]

Process Tracing

Process tracing involves the collection of sequential behavioral events and the analysis of the resulting event protocols so that inferences can be made about underlying cognitive processes. Thus, these methods are most often used to elicit procedural information, such as conditional rules used in decision making, or the order to which various cues are attended. The popular "think-aloud" technique in which verbal reports associated with task performance are collected and analyzed using protocol analysis is one variation on this general theme (vanSomeren, Barnard, & Sandberg, 1994). However, in addition to verbal events, events can also take the form of eye movements, gestures, and other nonverbal behaviors (Altmann, 1974; Sachett, 1977; 1978; Scherer & Ekman., 1982, VanHoof, 1982; Sanderson, James, & Seidler, 1989).

Verbal reports can vary in terms of their timing with the task, with concurrent reports occurring in conjunction with the task and retrospective reports occurring after the task (Elstein, Shulman, & Sprafka, 1978; Johnson, Zualkerman & Garber, 1987). Ericsson and Simon (1996) recommend concurrent verbal reports over retrospective ones. A possible problem with retrospective reports is that the conditions associated with verbalization are likely to differ from those associated with task performance and as a result, information processing may differ in the two cases. It is assumed that, the longer the interval between performance and reporting, the more prone the report is to this problem, with immediate retrospective reports being most similar to concurrent reports. Unfortunately, in applied settings, it is often difficult to obtain the report during task performance (e.g., in the case of air-to-air combat flight maneuvers or in the case of a task that is highly verbal such as our advising example), and in cases like these, practitioners have often attempted to re-enact the performance while collecting verbal reports, often with the aid of video. According to Ericsson and Simon's (1996) position, this practice should produce meaningful reports to the degree that the re-enactment captures the conditions and cognitive processing of the actual task.

Just as important as when the report is collected is how it is collected. Ericsson and Simon (1996) provide detailed procedures for collecting, analyzing, and interpreting verbal reports, including examples of instructions. Reports can be made by the person performing the task or by another individual who provides commentary on the task (Clarke, 1987). In addition, reports can differ in terms of instructions on what to report, and it is argued that only the current contents of consciousness can be accurately reported (Ericsson & Simon, 1996). This would rule out explanations or interpretations of specific thoughts or behaviors. In fact, many of the empirical results that demonstrate interference with verbal reports can be explained in terms of requiring individuals to do more than verbalize the current contents of consciousness (Ericsson & Simon, 1996). The general goal is to avoid requiring individuals to provide anything in the report that could interfere with or change their thought processes. Instead, individuals should be asked only to verbalize that information that they attend to--that information that is currently heeded. Thus, information regarding perceptual and retrieval processes will not be directly elicited; neither will processes that are compiled or automated. Instead, if this information is of interest, then it would need to be inferred from the information that is elicited.

The data collected using process tracing methods, like observations and interview data, can be costly to analyze, although the results are typically rich in information. The data from verbal reports, for instance, require transcription, segmentation, coding, and summary. Other forms of event data such as on-line event logs collected from computer users also require coding and summary. It is the coding process that is especially labor-intensive. Coding involves categorizing units of the transcribed and segmented protocol. The nature of the categories or code depends on the purpose of the analysis. If the analysis was done to identify procedural rules underlying task performance, then the categories may first consist of condition and action, with subcategories based on type of condition or action under each. Codes can be hierarchical with different levels of abstraction being useful for different analytic purposes (see example in Table 6).

Recent advancements have been made to facilitate the analysis of process tracing data. Tools have been developed to facilitate the coding and later summary of verbal transcripts. The field of Exploratory Sequential Data Analysis or ESDA (Sanderson & Fisher, 1994) has recently emerged as a result of a need for better methods for analyzing sequential data of the type collected from video observation or verbal reports. For example, PRONET (PROcedural NETworks: Cooke, Neville, & Rowe, 1996) is a method based on Pathfinder network scaling, in which sequential data can be reduced and represented in a graph in which nodes are events and links occur between contiguous events. Others have focused specifically on the identification of repeating patterns in the data (Siochi & Ehrich, 1991). The general goal of such methods is to summarize or reduce the data in a way that preserves its meaning, often providing some graphical way to visualize the data.

It is important to bear in mind that "process" is central to process tracing. That is, the methods are suited for identifying underlying process from data that are thought to reflect it. Therefore, although interviews can be transcribed and coded and frequencies of coded responses examined, the sequential nature of the results would not trace the interviewee's thought processes, but rather the process inherent in the interview itself (i.e., who said what when). Even frequencies with which concepts are mentioned in an interview may simply reflect idiosyncrasies of the interview. This is because think-aloud verbal reports are primarily monologues on that individual's thoughts, whereas interviews are more dialogues between elicitors and the experts in which responses are elicited by elicitor prompts.

Process tracing is illustrated in the context of the advising example using a retrospective think-aloud report procedure, prompted with video tape of performance. This retrospective approach is made necessary by the highly verbal nature of the advising session itself. That is, it is likely that concurrent verbal reports would interfere with the advising process. It is expected that the expert advisor would report the information currently heeded while viewing a video tape of the advising interview. In particular, this method would target conditional (if-then) rules, as well as general strategies applied in advising students.

A procedure for implementing the retrospective think aloud method in the context of the advising example is presented in Table 5. In addition, some hypothetical data that may be collected in the course of applying this procedure are listed in Table 6.

[Insert Tables 5 and 6 about here]

Conceptual Methods

Conceptual methods elicit and represent conceptual structure in the form of domain-related concepts and their interrelations. Several steps, are generally required, each associated with a variety of methods (Cooke, 1994). The steps are: (1) elicitation of concepts through interviews or analysis of documentation, (2) collection of relatedness judgments from one or more experts, (3) reduction and representation of relatedness data, and (4) interpretation of the resulting representation.

Concept elicitation is a critical step upon which the others depend. Cooke (1989) identified four methods of identifying concepts (i.e., concept listing, step listing, chapter listing, and interview transcription) and found that each differed in terms of the quantity and type of concepts elicited. One of the best ways to determine whether the concepts are adequate is to construct a hypothetical structure or structures using the concepts. If, for instance, meaningful distinctions between expert and novice structures, cannot be hypothesized using the concept set, then it is most likely inadequate.

Relatedness judgments can be collected from domain experts in a number of ways including pairwise relatedness ratings, sorting techniques, repertory grid, and frequency of co-occurrence (Zachary, Ryder, & Purcell, 1990). Relatedness ratings involve presenting pairs of concepts to the expert and requesting a quantitative estimate of the relatedness of the two concepts, usually using a scale that ranges from slightly to very related. This method can become costly in terms of expert time when the number of concepts exceeds 30. In these cases, a sorting method is advised in which concepts are grouped by the expert into piles based on relatedness and relatedness estimates are then derived in terms of co-occurrence of concepts in the piles (Cordingley, 1989; Geiwitz, et al., 1990; Miller, 1969; Schweikert, Burton, Taylor, Cortlett, Shadbolt, & Hedgecock, 1987).

An alternative method is the repertory grid approach in which a set of dimensions focuses the ratings. (Boose, 1986; Bradshaw, Ford, Adams-Webber, & Boose, 1993; Cordingley, 1989; Fransella & Bannister, 1977; Gaines & Shaw, 1992; Mancuso & Shaw, 1988; Shaw, 1980; Shaw & Gaines, 1987; 1989; Zachary, et al., 1990). That is, ratings are given for each concept (or element) along each of a set of dimensions (or constructs). So, for instance, the elements may be cars and the constructs along which the cars are rated may be dimensions such as gas mileage, maintenance, and cost. Similarity between a pair of concepts can then be derived by computing the summed difference or correlation between the ratings for each concept.

Once relatedness estimated have been collected they can be summarized using a number of psychometric scaling methods such as MDS (multidimensional scaling; Kruskal, 1977; Kruskal & Wish, 1978; Shepard 1962a,b), Pathfinder network scaling (Schvaneveldt, et al., 1989; Schvaneveldt, 1990) or cluster analysis (Corter & Tversky, 1986; Johnson, 1967; Lewis, 1991, Shepard & Arabie, 1979). MDS results in a spatial layout of concepts along dimensions thought to represent features which differentiate the concepts. Pathfinder on the other hand, results in a graphical structure in which concepts are represented as nodes, and relations as links connecting the nodes. In addition, representations similar to those derived using the methods described above can be generated more directly by having the expert draw the graph or some other representation of a set of concepts (e.g., Olson & Rueter, 1987; Thordsen, 1991). In general, these methods reduce the set of relatedness judgments to a graphical form that is easier to visualize. The resulting representations can then be interpreted qualitatively and quantitatively. For instance, the dimensions represented by MDS can be interpreted qualitatively as features which distinguish concepts, or quantitatively in order to compare two or more individuals according to how concepts are weighted differently along the dimensions.

Conceptual methods have been used to elicit knowledge in order to improve user interface design (McDonald, Dayton, & McDonald, 1988; Roske-Hofstrand & Paap, 1986), guide the development of training programs (Rowe, et al., 1996), and understand expert-novice differences (Cooke & Schvaneveldt, 1988; Gillan, et al., 1992; Housner, et al., 1993a; Schvaneveldt, et al., 1985). They are considered indirect in that experts are not asked to comment directly on domain facts and rules, but instead, this information is inferred through their judgments of conceptual relatedness. Some have argued that these methods result in an overly narrow focus that may not relate to performance (Geiwitz, et al., 1990). However, recent research (Rowe, et al., 1996) has indicated that distinctions between avionics technicians based on Pathfinder network structures of system components corresponded to performance on a verbal troubleshooting task. Other work has investigated the validity and stability of the outcome of these types of measures with generally favorable results (Cooke, 1992a; Cooke, Durso, & Schvaneveldt, 1986; Gammack, 1990; Ricci, Blickensderfer, Cannon-Bowers, & Sagi, 1996; Rowe, et al., 1996; Rowe, Cooke, Neville, & Schacherer, 1992)

Recent research in this area has focused on comparison and assessment using the conceptual structures. For instance, Goldsmith and Davenport (1990) have developed a measure of Pathfinder network similarity based on proportion of shared links and this measure has been used to compare student structures to instructor structures in a classroom context. It is assumed that students who are most like their more experienced counterparts would also be most likely to excel at task performance. Indeed, this assumption has been verified in various classroom domains (Goldsmith, Johnson, &

Acton, 1991; Housner, Gomez, & Griffey, 1993b), and in avionics troubleshooting (Rowe, et al., 1996).

In order to make an assessment based on conceptual structures, it is first necessary to derive a referent structure, an expert or ideal structure against which other structures can be compared. Referents can be derived logically by constructing an ideal network structure based on an analysis of the task or an expert's understanding of the task (e.g., Cooke, et al., 1996). Unfortunately, not all domains lend themselves to a logical analysis and in these cases, it may be best to derive an empirical referent using relatedness judgments from one or more high performers or experts (e.g., Cooke, et al., 1996). Interestingly, there are cases in which knowledge structures are most predictive of performance when assessed by comparison to a high performing intermediate than an expert referent (Rowe, et al. 1996). In other words, the best avionics troubleshooters at beginning levels have knowledge structures that look more like a very good intermediate than an expert troubleshooter

Interpretations of conceptual structures of groups of individuals can be based on an average of the relatedness judgments of the individuals in that group (as long as interparticipant correlations indicate that the group is cohesive). In cases in which there are discrepancies among individuals, major distortions in the representation can result (Ashly, Maddox, & Lee, 1994). In circumstances such as these, the INDSCAL (individual differences scaling) MDS procedure (Carroll & Chang, 1970) can be used, or in the case of network representations, aggregates can be formed by adding or deleting links in the referent network on the basis of that link's presence or absence in the majority of individual networks. Aggregate links can also be weighted according to number of experts who have that link.

Conceptual methods are illustrated in the advising example using pairwise relatedness ratings and Pathfinder network analysis. The resulting network structure is expected to yield information about an advisor's conceptual structure for a set of university courses. In particular, the analysis should reveal courses that are associated to one another and more general structural features that characterize the set.

A procedure for implementing the relatedness ratings and Pathfinder analysis in the context of the advising example is presented in Table 7. In addition, some hypothetical data that may be collected in the course of applying this procedure are listed in Table 8 and Figure 1.

[Insert Tables 7, 8, and Figure 1 about here]

Summary

The groupings described in this section embrace the majority of knowledge elicitation techniques available. However, there are many more techniques per grouping and variations on techniques than could be described in the space of this chapter. The procedures illustrated in the context of the advising application represent only one of many potential approaches.

In general, knowledge elicitation techniques are capable of providing rich information regarding the concepts, relations, facts, rules, and strategies relevant to the domain in question. The techniques differ in terms of their procedures, as well as their emphases on one type of knowledge or another. No technique is guaranteed to result in a complete and accurate representation of an expert's knowledge, although the goal is to model the expert's knowledge, not to extract or reproduce it in its entirety. The major drawback of these methods is that they can be costly. Rich data are associated with lengthy data collection sessions, unwieldy data analysis, and interpretation difficulties. Fortunately, recent methodological developments facilitate the process so that it can be more readily applied to time-critical settings. Additional recent developments in knowledge elicitation are discussed in the next section.

New Directions

Having amassed methods for eliciting knowledge, the field of knowledge elicitation has recently progressed along two fronts. The first involves research and tool development directed at facilitating the process. For example, there has been recent work on integrating multiple knowledge elicitation methods and evaluating existing methods. Along the second front has been work that extends the traditional role of knowledge elicitation to other problems or applications. Specifically, there has been a recent focus on task performance with task analytic and cognitive task analytic methods taking center stage. The most recent work along these two fronts is described in the next two sections.

Enhancing Existing Methods

The review and cataloging efforts of the late 80's and early 90's served to identify the number and variety of knowledge elicitation methods and tools available. It also revealed areas in which additional research and methods were needed. It became increasingly clear that due to the complexity of knowledge and even greater complexity of cognitive skill, that multiple knowledge elicitation methods were probably required for any single application. As a result, research has been directed toward evaluating and comparing methods and devising techniques for managing the results of multiple methods (e.g., Burton, Shadbolt, Rugg, & Hedgecock, 1990; Gaines & Shaw, 1997; Hoffman, 1987).

Evaluative efforts in which individual methods are assessed for reliability and validity, and in which two or more methods are compared have increased in the last decade (e.g., Gammack, 1990; Rowe, et al., 1996). Dhaliwal and Benbasat (1990) describe a framework which places such evaluation in the context of techniques and tools (the independent variables), the quality of the resulting interface and the efficiency of the process (the dependent variables) and various moderator variables. Then, in the context of this framework, they review the evaluative literature and point out difficulties associated with evaluation. The identification of a satisfactory criterion for the evaluation of a knowledge base is not trivial. Some have proposed that knowledge be evaluated in terms that go beyond reliability and validity, and instead examine knowledge structure and function (Reich, 1995) or utility of that knowledge to the target application (Cooke & Rowe, 1997).

There have been several developments directed toward integrating the results from multiple knowledge elicitation techniques. Mengshoel (1995) describes a Knowledge Reformulation Tool (KRF) that makes use of an intermediary language to translate between two techniques. This procedure is illustrated using the repertory grid and card sorting procedures discussed previously. Alternately, Gaines and Shaw (1997) approach the knowledge interchange goal through the World Wide Web. They propose that the Web be used to enhance the integration of techniques typically tied to individual laboratories. They illustrate their approach using repertory grid and conceptual network methods also described previously.

Others have addressed some specific shortcomings of existing knowledge elicitation methods. For instance, many of the methods tend to be biased from the perspective of the elicitor or knowledge engineer, neglecting the perspective of the user or expert (e.g., Hale, Sharpe, & Haworth, 1996; McNeese, Zaff, Citera, Brown, & Whitaker, 1995; Zaff, McNeese, & Snyder, 1993). It is thus argued that the result of elicitation efforts can be biased in the same way.

Zaff, et al. (1993), for instance, describe a methodology called AKADAM (Advanced Knowledge and Design Acquisition Methodology) which integrates three different knowledge elicitation methods, each revealing a distinct perspective of user requirements and intended to elicit knowledge in the form of concepts, rules, and design. The methodology is user-centered in that information is obtained directly from the expert and elaborated by the expert. The goals of AKADAM include (1) shared communication between the knowledge elicitor and the expert, (2) the facilitation of unconstrained knowledge expression, and (3) resulting knowledge representations compatible with needs, capabilities, and limitations of the stakeholders (McNeese, et al., 1995). Similar developments along these lines allow nonprogrammers to edit knowledge structures through the automatic generation of domain-specific knowledge acquisition tools (Eriksson, Puerta, & Musen, 1994).

In sum recent research has provided information for the selection and combination of knowledge elicitation methods from a larger palette. Other efforts have addressed the quality and perspective of the result of elicitation. In addition, new questions have been raised and new methods proposed in response to a broader view of knowledge elicitation.

<u>Broadening the Scope of Knowledge Engineering</u>

As mentioned earlier in this chapter, recent knowledge elicitation efforts have gone beyond the original mission of modeling the knowledge of a single expert, to include cognition and behavior embedded in the context of the actual task. These directions blur distinctions between knowledge elicitation and cognitive engineering or cognitive task analysis, enterprises associated with revealing the cognition underlying complex task performance.

Some of the traditional knowledge elicitation methods (i.e., unstructured interviews conceptual methods, and contrived tasks in general) remove the expert from the task context, thereby focusing on knowledge at the expense of task and contextual information. They therefore run the risk of generating a knowledge-base that is insensitive to context, as opposed to methods that can be applied concurrently with task performance (i.e., observations, some structured interviews, process tracing). Not only have investigators begun to consider methods for eliciting information *in* a broader context, but they are also interested in methods that elicit information *about* the broader context. For instance, methods have been proposed for investigating team as opposed to individual cognition (Cooke, Stout, & Salas, 1997).

Interest in the context surrounding the task has been accompanied by interest in the task itself and in particular, the cognitive and behavioral elements of task performance. This interest was also motivated by the fact that the knowledge elicitation techniques that do consider the task require an initial understanding of it to generate necessary materials (e.g., structured interview questions, task scenarios to perform during think-aloud). Task analysis, which focuses on behavioral aspects of task performance, is one way to satisfy this goal and has long been a mainstay of human factors and applied psychology (Meister, 1989, Kirwan & Ainsworth, 1992). There are numerous variations, but in general, task analysis involves decomposition of the task goal into subgoals (Cooke, 1994; Hoffman, et al., 1995; Wilson, 1989). Investigators have continued to develop new forms of task analysis to meet new goals. For instance, Sutcliffe (1997) has incorporated information needs into the traditional task analysis in order to aid the design of information displays.

Task analysis has also benefited from much of the work associated with the ESDA movement, described previously in the context of process tracing. ESDA is geared toward the understanding of sequential behavior in general such as the actions taken by surgeons in the operating room or the keystrokes and mouse clicks of a new computer user. ESDA methods can help to reveal subtle patterns and contingencies in sequential behavioral data associated with tasks. These data can be composed of verbal behaviors, as well as nonverbal ones such as gestures, eye movements, and user actions recorded by computer logging software. Unlike traditional knowledge elicitation and task analytic methods, methods that focus on computer-recorded events can amass data in the background, posing little threat of interference to task performance. However, the relation between these kinds of behavioral events and knowledge has been questioned. Rowe, et al. (1996) have proposed an approach for exploring the relationship between behavioral patterns and system knowledge (see also, Bailey & Kay, 1987).

Knowledge elicitation methods have also moved beyond the traditional conceptualization of knowledge in terms of concepts, relations, rules and strategies. When attempting to build applications using this "knowledge," it becomes clear that there is more to cognition than "knowledge," and probably much more to knowledge than what is captured by traditional knowledge elicitation. In short, much of cognition (e.g., perception) can be left untapped. Recent methods have been adapted or developed to capture other aspects of cognition such as decision making rules (e.g., Klein, 1989) and communication processes (e.g., Bowers, Braun, & Klein, 1994). The term "cognitive task analysis" also reflects this more general emphasis on cognition.

Conclusion

Knowledge elicitation, once a stage of development of knowledge-bases for expert systems, has evolved into a much more ambitious enterprise. Applications continue to include knowledge-bases in addition to a variety of other applications in training and software design. Methods no longer focus solely on knowledge, but encompass cognitive processes, task-associated behaviors, and task context as well. This new enterprise is more appropriately labeled cognitive engineering and the methods that are now used may be best referred to as cognitive task analysis methods. These methods continue to be useful, if not critical for solving many applied problems and additionally continue to have an impact on more basic research endeavors.

.

References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behavior, 49,* 227-65.

Anderson, J. R. (1995). *Cognitive Psychology and its Implications.* NY: W.H. Freeman & Co.

Ashly, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science, 5,* 144-151.

Bailey, W. A., & Kay, D. J. (1987). Structural analysis of verbal data. In J. M. Carroll & P.P. Tanner (Eds.), *Human Factors in Computing Systems and Graphics Interface* (pp. 297-301). New York: Association for Computing Machinery.

Bainbridge, L. (1979). Verbal reports as evidence of the process operator's knowledge. *International Journal of Man-Machine Studies, 11,* 411-36.

Benysh, D. V., Koubek, R. J., & Calvez, V. (1993). A comparative review of knowledge structure measurement techniques for interface design. *International Journal of Human-Computer Interaction, 5*, 211-237.

Best, J. B. (1995). *Cognitive Psychology.* St. Paul, MN: West Publishing Co.

Boose, J. H. (1986). *Expertise Transfer for Expert System Design.* Amsterdam: Elsevier.

Boose, J. H. (1989). A survey of knowledge acquisition techniques and tools,. *Knowledge Acquisition, 1,* 3-37.

Boose, J. H. & Bradshaw, J. M. (1987) Expertise transfer and complex problems: Using Aquinas as a knowledge-acquisition workbench for knowledge-based systems. *International Journal of Man-Machine Studies, 26,* 3-28.

Boose, J. H., & Gaines, B. R., (Eds.) (1990). *The Foundations of Knowledge Acquisition, Knowledge Based Systems, Vol. 4* San Diego, CA: Academic Press.

Boose, J. H., & Gaines, B. R., (Eds.) (1988). *Knowledge Acquisition Tools for Expert Systems, Knowledge Based Systems Vol. 2.* San Diego, CA: Academic Press.

Bowers, C. A., Braun, C., & Kline, P. B. (1994). Communication and team situational awareness. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.), *Situational awareness in complex systems* (p. 305-311). Daytona Beach, FL: Embry - Riddle Aeronautical University Press.

Bradshaw, J. M., Ford, K. M., Adams-Webber, J.R., & Boose, J. H. (1993). Beyond the repertory grid: New approaches to constructivist knowledge acquisition tool development. *International Journal of Intelligent Systems, 8,* 287-333.

Breuker, J. & Wielinga, B. (1987). Use of models in the interpretation of verbal data. In A. Kidd (Ed.), *Knowledge Acquisition for Expert Systems,* 17-43. NY: Plenum Press.

Broadbent, D. E., Fitzgerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the context of complex systems. *British Journal of Psychology, 77,* 33-50.

Burton, A.M., Shadbolt, N. R., Rugg, G., & Hedgecock, A. P. (1990). The efficacy of knowledge elicitation techniques: A comparison across domains and levels of expertise. *Journal of Knowledge Acquisition, 2,* 167-178.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckhart-Young decomposition. *Psychometrika, 35,* 283-319.

Charness, N. & Schultetus, S. (this volume),. *Knowledge and expertise.* In F. T. Durso (Ed.), *Handbook of Applied Cognition.* London: Wiley.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Cognitive skills and their acquisition* (pp. 141-189). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Feltovich, P. J., & Glaser (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Clarke, B. (1987). Knowledge acquisition for real-time knowledge-based systems. *Proceedings of the First European Workshop on Knowledge Acquisition for Knowledge Based Systems.* September 2-3, Reading University, UK.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247.

Compton, P. & Jansen, R. (1990). A philosophical basis for knowledge acquisition. *Knowledge Acquisition, 2,* 241-257.

Cooke, N.M. (1989). The elicitation of domain-related ideas: Stage one of the knowledge acquisition process. In C. Ellis (Ed.), *Expert Knowledge and Explanation:.* England: Ellis Horwood Limited, 58-75.

Cooke, N. J. (1992a). Predicting judgment time from measures of psychological proximity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 640-653.

Cooke, N. J. (1992b). Eliciting semantic relations for empirically derived networks. *International Journal of Man-Machine Studies, 37,* 721-750.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41,* 801-849.

Cooke, N. J., & Breedin, S. D. (1994). Constructing naive theories of motion on-the-fly. *Memory and Cognition, 22,* 474-493.

Cooke, N.M, Durso, F.T., & Schvaneveldt, R.W. (1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12 (4)*, 538-549.

Cooke, N. J., & Neville, K. J. , & Rowe, A. L. (1996). Procedural network representations of sequential data. *Human-Computer Interaction, 11,* 29-68.

Cooke, N. J., & Rowe, A. L. (1997). Measures of mental models: A synthesis of evaluative data. *Training Research Journal, 3,* 185-207.

Cooke, N. M. & Schvaneveldt, R. W. (1988). Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies, 29,* 407-427.

Cooke, N. J., Stout, R., & Salas, E. (1997) Expanding the measurement of situation awareness through cognitive engineering methods, *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting,* 215-219.

Cordingley, E. S. (1989). Knowledge elicitation techniques for knowledge-based systems. In D. Diaper (Ed.), *Knowledge Elicitation: Principles, Techniques, and Applications* (pp. 89-175). NY: John Wiley & Sons.

Corter, J. E., & Tversky, A. (1986). Extended similarity trees. *Psychometrika, 51,* 429-451.

Cullen, J. & Bryman, A. (1988). The knowledge acquisition bottleneck: A time for reassessment? *Expert Systems, 5,* 216-225.

deGroot, A. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem solving.* New York: Wiley.

Dhaliwal, J. S., & Benbasat, I. (1990). A framework for the comparative evaluation of knowledge acquisition tools and techniques. *Knowledge Acquisition, 2,* 145-166.

Diaper, D. (Ed.) (1989a). *Knowledge Elicitation: Principles, Techniques, and Applications*. England: Ellis Horwood Limited.

Diaper, D. (1989b). Designing expert systems--From Dan to Beersheba. In D. Diaper (Ed.), *Knowledge Elicitation: Principles, Techniques, and Applications* (pp. 17-46). England: Ellis Horwood Limited.

Diederich, J., Ruhmann, I., & May, M. (1987). KRITON: A knowledge acquisition tool for expert systems. *International Journal of Man-Machine Studies, 26,* 29-40.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning.* Cambridge, MA: Harvard Press.

Engle, R. W. & Bukstel, L. (1978). Memory processes among bridge players of differing expertise. *American Journal of Psychology, 91,* 673-689.

Ericsson, K. A., & Simon, H. A. (1996). *Protocol Analysis: Verbal Reports as Data* (revised edition). Cambridge, MA: MIT press.

Eriksson, H., Puerta, A. R., & Musen, M. A. (1994). Generation of knowledge acquisition tools from domain ontologies. *International Journal of Human-Computer Studies, 41,* 425-453.

Feigenbaum, E. A. (1989). What hath Simon wrought? In D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon* (pp. 165-182). Hillsdale, NJ: Erlbaum.

Ford, K. M., & Adams-Webber, J. R. (1992). Knowledge acquisition and constructivist epistemology. In R. R. Hoffman (Ed.), *The Psychology of Expertise: Cognitive Research and Empirical AI* (pp. 121-136). New York: Springer-Verlag.

Ford, J. M., & Wood, L. E. (1992). Structuring and documenting interactions with subject-matter experts. *Performance Improvement Quarterly, 5,* 2-24.

Fransella, F. & Bannister, D. (1977). *A manual for repertory grid technique.* London: Academic Press.

Gaines, B. R., & Shaw, M. L. G. (1992). Knowledge acquisition tools based on personal construct psychology. *Knowledge Engineering Review: Special Issue on Automated Knowledge Acquisition Tools.*

Gaines, B. R., & Shaw, M. L. G. (1997). Knowledge acquisition, modelling, and inference through the World Wide Web. *International Journal of Human-Computer Studies, 46,* 729-759.

Gammack, J. G. (1990). Expert conceptual structure: The stability of Pathfinder representations. In R. Schvaneveldt, Ed., *Pathfinder Associative Networks: Studies in Knowledge Organization,* pp. 213-226. Norwood, NJ: Ablex.

Geiwitz, J., Klatsky, R. L., & McCloskey, B. P. (1988). *Knowledge Acquisition for Expert Systems: Conceptual and Empirical Comparisons.* Santa Barbara, CA: Anacapa Sciences, Inc.

Geiwitz, J., Kornell, J., & McCloskey, B. (1990). *An Expert System for the Selection of Knowledge Acquisition Techniques.* (Technical Report 785-2) Santa Barbara, CA: Anacapa Sciences, Inc.

Gillan, D. J., Breedin, S. D., & Cooke, N. J. (1992). Network and multidimensional representations of the declarative knowledge of human-computer interface design experts. *International Journal of Man-Machine Studies, 36,* 587-615.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, and M. J. Farr (Eds.), *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Erlbaum.

Goldsmith, T. E., & Davenport, D. M. (1990). Assessing structural similarity of graphs. In R. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organization* (pp. 75-87). Norwood, NJ: Ablex.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83,* 88-96.

Greeno, J. G. (1978). Natures of problem-solving abilities. In W. K. Estes (Ed.), *Handbook of Learning and Cognitive Processes* (Vol. 5, pp. 239-270). Hillsdale, NJ: Erlbaum.

Grover, M. D. (1983). A pragmatic knowledge acquisition methodology. *Proceedings of the 8th International Joint Conference on Artificial Intelligence* (pp. 436-438), Karlsruhe, West Germany, August

Hale, D. P., Sharpe, S., & Haworth, D. A. (1996). Human-centered knowledge acquisition: a structural learning theory approach. *International Journal of Human-Computer Studies, 45,* 381-396.

Hall, E. M., Gott, S. P. & Pokorny, R. A. (1994). *A Procedural Guide to Cognitive Task Analysis: The PARI Methodology.* Manuscript submitted for publication. Brooks AFB, TX.

Harrison, B. L., & Baecker, R. M. (1991). Designing video automation and analysis systems. *Proceedings Graphics Interface '92,* 209-219.

Hart, A. (1986). *Knowledge Acquisition for Expert Systems.* London: Kogan Page.

Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine, 8,* 53-67.

Hoffman, R. R. (1989). A survey of methods for eliciting the knowledge of experts. *SIGART Newsletter, 108.*

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes, 62*, 129-158.

Housner, L. D., Gomez, R. L., & Griffey, D. C. (1993a). A Pathfinder analysis of expert-novice knowledge. *Proceedings of the Meeting on the National American Alliance for Health, Physical Education, Recreation and Dance, 64,* Supplement, A-88 - A-89.

Housner, L. D., Gomez, R. L., & Griffey, D. C. (1993b). Pedagogical knowledge growth in prospective teachers: Relationships to performance in a teaching methodology course. *Research Quarterly for Exercise and Sport, 64,* 167-177.

Howell, W. C., & Cooke, N. J. (1989). Training the human information processor: A look at cognitive models. In I. Goldstein (Ed.), *Training and Development in Work Organizations: Frontier Series of Industrial and Organizational Psychology, Volume 3,* New York: Jossey Bass, 121-182.

Hutchins, E. (1995). *Cognition in the Wild.* Cambridge, MA: MIT Press.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika, 32,* 241-254.

Johnson, L. & Johnson, N. (1987). Knowledge elicitation involving teachback interviewing. In Kidd, A. (Ed.) *Knowledge Elicitation for Expert Systems: A Practical Handbook* (pp. 91-108). NY: Plenum Press.

Johnson, P. E., Zualkerman, I, & Garber, S. (1987). Specification of expertise. *International Journal of Man-Machine Studies, 26,* 161-81.

Kirwan, B., & Ainsworth, L. K. (1992). *A Guide to Task Analysis.* London: Taylor & Francis.

Kitto, C. M., & Boose, J. H. (1989). Selecting knowledge acquisition tools and strategies based on application characteristics, *International Journal of Man-Machine Studies, 31,* 149-160.

Klein, G. A. (1989). Recognition-primed decisions. In W. Rouse (Ed.), *Advances in Man-Machine Systems Research, 5,* 47-92.

Klein, G. A., Calderwood, R., & MacGregor, D. (1989) Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, & Cybernetics, 19,* 462-472.

Kruskal, J. B. (1977). Multidimensional scaling and other methods for discovering structure. In Enslein, Ralston, & Wilf (Eds.), *Statistical Methods for Digital Computers*. NY: Wiley.

Kruskal, J. B. & Wish, M. (1978). *Multidimensional Scaling*. Sage University. Paper Series on Quantitative Applications in the Social Sciences, #07-011. London: Sage Publications.

LaFrance, M. (1987). The knowledge acquisition grid: A method for training knowledge engineers. *International Journal of Man-Machine Studies, 26,* 245-255.

LaFrance, M. (1992). Excavation, capture, collection, and creation: Computer scientists' metaphors for eliciting human expertise. *Metaphor and Symbolic Activity, 7,* 135-156.

Lewis, S. (1991). Cluster analysis as a technique to guide interface design. *International Journal of Man-Machine Studies, 35,* 251-265.

Mancuso, J. C., & Shaw, M. L. G. (1988). *Cognition and Personal Structure: Computer Access and Analysis.* New York: Praeger.

McDonald, J.E., Dayton, T., & McDonald, D. R. (1988). Adapting menu layout to tasks. *International Journal of Man-Machine Studies, 28,* 417-436.

McGraw, K. L., & Harbison-Briggs, K. (1989). *Knowledge Acquisition: Principles and Guidelines.* Englewood Cliffs, NJ: Prentice Hall.

McNeese, M. D., Zaff, B. S., Citera, M., Brown, C. E., & Whitaker, R. (1995). AKADAM: Eliciting user knowledge to support participatory ergonomics. *International Journal of Industrial Ergonomics, 15,* 345-363.

Meister, D. (1989). *Conceptual Aspects of Human Factors.* Baltimore, MD: The Johns Hopkins University Press.

Mengshoel, O. J. (1995). A reformulation technique and tool for knowledge interchange during knowledge acquisition. *International Journal of Human-Computer Studies, 43,* 177-212.

Meyer, M. A. (1992). How to apply the anthropological technique of participant observation to knowledge acquisition for expert systems. *IEEE Transactions on Systems, Man, and Cybernetics, 22,* 983-990.

Meyer, M. A., & Booker, J. M. (1990). *Eliciting and Analyzing Expert Judgment: A Practical Guide.* Technical report no. NUREG/CR-5424; LA-11667-MS. Los Alamos National Laboratory, Los Alamos, NM.

Miller, G. A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology, 6,* 169-91.

Minsky, M. (1975). A framework for representing knowledge. In J. Haugeland (Ed.), *Mind Design*, (pp. 95-128). Cambridge, MA: The MIT Press.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on x-ray image processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14,* 553-557.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Norman, D. A. (1986). Cognitive engineering. In D.A. Norman & S. W. Draper (Eds.), *User Centered System Design* (pp. 31-61). Hillsdale, NJ: Erlbaum.

Olson, J. R., & Biolsi, K. J. (1991). Techniques for representing expert knowledge. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise* (pp. 240-285). Cambridge: Cambridge University Press.

Olson, J. R. & Rueter, H. H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems, 4,* 152-168.

Pidgeon, N. F., Turner, B. A., & Blockley, D. I. (1991). The use of Grounded Theory for conceptual analysis in knowledge elicitation. *International Journal of Man-Machine Studies, 35,* 151-173.

Quillian, M. R. (1969). The teachable language comprehender. *Communications of the Association for Computing Machinery, 12,* 459-476.

Randel, J. M., Pugh, H. L., & Wyman, B. G. (1996). *Methods for Conducting Cognitive Task Analysis for a Decision Making Task.* Technical report Navy Personnel Research and Development Center, TN-96-10, San Diego, CA.

Regoczei, S. B., & Hirst, G. (1992). Knowledge and knowledge acquisition in the computational context. In R. R. Hoffman (Ed.), *The Psychology of Expertise* (pp. 12-25), New York: Springer-Verlag.

Reich, Y. (1995). Measuring the value of knowledge. *International Journal of Human-Computer Studies, 42,* 3-30.

Reitman, J. (1976). Skilled perception in GO: Deducing memory structures from inter-response times. *Cognitive Psychology, 8,* 336-356.

Ricci, K. E., Blickensderder, E., Cannon-Bowers, J. A., & Sagi, C. (1996). Validity and stability of Pathfinder Indices: A replication and extension. Paper presented at the Annual meeting of the American Psychological Association, Toronto., Canada.

Roediger, H. L. (1990). Implicit memory. *American Psychologist, 45*, 1043-1056.

Roske-Hofstrand, R. J., & Paap, K. R. (1986). Cognitive networks as a guide to menu organization: An application in the automated cockpit. *Ergonomics, 39,* 1301-1311.

Roske-Hofstrand, R. J., & Paap, K. R. (1990). Discriminating between degrees of low or high similarity: Implications for scaling techniques using semantic judgments. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organization* (pp. 61-73). Norwood, NJ: Ablex.

Rowe, A. L., Cooke, N. J., Hall, E. P., & Halgren, T. L. (1996). Toward an on-line knowledge assessment methodology: Building on the relationship between knowing and doing. *Journal of Experimental Psychology: Applied, 2,* 31-47.

Rowe, A. L., Cooke, N. J., Neville, K. J., Schacherer, C. W. (1992). Mental models of metal models: A comparison of mental model measurement techniques. *Proceedings of the Human Factors Society 36th Annual Meeting,* 1195-1199.

Sachett, G. P. (1977). *Observing Behavior* (vol.1) Baltimore: University Park Press.

Sachett, G. P. (1978). *Observing Behavior* (vol.2) Baltimore: University Park Press.

Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human-Computer Interaction, 9,* 251-317.

Sanderson, P. M., James, J. M., Seidler, K. (1989). *SHAPA: An interactive Software Environment for Protocol Analysis.* EPRL Technical Report no. EPRL-89-08. University of Illinois at Urbana-Champaign, Urbana, IL.

Sanderson, P. M., McNeese, M. D., & Zaff, B. S. (1994). Handling complex real-world data with two cognitive engineering tools: COGENT and MacSHAPA. *Behavior Research Methods, Instruments, and Computers, 26,* 117-124.

Sanderson, P. Scott, J., Johnston, T., Mainzer, J., Watanabe, L., & James, J. (1994). MacSHAPA and the enterprise of exploratory sequential data analysis (ESDA). *International Journal of Human-Computer Studies, 41,* 633-681.

Schacter, R. D., & Heckerman, D. E. (1987). Thinking backward for knowledge acquisition. *The AI Magazine*, 55-61.

Scherer, K. R., & Ekman, P. (1982). *Handbook of Methods in Nonverbal Behavior Research.* Cambridge, UK: Cambridge University Press.

Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1989). Network structures in proximity data. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 24, pp. 249-284). New York: Academic Press.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., & DeMaio, J. C. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies, 23,* 699-728.

Schweikert, R., Burton, A. M., Taylor, N. K., Cortlett, E. N., Shadbolt, N. R., & Hedgecock, A. P. (1987). Comparing knowledge elicitation techniques: A case study. *Artificial Intelligence Review, 1*, 245-53.

Shadbolt, N. & Burton, M.(1990). Knowledge elicitation. In J.R. Wilson & E. N. Corlett (Eds.), *Evaluation of Human Work: A Practical Ergonomics Methodology* (pp. 321-345). London: Taylor & Francis.

Shaw, M. L. G. (1980). *On Becoming a Personal Scientist.* London: Academic Press.

Shaw, M. L. G., & Gaines, B. R. (1987). An interactive knowledge elicitation technique using personal construct technology. In Kidd, A. (Ed.), *Knowledge Acquisition for Expert Systems: A Practical Handbook.* NY: Plenum Press.

Shaw, M. L. G., & Gaines, B. R. (1989). Comparing conceptual structures: Consensus, conflict, correspondence, and contrast. *Knowledge Acquisition, 1,* 341-363.

Shaw, M. L. G., & Woodward, J. B. (1989). Mental models in the knowledge acquisition process. *Proceedings of the 4th AAAI-sponsored knowledge acquisition for knowledge-based systems workshop,* 29/1-29/24. Banff, Canada, October.

Shaw, M. L. G., & Woodward, J. B. (1990). Modeling expert knowledge. *Knowledge Acquisition, 2,* 179-206.

Shepard, R. N. (1962a). Analysis of proximities: Multidimensional scaling with an unknown distance function. I *Psychometrika*, *27,* 125-140.

Shepard, R. N. (1962b). Analysis of proximities: Multidimensional scaling with an unknown distance function. II *Psychometrika*, *27,* 219-246.

Shepard, R. N., & Arabie, P. (1979). Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review, 86,* 87-123.

Siochi, A. C., & Ehrich, R. W. (1991). Computer analysis of user interfaces based on repetition in transcripts of user sessions. *ACM Transactions on information systems, 9,* 309-335.

Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28,* 1-16.

Smith, E. E., Shoben, E. J., & Rips, L. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214-241.

Suchman, L. & Trigg, R. (1991). Understanding practice: Video as a medium for reflection and design. In J. Greenbaum & M. Kyng (Eds.), *Design at work.* Hillsdale, NJ: LEA.

Suen, H. K., & Ary, D. (1989). *Analyzing Quantitative Behavioral Observation Data.* Hillsdale, NJ: Erlbaum.

Sundstrom, G. A. (1991). Process tracing of decision making: An approach for analysis of human-machine interactions in dynamic environments. *International Journal of Man-machine Studies, 35,* 843-858.

Sutcliffe, A. (1997). Task-related information analysis. *International Journal of Human-Computer Studies, 47,* 223-257.

Thordsen, M. L. (1991). A comparison of two tools for cognitive task analysis: Concept mapping and the critical decision method. *Proceedings of the Human Factors Society 35th Annual Meeting,* San Francisco, CA, 283-285.

VanHooff. J.A.R.A.M. (1982). Categories and sequences of behavior: Methods of description and analysis. In K.R. Scherer & P. Ekman (Eds.), *Handbook of Methods in Nonverbal Behavior Research,* Cambridge: Cambridge University Press.

vanSomeren, M. W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes.* London: Academic Press.

Vicente, K. J. (1997). Heeding the legacy of Meister, Brunswick, and Gibson: Toward a broader view of human factors research. *Human Factors, 39,* 323-328.

Welbank, M. (1990). An overview of knowledge acquisition methods. *Interacting with computers, 2,* 83-91.

Wielinga, B. J., Schreiber, A. Th., & Breuker, J. A. (1992). KADS: A modelling approach to knowledge engineering. *Knowledge Acquisition, 4,* 5-53.

Wilson, M. (1989). Task models for knowledge elicitation. In D. Diaper, *Knowledge elicitation: Principles, techniques, and applications* (pp. 197-219). Chichester: Ellis Horwood Limited.

Woods, D. D., & Roth, E. M. (1988). Cognitive systems engineering. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 3-43). Amsterdam: Elsevier.

Zachary, W. W., Ryder, J. M., & Purcell, J. A. (1990). A computer-based tool to support mental modeling for human-computer interface design. *CHI systems technical report no. 900831-8908*, August 31.

Zaff, B. S., McNeese, M. D., & Snyder, D. E. (1993). Capturing multiple perspectives: A user-centered approach to knowledge and design acquisition. *Knowledge Acquisition, 5,* 79-116.

Footnotes

1 This chapter benefited from the thoughtful comments of Frank Durso, Robert Hoffman, Thomas Seamster, and several students in Frank Durso's Fall 1997 course on ****.

2 The term "cognitive engineering" is not new. It was introduced by Don Norman (1986) in the context of designing human-computer interfaces. More recently this term has been broadly adopted by those who address applied problems in design and training in which issues of human cognition are critical. This work is also referred to as "applied cognitive psychology" and "cognitive ergonomics," however the Human Factors and Ergonomics Society technical group formed in 1996 refers to itself as the "Cognitive Engineering and Decision Making" group.

Table 1. Procedure for using naturalistic passive observation to elicit advising information.

---

1) Select advising experts and advisees and obtain consent from both to observe the advising sessions.

2) Identify a room suitable for observations (i.e., natural like the advisor's office). In addition the setting should allow the observer to be positioned nonintrusively (e.g., in the back corner outside the field of view, or behind a one-way mirror).

3) Observe 2-3 advising sessions for each of 2-3 different advising experts.

4) Take notes during the sessions. Record the basic events comprising the session with particular attention to topics discussed, questions raised, and problems encountered.

5) Summarize notes by listing events, topics, questions, and problems and any other type of item that may help define the scope and functionality of the expert system.

6) For each item, recorded (e.g., topic: prerequisites, career guidance), note the frequency with which it was mentioned. This could be the overall number of times it was mentioned or number of sessions in which it was mentioned. The latter measure better controls for talkative dyads.

7) Generate functional requirements for the expert system on the basis of these results.

Table 2. A hypothetical sample of data collected using the procedure outlined in Table 1.

Notes taken during a segment of one advising session (Step #4):

• Professor D. of History greets the student.

• The student requests some help with selecting from three potential history electives.

• Professor D. refreshes his memory on HIST 250 by reviewing a recent memo from the instructor of that course.

• Professor D. summarizes the content of the three courses and asks the student to state how each corresponds to her interests in European History.

Summary of Notes (Steps # 5 & 6):

• Professor Events: greeting (1), information seeking (1), provide course summary (1), probes student interests (1)

• Questions: Selecting electives (1)

Table 3. Using a forward scenario simulation to elicit advising information.

---

1) Develop a set of scenarios that represent cases with which the expert advisor typically deals. Each scenario should specify the case in as much detail as possible (i.e., student's background, scheduling details, course availability, etc.). This information may be constructed with the aid of another expert or from information recorded from actual advising sessions (the procedure in Table 1, for example).
2) For each scenario, create an initial problem statement in which only some of the case-relevant information is presented. The remaining information will be available only upon request by the expert.
3) Pre-test the scenarios with other experts to determine (as much as possible) whether any critical information has been left unspecified.
4) Enlist the participation and consent of several expert advisors.
5) Describe the forward scenario simulation method to each expert using an example from another area (e.g., career counseling, tax advising).
6) Present the initial information from one scenario to the expert.
7) Record the expert's comments (video, audio, or pen and paper), explicitly noting the information requested from the elicitor.
8) Present information to the expert as requested, recording the information presented.
9) Repeat the steps 6-8 across the entire set of scenarios (the number depends on the scope of the cases that are targeted).
10) Repeat interviews across all experts. The number of experts depends on their availability and the degree of variability in the responses. If experts are generally in agreement regarding information requested, then little will be gained from additional interviews.

11) List the information requested and the advice given across all experts and interviews.

12) Organize the list into categories of information separating information requested from advice and noting any sequential dependencies.

13) If there are any questions about certain categories, interview additional experts, focusing on these issues.

14) Generate if-then rules.

15) Show these rules to an expert for verification.

Table 4. A hypothetical sample of data collected using the procedure outlined in Table 3.

Interview Segment (Steps 6 & 8)

Elicitor: The advisee would like to switch majors from history to psychology and would like to know what courses are now required for a B.A.

Expert: Well...I need to know the year in which the advisee entered the University.

Elicitor: Why?

Expert: Because the requirements have changed over the years and the year of admission determines the requirements for each individual.

Elicitor: OK, The advisee was admitted in 1997. She is currently finishing her first year.

Expert: Then I need to know what courses she has had taken this year.

List of information (Step 11)

Information requested: year of admission, courses taken

Advice given: none at this point

Table 5. Using retrospective think-aloud verbal reports to elicit advising information.

---

1) Determine how many sessions will be recorded. For the most complete coverage of domain rules, the number of advising sessions is not as critical as the range of cases represented in the set of sessions. With inadequate foresight into this issue it may be most prudent to video tape more sessions than are ultimately analyzed.

2) Select advisors and advisees and obtain consent to be tested and video-taped.

3) Arrange equipment and room for advising sessions to be videotaped.

4) For each session, explain purpose to advisor and advisee and then record on video the advising session. The experimenter should remain nonintrusive, but may take notes on obvious rules or strategies used if possible.

5) Invite advisor back at a later time to view the tape while thinking aloud.

6) Give each advisor practice in the think-aloud technique using an unrelated task like mental addition, emphasizing the difference between reporting thoughts and reporting explanations of those thoughts.

7) Start the video tape.

8) Be prepared to remind the advisor to keep talking, but in all other ways the experimenter should remain nonintrusive. Simple prompts like "keep talking" or "think aloud please" are usually adequate.

9) Record the verbal protocol (audio tape is sufficient).

10) Repeat Steps 6-9 with the appropriate advisors for the set of representative interviews.

11) Using a protocol analysis tool like MacSHAPA (Sanderson, Scott, et al., 1994), enter the verbal protocol text and begin to generate labels for categories that focus on variations in rules and strategies.

12) Iteratively refine the coding scheme using progressively more segments of verbal protocol.

13) Identify rules and strategies that were applied. The analysis software may reveal frequent event transitions that may also reflect rules.

14) Compile the list of rules and have an expert advisor verify their accuracy.

Table 6. A hypothetical sample of data collected using the procedure outlined in Table 5.

Transcribed Protocol Segment (Steps 7-9)

OK...This is the part where the student tells me what courses they plan to take. I usually ask the student to back up and tell me about their goals (that is, their major, expected degree date, career visions, etc). So...now the student tells me that they are a computer science major interested in a high-paying career in software design and that they hope to graduate in one year. The "one-year" plan immediately suggests to me that I better do a quick degree check to make sure that the requirements for the degree have been met. I do this on my computer. This is what I'm doing now ....meanwhile, I'm thinking that this student needs to first complete the requirements for their degree before pursuing the electives they've planned to take next semester. Here...this is really typical...the student has forgotten the foreign language requirement and therefore needs to enroll in an introductory foreign language course. The student looks despondent and I'm thinking that maybe there is some other way to satisfy this requirement....

Segmented and Coded Protocol Sample (Step 11, code is capitalized)

CONDITION: STUDENT GIVES INFO --> COURSE PLAN

OK...This is the part where the student tells me what courses they plan to take.

ACTION: ASK ABOUT STUDENT GOALS

I usually ask the student to back up and tell me about their goals (that is, their major, expected degree date, career visions, etc).

CONDITION: STUDENT GIVES INFO-->CS MAJOR, SOFTWARE DESIGN CAREER, ONE YEAR ANTICIPATED GRADUATION

So...now the student tells me that they are a computer science major interested in a high-paying career in software design and that they hope to graduate in one year.

ACTION: DO DEGREE CHECK (cued by one-year for expected degree)

The "one-year" plan immediately suggests to me that I better do a quick degree check to make sure that the requirements for the degree have been met. I do this on my computer. This is what I'm doing now ....meanwhile,

THOUGHTS: NEED FOR DEGREE CHECK

I'm thinking that this student needs to first complete the requirements for their degree before pursuing the electives they've planned to take next semester.

CONDITION: INCOMPLETE DEGREE CHECK --> MISSING LANGUAGE REQUIREMENT

Here...this is really typical...the student has forgotten the foreign language requirement and therefore needs to enroll in an introductory foreign language course.

ACTION (IMPLIED): RECOMMEND COURSE -->INTRO FOREIGN LANGUAGE

CONDITION: STUDENT NONVERBAL -->DESPONDENT

The student looks despondent and

THOUGHTS: POTENTIAL ALTERNATIVES

I'm thinking that maybe there is some other way to satisfy this requirement....

<u>Rules and Strategies (Step 13)</u>

If student initially describes course plan then ask student about goals.

If the student expresses an anticipated degree in a year or less do a degree check.

If degree check is incomplete, then suggest a course to complete it.

If the student is despondent try to find alternatives.

Table 7. Using relatedness ratings and Pathfinder network scaling to elicit advising information.

---

1) Locate and obtain consent from several advisors (around 6-10). Although this analysis can be performed on ratings obtained from a single advisor, it may be interesting to collect ratings from multiple advisors and examine cross-advisor differences.
2) Generate a list of 25-30 courses that are representative of the university courses within the advisors' domains of expertise with the aid of another advisor or from records of advising sessions.
3) Obtain or write a computer program to randomly present pairs of courses to each advisor. It is typical to present each pair once (in one direction only) resulting in a total of (N(N-1))/2 pairs, where N is equal to the total number of courses. The presentation order of courses in each pair should also be counterbalanced across advisors.
4) For each advisor, first present the complete list of courses to the advisor so that the scope of the courses is clear. During this step, identify any courses with which the advisor is unfamiliar. Several unfamiliar courses, especially across multiple advisors may indicate nonrepresentative courses.
5) Have advisors each rate the pairs for relatedness on a scale that runs from unrelated to related (a 5 to 10 point scale is typical). Sometimes a discrete option of "unrelated" is also included, as it has been shown that individuals do not discriminate well at the unrelated end of the scale (Roske-Hofstrand & Paap, 1990).
6) For each advisor (and as interparticipant correlation warrants, for the advisors as a group), submit the relatedness ratings (or the mean ratings in the case of the group) to Pathfinder (KNOT software).
7) Use default parameters that relate to minimal network complexity (i.e., *r*=infinite and *q*=number of concepts -1). These can be altered under specific conditions (see Schvaneveldt, 1990).

8) Specify data type (similarities or distances depending on rating program data format).

9) Examine the resulting network (s), moving nodes as necessary using the KNOT tools to make the graph more legible.

10) Using the similarity metric in the KNOT program, quantitatively compare the advisor networks. In addition compare them visually for qualitative differences.

11) With the aid of an expert advisor, attempt to label the links with a specific relation

12) Identify common features that relate courses and any structural properties of the set of courses.

Table 8. A hypothetical sample of data collected using the procedure outlined in Table 7.

A Sample of courses (Step 2)

HIS 201: American History

LANG 150: Spanish

HIS 301: European History

PSY 310: Experimental Psychology

PSY 325: Abnormal Psychology

CHEM 300: Organic Chemistry

PSY 201: Intro to Psychology

CS 151: Intro to Computer Science

MATH 315: Calculus

PSY 390: Human-Computer Interaction

An Example of a Pair and Rating Scale (Step 5)

HIS 201: Amer His

PSY 201: Intro Psy

slightly related								highly related	
1	2	3	4	5	6	7	8	9	10

|^|

Figure Captions

Figure 1. Sample Pathfinder network based on concepts in Table 8. In it you can see some prerequisite structure in the node-link sequences. In addition, some of the more central nodes are associated with courses that are more interdisciplinary.