Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. (Eds.). *Cambridge handbook of expertise and expert performance* (pp. 203-222). New York: Cambridge University Press.

CHAPTER 12

Eliciting and Representing the Knowledge of Experts

Robert R. Hoffman & Gavan Lintern

Keywords: knowledge elicitation, expert systems, intelligent systems, methodology, Concept Maps, Abstraction-Decomposition, critical decision method

Introduction

The transgenerational transmission of the wisdom of elders via storytelling is as old as humanity itself. During the Middle Ages and Renaissance, the Craft Guilds had well-specified procedures for the transmission of knowledge, and indeed gave us the developmental scale that is still widely used: initiate, novice, apprentice, journeyman, expert, and master (Hoffman, 1998). Based on interviews and observations of the workplace, Denis Diderot (along with 140 others, including Emile Voltaire) created one of the great works of the Enlightenment, the 17 volume *Encyclopedie* (Diderot, 1751–1772), which explained many "secrets" – the knowledge and procedures in a number of tradecrafts. The emergent science of psychology of the 1700s and 1800s also involved research that, in hindsight, might legitimately be regarded as knowledge elicitation (KE). For instance, a number of studies of reasoning were conducted in the laboratory of Wilhelm Wundt, and some of these involved university professors as the research participants (Militello & Hoffman, forthcoming). In the decade prior to World War I, the stage was set in Europe for applied and industrial psychology; much of that work involved the systematic study of proficient domain practitioners (see Hoffman & Deffenbacher, 1992).

The focus of this chapter is on a more recent acceleration of research that involves the elicitation and representation of expert knowledge (and the subsequent use of the representations, in design). We lay out recent historical origins and rationale for the work, we chart the developments during the era of first-generation expert systems, and then we proceed to encapsulate our modern understanding of and approaches to the elicitation, representation, and sharing of expert knowledge. Our emphasis in this chapter is on methods and methodological issues.

Where This Topic Came From

The Era of Expert Systems

The era of expert systems can be dated from about 1971, when Edward Feigenbaum and his colleagues (Feigenbaum, Buchanan, & Lederberg, 1971) created a system in which a computable knowledge base of domain concepts was integrated with an inference engine of procedural (if-then) rules. This "expert system" was intended to capture the skill of expert chemists regarding the interpretation of mass spectrograms. Other seminal systems were MYCIN (Shortliffe, 1976), for diagnosing bacterial infections and PROSPECTOR (Duda, Gaschnig, & Hart, 1979), for determining site potential for geological exploration.

It seemed to take longer for computer scientists to elicit knowledge from experts than to write the expert system software. This "knowledge acquisition bottleneck" became a salient problem (see Hayes-Roth, Waterman, & Lenat, 1983). It was widely discussed in the computer science community (e.g., McGraw & Harbison-Briggs, 1989; Rook & Croghan, 1989). An obvious suggestion was that computer systems engineers might be trained in interview techniques (Forsyth & Buchanan, 1989), but the bottleneck also spawned the development of automated knowledge acquisition "shells." These were toolkits for helping domain experts build their own prototype expert systems (for a bibliography, see Hoffman, 1992).

By use of a shell, experts entered their expert knowledge about domain concepts and reasoning rules directly into the computer as responses to questions (Gaines & Boose, 1988). Neale (1988) advocated "eliminating the knowledge engineer and getting the expert to work directly with the computer" (p. 136) because human-on-human KE methods (interviews, protocol analysis) were believed to place an "unjustified faith in textbook knowledge and what experts say they do" (p. 135).

The field of expert systems involved literally thousands of projects in which expert knowledge was elicited (or acquired) (Hoffman, 1992), but serious problems soon arose. For example, software brittleness (breakdowns in handling atypical cases) and explanatory insufficiency (a printout of cryptic procedural rules fails to clearly express to non-programmers the reasoning path that was followed by the software) were quickly recognized as troublesome (for reviews that convey aspects of the history of this field, see David, Krivine, & Simmons, 1993; Raeth, 1990). At the same time, there was a burgeoning of interest in expertise on the part of cognitive psychologists.

Expertise Studies in Psychology

The application of cognitive science and the psychology of learning to topics in instructional design led to studies of the basis for expertise and knowledge organization at different stages during acquisition of expertise (Lesgold, 1994; Means & Gott, 1988). In the early 1970s, a group of researchers affiliated with the Learning Research and Development Center at the University of Pittsburgh and the Psychology Department at Carnegie-Mellon University launched a number of research projects on issues of instructional design in both educational contexts (e.g., elementary-school-level mathematics word problems; college-level physics problems) and technical contexts of military applications (e.g., problem solving by electronics technicians) (e.g., Chi, Feltovich, & Glaser, 1981) Lesgold et al., 1981. The research emphasized problem-solving behaviors decomposed as "learning hierarchies" (Gagné & Smith, 1962), that is, sequences of learning tasks arranged according to difficulty and direction of transfer.

Interest in instructional design quickly became part of a larger program of investigation that generated several foundational notions about the psychology of expertise (see Glaser, 1987). A number of researchers, apparently independently of one another, began to use the term "cognitive task analysis" both to refer to the process of identifying the knowledge and strategies that make up expertise for a particular domain

P1: JzG
052184097Xc12 CB1040B/Ericsson 0 521 84087 X March 23, 2006 9:43

and task as well as to distinguish the process from so-called behavioral task analysis (e.g., Glaser et al., 1985; see Schraagen, this volume). A stream of psychological research evolved that shifted emphasis from studies with naive, college-aged "subjects" who participated in artificial tasks using artificial materials (in service of control and manipulation of variables) to studies in which highly skilled, domain-smart participants engaged in tasks that were more representative of the complexity of the "real world" in which they practiced their craft (Chi, Glaser, & Farr, 1988; Hoffman, 1992; Knorr-Cetina & Mulkay, 1983; Shanteau, 1992).

Investigators began to shift their attention from cataloging biases and limitations of human reasoning in artificial and simple problems (e.g., statistical reasoning puzzles, syllogistic reasoning puzzles) to the exploration of human capabilities for making decisions, solving complex problems, and forming mental models (Gentner & Stevens, 1983; Klahr & Kotovsky, 1989; Klein & Weitzenfeld, 1982; Scribner, 1984; Sternberg & Frensch, 1991). The ethnographic research of Lave (1988) and Hutchins (1995) revealed that experts do not slavishly conduct "tasks" or adhere to work rules or work procedures but instead develop informal heuristic strategies that, though possibly inefficient and even counterintuitive, are often remarkably robust, effective, and cognitively economical. One provocative implication of this work is that expertise results in part from a natural convergence on such strategies during engagement with the challenges posed by work.

Studies spanned a wide gamut of topics, some of which seem more traditional to academia (e.g., physics problem solving), but many that would traditionally not be fair game for the academic experimental psychologist (e.g., expertise in manufacturing engineering, medical diagnosis, taxicab driving, bird watching, grocery shopping, natural navigation). Mainstream cognitive psychology took something of a turn toward applications (see Barber, 1988), and today the phrase "real world" seems to no longer require scare quotes (see Hollnagel, Hoc, & Cacciabue, 1996), although there are remnants of debate about the utility and scientific foundations of research that is conducted in uncontrolled or non-laboratory contexts (e.g., Banaji & Crowder, 1989; Hoffman & Deffenbacher, 1993; Hoffman & Woods, 2000).

The Early Methods Palette

Another avenue of study involved attempts to address the knowledge-acquisition bottleneck, the root cause of which lay in the reliance on unstructured interviews by the computer scientists who were building expert systems (see Cullen & Bryman, 1988). Unstructured interviews gained early acceptance as a means of simultaneously "bootstrapping" the researcher's knowledge of the domain, and establishing rapport between the researcher and the expert. Nevertheless, the bottleneck issue encouraged a consideration of methods from psychology that might be brought to bear to widen the bottleneck, including methods of structured interviewing (Gordon & Gill, 1997). Interviews could get their structure from preplanned probe questions, from archived test cases, and so forth.

In addition to interviewing, the researcher might look at expert performance while the expert is conducting their usual or "familiar" task and thinking aloud, with their knowledge and reasoning revealed subsequently via a protocol analysis (see Chi et al., 1981; Ericsson & Simon, 1993, Chapter 38, this volume). In addition, one could study expert performance at "contrived tasks," for example, by withholding certain information about the case at hand (limited-information tasks), or by manipulating the way the information is processed (constrained-processing tasks). In the "method of tough cases" the expert is asked to work on a difficult test case (perhaps gleaned from archives) with the idea that tough cases might reveal subtle aspects of expert reasoning, or particular subdomain or highly specialized knowledge, or aspects of experts' metacognitive skills, for example, the ability to reason about their own reasoning or create new procedures or conceptual categories "on the fly."

Empirical comparisons of KE methods, conducted in the late 1980s, were premised on the speculation that different methods might yield different "kinds" of knowledge – the "differential access hypothesis." (These studies are reviewed at greater length in Hoffman et al., 1995, and Shadbolt & Burton, 1990.) Hoffman worked with experts at aerial photo interpretation for terrain analysis, and Shadbolt and Burton worked with experts at geological and archaeological classification. Both research programs employed a number of knowledge-elicitation methods, and both evaluated the methods in terms of their yield (i.e., the number of informative propositions or decision/classification rules elicited as a function of the task time).

The results were in general agreement. Think-aloud problem solving, combined with protocol analysis, proved to be relatively time-consuming, having a yield of less than one informative proposition per total task minute. Likewise, an unstructured interview yielded less than one informative proposition per total task minute. A structured interview, a constrained processing task, and an analysis of tough cases were the most efficient, yielding between one and two informative propositions per total task minute.

The results from the studies by and Shadbolt and Burton and also showed that there was considerable overlap of knowledge elicited by two of the main techniques they used – a task in which domain concepts were sorted into categories and a task in which domain concepts were rated on a number of dimensions. Both of the techniques elicited information about domain concepts and domain procedures. Hoffman as well as Shadbolt and Burton concluded that interviews need to be used in conjunction with ratings and sorting tasks because contrived techniques elicit specific knowledge and may not yield an overview of the domain knowledge.

An idea that was put aside is that the goal of KE should be to "extract" expert knowledge. It is far more appropriate to refer to knowledge elicitation as a collaborative process, sometimes involving "discovery" of knowledge (Clancey, 1993; Ford & Adams-Webber, 1992; Knorr-Cetina, 1981; LaFrance, 1992). According to a transactional view, expert knowledge is created and maintained through collaborative and social processes, as well as through the perceptual and cognitive processes of the individual. By this view, a goal for cognitive analysis and design is to promote development of a workplace in which knowledge is created, shared, and maintained via natural processes of communication, negotiation, and collaboration (Lintern, Diedrich, & Serfaty, 2002).

The foundation for this newer perspective and set of research goals had been laid by the work of Gary Klein and his associates on the decision making of proficient practitioners in domains such as clinical nursing and firefighting (See Ross, Shafer, & Klein, Chapter 23; Klein et al., 1993). They had laid out some new goals for KE, including the generation of cognitive specifications for jobs, the investigation of decision making in domains involving time pressure and high risk, and the enhancement of proficiency through training and technological innovation. It became clear that the methodology of KE could be folded into the broader methodology of "cognitive task analysis" (CTA) (Militello & Hoffman, forthcoming; Schraagen, Chapter 11), which is now a focal point for human-factors and cognitive-systems engineering.

The Era of Cognitive Task Analysis

Knowledge engineering (or cognitive engineering) typically starts with a problem or challenge to be resolved or a requirement to be satisfied with some form of information processing technology. The design goal influences the methods to be used, including the methods of knowledge elicitation, and the manner in which they will be adapted. One thing that all projects must do is identify who is, and who is not, an expert.

Psychological research during the era of expert systems tended to define expertise somewhat loosely, for instance, "advanced

Table 12.1. Some Alternative Methods of Proficiency Scaling

Method	*Yield*	*Example*
In-depth career interviews about education, training, etc.	Ideas about breadth and depth of experience; Estimate of hours of experience	Weather forecasting in the armed services, for instance, involves duty assignments having regular hours and regular job or task assignments that can be tracked across entire careers. Amount of time spent at actual forecasting or forecasting-related tasks can be estimated with some confidence (Hoffman, 1991).
Professional standards or licensing	Ideas about what it takes for individuals to reach the top of their field.	The study of weather forecasters involved senior meteorologists of the US National Atmospheric and Oceanographic Administration and the National Weather Service (Hoffman, 1991). One participant was one of the forecasters for Space Shuttle launches; another was one of the designers of the first meteorological satellites.
Measures of performance at the familiar tasks	Can be used for convergence on scales determined by other methods.	Weather forecasting is again a case in point since records can show for each forecaster the relation between their forecasts and the actual weather. In fact, this is routinely tracked in forecasting offices by the measurement of "forecast skill scores" (see Hoffman & Trafton, 2006).
Social Interaction Analysis	Proficiency levels in some group of practitioners or within some community of practice (Mieg, 2000; Stein, 1997)	In a project on knowledge preservation for the electric power utilities (Hoffman & Hanes, 2003), experts at particular jobs (e.g., maintenance and repair of large turbines, monitoring and control of nuclear chemical reactions, etc.) were readily identified by plant managers, trainers, and engineers. The individuals identified as experts had been performing their jobs for years and were known among company personnel as "the" person in their specialization: *"If there was that kind of problem I'd go to Ted. He's the turbine guy."*

graduate students" in a particular domain. In general, identification of experts was not regarded as either a problem or an issue in expert-system development. (For detailed discussions, see Hart, 1986; Prerau, 1989.) The rule of thumb based on studies of chess (Chase & Simon, 1973) is that expertise is achieved after about 10,000 hours of practice. Recent research has suggested a qualification on this rule of thumb. For instance, Hoffman, Coffey, and Ford (2000) found that even junior journeymen weather forecasters (individuals in their early 30s) can have had as much as 25,000 hours of experience. A similar figure seems appropriate for the domain of intelligence analysis (Hoffman, 2003a).

Concern with the question of how to define expertise (Hoffman, 1998) led to an awareness that determination of who an expert is in a given domain can require effort. In a type of *proficiency-scaling* procedure, the researcher determines a domain and organizationally appropriate scale of proficiency levels. Some alternative methods are described in Table 12.1.

Social Interaction Analysis, the result of which is a sociogram, is perhaps the lesser known of the methods. A sociogram, which represents interaction patterns between people (e.g., frequent interactions), is used to study group clustering, communication patterns, and workflows and processes. For Social Interaction Analysis,

P1: JzG
052184097Xc12 CB1040B/Ericsson 0 521 84087 X March 23, 2006 9:43

multiple individuals within an organization are interviewed. Practitioners might be asked, for example, "If you have a problem of type x, who would you go to for advice?" Or they might be asked to sort cards bearing the names of other domain practitioners into piles according to one or another skill dimension or knowledge category.

Hoffman, Ford, and Coffey (2000) suggested that proficiency scaling for a given project should be based on at least two of the methods listed in Table 12.1. It is important to employ a scale that is both domain and organizationally appropriate, and that considers the full range of proficiency. For instance, in the project on weather forecasting (Hoffman, Coffey, & Ford, 2000), the proficiency scale distinguished three levels: experts, journeymen, and apprentices, each of which was further distinguished by three levels of seniority.

The expanded KE methods palette and the adoption of proficiency scaling represented the broadening of focus beyond expert systems to support for the creation of intelligent or knowledge-based systems of a variety of forms.

Foundational Methods of Cognitive Engineering

In North America, methods for CTA were developed in reaction to limitations of traditional "behavioral task analysis," as well as to limitations of the early AI knowledge acquisition techniques (Hollnagel & Woods, 1983; Rasmussen, 1986). CTA also emerged from the work of researchers who were studying diverse domains of expertise for the purpose of developing better methods for instructional design and enhancing human learning (see the chapters by Greeno, Gregg, Resnick, and Simon & Hayes in Klahr, 1976). Ethnographers, sociologists of science, and cognitive anthropologists, working in parallel, began to look at how new technology influences work cultures and how technology mediates cooperative activity (e.g., Clancey, Chapter 8; Hutchins, 1995, Knorr-Cetina & Mulkay, 1983; Suchman, 1987).

The field of "Work Analysis," which has existed in Europe since the 1960s, is regarded as a branch of ergonomics, although it has involved the study of cognitive activities in the workplace. (For reviews of the history of the research in this tradition see De Keyser, Decortis, & Van Daele, 1998 Militello & Hoffman, forthcoming; Vicente, 1999.) Work Analysis is concerned with performance at all levels of proficiency, but that of course entails the study of experts and the elicitation of their knowledge. Seminal research in Work Analysis was conducted by Jens Rasmussen and his colleagues at the Risø National Laboratory in Denmark (Rasmussen, Petjersen, & Goodstein, 1994; Rasmussen, 1985). They began with the goal of making technical inroads in the safety-engineering aspects of nuclear power and aviation but concluded that safety could not be assured solely through technical engineering (see Rasmussen & Rouse, 1981). Hence, they began to conduct observations in the workplace (e.g., analyses of prototypical problem scenarios) and conduct interviews with experts.

The theme to these parallel North American and European efforts has been the attempt to understand the interaction of cognition, collaboration, and complex artifacts (Potter, Roth, Woods, & Elm, 2000). The reference point is the field setting, wherein teams of expert practitioners confront significant problems aided by technological and other types of artifacts (Rasmussen, 1992; Vicente, 1999).

The broadening of KE, folding it into CTA, has resulted in an expanded palette of methods, including, for example, ethnographic methods (Clancey, 1993, Hutchins, 1995; Orr, 1996; Spradley, 1979). An example of the application of ethnography to expertise studies appears in Dekker, Nyce, and Hoffman (2003). In this chapter we cannot discuss all of the methods in detail. Instead, we highlight three that have been widely used, with success, in this new era of CTA: the Critical Decision Method, Work Domain Analysis, and Concept Mapping.

Table 12.2. A Sample of a Coded CDM Protocol (Adapted from Klein et al., 1989)

Appraisal	This is going to be a tough fire,
Cue	and we may start running into heat exhaustion problems.
Cue	It is 70 degrees now and it is going to get hotter.
Action	The first truck, I would go ahead and have them open the roof up,
Action	and the second truck I would go ahead and send them inside and
Action	have them start ventilating, start knocking the windows out and working
Elaboration	with the initial engine crew, false ceilings, and get the walls opened up.
Action	As soon as I can, order the second engine to hook up to supply and pump to engine 1.
Anticipation	I am assuming engine 2 will probably be there in a second.
Cue-deliberation	I don't know how long the supply lay line is,
Anticipation	but it appears we are probably going to need more water than one supply line is going to give us.
Metacognition	So I would keep in mind,
Contingency	unless we can check the fire fairly rapidly.
Contingency	So start thinking of other water sources.
Action-Deliberation	Consider laying another supply line to engine 1.

The Critical Decision Method

The Critical Decision Method (CDM) involves multi-pass retrospection in which the expert is guided in the recall and elaboration of a previously experienced case. The CDM leverages the fact that domain experts often retain detailed memories of previously encountered cases, especially ones that were unusual, challenging, or in one way or another involved "critical decisions." The CDM does not use generic questions of the kind "Tell me everything you know about x," or "Can you describe your typical procedure?" Instead, it guides the expert through multiple waves of retelling and prompts through the use of specific probe questions (e.g., "What were you seeing?") and "what-if" queries (e.g., "What might someone else have done in this circumstance?"). The CDM generates rich case studies that are often useful as training materials. It yields time-lined scenarios, which describe decisions (decision types, observations, actions, options, etc.) and aspects of decisions that can be easy or difficult. It can also yield a list of decision requirements and perceptual cues – the information the expert needs in order to make decisions.

An example of a coded CDM transcript appears in Table 12.2. In this example, events in the case have been placed into a timeline and coded into the categories indicated in the leftmost column. As in all methods for coding protocols, multiple coders are used and there is a reliability check.

Given its focus on decision making, the strength of the CDM is its use in the creation of models of reasoning (e.g., decisions, strategies). Detailed presentations of the method along with summaries of studies illustrating its successful use can be found in Crandall, Klein, and Hoffman (2006) and Hoffman, Crandall, and Shadbolt (1998).

Work Domain Analysis

Unlike the CDM, which focuses on the reasoning and strategies of the individual practitioner, Work Domain Analysis (WDA) builds a representation of an entire work domain. WDA has most frequently been used to describe the structure of human-machine systems for process control, but it is now finding increasing use in the analysis and design of complex, systems (Burns & Hajdukiewicz, 2004; Chow & Vicente, 2002; Lintern, Miller, & Baker, 2002; Naikar & Sanderson, 2001).

An Abstraction-Decomposition matrix represents a work domain in terms of

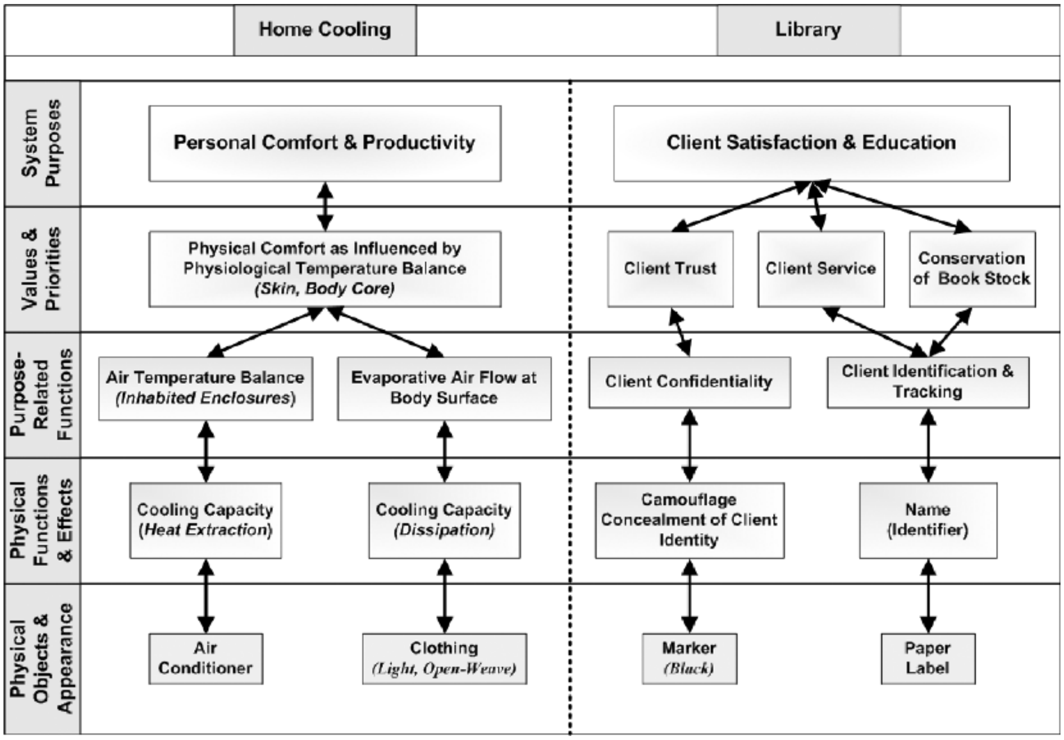

Figure 12.1. Two tutorial examples of the Abstraction-Decomposition representation, a primarily technical system (Home Cooling, left panel) and a sociotechnical system (Library Client Tracking, right panel).

"levels of abstraction," where each level is a distinctive type of constraint. Figure 12.1 presents matrices for two systems, one designed predominantly around physical laws and the other designed predominantly around social values. The library example grapples with a pervasive social issue, the need for individual identification balanced against the desire for personal confidentiality. These tutorial examples demonstrate that the Abstraction-Decomposition format can be used with markedly different work domains. In both of these matrices, entries at each level constitute the means to achieve ends at the level above. The intent is to express means-ends relations between the entries of adjacent levels, with lower levels showing how higher-level functions are met, and higher levels showing why lower-level forms and functions are necessary.

Work domains are also represented in terms of a second dimension: "levels of decomposition," from organizational context, down to social collectives (teams), down to individual worker or individual component (e.g., software package residing on a particular workstation).

Typically, a work-domain analysis is initiated from a study of documents, although once an Abstraction-Decomposition matrix is reasonably well developed, interviews with domain experts will help the analyst extend and refine it. Vicente (1999) argues that the Abstraction-Decomposition matrix is an activity-independent representation and should contain only descriptions of the work domain (the tutorial examples of Figure 12.1 were developed with that stricture in mind). However, Vicente's advice is not followed universally within the community that practices WDA; some analysts include processes in their Abstraction-Decomposition matrices (e.g., Burns & Hajdukiewicz, 2004).

It is possible to add activity to the representation yet remain consistent with Vicente (1999) by overlaying a trajectory derived from a description of strategic reasoning undertaken by experts. Figure 12.2 presents a fragment of a structural description of a weather-forecasting work domain, and Figure 12.3 presents the same structural description with an activity overlay developed from a transcript of an expert forecaster's description of jobs, roles, and tools involved in forecasting (Hoffman, Coffey, & Ford, 2000). Activity statements (shown as callouts in Figure 12.3) were coded as falling into one or another of the cells, and the temporal sequence of the activity was represented by the flow of arrows as connectors to show how forecasters navigate opportunistically through an abstraction-decomposition space as they seek the information to diagnose and solve the problems.

When used in this manner, the matrix captures important propositions as elicited from domain experts concerning their goals and reasoning (see, e.g., Burns, Bryant, & Chalmers, 2001; Rasmussen, 1986; Schmidt & Luczak, 2000; Vicente, Christoffersen, & Pereklita, 1995) within the context of collaboration with larger collectives and organizational goals.

Concept Mapping

The third CTA method we will discuss is also one that has been widely used and has met with considerable success. Unlike Abstraction-Decomposition and its functional analysis of work domains, and unlike the CDM and its focus on reasoning and strategies, Concept Mapping has as its great strength the generation of models of knowledge.

Concept Maps are meaningful diagrams that include concepts (enclosed in boxes) and relationships among concepts or propositions (indicated by labeled connections between related concepts). Concept Mapping has foundations in the theory of Meaningful Learning (Ausubel, Novak, & Hanesian, 1978) and decades of research and application, primarily in education (Novak, 1998). Concept Maps can be used to show gaps in student knowledge. At the other end of the proficiency scale, Concept Maps made by domain experts tend to show high levels of agreement (see Gordon, 1992; Hoffman, Coffey, & Ford, 2000). (Reviews of the literature and discussion of methods

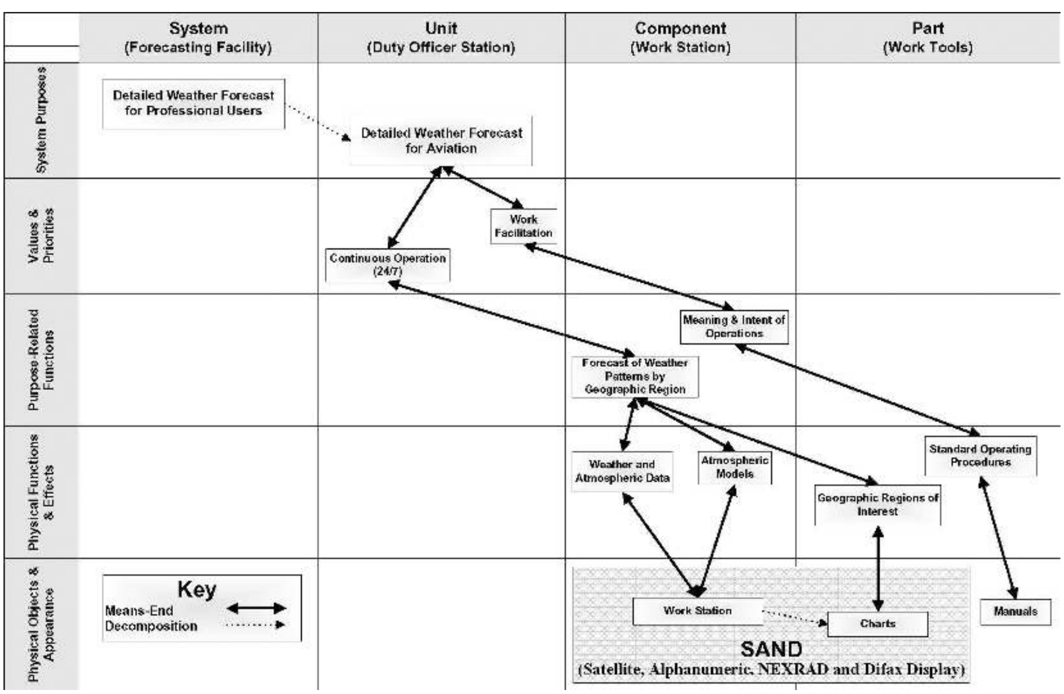

Figure 12.2. An Abstraction-Decomposition matrix of a fragment of a weather-forecasting work domain.

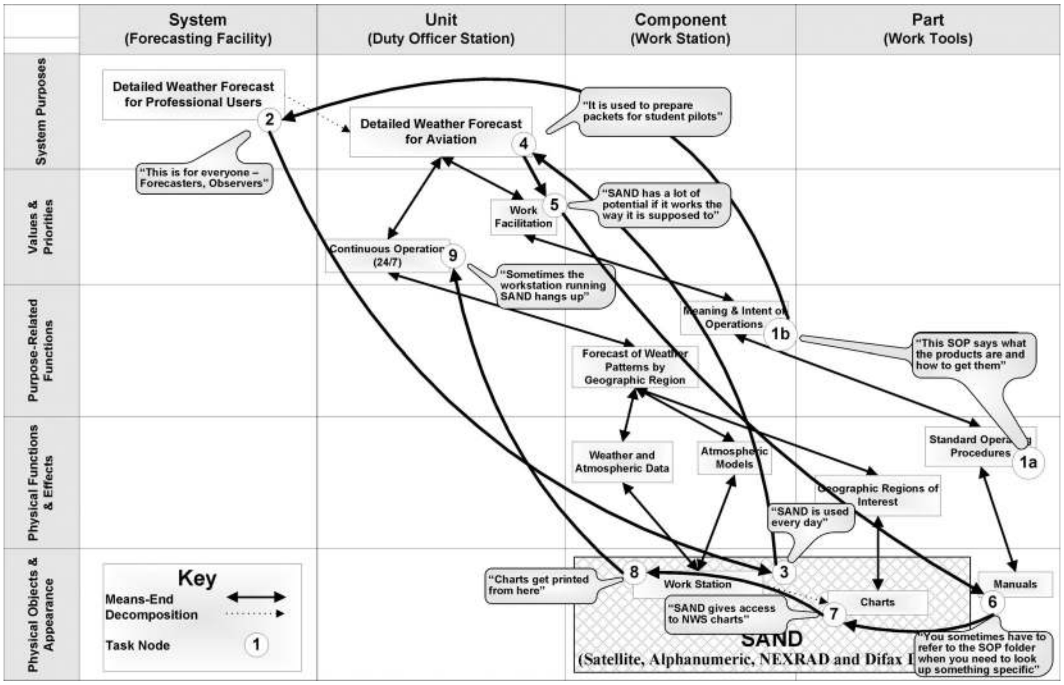

Figure 12.3. An Abstraction-Decomposition matrix of a fragment of a weather-forecasting work domain with an activity overlay (statements in callouts are quotes from a forecaster).

for making Concept Maps can be found in Cañas et al., 2004; and Crandall, Klein, & Hoffman, 2006.) Figure 12.4 is a Concept Map that lays out expert knowledge about the role of cold fronts in the Gulf Coast (Hoffman, Coffey, & Ford, 2000).

Although Concept Maps can be made by use of paper and pencil, a white board, or Post-Its, the Concept Maps presented here were created by use of CmapTools, a software suite created at the Institute for Human and Machine Cognition (free download at http://ihmc.us). In the KE procedure involving an individual expert, one researcher stands at a screen and serves as the facilitator while another researcher drives the laptop and creates the Concept Map that is projected on the screen. The facilitator helps the domain expert build up a representation of their domain knowledge, in effect combining KE with knowledge representation. (This is one reason the method is relatively efficient.) Concept Mapping can also be used by teams or groups, for purposes other than KE (e.g., brainstorming, consensus formation). Teams can be structured in a variety of ways and can make and share Concept Maps over the world-wide web (see Cañas et al., 2004).

The ability to hyperlink digital "resources" such as text documents, images, video clips, and URLs is another significant advantage provided by computerized means of developing Concept Maps (CmapTools indicate hyperlinks by the small icons underneath concept nodes). Hyperlinks can connect to other Concept Maps; a set of Concept Maps hyperlinked together is regarded as a "knowledge model." Figure 12.5 shows a screen shot from the top-level Concept Map in the System To Organize Representations in Meteorology (STORM), in which a large number of Concept Maps are linked together. In Figure 12.5, some of the resources have been opened for illustrative purposes (real-time satellite imagery, computer weather forecasts, and digital video in which the domain expert provides brief explanatory statements for some of the concepts throughout the model).

All of the STORM Concept Maps and resources can be viewed at http://www.ihmc.us/research/projects/STORMLK/

Knowledge models structured as Concept Maps can serve as living repositories of expert knowledge to support knowledge sharing as well as knowledge preservation. They can serve also as interfaces for intelligent systems where the model of the expert's knowledge becomes the interface for a performance support tool or training aid. (Ford et al., 1996).

Methodological Concepts and Issues

Research and various applied projects conducted since the seminal works on KE methodology have left some ideas standing and have led to some new and potentially valuable ideas. One recent review of CTA methods (Bonacteo & Burns, forthcoming) lists dozens of methods. Although not all of them are methods that would be useful as knowledge-elicitation or knowledge-representation procedures, it is clear that the roster of tools and methods available to cognitive engineers has expanded considerably over the past two decades. We look now to core ideas and tidbits of guidance that have stood the test of time.

Where the Rubber Meets the Road

(1). In eliciting expert knowledge one can: (a) Ask people questions, and (b) Observe performance. Questions can be asked in the great many forms and formats for interviewing, including unstructured interviews, the CDM procedure, and Concept Mapping, as well as many other techniques (e.g., Endsley & Garland, 2000). Performance can be observed via ethnographic studies of patterns of communication in the workplace,

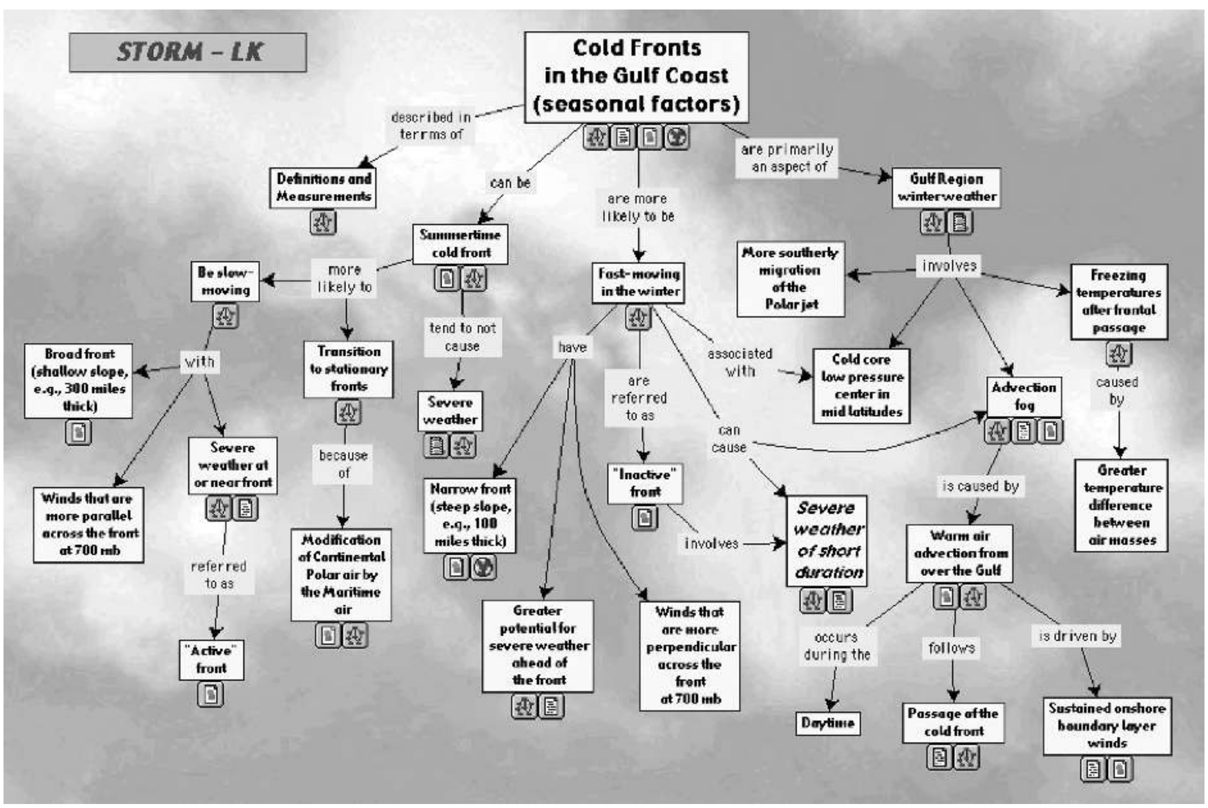

Figure 12.4. A Concept Map about cold fronts in Gulf Coast weather.

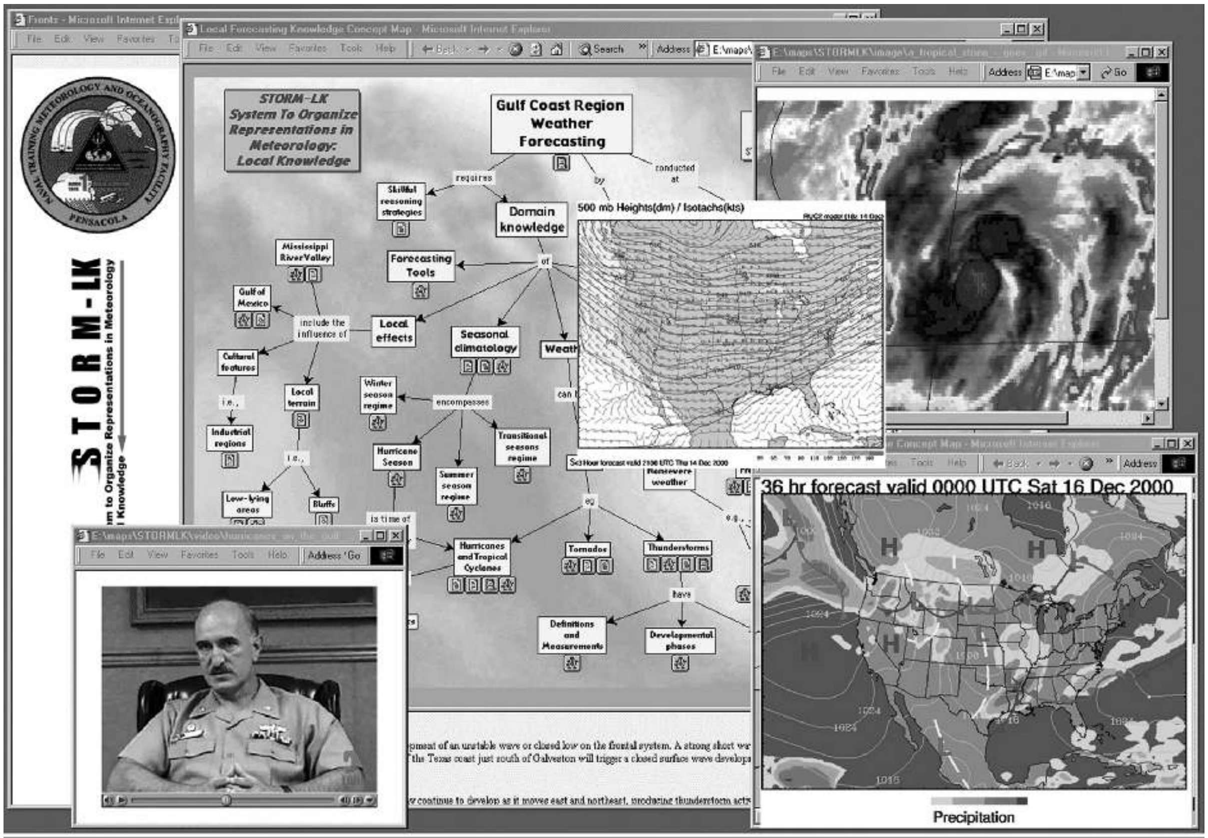

Figure 12.5. A screen shot of a Concept Map with some opened resources.

evaluations in terms of performance measures (e.g., the accuracy of weather forecasts), or evaluations of recall, recognition, or reaction-time performance in contrived tasks or think-aloud problem solving tasks.

(2). In eliciting expert knowledge one can attempt to create models of the work domain, models of practitioner knowledge of the domain, or models of practitioner reasoning. Models of these three kinds take different forms and have different sorts of uses and applications. This is illustrated roughly by the three methods we have described here. The CDM can be used to create products that describe practitioner reasoning (e.g., decision types, strategies, decision requirements, informational cues). The Abstraction-Decomposition matrix represents the functional structure of the work domain, which can provide context for an overlay of activity developed from interview protocols or expert narratives. Concept Mapping represents practitioner knowledge of domain concepts such as relations, laws, and case types.

(3). Knowledge elicitation methods differ in their relative efficiency. For instance, the think-aloud problem solving task combined with protocol analysis has uses in the psychology laboratory but is relatively inefficient in the context of knowledge elicitation. Concept Mapping is arguably the most efficient method for the elicitation of domain knowledge (Hoffman, 2002).We see a need for more studies on this topic.

(4). Knowledge-elicitation methods can be combined in various ways. Indeed, a recommendation from the 1980s still stands – that any project involving expert knowledge elicitation should use more than one knowledge-elicitation method. One combination that has recently become salient is the combination of the CDM with the two other procedures we have discussed.

Concept-Mapping interviews almost always trigger in the experts the recall of previously encountered tough cases. This can be used to substitute for the "Incident Selection" step in the CDM. Furthermore, case studies generated by the CDM can be used as resources to populate the Concept-Map knowledge models (see Hoffman, Coffey, & Ford, 2000). As another example, Naikar and Saunders (2003) conducted a Work Domain Analysis to isolate safety-significant events from aviation incident reports and then employed the CDM in interviews with authors of those reports to identify critical cognitive issues that precipitated or exacerbated the event.

(5). The gold is not in the documents. Document analysis is useful in bootstrapping researchers into the domain of study and is a recommended method for initiating Work Domain Analysis (e.g., Lintern et al., 2002), but experts possess knowledge and strategies that do not appear in documents and task descriptions. Cognitive engineers invariably rely on interactions with experts to garner implicit, obscure, and otherwise undocumented expert knowledge. Even in Work Domain Analysis, which is heavily oriented towards Document Analysis, interactions with experts are used to confirm and refine the Abstraction-Decomposition matrices.

In the weather-forecasting project (Hoffman, Coffey, & Ford, 2000), an expert told how she predicted the lifting of fog. She would look out toward the downtown and see how many floors above ground level she could count before the floors got lost in the fog deck. Her reasoning relied on a heuristic of the form, "If I cannot see the 10th floor by 10 AM, pilots will not be able to take off until after lunchtime." Such a heuristic has great value but is hardly the sort of thing that could be put into a formal standard operating procedure. Many observations have been made of how engineers in process control bend rules and deviate from mandated procedures so that they can do their jobs more effectively (see Koopman & Hoffman, 2003). We would hasten to generalize by saying that all experts who work in complex sociotechnical contexts possess knowledge and reasoning strategies that are not captured in existing procedures or documents, many of which represent (naughty) departures from what those experts are supposed to do or to believe (Johnston, 2003; McDonald, Corrigan, & Ward, 2002).

Discovery of these undocumented departures from authorized procedures represents a window on the "true work" (Vicente, 1999), which is cognitive work independent of particular technologies, that is, it is governed only by domain constraints and by human cognitive constraints. Especially after an accident, it is commonly argued that experts who depart from authorized procedures are, in some way, negligent. Nevertheless, the adaptive process that generates the departures is not only inevitable but is also a primary source of efficient and robust work procedures (Lintern, 2003). In that these windows are suggestive of leverage points and ideas for new aiding technologies, cognitive engineers need to pay them serious attention.

(6). Differential access is not a salient problem. The first wave of comparative KE methodology research generated the hypothesis that different "kinds" of knowledge might be more amenable to elicitation by particular methods (Hoffman, 1987), and some studies suggested the possibility of differential access (Cooke & MacDonald, 1986, 1987; Evans, Jentsch, Hitt, Bowers, & Salas, 2001). Tasks involving the generation of lists of domain concepts can in fact result in lists of domain concepts, and tasks involving the specification of procedures can in fact result in statements about rules or procedures. However, some studies have found little or no evidence for differential access (e.g., Adelman, 1989; Shadbolt & Burton, 1990), and we conclude that a strong version of the differential-access hypothesis has not held up well under scrutiny. All of the available methods *can* say things about so-called declarative knowledge, so-called procedural knowledge, and so forth.

All KE methods can be used to identify leverage points – aspects of the organization or work domain where even a modest infusion of supporting technologies might have positive results (e.g., redesign of interfaces, redesign of the workspace layout, creation of new functionalities for existing software, and ideas about entirely new software systems.) Again we can use the project on expert weather forecasting as an example (Hoffman, Coffey, & Ford, 2000). That project compared a number of alternative knowledge-elicitation methods including protocol analysis, the CDM, the Knowledge Audit (Militello & Hutton, 1998; see Ross, Shafer and Klein, this volume), an analysis of "Standard Operating Procedures" documents, the Recent Case Walkthrough method (Militello & Hutton, 1998), a Workspace and Workpatterns analysis (Vicente, 1999), and Concept Mapping. All methods yielded data that spoke to practitioner knowledge and reasoning and all also identified leverage points.

(7). "Tacit" knowledge is not a salient problem. Without getting into the philosophical weeds of what one means by "kinds" of knowledge, another concern has to do with the possibility that routine knowledge about procedures or task activities might become "tacit," that is, so automatic as to be inexpressible via introspection or verbal report. This hangover issue from the heyday of Behaviorism remains to this day a non-problem in the practical context of eliciting knowledge from experts. For one thing, it has never been demonstrated that there exists such a thing as "knowledge that cannot be verbalized in principle," and the burden of proof falls on the shoulders of those who make the existence claim. Again sidestepping the philosophical issues (i.e., if it cannot be articulated verbally, is it really knowledge?), we maintain that the empirical facts mitigate the issue. For instance, in Concept-Mapping interviews with domain experts, experience shows that almost every time the expert will reach a point in making a Concept Map where s/he will say something like, *"Well, I've never really thought about that, or thought about it in this way, but now that you mention it . . . ,"* and what follows will be a clear specification on some procedure, strategy, or aspect of subdomain knowledge that had not been articulated up to that point.

(8). Good knowledge elicitation procedures are "effective scaffolds." Although there may be phenomena to which one could legitimately, or at least arguably, append the designation "tacit knowledge," there is no indication that such knowledge lies beyond the reach of science in some unscientific netherworld of intuitions or unobservables. Over and over again, the lesson is not that there is knowledge that experts literally cannot articulate, nor is it the hangover issue of whether verbalization "interferes" with reasoning. Rather, the issue is whether the KE procedure provides sufficient scaffolding to support the expert in articulating what they know. Support involves the specifics of the procedure (e.g., probe questions), but it also involves the fact that knowledge elicitation is a collaborative process. There is no substitute for the skill of the elicitor (e.g., in framing alternative suggestions and wordings). Likewise, there is no substitute for the skill of the participating practitioner. Some experts will have good insight, but others will not. Though it might be possible for someone to prove the existence of "knowledge" that cannot be uncovered, knowledge engineers face the immediate, practical challenges of designing new and better sociotechnical systems. They accomplish something when they uncover useful knowledge that might have otherwise been missed.

New Ideas

Recent research and application efforts have also yielded some new ideas about the knowledge elicitation methods palette.

(1). The (hypothetical) problem of differential access has given way to a practical consideration of "differential utility." Any given method might be more useful for certain purposes, might be more applicable to certain domains, or might be more useful with experts having certain cognitive styles. In other words, each knowledge-elicitation method has its strengths and weaknesses. Some of these are more purely

methodological or procedural (e.g., transcription and protocol analysis takes a long time), but some relate to the content of what is elicited. The CDM has as its strength the elicitation of knowledge about perceptual cues and patterns, decision types, and reasoning strategies. The strength of Concept Mapping lies in the creation of knowledge models that can be used in the creation of knowledge bases or interfaces. Work Domain Analysis, which maps the functional structure of the work domain, can provide a backdrop against which the knowledge and skills of the individual expert can be fitted into the larger functional context of the organization and its purposes. Products from any of these procedures can support the design of new interfaces or even the redesign of workplaces and methods of collaboration.

(2). Methodology benefits from opportunism. It can be valuable during a knowledge-elicitation project to be open to emerging possibilities and new opportunities, even opportunities to create new methods or try out and evaluate new combinations of methods. In the weather-forecasting project (Hoffman, Coffey, & Ford, 2000), Concept-Mapping interviews demonstrated that practitioners were quite comfortable with psychologists' notion of a "mental model" because the field has for years distinguished forecaster reasoning ("conceptual models") from the outputs of the mathematical computer models of weather. Indeed, the notion of a mental model has been invoked as an explanatory concept in weather forecasting for decades (see Hoffman, Trafton, & Roebber, forthcoming). Practitioners were quite open to discussing their reasoning, and so a special interview was crafted to explore this topic in detail (Hoffman, Coffey, & Carnot, 2000).

(3). Knowledge elicitation is not a one-off procedure. Historically, KE was considered in the context of creating intelligent systems for particular applications. The horizons were expanded by such applications as the preservation of organizational or team knowledge (Klein, 1992). This notion was recently expanded even further to the idea of "corporate knowledge management," which includes capture, archiving, application to training, proprietary analysis, and other activities (e.g., Becerra-Fernandez, Gonzalez, & Sabherwal, 2004; Davenport & Prusak, 1998). A number of government and private sector organizations have found a need to capture expert knowledge prior to the retirement of the experts and also the need, sometimes urgent, to reclaim expertise from individuals who have recently retired (Hoffman & Hanes, 2003). Instantiation of knowledge capture as part of an organizational culture entails many potential obstacles, such as management and personnel buy-in. It also raises many practical problems, not the least of which is how to incorporate a process of ongoing knowledge capture into the ordinary activities of the experts without burdening them with an additional task.

Recognition of the value of the analysis of tough cases led to a recommendation that experts routinely make notes about important aspects of tough cases that they encounter (Hoffman, 1987). This idea has been taken to new levels in recent years. For instance, because of downsizing in the 1980s, the electric power utilities face a situation in which senior experts are retiring and there is not yet a cohort of junior experts who are primed to take up the mantle (Hoffman & Hanes, 2003). At one utility, a turbine had been taken off line for total refitting, an event that was seen as an opportunity to videotape certain repair jobs that require expertise but are generally only required occasionally (on the order of once every 10 or more years).

Significant expertise involves considerable domain and procedural knowledge and an extensive repertoire of skills and heuristics. Elicitation is rarely something that can be done easily or quickly. In eliciting weather-forecasting knowledge for just the Florida Gulf Coast region of the United Sates, about 150 Concept Maps were made about local phenomena involving fog, thunderstorms, and hurricanes. And yet, dozens more Concept Maps could have been made on additional topics, including the use of the new weather radar systems and the use of the many computer models for weather forecasting (Hoffman, Coffey, & Ford, 2000). A current project on "knowledge recovery" that involves reclamation of expert knowledge about terrain analysis from existing documents such as the Terrain Analysis Database (Hoffman, 2003b) has generated over 150 Concept Maps containing more than 3,000 propositions.

Although knowledge elicitation on such a scale is daunting, we now have the technologies and methodologies to facilitate the elicitation, preservation, and sharing of expert knowledge on a scale never before possible. This is a profound application of cognitive science and is one that is of immense value to society.

Practice, Practice, Practice

No matter how much detail is provided about the conduct of a knowledge-elicitation procedure, there is no substitute for practice. The elicitor needs to adapt on the fly to individual differences in style, personality, agenda, and goals. In "breaking the ice" and establishing rapport, the elicitor needs to show good intentions and needs to be sensitive to possible concerns on the part of the expert that the capture of his/her knowledge might mean the loss of their job (perhaps to a machine). To be good and effective at knowledge-elicitation, one must attempt to become an "expert apprentice" – experienced at, skilled at, and comfortable with going into new domains, boostrapping efficiently and then designing and conducting a series of knowledge-elicitation procedures appropriate to project goals. The topic of how to train people to be expert apprentices is one that we hope will receive attention from researchers in the coming years (see Militello & Quill, forthcoming).

Acknowledgments

The senior Author's contribution to this chapter was supported through his participation in the National Alliance for Expertise Studies, which is supported by the "Sciences of Learning" Program of the National Science Foundation, and his participation in The Advanced Decision Architectures Collaborative Technology Alliance, which is sponsored by the US Army Research Laboratory under cooperative agreement DAAD19-01-2-0009.

References

Adelman, L. (1989). Management issues in knowledge elicitation. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 483–488.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view, 2nd ed*. New York: Holt, Rinehart and Winston.

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist*, 44, 1185–1193.

Barber, P. (1988). *Applied cognitive psychology*. London: Methuen.

Becerra-Fernandez, I., Gonzalez, A., & Sabherwal, R. (2004). *Knowledge management: challenges, solutions, and technologies*. Upper Saddle River, NJ: Prentice-Hall.

Bonaceto, C., & Burns, K. (forthcoming). Mapping the mountains: A survey of cognitive engineering methods and uses. In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*. Mahwah, NJ: Erlbaum.

Burns, C., & Hajdukiewicz, J. R. (2004). *Ecological interface design*. Boca Raton, FL: CRC Press.

Burns, C. M., Bryant, D., & Chalmers, B. (2001). Scenario mapping with work domain analysis. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 424–428). Santa Monica, CA: Human Factors and Ergonomics Society.

Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., Gómez, G., Arroyo, M., & Carvajal, R. (2004). CmapTools: A Knowledge Modeling and Sharing Environment. In A. J. Cañas, J. D. Novak, & F. M. González (Eds.), *Concept Maps: Theory, Methodology, Technology, Proceedings of the 1st International Conference on Concept Mapping* (pp. 125–133). Pamplona, Spain: Universidad Pública de Navarra.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.

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

Chi, M. T. H., Glaser, R., & Farr, M. L. (Eds.) (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Chow, R., & Vicente, K. J. (2002). A field study of emergency ambulance dispatching: Implications for decision support. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 313–317). Santa Monica, CA: Human Factors and Ergonomics Society.

Clancey, W. J. (1993). The knowledge level reinterpreted: Modeling socio-technical systems. In K. M. Ford & J. M. Bradshaw (Eds.), *Knowledge acquisition as modeling* (pp. 33–49, Pt. 1). New York: Wiley.

Cooke, N. M., & McDonald, J. E. (1986). A formal methodology for acquiring and representing expert knowledge. *Proceedings of the IEEE* 74, 1422–1430.

Cooke, N. M., & McDonald, J. E. (1987). The application of psychological scaling techniques to knowledge elicitation for knowledge-based systems. *International Journal of Man-Machine Studies*, 26, 533–550.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working Minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

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

David, J.-M., Krivine, J.-P., & Simmons, R. (Eds.) (1993). *Second-generation expert systems*. Berlin: Springer Verlag.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Cambridge, MA: Harvard Business School Press.

De Keyser, V., Decortis, F., & Van Daele, A. (1998). The approach of Francophone ergonomy: Studying new technologies. In V. De Keyser, T. Qvale, & B. Wilpert (Eds.), *The meaning of work and technological options* (pp. 147–163). Chichester, England: Wiley.

Dekker, S. W. A., Nyce, J. M., & Hoffman, R. R. (March–April 2003). From contextual inquiry to designable futures: What do we need to get there? *IEEE Intelligent Systems*, pp. 74–77.

Diderot, D. (with d'Alembert, J.) (Eds.) (1751–1772). *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers, par une Société de Gens de lettere*. Paris: Le Breton. *Compact Edition* published in 1969 by The Readex Microprint Corporation, New York. Translations of selected articles can be found at http://www.hti.umich.edu/d/did/.

Duda, J., Gaschnig, J., & Hart, P. (1979). Model design in the PROSPECTOR consultant system for mineral exploration. In D. Michie (Ed.), *Expert systems in the micro-electronic age* (pp. 153–167). Edinburgh: Edinburgh University Press.

Endsdley, M. R., & Garland, D. L. (2000). *Situation awareness analysis and measurement*. Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data, 2nd ed*. Cambridge, MA: MIT Press.

Evans, A. W., Jentsch, F., Hitt, J. M., Bowers, C., & Salas, E. (2001). Mental model assessments: Is there a convergence among different methods? In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 293–296). Santa Monica, CA: Human Factors and Ergonomics Society.

Feigenbaum, E. A., Buchanan, B. G., & Lederberg, J. (1971). On generality and problem solving: A case study using the DENDRAL program. In B. Meltzer & D. Michie (Eds.), *Machine intelligence 6* (pp. 165–190). Edinburgh: Edinburgh University Press.

Ford, K. M., & Adams-Webber, J. R. (1992). Knowledge acquisition and constructive epistemology. In R. Hoffman (Ed.), *The cognition of experts: Psychological research and empirical AI* (pp. 121–136). New York: Springer-Verlag.

Ford, K. M., Coffey, J. W., Cañas, A., Andrews, E. J., & Turne, C. W. (1996). Diagnosis and explanation by a nuclear cardiology expert system. *International Journal of Expert Systems*, 9, 499–506.

Forsyth, D. E., & Buchanan, B. G. (1989). Knowledge acquisition for expert systems: Some pitfalls and suggestions. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 345–442.

Gaines, B. R., & Boose, J. H. (Eds.) (1988). *Knowledge acquisition tools for expert systems*. London: Academic Press.

Gagné, R. M., & Smith, E. C. (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology*, 63, 12–18.

Gentner, D., & Stevens, A. L. (Eds.) (1983). *Mental models*. Mahwah, NJ: Erlbaum.

Glaser, R. (1987). Thoughts on expertise. In C. Schooler & W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 81–94). Norwood, NJ: Ablex.

Glaser, R., Lesgold, A., Lajoie, S., Eastman, R., Greenberg, L., Logan, D., Magone, M., Weiner, A., Wolf, R., & Yengo, L. (1985). "Cognitive task analysis to enhance technical skills training and assessment." Report, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA.

Gordon, S. E. (1992). Implications of cognitive theory for knowledge acquisition. In R. R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and empirical AI* (pp. 99–120). Mahwah, NJ: Erlbaum.

Gordon, S. E., & Gill, R. T. (1997). Cognitive task analysis. In C. Zsambok and G. Klein (Eds.), *Naturalistic decision making* (pp. 131–140). Mahwah, NJ: Erlbaum.

Hart, A. (1986). *Knowledge acquisition for expert systems*. London: Kogan Page.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. B. (1983). *Building expert systems*. Reading, MA: Addison-Wesley.

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

Hoffman, R. R. (1991). Human factors psychology in the support of forecasting: The design of advanced meteorological workstations. *Weather and Forecasting*, 6, 98–110.

Hoffman, R. R. (Ed.) (1992). *The psychology of expertise: Cognitive research and empirical AI*. Mahwah, NJ: Erlbaum.

Hoffman, R. R. (1998). How can expertise be defined?: Implications of research from cognitive psychology. In R. Williams, W. Faulkner, & J. Fleck (Eds.), *Exploring expertise* (pp. 81–100). New York: Macmillan.

Hoffman, R. R. (2002, September). An empirical comparison of methods for eliciting and modeling expert knowledge. In *Proceedings of the 46th Meeting of the Human Factors and Ergonomics Society* (pp. 482–486). Santa Monica, CA: Human Factors and Ergonomics Society.

Hoffman, R. R. (2003a). "Use of Concept Mapping and the Critical Decision Method to support Human-Centered Computing for the intelligence community." Report, Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R. (2003b). "Knowledge recovery." Report, Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R., Coffey, J. W., & Carnot, M. J. (2000, November). Is there a "fast track" into the black box?: The Cognitive Modeling Procedure. Poster presented at the 41st Annual Meeting of the Psychonomics Society, New Orleans, LA.

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). A case study in cognitive task analysis methodology: The Critical Decision Method for the elicitation of expert knowledge. *Human Factors*, 40, 254–276.

Hoffman, R. R., & Deffenbacher, K. (1992). A brief history of applied cognitive psychology. *Applied Cognitive Psychology*, 6, 1–48.

Hoffman, R. R., & Deffenbacher, K. A. (1993). An analysis of the relations of basic and applied science. *Ecological Psychology*, 5, 315–352.

Hoffman, R. R., Coffey, J. W., & Ford, K. M. (2000). "A case study in the research paradigm of Human-Centered Computing: Local expertise in weather forecasting." Report to the National Technology Alliance on the Contract, "Human-Centered System Prototype." Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R., Ford, K. M., & Coffey, J. W. (2000). "The handbook of human-centered computing." Report, Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R., & Hanes, L. F. (2003/July–August). The boiled frog problem. *IEEE: Intelligent Systems*, pp. 68–71.

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.

Hoffman, R. R., Trafton, G., & Roebber, P. (2006). *Minding the weather: How expert forecasters think*. Cambridge, MA: MIT Press.

Hoffman, R. R., & Woods, D. D. (2000). Studying cognitive systems in context. *Human Factors*, 42, 1–7.

Hoc, J.-M., Cacciabue, P. C., & Hollnagel, E. (1996). Expertise and technology: "I have a feeling we're not in Kansas anymore." In J.-M. Hoc, P. C. Cacciabue, & E. Hollnagel (Eds.), *Expertise and technology: Cognition and human-computer cooperation* (pp. 279–286). Mahwah, NJ: Erlbaum.

Hollnagel, E., & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583–600.

Johnston, N. (2003). The paradox of rules: Procedural drift in commercial aviation. In R. Jensen (Ed.), *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 630–635). Dayton, OH: Wright State University.

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

Klahr, D. (Ed.) (1976). *Cognition and instruction*. Hillsdale, NJ: Erlbaum.

Klahr, D., & Kotovsky, K. (Eds.) (1989). *Complex information processing: The impact of Herbert A. Simon*. Mahwah, NJ: Erlbaum.

Klein, G. (1992). Using knowledge elicitation to preserve corporate memory. In R. R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and empirical AI* (pp. 170–190). Mahwah, NJ: Erlbaum.

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

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.) (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Publishing Corporation.

Klein, G., & Weitzenfeld, J. (1982). The use of analogues in comparability analysis. *Applied Ergonomics, 13*, 99–104.

Knorr-Cetina, K. D. (1981). *The manufacture of knowledge*. Oxford: Pergamon.

Knorr-Cetina, K. D., & Mulkay, M. (1983). *Science observed*. Berkeley Hills, CA: Sage.

Koopman, P., & Hoffman, R. R. (2003/ November–December). Work-arounds, make-work, and kludges. *IEEE: Intelligent Systems*, pp. 70–75.

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

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge: Cambridge University Press.

Lesgold, A. M. (1984). Acquiring expertise. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower* (pp. 31–60). San Francisco, CA: W. H. Freeman.

Lesgold, A., Feltovich, P. J., Glaser, R., & Wang, M. (1981). "The acquisition of perceptual diagnostic skill in radiology." Technical Report No. PDS-1, Learing Research and Development Center, University of Pittsburgh, Pittsburgh, PA.

Lintern, G. (2003). Tyranny in rules, autonomy in maps: Closing the safety management loop. In R. Jensen (Ed.), *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 719–724). Dayton, OH: Wright State University.

Lintern, G., Diedrich, F. J. & Serfaty, D. (2002). Engineering the community of practice for maintenance of organizational knowledge. *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants* (pp. 6.7–6.13). New York: IEEE.

Lintern, G., Miller, D., & Baker, K. (2002). Work centered design of a USAF mission planning system. In *Proceedings of the 46th Human Factors and Ergonomics Society Annual Meeting* (pp. 531–535). Santa Monica, CA: Human Factors and Ergonomics Society.

McDonald, N., Corrigan, S., & Ward, M. (2002). Cultural and organizational factors in system safety: Good people in bad systems. *Proceedings of the 2002 International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero 2002)* (pp. 205–209). Menlo Park, CA: American Association for Artificial Intelligence Press.

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

Means, B., & Gott, S. P. (1988). Cognitive task analysis as a basis for tutor development: Articulating abstract knowledge representations. In J. Psotka, L. D. Massey, & S. A. Mutter (Eds.), *Intelligent tutoring systems: Lessons learned* (pp. 35–57). Hillsdale, NJ: Erlbaum.

Mieg, H. (2000). *The social psychology of expertise*. Mahwah, NJ: Erlbaum.

Militello, L. G., & Hutton, R. J. B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics, 41*, 1618–1641.

Militello, L., & Hoffman, R. R. (forthcoming). *Perspectives on cognitive task analysis.* Mahwah NJ: Erlbaum.

Militello, L., & Quill, L. (2006). Expert apprentice strategies. In R. R. Hoffman (Ed.), *Expertise out of context*. Mahwah, NJ: Erlbaum.

Naikar, N., & Sanderson, P. M. (2001). Evaluating system design proposals with work domain analysis. *Human Factors*, 43, 529–542.

Naikar, N., & Saunders, A. (2003). Crossing the boundaries of safe operation: A technical training approach to error management. *Cognition Technology and Work*, 5, 171–180.

Neale, I. M. (1988). First generation expert systems: A review of knowledge acquisition methodologies. *Knowledge Engineering Review*, 3, 105–146.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps® as facilitative tools in schools and corporations*. Mahweh, NJ: Lawrence Erlbaum Associates.

Orr, J. (1996). *Talking about machines: An ethnography of a modern job*. Ithaca, NY: Cornell University Press.

Potter, S. S., Roth, E. M., Woods, D. D., and Elm, W. C. (2000). Bootstrapping multiple converging cognitive task analysis techniques for system design. In J. M. Schraagen and S. F. Chipman (Eds.), *Cognitive task analysis* (pp. 317–340). Mahwah, NJ: Erlbaum.

Prerau, D. (1989). *Developing and managing expert systems: Proven techniques for business and industry*. Reading, MA: Addison Wesley.

Raeth, P. G. (Ed.) (1990). *Expert systems*. New York: IEEE Press.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision-making and system-management. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-15*, 234–243.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. Amsterdam: North-Holland.

Rasmussen, J. (1992). Use of field studies for design of workstations for integrated manufacturing systems. In M. Helander & N. Nagamachi (Eds.), *Design for manufacturability: A systems approach to concurrent engineering and ergonomics* (pp. 317–338). London: Taylor and Francis.

Rasmussen, J., Petjersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. New York: John Wiley.

Rasmussen, J., & Rouse, W. B. (Eds.) (1981). *Human detection and diagnosis of system failures*. New York: Plenum Press.

Rook, F. W., & Croghan, J. W. (1989). The knowledge acquisition activity matrix: A systems engineering conceptual framework. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 586–597.

Schmidt, L., & Luczak, H. (2000). Knowledge representation for engineering design based on a cognitive model. In *Proceedings of the IEA 2000/HFES 2000 Congress* (vol. 1) (pp. 623–626). Santa Monica, CA: Human Factors and Ergonomics Society.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff & S. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 9–40). Cambridge: Harvard University Press.

Shadbolt, N. R., & Burton, A. M. (1990). Knowledge elicitation techniques: Some experimental results. In K. L. McGraw & C. R. Westphal (Eds.), *Readings in knowledge acquisition* (pp. 21–33). New York: Ellis Horwood.

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, 53, 252–266.

Shortliffe, E. H. (1976). *Computer-based medical consultations: MYCIN*. New York: Elsevier.

Simon, H. A., & Hayes, J. R. (1976). Understanding complex task instructions. In D. Klahr (Ed.), *Cognition and instruction* (pp. 51–80). Hillsdale, NJ: Erlbaum.

Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart and Winston.

Stein, E. W. (1997). A look at expertise from a social perspective. In P. J. Feltovich, K. M. Ford, & R. R. Hoffman (Eds.), *Expertise in context* (pp. 181–194). Cambridge: MIT Press.

Sternberg, R. J., & Frensch, P. A. (Eds.) (1991). *Complex problem solving: Principles and mechanisms*. Mahwah, NJ: Erlbaum.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Vicente, K. J. (1999). *Cognitive work analysis: Towards safe, productive, and healthy computer-based work*. Mahwah, NJ: Erlbaum.

Vicente, K. J., Christoffersen, K., & Pereklita, A. (1995). Supporting operator problem solving through ecological interface design. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-25*, 529–545.